Updating function model: Model updating method transferable in a wider range of data sizes

Sanko, Nobuhiro

(Citation)
Asian Transport Studies,8:100071

(Issue Date)
2022

(Resource Type)
journal article

(Version)
Version of Record

Asian Transport Studies 8 (2022) 100071

Contents lists available at ScienceDirect

Asian Transport Studies

journal homepage: www.sciencedirect.com/journal/asian-transport-studies

Updating function model: Model updating method transferable in a wider range of data sizes

Nobuhiro Sanko

Graduate School of Business Administration, Kobe University, Japan, 2-1 Rokkodai-cho, Nada-ku, Kobe, 657-8501, Japan

A R T I C L E I N F O

Keywords:
Model updating
Temporal transferability
Number of observations
Data collection time points
GDP per capita

A B S T R A C T

When data are available from two time points—older data with a larger number of observations and more recent data with a smaller number of observations—then model updating is utilised to take advantage of the different merits of each data set. However, the author's previous study demonstrated that conventional model updating methods—transfer scaling, joint context estimation, Bayesian updating, and combined transfer estimation—were inferior to models using only the more recent data. The present study examines an updating method that the author calls an 'updating function model' in which the parameters are assumed to follow the functions of gross domestic product per capita. The present study demonstrates that the updating function model often produces statistically significantly better forecasts than models using only the more recent data. The study extends the applicability of the model updating to cases in which the more recent time point has more observations than the older time point.

1. Introduction

The forecasting performance of transport models was studied because they were frequently applied to forecasting. The success of forecasting depends on the transferability of the models. One approach for improving transferability is model updating. Transport planners sometimes use old data sets. Although old data sets can be used for developing models, when the data sets are very old, planners often update their models using newer but smaller data sets in an attempt to combine the merits of large sample sizes from the older data sets with the merits of newer but smaller data sets. Conventional model updating methods include transfer scaling, joint context estimation, Bayesian updating, and combined transfer estimation.

Despite their excellent potential, these conventional model updating methods have failed to achieve their expected performance. When these updating methods were first introduced, the temporal transferability of updated models was rarely evaluated due to the lack of data availability. Studies for temporal transferability need data from at least three time points: one for estimating the models, a second for updating the models, and a third for validating the updated models. Some earlier studies that had limited data availability conducted incomplete evaluations by utilising data from only two time points: the older data for modelling and the more recent data for both updating and evaluation. However, the updated models were, at best, comparable to models utilising only the smaller data set.

Data availability is now less of a problem, and Sanko (2020) clearly showed that none of the conventional updating methods produced statistically significantly better forecasts than models utilising only the smaller data. The study comprehensively examined numerous combinations of observation numbers and data collection time points.

Sanko (2014) proposed a new method that combined data sets from multiple time points even when the data set at each time point was large. This method contradicted the consensus of the research community that only the most recent data set should be used when it is large. The resulting models produced better forecasting than models that utilised only the most recent data set. A key idea behind the method was expressing parameters as functions of variables that change over time. Since the parameters are updated following the function, the present study calls this function an updating function.

The reasoning behind these methods differs. Sanko (2014) used older data sets that had been ignored, even when large data sets existed for the most recent time point. On the other hand, following the consensus of the research community, conventional model updating methods utilise large data sets from the most recent time point, but collect a smaller data set for updating, which turns the large data set from the most recent time point into an older data set. Although the reasoning differs, both methods have the same idea: use data sets from multiple time points to improve forecasting. The updating function models were already

E-mail address: sanko@kobe-u.ac.jp.

https://doi.org/10.1016/j.eastsj.2022.100071
Received 23 March 2021; Received in revised form 25 January 2022; Accepted 24 March 2022
Available online 2 April 2022

formulated in Sanko (2014); the present study is the first study that applies the updating function models to a context where conventional model updating methods have been applied. If the updating function models outperform models utilising the most recent data, then they would be unique in producing better forecasts than the most recent data models. Furthermore, this study aims to extend the applicability of model updating to cases in which the more recent data has a larger number of observations than an older time point to which the model updating methods have not been applied. This follows from Sanko (2014), which found that models utilising older data sets produce better forecasts even when the most recent data sets are large.

This paper is organised as follows. Section 2 presents studies that describe the limitations of conventional updating methods and that propose and apply updating function models. Section 3 describes the data sets. Section 4 presents the research question, describes a multinomial logit model, a more recent data model, an updating function model, the bootstrapping procedure, and hypothesis testing. Section 5 reports the estimated parameters and the forecasting performance of the models using statistical tests and discusses cases in which the proposed updating function models outperformed the more recent data model. Section 6 presents the concluding remarks.

2. Literature review

2.1. Comparing models updated with conventional methods and a more recent data model

When disaggregate choice models emerged as a forecasting method, much attention was given to developing methods for improving forecasting performance. Atherton and Ben-Akiva (1976) proposed transfer scaling, in which only alternative-specific constants and utility scales are updated by the more recent data since no theory supports their transferability. Badoe and Miller (1995) proposed joint context estimation, in which alternative-specific constants and utility scales are specific to each time point while the other parameters are shared across years. Bayesian updating (Atherton and Ben-Akiva, 1976) calculates the weighted averages of parameters from multiple time points, in which variance-covariance matrices are utilised as weights. Combined transfer estimation (Ben-Akiva and Bolduc, 1987) is an extension of Bayesian updating, but explicitly takes transfer bias into account. Transfer scaling and joint context estimation both assume that alternative-specific constants and utility scales are less transferable. In these two studies, however, data utilised for updating does not contribute to the other parameter estimates in transfer scaling but does contribute to the other parameter estimates in joint context estimation in which the other parameters are shared across years. Bayesian updating simply utilises weighted averages and does not consider which parameters are updated by which data set; combined transfer estimation corrected this problem. These updating methods are compared not only to each other but also to models utilising only the more recent data set. Badoe and Miller (1995) and Karasmaa and Pursula (1997) found that when a small number of observations (at least 400–500) is available from the more recent time point, the four conventional updating methods provided little to no improvement in the model's forecasting performance. However, they did not consider statistical tests.[1]

Sanko (2020) is the first study to thoroughly and statistically compare four conventional updating models and a more recent data model in terms of their forecasting performance when the data collection time points and the numbers of observations from each time point differ. Statistical tests were conducted using bootstrapping. The data sets used in Sanko (2020) are identical to those used in the present study, so a direct comparison is possible. The research question in Sanko (2020) was as follows: *Suppose that two data sets are available: a large data set from an older time point and a small data set from a more recent time point. Which conventional updating method or more recent data model produces better forecasting when the data collection time points and the numbers of observations from each time point differ?* Sanko (2020) investigated a total of 234 combinations: three combinations relating to data collection time points multiplied by 78 combinations relating to the numbers of observations from each time point. Sanko (2020) showed that the more recent data models often produced better forecasts than the conventional updating methods with or without statistical significance. On the other hand, while conventional updating methods sometimes produce better forecasting without statistical significance, they never produce better forecasting with statistical significance. Moreover, conventional updating methods are superior only in very limited cases. Transfer scaling and joint context estimation have some value when the number of observations from the more recent time point is extremely small. However, this is not supported by any statistical tests. Bayesian updating and combined transfer estimation have little value. In conclusion, no statistical tests support the use of any conventional updating methods. The findings of Sanko (2020) are consistent with those of Badoe and Miller (1995) and Karasmaa and Pursula (1997).

2.2. Updating function models

Updating function models can efficiently use repeated cross-sectional data. When researchers develop models for forecasting, they use only the most recent data, even when cross-sectional data are available from multiple time points. Ignoring older data is not a good use of data from the viewpoint of efficient data usage. Using older data together with the most recent data would be a great way to improve forecasting performance, since no additional survey costs are incurred.

Questioning one of the assumptions made regarding joint context estimation—one of conventional updating methods—reveals a key concept of the updating function model. The joint context estimation utilises data from two points in time under the assumption that some parameters are identical between the time points. However, one reason for the lower transferability when updating using the joint context estimation method is that parameter changes over time are not modelled. On the other hand, the updating function model assumes that the parameters are expressed as functional forms, which allows the parameters to change over time and allows future parameter values to be predicted. Introducing functions to update parameters is the key concept, so this approach is called the updating function method.

The specifications for the updating functions are not fixed. Researchers must determine which parameter in the model (e.g., mode choice model) follows which functional form of which variable. It is important to have a basic understanding of not only which factors (variables) affect parameter changes but also how the factors (which functional forms of the factors) affect parameter changes. The present study uses not only the same data set but also the same model specifications as Sanko (2014, 2016, 2018), so a direct comparison is possible. The data sets in those previous studies consist of three time points: 1971, 1981, and 1991. In those studies, the updating function model was estimated by jointly utilising the data from all three time points, while the most recent model utilised data from only 1991. The forecasting performances for behaviours in 2001 were compared. When the method was first proposed by Sanko (2014), all the parameters for the mode choice models were assumed to follow the updating functions of time (year). Furthermore, Sanko (2018) assumed that all the parameters follow the updating functions of GDP per capita (constant price). In both studies, five functional forms—linear, square root, square, exponential, and log—were trialled. Sanko (2014) showed that the updating function model of time (year) outperformed the most recent data model, and Sanko (2018) further improved the forecasting performance.

Sanko (2014, 2018) assumed that all the parameters in the mode

[1] Bowman et al. (2013) utilised a joint context estimation and discussed the impact of the number of observations on transferability. However, the context of their analysis was spatial transferability.

choice models follow the same functional form of the same variable. It is reasonable, however, to consider that different parameters might be influenced by different factors in a different functional form. Sanko (2016) investigated 288 models by assuming that the parameters in mode choice models independently follow a time (year) (in linear form), GDP per capita (in linear, square, and square root forms),[2] female social participation (in linear form), and Nagoya City's subway length (in linear form). A notable difference was found between the functions of time and the functions of GDP per capita. The functions of time were too trained on the estimation data sets and produced poor forecasts. The functions of time ascribe parameter changes to the trends of the times without any economic reasons, while the functions of GDP per capita attribute them to economic factors. This demonstrates the importance of understanding the basic factors, such as economic factors, that are behind the parameter changes. Among three functional forms of GDP per capita, little difference can be found in forecasting performance, so the simplest linear form of GDP per capita is a good choice. Functions of female social participation and Nagoya City's subway length differ little in forecasting performance from the linear form of GDP per capita. Therefore, the choice of the function of GDP per capita (in linear form) is not inferior to any of the other forms examined. If just one had to be chosen, the function of GDP per capita (in linear form) would be the best choice. Although this study shows that using simple GDP per capita (in linear form) may improve transferability, the behavioural reasoning remains a topic for future study.

However, Sanko (2014, 2016, 2018) all utilise a large and identical number of observations (10000) from each time point. In this context, conventional updating methods were not applied. Moreover, the results were not verified statistically.

3. Data

Household travel survey data was collected in the metropolitan area of Nagoya (Japan), which is the third largest metropolitan area in Japan and is located between the two largest metropolises of Tokyo and Osaka. This study utilised repeated cross-sectional data from four points in time; data from 1971, 1981, and 1991 are used for modelling, while that from 2001 is used for validation. Respondents were recruited by random sampling. Staff employed for the survey visited the sampled households to hand out the survey form and ask them to complete it, and then visited them again to collect the form. Response ratios were highest in 1971 at more than 90% and lowest in 2001 at more than 75% (Takahashi et al., 2009). The sample size in all four surveys was more than 500,000 trips. No significant changes in the survey methods were reported.

The Nagoya metropolitan area is highly industrialised and includes Toyota City, the hometown of Toyota Motor Corporation. This industrial structure may be the reason for the area's higher level of car usage than in other large cities. Motorisation progressed at a different pace during the study period. For example, after cleaning the data for estimation purposes, the commuting modal shares among rail, bus, and car 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. Meanwhile, Nagoya City has invested in public transportation (Nagoya City, 2021). The first subway line started operation in 1957, and operating kilometres increased from 27.8 km in 1970 to 78.2 km in 2000. Nagoya City bus operating kilometres increased from 493 km to 702 km during the same period. In addition, some private railway companies, including the former Japan National Railways, and bus operators provide service. However, as in other cities, people continue to use cars.

Commuting-to-work modal choice models among the three transportation modes of rail, bus, and car, are estimated. Sanko (2014) provided more detailed information on the data sets, but an issue relating to travel cost is restated. The model does not include travel cost, since commuters usually do not pay the cost; companies usually provide allowances for commuting. The GDP per capita of Japan in constant 2005 million JPY was 1.73 in 1971, 2.39 in 1981, 3.54 in 1991, and 3.75 in 2001 (World Bank, 2013). The most recent data used in the study is from 2001, which is around 20 years ago, but this is not a concern.

4. Methodology

For the sake of simplicity, this study assumes utility maximisation utilising linear-in-parameters multinomial logit models and uses a single model specification throughout the paper.[3] This assumption allows the present study to focus on data collection time points and the numbers of observations. Other dimensions that might affect temporal transferability (Sikder, 2013) but are not considered in the present study include: (i) decision-making rules, (ii) mathematical model structures, and (iii) model specifications. Incorporating these dimensions is a topic for future study. Of course, the method is applicable to other models. This section presents a research question, followed by explanations of multinomial logit models, the more recent data model and updating function model, bootstrapping, and hypothesis testing.

4.1. Research question

This study aims to demonstrate that the author's updating function models outperform the more recent data model in terms of forecasting performance with statistical significance. The research question is stated below.

> Research Question: Suppose that data from two time points is collected, with n_1 representing the number of older observations from year y_1 and n_2 representing the number of more recent observations from year y_2. In which combinations of data collection time points and the numbers of observations from each time point—y_1, y_2, n_1, and n_2—do the updating function models produce statistically significantly better forecasts than the more recent data models? Note that both cases, $n_1 \geq n_2$ and $n_1 < n_2$, are considered.

This study examines not only the case of $n_1 \geq n_2$, where the conventional updating models have been applied, but also the case of $n_1 < n_2$. The latter case has not received any attention from the view of model updating, since the data set from the more recent time point is larger.

4.2. Multinomial logit models

In linear-in-parameters multinomial logit models, the total utility that an individual p has for alternative i, at t ($= t1$ or $t2$; Notations of $t1$ and $t2$ sometimes are utilised to refer to the older and more recent time points, respectively.) is decomposed into the deterministic component of V_{ip}^t, as expressed in Eq. (1), and the error components which follow independent and identical Gumbel distributions.

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

where α_i^t denotes an alternative-specific constant for alternative i at t;

[2] Of the five functional forms examined in Sanko (2014, 2018), only the linear form was examined for time (year) and only linear, square, and square root forms were examined for GDP per capita. The choice of these forms is discussed in Sanko (2016).

[3] The impact of model specifications on transferability is less of a concern. If different specifications were able to improve to the same extent both the more recent data models and the updating function models, then the impact of the model specifications would be cancelled out. If the specifications produced different degrees of improvement, then the model specifications would affect the results of the study. This is a topic for future study.

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. The probability that individual p chooses alternative i among alternative j's in p's choice set at t, P_{ip}^t, is written as Eq. (2), where scale parameter is fixed to a unity value for identification.

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

Eq. (3) expresses the total log-likelihood (L), which consists of the sum of log-likelihood from year t (L^t).

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

where $y_{jp}^t = 1$ if p chose j at t; $= 0$ otherwise.

4.3. More recent data model and updating function model with GDP per capita

4.3.1. More recent data model

The more recent data model follows procedure described in Section 4.2 by applying data from $t = t2$. The log-likelihood function is expressed as $L = L^{t2}$ in Eq. (3). A log-likelihood on 2001 data (L^{2001}) is calculated for evaluating forecasting performance, by applying $\widehat{\boldsymbol{\alpha}}^{t2}$, $\widehat{\boldsymbol{\beta}}^{t2}$, $\mathbf{x}^{2001}$, and $\mathbf{y}^{2001}$ to Eqs. (1)–(3). Note that a caret $\widehat{()}$ indicates an estimate.

4.3.2. Updating function model with GDP per capita

The updating function model with GDP per capita makes changes in Eq. (1) regarding α_i^t and β_{ik}^t. For the reason mentioned in Section 2, a linear form is adopted as shown in Eq. (4).

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

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

α_i and β_{ik} are independent of time and are called 'base parameters'. On the other hand, α_{di} and β_{dik} themselves are independent of time but are called 'historically changing parameters', since α_{di} and β_{dik} multiplied by GDP per capita vary over time. The second terms of the right-hand side of Eq. (4) express some temporal evolution in parameters relating to economic conditions. gdp^t denotes GDP per capita at t (units in 10 million in constant 2005 JPY). Eq. (4) is an updating function since it updates parameters for each t.

The updating function model applies data from $t = t1$ and $t2$ to Eqs. (1), (2) and (4). The log-likelihood function is expressed as $L = L^{t1} + L^{t2}$ in Eq. (3). A log-likelihood on 2001 data (L^{2001}) is calculated for evaluating forecasting performance by applying $\widehat{\boldsymbol{\alpha}}$, $\widehat{\boldsymbol{\beta}}$, $\widehat{\boldsymbol{\alpha}}_{\mathbf{d}}$, $\widehat{\boldsymbol{\beta}}_{\mathbf{d}}$, $\mathbf{x}^{2001}$, $\mathbf{y}^{2001}$, and gdp^{2001} to Eqs. (1)–(4).

The model formulation is identical to that of the model used for estimating the main effects and interactions between explanatory variables and the GDP per capita. When the updating function models specified in Eq. (4) are applied to data sets from two points in time, the parameter values for each time point calculated by the models are identical to those estimated with data set from each time point independently. (This does not generally apply to cases with data sets from three and more time points.) Suppose that the horizontal and vertical axes represent the GDP per capita (constant 2005 JPY) and the estimates, respectively. The estimates for the k-th variable using data from $t1$ and $t2$ independently are drawn at (gdp^{t1}, β_{ik}^{t1}) and (gdp^{t2}, β_{ik}^{t2}), respectively. Parameters in the updating function models are estimated by drawing a straight line connecting these points. Although the resulting parameter values are identical to those estimated in independent models, the main advantage of the proposed model is the ability to predict future parameter values using the updating function.

4.4. Bootstrap

In this study, bootstrapping (Efron and Tibshirani, 1993) is applied in the following manner.

10000 commuting trips were selected at random from each of the three time points of 1971, 1981, and 1991. Selecting the same number of commuting trips avoided any influences on the results that might occur from selecting a different number of observations from each year. In addition, 10000 commuting trips were selected at random from the 2001 data set for evaluation.

Three notations—y, n, and b—are introduced.

- y (= 1971, 1981, and 1991) denotes the year in which the data was collected.
- n (= 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, and 10000) denotes the number of observations.
- b (= 1, 2,…, 1000) denotes a repetition for bootstrapping.

From each y, n observations were selected at random 1000 times, with replacement from 10000 commuting trips that were already selected from each year. (For each of the b-th draw from the same y, the larger n observations contain all the records included in the smaller n observations.) In total, 36000 data sets (= 3 y's × 12 n's × 1000 b's) were generated.

Additional notations are y_1, y_2, n_1, and n_2, where

- y_1 and y_2 denote an older and a more recent time point, respectively, and
- n_1 and n_2 denote the number of observations from y_1 and y_2, respectively.

Therefore, 432 combinations are generated; that is, three (=3 × 2/2 relating to the choices of y_1 and y_2) multiplied by 144 (= 12 × 12 relating to the choices of n_1 and n_2). The study continues with the following procedure, requiring 456000 estimations and forecasts.

- The more recent data models are estimated 24000 times and applied to forecasting behaviours for 2001. The 24000 estimations are the result of multiplying y_2 (2), n_2 (12), and b (1000). (y_2 must be from 1981 or 1991.)
- The updating function models are estimated 432000 times and applied to forecasting behaviours for 2001. The 432000 estimations are the result of multiplying y_1 and y_2 (3), n_1 and n_2 (144), and b (1000).

4.5. Hypothesis testing

The forecasting performances of the updating function model and the more recent data model are compared. The present study evaluates the forecasting performance by utilising log-likelihood values on the 2001 data set when the parameter values are fixed to those estimated by the two models. A choice of log-likelihood on the 2001 is briefly explained. The log-likelihood is a basis for other transferability measures, e.g., TTS (Transferability Test Statistic), TI (Transfer Index), and transfer rho-squared. By definition, models superior in terms of log-likelihood values are always superior in TTS, TI, and transfer rho-squared. Transferability measures from other perspectives include tests of individual parameter equalities and REM (relative error measures). However, the transferability of specific parameters is not the main interest of this study, and REM is not suitable since it is an aggregate measure, and its transferability is good even if two wrong predictions are aggregated.

The reason for conducting a statistical test is explained. If two models provide forecasting with a reasonable level of precision, then both models are acceptable; neither a comparison nor a statistical test is required. However, this study does not examine this context. Transferability studies have received some criticism that the results are not supported by statistical tests. In addition, this study includes cases where the number of observations is small, and there is a huge possibility that results are obtained by chance.

This is a comparison of two non-nested models. Hence, a standard likelihood ratio test cannot be used. Therefore, this study uses the following approach as suggested by Ette (1996).

The null hypothesis (H_0) is described as follows:

H_0. The updating function model and the more recent data model have the same log-likelihood on the 2001 data set.

The differences in the log-likelihood values on the 2001 data set between the two models are considered. The bootstrapping approximates its sample distribution. x_b ($b = 1, 2, \ldots, 1000$) are defined for each combination of y_1, y_2, n_1, and n_2, as shown in Eq. (5).

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

where, $L1\ (\bullet, y_2, \bullet, n_2, b)$ denotes a log-likelihood on the 2001 data set for the more recent data model with the b-th repetition for n_2 from y_2, and $L2\ (y_1, y_2, n_1, n_2, b)$ denotes a log-likelihood on the 2001 data set for the updating function model with the b-th repetition for n_1 and n_2 from y_1 and y_2, respectively. $L1$ refers to models utilising data from *one* point in time, while $L2$ refers to models utilising *two* points in time. In calculating the $L1$ and $L2$, the same b is applied; therefore, the x_b is defined only when both $L1$ and $L2$ are calculated. If the updating function model produces better forecasts, then x_b is more positive.

When $x_{b\ (0.025)}$ and $x_{b\ (0.975)}$ represent the 2.5 and 97.5 percentiles of the x_b, respectively, then a 95 percent confidence interval of the x_b is expressed as $(x_{b\ (0.025)}, x_{b\ (0.975)})$. If $x_{b\ (0.025)} > 0.0$ or $x_{b\ (0.975)} < 0.0$ is satisfied, then the null hypothesis is rejected at the five percent level of significance.

5. Results and discussion

This section presents results in the following order: estimates, forecasting performance, and results of hypothesis testing.

5.1. Estimates

With respect to the choice of explanatory variables, this study examined numerous combinations of variables and reports the best results.[4] Dummy variables are defined as: male (1 for male, 0 otherwise), 20 years old or older (1 if 20 years old or older, 0 otherwise), 65 years old or older (1 if 65 years old or older, 0 otherwise), and Nagoya (1 if origin and/or destination of the trip are in Nagoya City, 0 otherwise). The models do not include car ownership to avoid the potential problem of endogeneity. The models also do not include travel cost for the reason mentioned in Section 3. Table 1 shows the descriptive statistics for the variables included in the models (Sanko, 2014). The shares of those aged 20 years old or older are large, while the shares of those aged 65 years old or older are small.

Estimates for the models independently utilising data sets from 1971, 1981, 1991, and 2001 are presented in Table 2 (Sanko, 2014). The columns of '1981' and '1991' represent estimates of the more recent data model when the more recent time is 1981 and 1991, respectively. Estimates in the '1971' columns are not for the more recent model, but for interpreting parameter values when GDP per capita in 1971 is applied to the 1971/1981 and 1971/1991 updating function models. The same discussion applies to the estimates in the '1981' and '1991' columns, and the values in the columns are identical to the values generated by applying the GDP per capita of the corresponding year to the updating function model when one of data points is the corresponding year. The estimates in the '2001' column are presented as a reference. (The 2001 data set is used solely for validation.) The models presented in Table 2 utilise all the 10000 commuting trips chosen from each year, while the present study utilises the same model specification, but with different subsamples for bootstrapping.

All parameters are estimated with statistical significance having reasonable signs. Travel time estimates are negative. Car alternative has positive estimates for the male dummy and the 20 years old or older dummy, and negative estimates for the Nagoya dummy (core of the metropolis). Bus alternative has positive estimates for the 65 years old or older dummy. Rail alternative has positive estimates for the male dummy. Two male dummies suggest the unattractiveness of buses for males. A more detailed interpretation can be found in Sanko (2014).

Table 3 produced estimates for the updating function models with the base parameters and the historically changing parameters at the top and bottom of the table, respectively. The choice of explanatory variables is the same as that for the more recent data models shown in Table 2. Models shown in Table 3 utilise all the 20000 commuting trips chosen from the two time points, while the present study utilises the same model specification but with different subsamples for bootstrapping.

Base parameter estimates represent parameter values when GDP per capita is zero, which is highly unlikely and so interpreting them is avoided. Interpretation should be given for parameter values when the actual GDP per capita is applied, but they are already presented in Table 2 and interpreted.

Historically changing parameters are interpreted. Bus and car constants are decreasing and increasing with statistical significance, respectively in any combinations of years, suggesting an impact from motorisation in accordance with the increase in GDP per capita. Nagoya dummy for car had negative estimates with statistical significance in all combinations of years, suggesting that the impact of motorisation was mitigated in the central area of the metropolis compared to other parts of the metropolis as the GDP per capita increased. This is probably due to unevenness in road congestion and investment in alternative transport modes in the metropolis. The 20 years old or older dummy, and the 65 years old or older dummy are statistically insignificant, suggesting that the effects of these variables are stable.

Other parameters have different signs and/or levels of statistical significance across combinations of years. Travel time has negative estimates with statistical significance in the 1971/1981 and 1971/1991 models. The disutility of travel time is amplified in accordance with economic development, probably due to time becoming a more valuable resource. However, it turns positive and insignificant in the 1981/1991 model. The decreasing trend in the travel time parameters might have stopped in 1981. Male dummies have positive estimates with statistical significance in 1971/1981. The male dummy for rail is insignificant in the 1971/1991 and 1981/1991 models; the male dummy for car is insignificant in 1971/1991 and negative and statistically significant in 1981/1991. Economic development has impacts on both males and females, but it may first affect males and then females. Female participation in an employment market is more apparent in the latter half of the study period, and gender differences that had emerged in the first half of the study period might have stopped in 1981 and possibly reversed direction in the second half of the study period. Sanko (2018) provided a more detailed explanation.

The rows labelled 'Log-likelihood on 2001 data' in Tables 2 and 3 present the forecasting performances of the models for the 2001 data set. The best forecasting performance was achieved by the 1971/1991 model, followed by the 1981/1991, 1991, 1971/1981, 1981, and 1971 models. When the more recent data set is from 1991, the updating function models improve the forecasting performance. The same applies to cases in which the more recent data set is from 1981.

[4] The model was developed solely for this study and seemed to have fewer variables. However, models estimated for developing transport policies by utilising household travel surveys for the Tokyo Metropolitan Area (MLIT, 2015) include not the same set of parameters but a similar number of parameters. Note that the Tokyo and Nagoya surveys have similar formats.

Table 1
Descriptive statistics.

Variables	1971		1981		1991		2001	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Rail travel time [hr][a]	0.855	0.303	0.857	0.296	0.843	0.278	0.803	0.273
Bus travel time [hr][a]	0.672	0.246	0.666	0.219	0.650	0.240	0.633	0.218
Car travel time [hr][a]	0.444	0.206	0.434	0.216	0.456	0.219	0.436	0.217
Male dummy	0.735	0.441	0.730	0.444	0.681	0.466	0.654	0.476
20 years old or older dummy	0.941	0.235	0.962	0.191	0.971	0.167	0.988	0.108
65 years old or older dummy	0.015	0.122	0.016	0.127	0.021	0.145	0.031	0.173
Nagoya dummy	0.585	0.493	0.487	0.500	0.466	0.499	0.448	0.497

[a] The author selected respondents who chose rail, and calculates mean and standard deviations of their rail travel times. The same applies to bus and car travel times.

Table 2
Estimates of more recent data models.

Variables	1971		1981		1991		2001[a]	
	Est.	*t*-stat.	Est.	*t*-stat.	Est.	*t*-stat.	Est.	*t*-stat.
Constant (B)	0.127	2.42	−0.392	−6.21	−0.638	−8.98	−1.03	−12.11
Constant (C)	−1.15	−9.84	−0.645	−4.65	0.301	1.96	0.560	2.23
Travel time [hr]	−0.606	−6.94	−1.81	−16.47	−1.59	−15.71	−2.60	−20.48
Male dummy (R)	0.577	8.59	0.787	8.70	0.812	7.53	0.511	3.89
Male dummy (C)	1.97	29.44	2.17	25.22	1.78	17.30	1.38	10.91
20 years old or older dummy (C)	0.900	8.28	0.764	5.78	0.776	5.18	0.511	2.06
65 years old or older dummy (B)	1.91	8.89	1.37	5.73	1.33	5.59	0.561	2.05
Nagoya dummy (C)	−1.12	−24.08	−1.77	−33.21	−2.18	−37.81	−2.21	−36.70
N (randomly drawn)	10000		10000		10000		10000	
L(**β**)	−7776.86		−5985.02		−5300.58		−4716.28	
L(**0**)	−8948.26		−8593.88		−8398.85		−8159.63	
Adj rho-squared	0.130		0.303		0.368		0.421	
Log-likelihood on 2001 data	−6521.95		−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 the purpose of comparison.

5.2. Forecasting performance

The nine panels in Fig. 1 present the characteristics of the forecasting performance. Notations (a)–(c) express the data collection time points: (a) 1971/1981, (b) 1971/1991, and (c) 1981/1991. Notations (1)–(3) describe characteristics. Panels (1)'s and (2)'s present averages and standard deviations for forecasting performances $L1$ and $L2$, respectively. When the sample sizes were small, out of 1000 trials, some cases were not estimable, and hence excluded before calculating the averages and standard deviations. (3)'s refers to the number of remaining repetitions after excluding the above cases. Large shares for 20 years old or older, small shares for 65 years old or older (see Section 5.1), and small bus shares (see Section 3) are the main reasons for this. For example, if the bootstrapped samples do not include bus users, then the models with two alternative-specific constants are theoretically unidentifiable. In each panel, the number of observations from the more recent time point (n_2) is plotted on the horizontal axis. The more recent model results are drawn in all panels. Each panel has thirteen lines: 12 *solid* lines for 12 n_1's for the updating function models and one *dashed* line for the more recent data models. The results of panels (a)'s, (b)'s, and (c)'s are interpreted below.

First, averages of the forecasting performance are interpreted in panels (a1), (b1), and (c1). All 13 lines increase to the right. When n_1 is fixed, the larger the more recent sample size, the better the forecast. The lines for the updating function models sometimes appear above those for the more recent data model, meaning that the updating function method is superior to the more recent data model. This indicates that parameter values extrapolated by updating function model are more transferable to the future context. On the other hand, the lines for the larger n_1 generally appear above those for the smaller n_1. When n_2 is fixed, the larger the older sample size, the better the forecasting for the updating function models. Thus, the updating function model benefits from largeness in sample size from both time points. This can be compared to the joint context estimation examined in Sanko (2020), which found that the larger the more recent sample size, the better the joint context estimation. However, this is not necessarily true for the older time point. While larger sample sizes from older time point are preferable, the joint context estimation imposes a constraint: parameter equalities between time points. Parameters estimated with older data usually are less transferable. Therefore, a larger sample size from the older time point, together with the constraint, reduces the transferability of the parameters of the joint context estimation. Joint context estimations do benefit from a larger sample size from the older time point when the sample size from the more recent time point is smaller, but not when it is larger.

Second, standard deviations of the forecasting performance are interpreted in panels (a2), (b2), and (c2). All 13 lines generally decrease to the right. For fixed n_1, the larger the more recent sample size, the smaller the standard deviations of the forecasts. In panel (a2), the lines for the more recent data models appear below those for the updating function models, implying that the standard deviation of the more recent data model is smaller. However, the relationship is not clear-cut in panels (b2) and (c2). In general, however, the lines for larger n_1 generally appear below those for smaller n_1. For fixed n_2, the larger the older sample size, the smaller the variance for the updating function model.

Third, the numbers of remaining repetitions are interpreted in panels (a3), (b3), and (c3). All 13 lines generally increase to the right. When n_1 is fixed, the larger the more recent sample size, the fewer the number of poor estimations. The lines for the more recent data models appear above those for the updating function models, implying that the more recent data models have fewer numbers of poor estimations. On the other hand, the lines for larger n_1 generally appear above those for smaller n_1. This is interpreted from model specification perspective. The models (see Table 2) include two alternative-specific constants among

Table 3
Estimates of updating function models.

Variables	1971/1981		1971/1991		1981/1991	
	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. The number of observations is the sum of 10000 from each of the two time points.

three alternatives. Therefore, the data set must contain at least one observation from each alternative. A similar constraint applies to dummy variables.[5] However, in the case of a small number of observations, the data sets often fail to satisfy this requirement.

To sum up, the updating function model benefits from larger data set sizes not only in the older time point but also in the more recent time point. Therefore, this method makes use of both data sets. Updating function models sometimes have advantages in terms of average forecasting performance, but this comes at the expense of the remaining numbers of repetitions.

5.3. Results of hypothesis testing

Fig. 2 presents three separate panels for different combinations of older and more recent time points: 1971/1981 in panel (a), 1971/1991 in panel (b), and 1981/1991 in panel (c). Note that the number of b's utilised here differs from that utilised in panels (a3), (b3), and (c3) of Fig. 1. While Fig. 1 excludes b's when each model produced poor estimates, Fig. 2 excludes b's when at least one of two models produced poor estimates.

The number of observations is plotted in each panel: n_1 from the older time point on the horizontal axis and n_2 from the more recent time point on the vertical axis. The level of shading in the cells expresses the test results. The black and dark grey cells indicate that the updating function model (black cells) and the more recent data model (dark grey cells) produced better forecasting than the other model at a five percent level of significance ($x_{b\ (0.025)} > 0.0$ for the black cells; $(-x_b)_{\ (0.025)} > 0.0$ for the dark grey cells; Percentiles for $(-x_b)$ are defined similarly for x_b.) The light grey and white cells indicate that the updating function model (light grey cells) and the more recent data model (white cells) more frequently produced better forecasting than the other model ('$x_{b\ (0.025)} < 0.0 < x_{b\ (0.5)}$' for the light grey cells; '$(-x_b)_{\ (0.025)} < 0.0 < (-x_b)_{\ (0.5)}$' for the white cells) without a five percent level of statistical significance. Diagonal cells from the lower left to the upper right of the panels, where dashed lines are passing through, indicate that the two time points have the same number of observations ($n_1 = n_2$).

First, when the two time points have the same number of observations ($n_1 = n_2$) and $n_1 = n_2 = 10000$ or 2000, then the updating function models produce better forecasts with statistical significance. Second, in some cells where $n_1 > n_2$ (parts below the diagonal cells), the updating function models produce better forecasts with statistical significance. Although this is an area of interest for conventional model updating methods, Sanko (2020) demonstrated that the four conventional model updating methods never produced statistically significantly better forecasts than the more recent data model. Therefore, the proposed updating function model is unique. Third, $n_1 < n_2$ (parts above the diagonal cells) is not of interest in conventional model updating methods. However, the updating function models produced statistically significantly better forecasts in some cells in panels (b) and (c). Therefore, the updating function models have another unique feature that extends the possibilities for model updating. Note that the more recent data model produces statistically significantly better forecasting than the updating function model when $n_1 = 100$ and $n_2 = 200$ in panel (a). This is not surprising, however, since $n_1 < n_2$ is a case where the more recent data model has been considered to be more appropriate.

Panels (a) and (c), in which the time points are 10 years apart, and panel (b), in which then time points are 20 years apart, are compared. Panel (b) has more cells shaded in black and light grey, implying that the updating function models contributes more when the data come from a wider period. An intuitive interpretation is that historically changing parameters are estimated more reliably when the domain of GDP per capita is wider.

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 (2018) described a sensitivity analysis in which 10000 observations were taken from each of three time points—1971, 1981, and 1991—and found that the proposed model was practical. The present study does not repeat the analysis, but the same analysis is applicable.

In closing the analysis, to provide a rough estimate of predictive accuracy in terms of market share, the absolute predictive error (*abserr*) is defined in Eq. (6) for mode i (= rail, bus, and car) by using model M (•, •, •, •, b). The M (•, •, •, •, b) is either M (•, y_2, •, n_2, b) for the more recent data model or M (y_1, y_2, n_1, n_2, b) for the updating function model.

$$abserr\ (i, M\ (\bullet, \bullet, \bullet, \bullet, b)) = |S\ (i, M\ (\bullet, \bullet, \bullet, \bullet, b)) - S\ (i)| \quad (6)$$

where, S (i, M (•, •, •, •, b)) denotes predictive share (%) in the 2001 data set for mode i by using M (•, •, •, •, b) and S (i) denotes actual share (%) in the 2001 data set for mode i.

The averages of *abserr* are calculated and presented in Table 4. Note that the number of repetitions used for the calculation is the same as that used in Fig. 1. Presented in Table 4 are four extreme cases, corresponding to the four corners in each panel of Fig. 2: (n_1, n_2) are *i* (100, 100) in the lower-left corner, *ii* (10000, 100) in the lower-right corner, *iii* (100, 10000) in the upper-left corner, and *iv* (10000, 10000) in the upper-right corner. Notations (a)–(c) in Table 4 correspond to panels

[5] The two requirements of (a) alternative-specific constants and (b) dummy variables apply differently to the two methods. For the more recent data model, the more recent data set must satisfy both requirements. For the updating function model, the older and more recent data sets independently satisfy both requirements.

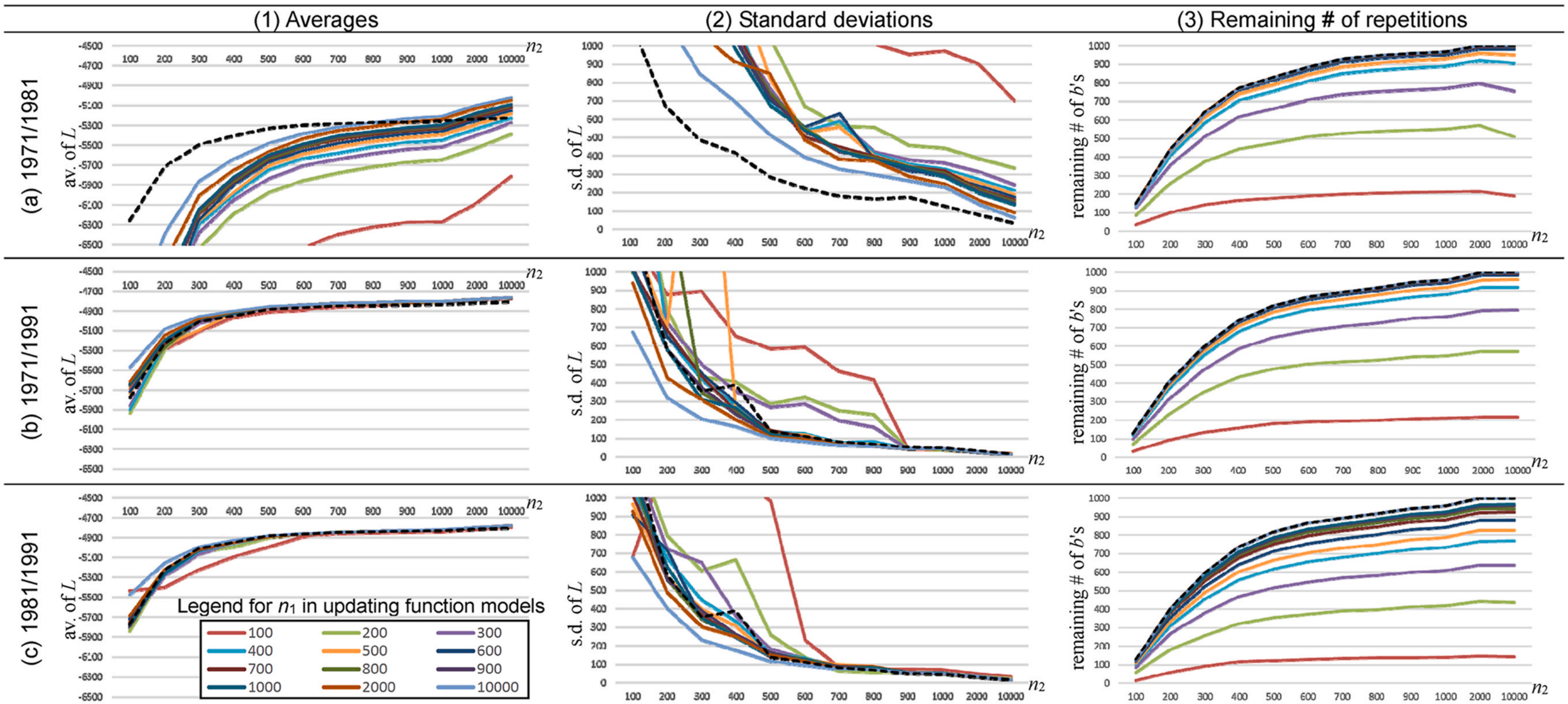
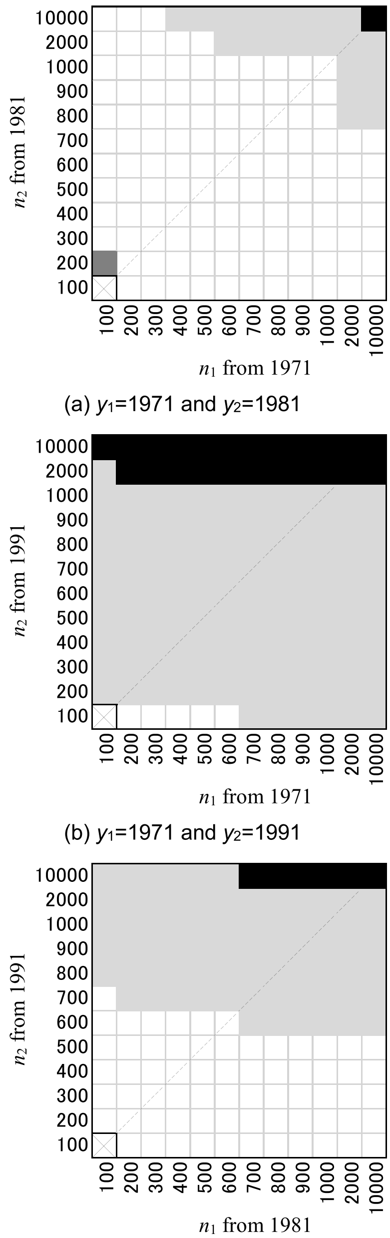


Fig. 1. Characteristics of forecasting performance.

Note: The horizontal axis, representing the number of observations from the more recent time point (n_2), does not have equal intervals. Solid lines indicate updating function models, and the number of observations from the older time point (n_1) is provided in the figure legend. Dashed lines indicate the more recent data models. The dashed lines are identical in (b1) and (c1), (b2) and (c2), and (b3) and (c3).



(a) y_1=1971 and y_2=1981

(b) y_1=1971 and y_2=1991

(c) y_1=1981 and y_2=1991

Fig. 2. Statistical tests.
Note: The black and dark grey cells indicate that the updating function model (black cells) and the more recent data model (dark grey cells) produced better forecasting than the other model at a five percent level of significance ($x_{b\ (0.025)} > 0.0$ for the black cells; $(-x_b)_{(0.025)} > 0.0$ for the dark grey cells). The light grey and white cells indicate that the updating function model (light grey cells) and the more recent data model (white cells) more frequently produced better forecasting than the other model ('$x_{b\ (0.025)} < 0.0 < x_{b\ (0.5)}$' for the light grey cells; '$(-x_b)_{(0.025)} < 0.0 < (-x_b)_{(0.5)}$' for the white cells) without a five percent level of statistical significance. Cells crossed out indicate that less than 40 x_b's were successfully calculated and are unsuitable to produce 2.5 percentiles. Therefore, they are excluded from the analysis. The cells, where the dashed lines are passing through, mean $n_1 = n_2$.

(a)–(c) in Fig. 2, and notations *i*–*iv*, which are defined here, also are used there. The absolute predictive errors by the more recent data model also are presented: (n_1, n_2) are *I&II* (•, 100) for comparison with *i* and *ii* and *III&IV* (•, 10000) for comparison with *iii* and *iv*. A significant improvement was observed in case (a); for example, *iv* has a smaller error than *III&IV* in car share by about 8.1 percentage points.

6. Conclusions

This study is the first application of updating function models proposed by Sanko (2014) to the context of model updating. The reasoning behind the updating function models and conventional model updating methods differ. While the former utilises older data for updating even when the data from the most recent time is large, the latter utilises newly collected small data for updating. Although the reasons differ, both models share the idea of utilising data from multiple time points to improve forecasting performance. Therefore, the present study considers the updating function models to be a model updating technique and compares their forecasting performance with that of the more recent data model. If the proposed models are superior to the more recent data model, then the updating function model is novel, since conventional updating methods never produce better forecasting than the more recent data models. In addition, the present study extends the applicability of the model updating to cases where the older data set is smaller in size than the more recent data set. This follows from the author's finding that updating function models were successful even when the data set from each time point is large.

Based on the author's previous studies, the updating functions adopted in this study are functions of GDP per capita (in linear form) for all the parameters. The present study examined commuting-to-work mode choice behaviours in Nagoya, Japan. A bootstrapping technique was applied to determine for which combinations of data collection time points (two of three data collection time points (1971, 1981, and 1991)) and numbers of observations from each time point (12 different numbers of observations ranging from 100 to 10000) the updating function models produce statistically significantly better forecasts than the more recent data models.

Although not tested statistically (see Fig. 1), by applying the bootstrapping technique, the updating function models produced forecasts with the following characteristics.

- When the sample size from the older time point is fixed (i.e., 100, 200,…, 10000), then the larger the sample size from the more recent time point, the better the forecasting.
- When the sample size from the more recent time point is fixed (i.e., 100, 200,…, 10000), then the larger the sample size from the older time point, the better the forecasting.
- The updating function model benefits from a large sample size from both time points. This differs from joint context estimation, which benefits from a large sample size from the more recent time point but not necessarily from the older time point.
- For some combinations of years and the numbers of observations, the updating function models produce better forecasting (measured as average log-likelihood values in Fig. 1) than models that use more recent data.

Statistical tests produced the following findings (see Fig. 2).

- When the number of observations from the more recent time point is equal to or smaller than that from the older time point, for some combinations of years and the numbers of observations, the updating function models produced statistically significantly better forecasts than the more recent data models. The four conventional updating methods—transfer scaling, Bayesian updating, joint context estimation, and combined transfer estimation—never produced statistically significantly better forecasts than the more recent data models (Sanko, 2020). Therefore, the updating function model is a unique model updating method.
- When the number of observations from the more recent time point is larger than that from the older time point (which is not of interest for conventional model updating), then for some combinations of years and the numbers of observations, the updating function models produced statistically significantly better forecasts than the more

Table 4
Absolute predictive errors (%).

	(a) 1971/1981			(b) 1971/1991			(c) 1981/1991		
	Rail	Bus	Car	Rail	Bus	Car	Rail	Bus	Car
I&II. Av. of *abserr* (*i*, *M* (•, •, •, 100, *b*))	6.1	4.8	10.7	3.2	2.5	4.1	3.2	2.5	4.1
i. Av. of *abserr* (*i*, *M* (•, •, 100, 100, *b*))	9.8	2.1	10.0	3.3	1.9	3.8	3.2	3.1	3.3
ii. Av. of *abserr* (*i*, *M* (•, •, 10000, 100, *b*))	6.6	2.3	7.2	3.3	2.1	3.9	3.5	2.4	4.1
III&IV. Av. of *abserr* (*i*, *M* (•, •, •, 10000, *b*))	5.7	4.1	9.8	0.8	1.4	2.3	0.8	1.4	2.3
iii. Av. of *abserr* (*i*, *M* (•, •, 100, 10000, *b*))	5.6	1.4	6.0	0.5	0.8	1.1	0.6	1.0	1.1
iv. Av. of *abserr* (*i*, *M* (•, •, 10000, 10000, *b*))	2.6	1.0	1.7	0.3	0.9	1.1	0.3	1.0	1.1

recent data models. This means that the updating function models extend the possibilities for model updating.

- The updating function models provide better forecasting performance when the data come from a wider period.

This is the first study to compare the forecasting performances of the updating function model and the more recent data model. Since a good use of data from multiple time points is a striking advantage of the updating function model, two interests of the study were the data collection time points and the number of observations. Other dimensions were less of a concern, and the present study assumes utility maximisation utilising linear-in-parameters multinomial logit models and applies a single model specification throughout the paper. Extending the analysis by relaxing these assumptions is a topic for future study. Including an updating function model utilising data from more than two time points is another direction of extension, where trajectory of parameter changes can be modelled. The most recent data in this study are from 2001, which means the data are quite old. While this is not a concern, a topic for future study is the applicability of the methodology to more recent data that contain unique characteristics, such as ‘peak-car’. If the approach proposed in the study—the use of older data without the additional cost of surveying—proves to be successful in practice, then forecasting performance can be improved in this era of budget cuts. Further investigation into behavioural theory, under which the updating function is defined, is one direction for improvement.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers 25380564, 16K03931, and 19H01538. 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. This paper is based on presentations at the 96th Annual Meeting of the Transportation Research Board in Washington, D.C., U.S.A., in January 2017 and the 53rd Infrastructure Planning Conference of the Japan Society of Civil Engineers, Sapporo, Japan, in May 2016. Comments and suggestions from three anonymous reviewers were useful in improving the paper.

References

Atherton, T.J., Ben-Akiva, M.E., 1976. Transferability and updating of disaggregate travel demand models. Transport. Res. Rec. 610, 12–18.
Badoe, D.A., Miller, E.J., 1995. Comparison of alternative methods for updating disaggregate logit mode choice models. Transport. Res. Rec. 1493, 90–100.
Ben-Akiva, M., Bolduc, D., 1987. Approaches to model transferability and updating: the combined transfer estimator. Transport. Res. Rec. 1139, 1–7.
Bowman, J.L., Bradley, M., Castiglione, J., 2013. Making Advanced Travel Forecasting Models Affordable through Model Transferability. Prepared for Federal Highway Administration.
Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York.
Ette, E.I., 1996. Comparing non-hierarchical models: application to non-linear mixed effects modeling. Comput. Biol. Med. 26 (6), 505–512.
Karasmaa, N., Pursula, M., 1997. Empirical studies of transferability of Helsinki metropolitan area travel forecasting models. Transport. Res. Rec. 1607, 38–44.
MLIT (Ministry of Land, Infrastructure, Transport and Tourism). https://www.mlit.go.jp/common/001091184.pdf, 2015–. (Accessed 17 September 2019).
Nagoya City, Bureau of Transportation, 2021. City Bus and Subway (in Japanese). https://www.kotsu.city.nagoya.jp/jp/pc/ABOUT/TRP0004327/37-0364%E5%B8%82%E3%83%90%E3%82%B9%E5%9C%B0%E4%B8%8B%E9%89%84%E5%86%8A%E5%AD%90.pdf. (Accessed 22 December 2021).
Sanko, N., 2014. Travel demand forecasts improved by using cross-sectional data from multiple time points. Transportation 41 (4), 673–695.
Sanko, N., 2016. Factors affecting temporal changes in mode choice model parameters. Transport. Plann. Technol. 39 (7), 641–652.
Sanko, N., 2018. Travel demand forecasts improved by using cross-sectional data from multiple time points: enhancing their quality by linkage to gross domestic product. Transportation 45 (3), 905–918.
Sanko, N., 2020. Criteria for selecting model updating methods for better temporal transferability. Transportmetrica A: Transport. Sci. 16 (3), 1310–1332.
Sikder, S., 2013. Spatial Transferability of Activity-Based Travel Forecasting Models. Ph. D. Dissertation. University of South Florida, Tampa, U.S.A.
Takahashi, K., Hirami, K., Morio, J., Nishino, H., 2009. Non-response in person-trip survey in Japan (in Japanese). Proceedings of Infrastructure Planning. Japan Society for Civil Engineers, 39. No. 283, CD-ROM.
World Bank. http://data.worldbank.org/country/japan, 2013–. (Accessed 17 April 2013).