

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 6 日現在

機関番号：14501

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25380564

研究課題名(和文)パラメータの時点変化を考慮した長期交通需要予測モデルの構築

研究課題名(英文)Long-term travel demand forecasting models considering parameter changes over time

研究代表者

三古 展弘(SANKO, Nobuhiro)

神戸大学・経営学研究科・准教授

研究者番号：00403220

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：複数時点の断面データが得られている場合には、直近の1時点のデータのみを用いる従来の方法よりも、複数時点のデータを用いてパラメータを1人当たり実質国内総生産の関数としたほうが、予測精度が高い。10年違いの2時点のデータが得られていた場合、新しい時点のデータに加えて古い時点のデータを用いる意義は極めて小さい。古い時点の10000サンプルよりも新しい時点の数百サンプルを用いて構築したモデルのほうが有意に良い予測を行える。また、古い時点と新しい時点のデータを両方利用するモデル更新法は、新しい時点のデータのみを用いる場合よりも有意に良い予測精度を示すことはなかった。

研究成果の概要(英文)：When cross-sectional data from multiple time points are available for modelling, models estimated with data from multiple time points, where parameters are expressed as functions of constant gross domestic product per capita, produced better forecasting performance than conventional models estimated with data from only the most recent time point. When data are available from two points in time with ten year interval, a use of older data in addition to more recent data contributes little to improve forecasting performance. Models estimated with several hundred observations from the more recent time point produced statistically significantly better forecasts than models estimated with 10000 observations from the older time point. Model updating methods, which utilise data from both the older and more recent time points, never produced statistically significantly better forecasts than models utilising data from only the more recent time point.

研究分野：社会科学

キーワード：消費者行動 交通行動 交通手段選択 需要予測 移転可能性 断面調査 サンプル数 モデル更新

1. 研究開始当初の背景

交通行動分析の目的の1つに将来の交通需要を予測することがあるが、そのためのモデルは高い移転性を持つ必要がある。「移転性を持つ」(「予測精度が高い」と同義)とは、ある地点や時点で構築されたモデルが別の地点や時点でも利用可能であるということであり、将来予測の文脈で言えば、過去の時点のデータで構築したモデルが将来の時点でも十分に使用できるということである。交通手段選択や自動車保有などの交通行動を個人や世帯単位で分析するのに適した非集計モデルでも、そのモデルの開発された1970～80年代においては移転性に関する多くの研究成果が発表されてきた。しかし、当時の移転性の研究はモデル構築年と予測対象年が数年程度という短期予測の問題に限定されていた。ところが、交通プロジェクトは構想から完成までに数十年もかかることやその耐用年数の長さを考えると、長期予測における移転性の研究も必要である。

長期の移転性に関する研究は、1990年代にカナダのトロント大学が1964年と1986年のデータを用いて1991年の予測を行ったものなどがあるが、その後は移転性に関する研究自体あまり行われなくなってきた。しかし、最近になって長期予測を対象とした移転性分析が注目を集めている。長期にわたって得られたデータを分析可能な形で整備するのは必ずしも容易ではないことから、筆者が日本の1971年～2001年のデータを用いて分析することは、世界の交通行動分析研究の中でも大きな意味を持つ。

モデルを予測に用いる場合、パラメータが時点間で安定していることを仮定している。しかし、長期予測においては、パラメータが変化していることが考えられるため、そのことを考慮できるモデルであることが望ましい。本研究課題においては複数時点の断面データを用いることでこの問題に取り組むことにする。

現在、多くの都市圏においては断面データが複数時点において収集されているが、需要予測においては直近の1時点のデータのみが用いられることが多い。これは貴重な過去のデータを無駄にしている。しかし、筆者は複数時点の断面データを同時に用いる方法を提案し、直近の1時点のデータのみを用いるよりも高い予測精度が得られることを示した。筆者の提案した方法では、パラメータが分析者の設定した関数に従って経年変化することを許容している。そのため、将来時点のパラメータを予測することができ、そのために将来の需要予測がより正確になるという構造になっている。

これまでの筆者の研究ではパラメータを西暦年の関数で表現していた。西暦年の関数の場合、例えば高度経済成長期の10年間のパラメータの変化とバブル経済崩壊後の10

年間のパラメータ変化が同等であると仮定している。しかし、社会は常に同じスピードで発展するわけではなく、この仮定は適切ではない可能性がある。もし、交通行動が経済状況とも密接に関係しているのであれば、経済状況を表現する1人当たり実質国内総生産(GDP)の関数によって表現するほうが適切な可能性がある。実際、自動車保有や利用は経済状況と密接な関係があると考えられている。

2. 研究の目的

(1) パラメータの時点変化を考慮したモデル

1971, 1981, 1991年のデータがモデル構築に利用可能と仮定し、2001年の行動を予測する場合を想定する。このとき、直近の1991年のデータのみを用いるのが従来の方であった。しかし、筆者の研究によって、3時点のデータを同時に用いて、パラメータの時点変化を考慮するほうが、直近の1991年のデータのみを用いるよりも良い予測を行うことができた。これまでの筆者の研究ではパラメータが西暦年の関数によって変化することを仮定していたが、1人当たり実質GDPの関数によって表現するほうが適切な可能性があるため、それを実証する。

(2) データの新しさとサンプル数のトレードオフ

複数時点の断面データが利用可能である場合、新しい時点のデータを用いるほうが良い需要予測を行えることが知られている。このことを実証した研究は多くあるが、それらの研究では各時点において同程度のサンプル数が得られていることを仮定している。しかし、この仮定の成り立たないことも多い。

ここでは、古い時点からは多数のデータが得られているが新しい時点からは少数のデータしか得られていない場合の予測精度を検討する。これは実務的には次の問いに答えるものである。

- 手元に古い時点の多数のサンプルと新しい時点の少数のサンプルがある場合、どちらのデータを用いて需要予測をするのが良いか。
- 手元にあるデータは古いので新しく調査を行う場合、新しい調査のサンプル数をいくらにすれば、古いデータを使う場合よりも良い需要予測を行えるか。

(3) モデル更新法の選択基準

需要予測において、古い時点の多数のサンプルと新しい時点の少数のサンプルが利用可能な場合、2時点のサンプルを共に用いるモデル更新法を利用するか、新しい時点のサ

ンプルのみを利用するかが問題となる。しかし、2 時点のサンプル数がどのような場合に、どの方法を採用するべきかの基準は明確ではない。

本研究では、2 時点のサンプル数がどのような場合に、どのモデル更新法あるいは新しい時点のデータのみを用いるのが適切に判断基準を提供することを目的とする。

3. 研究の方法

以下の分析は、出勤時における交通手段選択行動を対象としている。1971, 1981, 1991 年のデータが利用可能と仮定し、2001 年の行動を予測する。

(1) パラメータの時点変化を考慮したモデル

図 1 に示すように、 y_1 , y_2 の 3 種類のモデルを構築し、その予測精度を比較する。

直近の 1 時点である 1991 年のデータのみを使用

1971, 1981, 1991 年のデータを用いてパラメータを西暦年の関数で表現

1971, 1981, 1991 年のデータを用いてパラメータを 1 人当たり実質 GDP の関数で表現

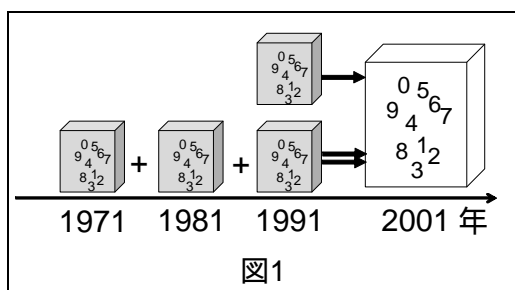


図 1

また、手段選択モデルのパラメータは定数項を含めて 8 つ推定されているが、それぞれのパラメータが異なる要因の異なる関数形によって表現される可能性もある。そこで、8 つのパラメータに関して西暦年（線形）、1 人当たり実質 GDP（線形、2 乗、1/2 乗）、女性の社会参加（線形）、地下鉄路線長（線形）の関数の様々な場合を検討した。

(2) データの新しさとサンプル数のトレードオフ

まず、データ収集時点（1971, 1981, 1991 年の 3 時点）とサンプル数（50 ~ 10000 サンプルの範囲の 33 通り）の 99 の組み合わせの場合についてモデルを構築し、2001 年の行動を予測する。データ収集時点とサンプル数によって予測精度にどの程度の違いがあるか比較する。ここで、各時点からサンプルをランダムに抽出するが、1 回の抽出では偶然性

に左右される。そのため、99 の組み合わせのそれぞれについて 200 回無作為抽出し（ブートストラップ法）、 $99 \times 200 = 19800$ のモデルを構築し、それらを用いて 2001 年の行動を予測した。

次に、図 2 に示すように、以下の y_1 と y_2 のモデルの予測精度を比較する。

古い時点 y_1 の n_1 サンプルを用いたモデル
新しい時点 y_2 の n_2 サンプルを用いたモデル

上で推定した 19800 通りのモデルの中から、 $y_1 < y_2$ （3 時点から 2 時点を選ぶ 3 通り）かつ $n_1 \geq n_2$ （ $33 \times 34 / 2 = 561$ 通り）となる組み合わせについて 200 回の計算結果に基づいて、どちらが統計的に有意に良い予測ができるかを検定した。

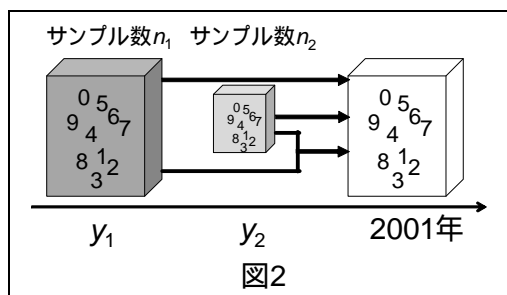


図 2

(3) モデル更新法の選択基準

図 2 の記号を使うと、ここでは y_1 と y_2 を比較する。

新しい時点 y_2 の n_2 サンプルを用いたモデル

古い時点 y_1 の n_1 サンプルと新しい時点 y_2 の n_2 サンプルを用いたモデル（ y_1 の方法として、Transfer scaling, Bayesian updating, Joint context estimation, Combined transfer estimation という 4 種類のモデル更新法を検討した）

ここでも、 $y_1 < y_2$ かつ $n_1 \geq n_2$ となる組み合わせについて 200 回の計算結果に基づいて、どちらが統計的に有意に良い予測ができるかを検定した。なお、計算時間の短縮のため、 n_1 と n_2 については 100 ~ 10000 サンプルの範囲の 12 通りを考慮した。つまり、 $y_1 < y_2$ となる 3 通り、 $n_1 \geq n_2$ となる $12 \times 13 / 2 = 78$ 通りを考えた。

4. 研究成果

(1) パラメータの時点変化を考慮したモデル

1971, 1981, 1991 年の 3 時点のデータを用いてパラメータを 1 人当たり実質 GDP の関数で表現したモデルのほうが、同じ 3 時点のデータを用いてパラメータを西暦年の関数で表現したモデルよりも、良い予測を行うことができた。また、これらのモデルはいずれも直近の 1991 年のデータのみを用いたモデル

ルよりも予測精度が高かった。

つまり、パラメータの変化は年によって一定と仮定するのではなく、経済状況に従って変化すると仮定したほうがよりよい需要予測を行えることを示唆している。

また、手段選択モデルのパラメータに異なる要因の異なる関数形を仮定した結果によると、西暦年の関数は Over-fitting の可能性が指摘された。また、1人当たり実質 GDP の関数については、線形、2乗、1/2乗の間で予測精度に大差はなかった。1人当たり実質 GDP と女性の社会参加や地下鉄路線長の間にも予測精度に大差はなかった。

(2) データの新しさとサンプル数のトレードオフ

得られた知見は次の通り。

- サンプル数が同じなら、新しいデータを用いるほうが、予測精度が平均して高く、予測のばらつきは小さい。
- データ収集時点が同じなら、サンプル数が大きいほど、予測精度が平均して高く、予測のばらつきは小さい。
- 新しい時点の少数のサンプルによるモデルのほうが古い時点の多数のサンプルによるモデルよりも統計的に有意に良い予測を行える場合がある。例えば、1971年のサンプル数が700から10000サンプルの範囲にあるとき（この範囲でならサンプル数がいくらであっても）、1981年の400サンプルのモデル、1991年の250サンプルのモデルのほうが、有意に予測精度が高かった。一方、古い時点の多数のサンプルによるモデルのほうが新しい時点の少数のサンプルによるモデルよりも統計的に有意に良い予測を行える場合はいかなるデータ年とサンプル数の組み合わせにおいても存在しなかった。

(3) モデル更新法の選択基準

2時点のサンプルを用いる4種類のモデル更新法は、新しい時点のサンプルのみを用いるモデルよりも統計的に有意に良い予測を行える場合は全くなかった。しかし、統計的に有意ではないものの、最も良い予測を与えるモデルは次の通り。

- 新しい時点のサンプル数が多い場合は新しい時点のサンプルのみのモデル
- 新しい時点のサンプル数が少ない場合、2時点の文脈差（有・無）と古い時点のサンプル数（多・少）の組み合わせによって次の通り
 - 有・多：Joint context estimation
 - 有・少：Combined transfer estimation
 - 無・多：Transfer scaling
 - 無・少：Joint context estimation

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕(計 1件)

三古 展弘, 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, pp. 1-34, 2015, 査読無

〔学会発表〕(計 8件)

三古 展弘, Criteria for selecting model updating methods for better temporal transferability, The 95th Annual Meeting of the Transportation Research Board, 2016.1.13, ワシントン DC (米国)

三古 展弘, 時間移転性向上のためのモデル更新法の選択基準, 第52回土木計画学研究発表会, 2015.11.21, 秋田大学 (秋田県)

三古 展弘, Should small samples from recent time point be used with older data? applicability of updating models by transfer scaling, hEART 2015 - 4th Symposium of the European Association for Research in Transportation, 2015.9.9, コペンハーゲン (デンマーク)

三古 展弘, 新しい小さいサンプルは古い大きいサンプルと同時に使うべきか: 定数項の修正によるモデル更新の適用可能性, 第51回土木計画学研究発表会, 2015.6.6, 九州大学 (福岡県)

三古 展弘, 複数時点の断面データを同時に用いた交通需要予測の改善, 交通需要予測におけるデータの新しさとサンプル数のトレードオフ, 日本交通学会関西支部会「4月例会」, 2015.4.3, 中央電気倶楽部 213号室 (大阪府)

三古 展弘, Trade-off between data newness and number of observations for travel demand forecasting, The 94th Annual Meeting of the Transportation Research Board, 2015.1.14, ワシントン DC (米国)

三古 展弘, 交通需要予測におけるデータの新しさとサンプル数のトレードオフ, 第50回土木計画学研究発表会, 2014.11.3, 鳥取大学 (鳥取県)

三古 展弘, Travel demand forecasts by using repeated cross-sectional data: attempt to express parameters as functions of gross domestic product per capita, The 93rd Annual Meeting of the Transportation Research Board, 2014.1.14, ワシントン DC (米国)

〔図書〕(計 0件)

〔産業財産権〕

○出願状況（計 0件）

○取得状況（計 0件）

〔その他〕

学部「交通論」,「外国書講読」,「社会科学のフロンティア」や専門職大学院「現代経営学応用研究（公益企業論）」の講義でも学生に最新の研究成果として紹介している。また,航空に従事する政策担当者が読者に多い雑誌である『ていくおふ』においても研究成果を紹介した。土木学会・土木計画学研究委員会「都市間旅客交通研究小委員会」にも委員として参加し,研究成果を紹介した。

6. 研究組織

(1)研究代表者

三古 展弘 (SANKO, Nobuhiro)
神戸大学・大学院経営学研究科・准教授
研究者番号：00403220

(2)研究分担者

()

研究者番号：

(3)連携研究者

()

研究者番号：