

[特 集]

# 非 IIA 型選択モデルの選択肢集合と パラメータ特性

屋井 鉄雄<sup>1</sup>・清水 哲夫<sup>2</sup>・坂井 康一<sup>3</sup>・小林 亜紀子<sup>4</sup><sup>1</sup>正会員 工博 東京工業大学教授 大学院理工学研究科土木工学専攻(〒152-8552 東京都目黒区大岡山2-12-1)<sup>2</sup>正会員 工修 東京大学助手 大学院工学系研究科社会基盤工学専攻(〒113-8656 東京都文京区本郷7-3-1)<sup>3</sup>正会員 工修 国土交通省 住宅局住宅総合整備課住環境整備室(〒100-8918 東京都千代田区霞が関2-1-3)<sup>4</sup>正会員 工修 日本ユニシス㈱ W2Kテクノロジーセンター(〒135-8560 東京都江東区豊洲1-1-1)

本研究は、交通行動データを用いた実証的な分析を通じて、非 IIA 型離散選択モデルの適用時に問題となる、選択肢数が多い場合の推定の安定性、および選択肢集合の設定方法の妥当性を論じたものである。分析の対象として、東京都市圏の通勤鉄道経路選択行動、および日帰りの自動車観光目的地選択行動を取り上げ、選択肢間の類似性に関する考え方とその定式化を行い、構造化プロビットモデルと Mixed Logit Model の適用により 10 脈程度までのモデル推定を可能とした。さらに分析を通じて、選択肢集合に含まれる選択肢効用レベル、および選択された選択肢と代替選択肢の類似性により推定パラメータが異なる様子を明らかにした。

*Key Words:* non-IIA discrete choice model, choice set, model estimation

## 1. はじめに

我が国の交通需要予測に非集計型の確率選択モデルが実用的に用いられるようになってから 15 年ほどが経過する。特に、選択肢相互の独立性を仮定したロジットモデルは、そのパラメータ推定の容易さから多くの予測対象に利用されてきた。しかし、大都市圏の鉄道経路選択、観光目的地選択など、選択肢の独立性が必ずしも保証されない対象では、プロビットモデル<sup>1),2)</sup>や Mixed Logit Model<sup>3)</sup>など、選択肢の類似性を表現する非 IIA 型モデルの研究が進められている。その中でも、プロビットモデルは 2000 年に策定された運輸政策審議会答申第 18 号（2015 年を目標とした東京首都圏の鉄道整備計画）における将来需要の予測手法として採用されている<sup>4)</sup>。

非 IIA 型選択モデルはロジットモデルと異なり、そのモデル構造の特性によって選択肢集合の設定の相違がモデルパラメータの推定結果に影響を及ぼす。このため、非 IIA 型選択モデルの適用に際しては、ロジットモデル以上に推定や予測の段階で選択肢集合の設定方法に留意しなければならない。しかし、通常分析者側は、トリップメーカーが最終的に選択

した選択肢しか知ることができず、選択肢の認識・選別過程を知ることは困難で、その過程を表現した選択肢集合形成モデルの作成も難しい。仮に現在の過程を知ることができても、将来時点での適用可能性は不明確である。結局、分析者は想定できる全ての選択肢集合を設定した後に、どの選択肢はモデル推定上削除して構わないか、どの選択肢とどの選択肢をモデル推定上統合しても良いか、といった機械的な選択肢設定を検討することが現実的である。

本研究では、以上の問題意識を背景に、交通需要予測の場で非 IIA 型の選択モデルを実用的に用いるために残されている、①多くの選択肢を有する場合の推定可能性、②選択肢を統合・削除できる条件の検討、の課題を解決するための基礎的分析を行うことを目的とした。なお、対象とする選択行動として、東京都市圏での通勤鉄道経路選択行動、および観光目的地選択行動を取り上げる。

## 2. 非 IIA 型モデルに関する研究の現状

交通需要予測における非 IIA 型モデルの研究の経緯および種類については兵藤・章が詳しく論じてい

る<sup>5)</sup>。しかし、その実用的な利用に向けて、取り扱うことのできる選択肢数に制限があるのが障害であると考えられる。すなわち、既存研究における分析対象は選択肢数の少ない例にほぼ限られており、選択肢数が多い場合の適用可能性については、いずれのモデルにおいても今後の継続的な研究が必要である。特に、HEV モデル<sup>6)</sup>や PCL モデル<sup>7)</sup>などは、乱数シミュレーションによる推定が不要であるものの、選択肢数の増加に伴い推定パラメータ数が増加するという問題を抱えている。そのため、選択肢数が多い対象（例えば、経路選択、目的地選択モデル）を扱う場合には、現時点では構造化プロビットモデル<sup>8)</sup>や Mixed Logit Model など、乱数シミュレーションによる推定をベースとしたモデルの利用が現実的であると考えられる。

また、非 IIA 型モデルで重要な課題である選択肢集合の設定については、3 脇選択を例に、サンプル構成がパラメータ推定に及ぼす影響を分析した例<sup>9)</sup>はあるものの、モデルの基本的な推定特性の把握に留まり、他にほとんど検討されていないのが現状である。

本研究の動機は、まさに非 IIA 型モデルの研究が積み残してきたこのような課題について、いくつかの交通行動データを用いた実証的な分析を行うことである。

### 3. 非 IIA 型モデルの鉄道経路選択モデルへの適用

#### (1) モデルの定式化および推定方法

筆者らは、首都圏の都市鉄道ネットワークにおける通勤利用者の経路選択モデルについて、経路の重複距離を用いて選択肢の類似性を表現してきた。誤差項の展開に関する詳細な導出は文献 10)に譲り、ここではプロビットモデルを用いた場合の、経路  $r$  の選択確率  $P_r$  の算出式の最終型のみを示す。

$$P_r = \int_{-\infty}^{t_r + \varepsilon_r - l_1} \cdots \int_{-\infty}^{t_r + \varepsilon_r - l_k} \phi(\varepsilon) d\varepsilon \quad (1)$$

$$\phi(\varepsilon) = (2\pi)^{-\frac{R}{2}} \exp\left(-\frac{1}{2}\varepsilon^T \Sigma \varepsilon\right) \quad (2)$$

$$\Sigma = \sigma_0^2 \begin{pmatrix} \lambda L_1 + 1 & \cdots & \lambda L_R \\ \vdots & \ddots & \vdots \\ \lambda L_R & \cdots & \lambda L_R + 1 \end{pmatrix} \quad (3)$$

$$\lambda = \sigma_r^2 / \sigma_0^2 \quad (4)$$

ここで、 $R$  は総経路数、 $V_r$  は  $r$  の確定効用、 $L_r$  は  $r$  の経路長、 $L_{rq}$  は  $r$  と  $q$  の重複経路長、 $\sigma_0^2$  は選択肢固有の誤差分散、 $\sigma_r^2$  は単位経路長当たりに発生する誤差分散である。推定されるパラメータは効用関数の確定項の変数パラメータと誤差項の分散比  $\lambda$  である。

従来の推定特性に関する分析については 3 脇選択が中心であり、4 脇を越えるケースについては今後の検討課題となっている。推定には GAUSS を使用するが、3 重積分が汎用的な関数として用意されており、これを用いれば 4 脇選択までは数値積分による推定が可能である。5 脇以上は乱数シミュレーション法 (GHK 法など) を利用しなければならない。GHK を利用した推定方法やその特性については文献 10)を参照されたい。

#### (2) 選択肢数を増したモデル推定の結果

以上のモデルについて、平成 7 年大都市交通センサスの定期券調査データを用いて推定を行った。この調査では各サンプルに対して定期券を購入する経路のみが回答されるため、代替経路の設定については、付録に記載した方法を用いて分析者側が利用実績に基づいて設定している。全サンプルのうち、元データセットとして、ゾーン別 OD 交通量のシェアに応じて 2,999 サンプルをランダムに抽出し、これに代替経路のデータを加えた上で選択肢数別（最大 39 選択肢）に分類した。

プロビットモデルの推定では、非 IIA 特性を持つが故に、ロジットモデルとは異なり、各サンプルの選択肢数を一律とすることに理論上の問題がある。通常のプロビットモデルでは、式(2)の分散共分散行列  $\Sigma$  が全サンプルで同一としなければならないため、全選択肢集合を同一の選択肢集合を持つグループに分割し、グループごとに別々にモデルを推定する必要があったが、分散共分散行列を式(3)のように構造化すれば、異なる選択肢集合や選択肢数を持つサンプルを、同一のモデルとして推定できる。

表-1 は 3 脇から 9 脇までのサンプルを混在させたデータによる推定結果である。元データセットのうち、東京都区外から都区内へのトリップである 1,358 サンプルを抽出して推定を行った。ここで用いたデータの選択肢数の取り扱いであるが、例えば、3~9 脇のモデルでは、10 脇以上のサンプルについて選択経路と利用実績頻度の高い順に 8 つの経路を抽出して 9 脇としている。なお、GHK 法のアルゴリズムにおいて、ほぼ重複している経路が含まれているサンプルで分散共分散行列のコレスキーフ分解が不可能なケースが存在するため、3~8 脇および 3~

表-1 モデルの推定結果(サンプルにより選択肢数が異なる場合)

説明変数	3~5肢		3~6肢		3~7肢		3~8肢		3~9肢	
	パラメータ	t値								
運賃(円)	-0.00103	-3.33	-0.00133	-4.35	-0.00138	-4.63	-0.00148	-5.21	-0.00147	-5.31
アクセス時間(分)	-0.0571	-8.34	-0.0667	-9.33	-0.0692	-10.3	-0.0702	-10.7	-0.0726	-11.5
イグレス時間(分)	-0.0214	-3.53	-0.0261	-4.23	-0.0268	-4.53	-0.0281	-4.88	-0.0288	-5.10
乗車時間(分)	-0.00917	-1.82	-0.0125	-2.45	-0.0128	-2.68	-0.0125	-2.70	-0.0126	-2.80
乗換時間(分)	-0.0240	-3.36	-0.0264	-3.70	-0.0279	-4.16	-0.0290	-4.47	-0.0305	-4.77
待ち時間(分)	-0.0860	-4.29	-0.100	-4.96	-0.0992	-5.20	-0.100	-5.37	-0.0985	-5.54
乗換回数(回)	-0.147	-2.25	-0.182	-2.79	-0.194	-3.14	-0.194	-3.26	-0.202	-3.46
混雑指標(※)	-0.00227	-1.81	-0.00242	-1.89	-0.00258	-2.15	-0.00268	-2.33	-0.00270	-2.41
分散比	0.147	2.02	0.194	2.50	0.146	2.31	0.123	2.21	0.115	2.31
尤度比	0.0658		0.0908		0.107		0.119		0.128	
サンプル数	1358		1358		1358		1357		1357	
うち 3選択肢	195		195		195		195		195	
4選択肢	176		176		176		176		176	
5選択肢	987		181		181		181		181	
6選択肢	-		806		176		176		176	
7選択肢	-		-		630		147		147	
8選択肢	-		-		-		482		94	
9選択肢	-		-		-		-		388	

※混雑指標は、乗車区間の平均混雑率(%)を100で割った値を二乗し、乗車時間(分)を乗じたものである。例えば、乗車時間が短くても混雑率が高い、混雑率がそれほど大きくなくとも乗車時間が長い場合には車内の混雑に対する不快感が大きくなる状態を表現している。

9肢はこれに該当する1サンプルを除外した。説明変数のパラメータ値およびt値が選択肢数の増加に伴いその絶対値が大きくなっていることが見て取れる。選択肢数が多いサンプルでは、選択経路と効用レベルが大差ない代替経路が多く含まれる傾向があり、最尤法の推定過程においてパラメータ値を大きくして、より選択経路の効用を上昇させることができたからであると考えられる。分散比パラメータについては、サンプルの平均的な重複度合いにより、その大きさやt値が大きく影響を受けることが既存の研究<sup>11)</sup>から明らかとなっているが、多様な選択肢数や重複度合いを持つサンプルが混在しているために、有意な推定が行えたと考えられる。

本稿では、最大9肢選択までのモデルを推定したが、10肢以上のデータについても推定可能と考えている。しかし、次の2点の考慮により、これ以上の推定可能性の確認は実務的に不要であると考えている。第一に、現実的には利用者が認識する経路数はそれほど大きくない。第二に、仮に利用者の認識する経路数が大きくても、類似性の極めて高い経路が相互に多くなり、これらを同一の経路として取り扱うべきである。この点は(3)で述べる。

このように、選択肢の誤差を構造化し、推定パラメータ数を飛躍的に減少させる工夫により、また、GHK法といった効率的な推定アルゴリズムの登場により、従来には不可能であった10肢未満程度の選択肢数の推定が行えるようになった。

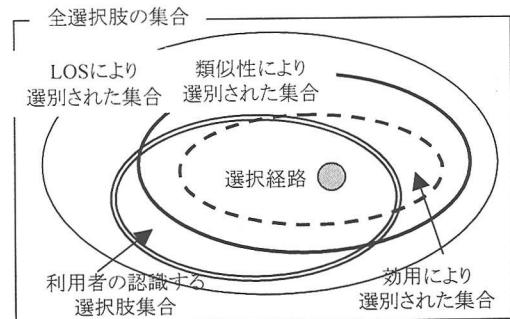


図-1 利用者の選択肢の認識と機械的選別の関係

### (3) 分析者側の選択肢集合設定の考え方

以上の議論を踏まえ、今度は逆に選択肢数を適切に減らしていく方法について検討する。分析者側が選択肢集合を設定するときには、極力トリップメーカーの想定している選択肢集合に近いものを設定しなければならないことは言うまでもない。

はじめにトリップメーカーのODから物理的に選択可能な全経路が設定される。完全情報を仮定するときに実際の選択肢集合が形成される(以下、選別と称する)条件には、1)LOS、2)類似性、3)効用、の3つが考えられる。1)は経路長、所要時間などによる選別を意味し、非補償的にその大小によって選択肢集合に入るか否かを決定すると考える。2)は、経路の重複による選別、識別可能性を意味し、ほとんど重複している経路は同一と見なす可能性が高いと考える。3)は1)とは異なり補償的な選別である。これらと利用者の認識している選択肢集合との関係を図-1に示す。LOSにより選別された集合は利用

者の認識している選択肢集合を包含するが、いまだ多くの選択肢を含んでおり、類似性による選別を行い、選択肢集合を絞り込む。このように設定された選択肢集合は、利用者の認識する選択肢集合とは完全に一致しないものの、それに近いものと想定される。このプロセスを経てもなお推定上選択肢が多い場合には、効用による選別を行い、選択肢集合を絞り込むという手順が考えられる。

#### (4) 選択肢集合がモデル推定に及ぼす影響

ここでは(3)の選択肢集合の選別方法について、LOS および類似性で選別されたデータセットによるパラメータ推定を行った。現時点でパラメータ推定が最大 9 肢まで対応しているため、全選択肢が 10 以上のサンプルは、選択経路と発着ゾーン間経路利用実績の多い順に選択経路を除いた選択肢の上位 8 肢までを選択肢集合とし、また選別の影響を見るために元々選択肢数が少ない 5 以下のサンプルを元データセットから除外した。以下ではこれを基準データセット A と呼ぶ。

はじめに、LOS による選別を以下の 2 つの基準で行った。

1) 所要時間が選択経路の 2 倍以上、もしくは 30 分以上長い経路は取り除く

2) 経路長が選択経路の 1.8 倍以上の経路は取り除く

なお、所要時間は「アクセス時間+イグレス時間+乗車時間+乗換時間+待ち時間」とする。このプロセスで作成されたデータセットを B とする。データセット A および B による推定結果を表-2 に示すが、B では乗換時間以外のパラメータで絶対値が有意に減少し、同時にパラメータの *t* 値も減少している。選択肢集合に含まれる選択肢は利用実績に基づいているため、B では効用を形成する LOS の値は大きな差がないと考えられるが、選択経路が必ずしも最も効用の高い経路でないサンプルが多く含まれている場合には、推定パラメータが小さくなることが考えられる。その結果、B の尤度比が A より大幅に低下している。このように、利用者がほとんど認識していないと考えられる遠回りの経路の存在により、モデルの尤度に影響を与える可能性があり、このような選択肢を集合に含めるか否かは従来の IIA モデル以上に注意が必要である。また、選別による効用関数パラメータを変化に注目したため、分散比パラメータの有意性についてはここでは深い議論はしないが、B では LOS による選別により残された経路の類似性が総じて高く、分散比パラメータの推定が不安定になった可能性が高い。なお、本来

表-2 LOS により選別された場合の推定結果

データセット名	A	B
説明変数	パラメータ	<i>t</i> 値
運賃(円)	-0.00197	-6.28
アクセス時間(分)	-0.0712	-11.0
イグレス時間(分)	-0.0336	-5.69
乗車時間(分)	-0.0136	-2.62
乗換時間(分)	-0.0208	-2.88
待ち時間(分)	-0.120	-6.13
乗換回数(回)	-0.109	-1.57
混雑指標(※)	-0.00213	-1.61
分散比	0.0507	0.929
尤度比	0.158	0.131
サンプル数	896	896
5選択肢	0	141
6選択肢	323	308
7選択肢	265	216
8選択肢	168	129
9選択肢	140	102

※定義は表-1を参照

正であるべきの分散比が負と推定されているが、このケースは選択肢の相関を考慮することに意味がない（ロジットモデルの方が尤度比が高い）ことを意味している。

次に、類似性による選別を考える。2 つの経路の類似性を表現する指標として、2 経路で共通する区間長（重複区間長）の総経路長に対する割合（重複率）を考える。重複率は 0 から 1 の間の値をとり、0 に近ければ 2 経路は類似（重複）していない状態、1 に近ければ類似（重複）している状態である。経路間の重複率が大きければ、経路間の相関も大きいと捉えることができる。選択肢間の相関を考慮できる点がプロビットモデルの特徴であるが、利用者は重複率が 1 に近い経路を別々に認識している可能性は小さいと考えられる。この場合、モデル推定上は両者を統合するか、あるいはどちらかの経路を選択肢集合に入れ、もう一方の経路を削除することになる。なお、どの程度重複していれば同一の経路と考えられるかについては明らかでないが、選択経路の選択確率の推計上、最も影響の小さい選別のルールを検討しておけば実用上問題はないと考える。以上の考え方から、ここでは先のデータセット B を基準として、重複率がある一定以上の水準の経路同士を 1 つの経路に統合したデータセットを作成する。その際、統合する重複率を 95%, 90%, 85%, 80% の 4 通りで考えた。これらのデータセットは C95, C90, C85, C80 と呼ぶことにする。

データセット C によるモデル推定の結果を表-3 に示す。なお、選別により選択肢数が 2 以下になつたサンプルについてはこれを除外しているため、データセットによりサンプル数が異なる。運賃、アクセス時間、イグレス時間、乗車時間、待ち時間のパ

表-3 類似性により選別された場合の推定結果

データセット名	C95		C90		C85		C80	
	パラメータ	t値	パラメータ	t値	パラメータ	t値	パラメータ	t値
説明変数								
運賃(円)	-0.00198	6.08	-0.00217	6.01	-0.00199	5.31	-0.00217	5.38
アクセス時間(分)	-0.0668	10.2	-0.0773	9.83	-0.0789	9.56	-0.0823	8.81
イグレス時間(分)	-0.0314	5.22	-0.0369	5.39	-0.0377	5.33	-0.0382	5.06
乗車時間(分)	-0.0110	2.14	-0.0149	2.42	-0.0202	3.18	-0.0196	2.91
乗換時間(分)	-0.0214	2.94	-0.0195	2.32	-0.0214	2.45	-0.0224	2.44
待ち時間(分)	-0.106	5.18	-0.132	5.47	-0.130	5.26	-0.1345	5.03
乗換回数(回)	-0.0793	1.15	-0.0983	1.25	-0.0388	0.479	-0.0317	0.386
混雑指標(※)	-0.00145	1.09	-0.00251	1.65	-0.00173	1.15	-0.00231	1.43
分散比	0.0127	0.205	0.155	1.67	0.129	1.42	0.166	1.35
尤度比		0.137		0.133		0.129		0.135
サンプル数	896		892		875		858	
3選択肢	7		27		59		97	
4選択肢	30		85		139		166	
5選択肢	178		224		249		263	
6選択肢	297		272		235		187	
7選択肢	206		159		120		101	
8選択肢	106		82		57		36	
9選択肢	72		43		16		8	

※定義は表-1を参照

パラメータの絶対値は統合される重複率が低くなるほど大きくなる傾向がある。これらは選択経路と類似している経路の統合により、選択経路の効用が相対的に高くなつたためであると想定される。一方、乗換回数については統合される重複率が小さくなるほど小さくなっているが、選別により乗換回数の少ない経路が選択肢集合に残り、乗換回数の相対的な差が選択肢間で広がり、パラメータが小さくなつたものと考えられる。分散比パラメータについては、C95 はパラメータ値、t 値ともに小さく、C90, C85, C80 は必ずしも有意でないものの、比較的安定したパラメータ値が得られている。分散比が大きく推定されれば、そのデータは選択肢の類似性をより考慮した方がよいと考えることができよう。しかし、既存研究<sup>10)</sup>の知見も考慮すれば、C95 のようにあまりに選択経路と重複している代替経路が存在すれば、分散比が小さく推定され、結果として、他の経路間の類似性も小さく見積もってしまうことになる。このように、経路の類似性を確保する観点からは、90%以上重複している経路は事前にこれらを統合しておく方がよいと考えられる。

#### 4. 非 IIA 型モデルの日帰り観光目的地選択モデルへの適用

##### (1) 実用面から見た観光目的地選択モデルの課題

本章では自動車による日帰り観光目的地選択モデルを取り上げる。観光目的地選択では従来から非集計型のモデルが研究されてきた<sup>12)</sup>。全国規模の

観光実態調査を契機に、岡本は非集計ロジットモデルにより観光目的地選択モデルを構築し、分析者が事前に設定した 20 程度の選択肢数を持つモデルを推定している<sup>13)</sup>。しかし、選択肢集合の変化による推定精度の確認、および分析上望ましい選択肢集合の設定方法に関する検討はなされていない。また、選択されている観光地の効用が他のそれと比べてはっきりと大きくなっている場合が多く、選択肢数の変化に応じてパラメータ推定値が変化しやすいと想定される。そのために一方では、選択肢集合を形成するモデルを構築し、選択肢数を事前に減らす工夫も行われてきているが<sup>14)</sup>、モデルで仮定している選別の過程がトリップメーカーのそれとは同じであるとは考えにくい。そのため、選択肢の選別と選択を同時に表現するようなモデルも提案されているが<sup>15)</sup>、選択肢間の類似性は考慮されていない。

以上のことから、観光目的地選択モデルを実用的に用いるために、分析者側が何らかの方法で選択肢集合を設定する方法を検討すべきであると考える。特に、非 IIA 型の選択モデルを利用する場合には、選択肢集合に含まれる観光地間の類似性によって推定結果が変化するため、その適切な設定のために、選択肢の類似性が推定特性に及ぼす影響を詳細に分析する必要がある。

なお、目的地選択モデルに非 IIA 型のモデルを適用した例として、Bolduc *et al.* の自己回帰型のプロビットモデル<sup>16)</sup>がある。ここでは、目的地の効用の誤差をその他の目的地からの位置関係によって構造化し、目的地間の類似性を表現している。その後目的地と出発地の位置関係によつても誤差を構造化

する工夫が行われた<sup>17)</sup>。しかし、モデルにおける類似性の表現は一般的であり、実際の適用には以下で定式化を行うような改良が必要である。一方、福田らは目的地間の類似性を目的別施設数に応じて定義している<sup>6)</sup>が、選択肢数の増加とともに推定パラメータ数が飛躍的に増加することが課題となっている。本研究ではこれら両方の概念を包含した類似性の定義を試みる。すなわち、利用アクセス経路が同じ、相互に近接している、などの空間的な類似性と、同一の施設を観光地が共有していることによる類似性との両者である。

## (2) モデルの定式化

個人が観光目的地を想起する方法に、「伊豆」、「箱根」などの漠然とした地域を想定して、その地域にある個別の施設を続けて想起するケースと、逆に、ある個別の施設を想起してからその周辺を含めて目的地として想起するケースが考えられる。後者のケースでは、個人が想起している施設が近接する2つの目的地の境界周辺に存在する場合、この個人はこれら2つの目的地を重複するものと認識しているかも知れない。そして類似性は隣接する目的地間で共有する施設数が多いほど高いと考えられる。このような考え方は、目的地間の距離や面積を用いて一般化して表現することが可能である。この場合、目的地の効用の誤差項は何らかの相関を持つことになる。本研究では、いずれのケースでも「観光目的地」を全旅行者に共通の固定された単一地域名称で認識する（例えば、伊豆と箱根は統合せず別の観光地として扱う）が、その範囲（以下、圏域と呼ぶ）が個人で異なると仮定する。

さらに、同じアクセス道路を用いる2つの観光目的地を考える。例えば、この道路の所要時間の不確実性が高い場合、両方のアクセス時間の不確実性が高くなる。この不確実性を観測できない変数として誤差と考えれば、両目的地を選択する際の誤差の傾向が似ていることを意味する。すなわち、アクセス経路によっても観光目的地の類似性を考慮することができる。

以下にこのような類似性の考え方を反映したモデル化を行う。各観光目的地*i*の効用*U<sub>i</sub>*を、以下のように確定項*V<sub>i</sub>*と確率項*ε<sub>i</sub>*で表現する。

$$U_i = V_i + \varepsilon_i \quad (5)$$

今、*ε<sub>i</sub>*が以下のように、互いに独立な目的地に含まれる施設による誤差η<sub>*i*</sub>、アクセス経路による誤差ω<sub>*i*</sub>、およびその他の誤差（目的地に固有）ξ<sub>*i*</sub>に分割されるとする。

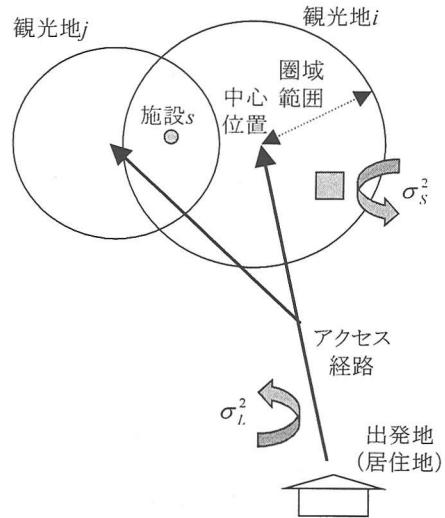


図-2 観光目的地間の類似性の概念

$$\varepsilon_i = \eta_i + \omega_i + \xi_i \quad (6)$$

次に、図-2 のように部分的に共通のアクセス経路を持ち、圏域が重複している2つの観光目的地*i*, *j*を考える。ここで、重複圏域にある施設*s*が立地しているとする。この施設の存在による目的地効用の誤差が正規分布  $N(0, \sigma_s^2)$  に従うと仮定すると、*i*, *j* の両方にこの誤差が反映される。逆に、*i*のみに含まれる施設については、その誤差は *i* のみに反映される。簡便のため、圏域を単位面積に分割し、それぞれに施設が1つ存在し、そこから発生する誤差は互いに独立であり、かつその分散は一定であると考えれば、*i* の誤差分散、および *i*, *j* の誤差共分散は以下のように表現できる。

$$\text{var}(\eta_i) = S_i \sigma_s^2 \quad (7)$$

$$\text{cov}(\eta_i, \eta_j) = S_{ij} \sigma_s^2 \quad (8)$$

ここで、*S<sub>i</sub>* は *i* の圏域面積、*S<sub>ij</sub>* は目的地 *i*, *j* の重複圏域面積、 $\sigma_s^2$  は単位面積当たりの誤差分散である。なお、圏域に含まれる施設の目的種別（温泉、ゴルフ等）別に単位誤差を設定することも可能である。アクセス経路による誤差については、鉄道経路選択モデルと同様に、経路上のリンク（単位距離）ごとに互いに独立かつ分散の等しい誤差が発生していると考えれば、*i* の誤差分散、および *i*, *j* の誤差共分散は以下のように表現できる。

$$\text{var}(\omega_i) = L_i \sigma_L^2 \quad (9)$$

$$\text{cov}(\omega_i, \omega_j) = L_{ij} \sigma_o^2 \quad (10)$$

ここで、 $L_i$  は  $i$  のアクセス経路長、 $L_{ij}$  は  $i, j$  のアクセス経路の重複距離、 $\sigma_o^2$  は単位距離当たりの誤差分散である。目的地に固有の誤差に関しては、その分散が  $\sigma_0^2$  で等しいと仮定すれば以下のようになる。

$$\text{var}(\xi_i) = \sigma_0^2 \quad (11)$$

$$\text{cov}(\xi_i, \xi_j) = 0 \quad (12)$$

式(7)～(12)をまとめて、分散共分散行列の形で表現すれば、

$$\Sigma = \sigma_0^2 \begin{pmatrix} 1 + \gamma S_1 + \delta L_1 & \cdots & \gamma S_{II} + \delta L_{II} \\ \vdots & \ddots & \vdots \\ \gamma S_{II} + \delta L_{II} & \cdots & 1 + \gamma S_I + \delta L_I \end{pmatrix} \quad (13)$$

$$\gamma = \sigma_s^2 / \sigma_0^2 \quad (14)$$

$$\delta = \sigma_L^2 / \sigma_0^2 \quad (15)$$

となる。

ここでは、非 IIA 型モデルとして Mixed Logit Model(MXL)を取り上げる。MXL では、観光地  $i$  の選択確率は以下の式で与えられる。

$$P_i = \frac{\int \int \exp(V_i + \eta_i + \omega_i)}{\sum_i \exp(V_i + \eta_i + \omega_i)} f(\eta) g(\omega) d\omega d\eta \quad (16)$$

式(16)は積分型であり、モデル推定にはモンテカルロ法などの乱数シミュレーションによる近似計算を行うことになる。現状では、MXL とプロビットモデルの推定特性に顕著な差はないと考えているが、MXL はプロビットモデルにおける GHK 法のように乱数発生回数を減少させる推定法がなく、計算時間の点ではプロビットモデルの方が優れている<sup>9)</sup>。一方、MXL はモデル構造がより柔軟であり、拡張性や発展性はより高いと考えられる。

最後に観光地の圏域の設定方法を考える。本研究で対象とする日帰りの観光行動では、これに要する全時間のうち、往復の所要時間を除いた時間が観光地内で活動可能な時間である。ここでは、観光地の圏域はこの活動可能時間で到達することが可能な範囲であると考える。この考え方によれば、アクセスに時間がかかる目的地を選択すれば、到達可能な範囲が小さくなる（圏域が小さくなる）。逆に、近い目的地では圏域が大きくなる。今、近接する 2 つの観光地  $i, j$  を考え、出発地からこれら観光地までの距離（アクセス時間）が短い場合と長い場合には、図-3 に示すように、目的地が遠い場合には、

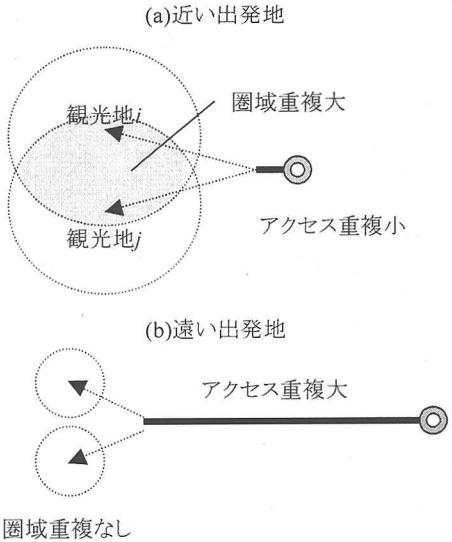


図-3 出発地の遠近と観光地間の類似性の関係

活動可能範囲が小さくなるため重複する圏域の面積は小さくなるが、アクセスは途中までほぼ同じ経路であると考えれば重複が大きくなる。逆に、近い場合には、活動可能範囲が大きくなるため重複する圏域の面積は大きくなるが、アクセス経路の重複は短くなることが考えられる。すなわち、近い出発地では重複圏域による類似性の認識が高くなり、遠い出発地ではアクセス経路の重複による類似性の認識が高くなることになる。先の「箱根」と「伊豆」の例では、横浜と千葉から出発する場合には、より近い前者の方がこれら両者を周遊しやすく、これら 2 つは相関の高い観光地ということになる。

### (3) モデルの推定

モデルのパラメータ推定に用いるデータは、1992 年に建設省土木研究所（当時）で行った全国観光交通実態調査のうち、関東地区（東京都、横浜市）および東北地区（仙台市、郡山市）の家庭訪問調査によるサンプルである。選択肢集合を形成する観光地群はこれらの出発地から選択割合の高かった 20（横浜、鎌倉、箱根、熱海、甲府、秩父、水戸、上高地、金沢、日光、草津、志賀、尾瀬、湯沢、会津、飯坂、いわき、蔵王、松島、十和田湖）を設定した。全サンプルからランダムに 269 を抽出し、各サンプルが実際に選択した観光地に残りの 19 観光地からランダムに 9 を抽出し選択肢集合に加え、10 脈選択のデータセットを作成した。表-4 にロジットモデル(MNL)と MXL による推定結果を示す。観光地までの所要時間と費用の相関が非常に大きく、これ

表-4 10肢選択モデルの推定結果

説明変数	MNL		MXL(1)		MXL(2)		MXL(3)	
	パラメータ	t値	パラメータ	t値	パラメータ	t値	パラメータ	t値
合成費用価値	-0.0409	-6.82	-0.0415	-6.71	-0.0426	-6.44	-0.0426	-6.46
観光地魅力度	0.0738	6.75	0.0755	6.63	0.0735	6.70	0.0741	6.63
分散比 $\gamma$			0.0111	0.570			0.0138	0.188
分散比 $\delta$					0.0286	0.674	0.0219	0.527
尤度比	0.117		0.118		0.118		0.119	

表-5 空間パターンと分散比の推定特性

	空間配置パターン	$\gamma$	$\delta$
A		×	○
B		○	×
C		×	△
D		△	×

○選択観光地 ○代替観光地 □出発地

○:有意性が高い △有意性が低い

×:有意性が極めて低い

らを「合成費用価値(円)」として統合(時間評価値を固定)し、サンプルのばらつきを確保するために  $\ln(\text{年収(万円)})$  で除している。観光地魅力度は室谷による定義<sup>18)</sup>により算出している。また、観光地の圏域面積を算出するための半径は、日帰り観光行動の総活動可能時間、アクセス平均速度および域内の平均移動速度を考慮して決定している。簡便のため、アクセス経路は直線距離とし、重複アクセス距離は出発地から、出発地と 2 つの目的地を頂点とする三角形の内心までの距離としている。説明変数のパラメータの t 値は十分有意であるが、分散比パラメータについては有意でない結果となった。その理由は次節で考察する。

#### (4) 選択肢集合が推定特性に及ぼす影響

はじめに、選択肢集合に含まれる選択肢間の類似度の構成により、モデル推定結果がどのような影響を受けるかを定性的に示す。簡便のため 3 脅選択を例に取る。本研究における観光地選択モデルでは、出発地と観光地の位置関係により類似性が定義されるが、実際に選択された観光地と残りの 2 観光地の空間パターンは、相互に独立した場合を除けば表-5

のように整理できる。鉄道経路選択の場合と同様に、各分散比パラメータ  $\gamma$  および  $\delta$  の推定特性に与える影響は表-5 のように想定される。例えば、A のケースでは選択観光地が出発地から近くかつ独立、代替観光地が遠くかつ近接(類似)しているが、効用関数の確定項のパラメータを固定した条件の下で、選択観光地の選択確率を上げようとすれば、分散比パラメータを大きくなることで、代替観光地間の相関が大きくなるように作用すればよいことになる。この時、代替観光地間の重複圏域は非常に小さく、逆に重複アクセス経路長は大きくなるため、 $\delta$  の効果が大きくなればよいことになる。すなわち、A のタイプの選択肢集合が卓越するサンプルによる推定では  $\delta$  が有意となる可能性が高い。一方、C の空間配置では、選択観光地の選択確率を上げようとすれば、近接する代替観光地との相関をできるだけ小さくする作用することになる。すなわち、 $\delta$  が有意とはならず、非常に小さい値が推定される可能性が高い。B および D のケースも同様に考えれば、 $\gamma$  の有意性が B のケースでより高くなることになる。

ここで(3)の推定において、分散比パラメータの t 値が小さい理由について考察する。分析に用いたサンプルは出発地に近接した観光地の選択シェアが大きいため、表-5 の A と D のケースが多く含まれていることになる。そのため  $\gamma$  と  $\delta$  の推定の有意性が、両者が混在することにより効果が相殺されてしまったため低下したと考えることができよう。

#### (5) 実験的な選択肢集合設定によるモデル推定

以上で定性的に考察した分散比パラメータの推定特性について、限定的ではあるが定量的に確認する。図-4 は出発地とその周辺の 8 つの仮想的な観光目的地の位置関係を表す。各目的地は図中に示すような選択実績シェアおよび一定の魅力度を持っており、出発地からのアクセス費用は距離をもとにランダムに与えている。サンプルの選択目的地は選択実績シェアに一致するようにランダムに与えた。変数のばらつきを確保するために、先の全国観光交通実態調査で得られた年収分布を基に年収をランダム

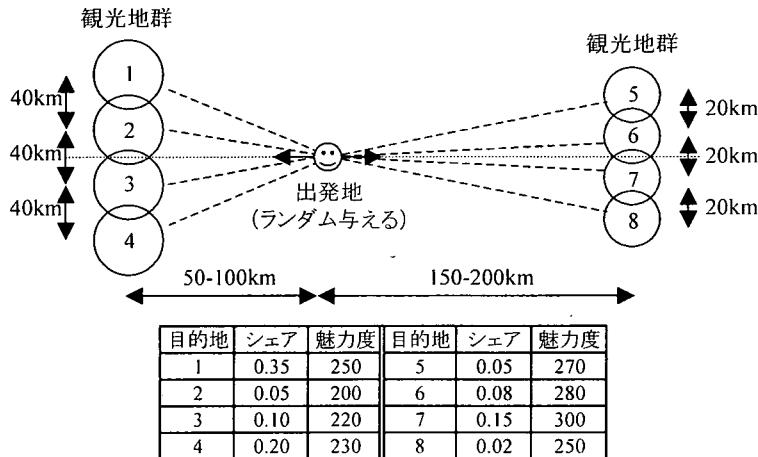


図-4 実験における出発地および観光地の位置関係

表-6 選択肢集合のタイプ分け

Type	選択肢集合の特性
A	目的地1~4全て&5~8から1つ
B	目的地1~4から3つ&5~8から2つ
C	目的地1~4から2つ&5~8から3つ
D	目的地1~4から1つ&5~8全て

に与え、かつ出発地は直線上の一定範囲内でランダムに与えている。なお、各重複の与え方については(3)と同様である。使用するモデルは(3)と同様にMXLであり、選択肢数は5で固定し、表-6のような4種類の選択肢集合タイプによる推定を行った。Type Aは表-5の空間配置パターンDのサンプルが多数を占めており、逆にType Dは空間配置パターンAのサンプルが多数を占める。そのため、Type Dから順にC、B、Aとなるにしたがって分散比パラメータの推定の有意性が低下することが想定される。なお、推定サンプル数は、MXLの推定時に発生させる乱数の安定性を考慮して600とした。

表-7に推定結果を示す。説明変数パラメータの推定値については、MNL、MXLともほぼ同一である。分散比パラメータについては必ずしも有意でないものの、事前の想定通りにType Dが一番大きいパラメータ値であり、かつ $\gamma$ 値も一番大きい結果となっている。Type AおよびBでは一部推定不可能であったが、Type Aでは観光地1~4のアクセス経路重複が相互に大きいために、 $\delta$ を小さくした方が選択目的地の選択確率を大きくできるため、推定の過程でこれを0に近づけようとした結果であり、Type Bでは観光地1~4の圏域重複が $\gamma$ を小さくしようとする一方、観光地5~8の圏域重複は $\gamma$ を大

きくしようとする作用がほぼ同じ効果を持っていたため、収束が進行しなかった結果である。これらの選択肢集合では各選択肢を独立と見なした方がよりよい尤度を与えるためと考えられる。推定が可能であったケースではいずれも、MXLの適用により尤度比が向上する結果となっている。なお、 $\gamma$ を考慮したMXL(I)の方が $\delta$ よりも尤度比がよいが、本稿で用いた重複アクセス距離の導出方法では、例えば観光地1と5にも重複アクセス距離が定義できてしまい、選択肢の独立性が弱くなってしまうためである。実際の適用では、観光地ごとに特定のアクセス経路を設定すれば、この問題が生じる可能性は小さいと考える。

## 5. おわりに

本研究は、近年交通需要予測の場で適用の進んでいる非IIA型選択モデルについて、課題となっている選択肢集合の設定方法を提案するための基礎的な研究として、東京都市圏の通勤鉄道経路選択行動、および日帰りの観光目的地選択行動を例に、選択肢集合の設定による推定特性の相違を詳細に分析した。その結果、①選択肢集合の設定を行う上で、多くの選択肢に対する推定可能性の検討が必要であったが、従来の研究では実証が困難であった概ね10程度の選択肢数に対応可能となった。②実際に選択された選択肢の効用が代替選択肢の中で最良でないサンプルが多く含まれている場合には、泡沢的な選択肢の存在により推定パラメータが大きく変動する

表-7 実験的な選択肢集合による推定結果

説明変数	Type A					
	MNL		MXL(1)		MXL(2)	
	パラメータ	t値	パラメータ	t値	パラメータ	t値
合成費用価値	-0.000901	-8.72	-0.000910	-8.74	---	---
観光地魅力度	0.0365	14.6	0.0367	14.5	---	---
分散比 $\gamma$			0.0000116	0.532		
分散比 $\delta$					---	---
尤度比	0.1379		0.1385		---	
説明変数	Type B					
	MNL		MXL(1)		MXL(2)	
	パラメータ	t値	パラメータ	t値	パラメータ	t値
合成費用価値	-0.00149	-13.9	---	---	-0.00150	-13.7
観光地魅力度	0.0341	13.5	---	---	0.0343	13.5
分散比 $\gamma$			---	---		
分散比 $\delta$					0.000499	0.678
尤度比	0.1201		---		0.1211	
説明変数	Type C					
	MNL		MXL(1)		MXL(2)	
	パラメータ	t値	パラメータ	t値	パラメータ	t値
合成費用価値	-0.00218	-17.6	-0.00222	-16.9	-0.00220	-17.2
観光地魅力度	0.0358	13.0	0.0369	12.5	0.0361	12.9
分散比 $\gamma$			0.0000711	1.01		
分散比 $\delta$					0.000441	0.592
尤度比	0.2121		0.2143		0.2128	
説明変数	Type D					
	MNL		MXL(1)		MXL(2)	
	パラメータ	t値	パラメータ	t値	パラメータ	t値
合成費用価値	-0.00276	-17.2	-0.00290	-15.9	-0.00286	-15.5
観光地魅力度	0.0330	9.13	0.0351	9.02	0.0346	8.83
分散比 $\gamma$			0.000268	1.58		
分散比 $\delta$					0.00233	1.09
尤度比	0.3965		0.4020		0.3991	

可能性があることが実証された。③非 IIA 型選択モデルにおいて選択肢を統合・削除できる条件の把握については、通勤鉄道経路選択モデルの例では、90%程度以上経路が重複していればこれらを統合して 1 つの選択肢とした方が、選択肢の類似性を考慮する意味があるモデルの推定が行えることが確認できた。

なお、観光目的地選択モデルについては、サンプルの発地の分布が偏っているなど、適用可能なデータ整備が不十分であり、実験的に行った限定的な条件での推定特性の把握に留まっているが、④本研究の鉄道経路選択モデルの推定特性と同様の傾向が確認できた。今後は本稿で提案したモデルに対応できる観光地選択データの整備を検討し、再度本モデルの推定特性を詳細に分析していく必要がある。

最後に、適用対象が限られており、かつ詳細な分析は今後の課題であるが、本稿の成果を踏まえて非 IIA 型モデルにおける推定結果の妥当性の判断方法の可能性について考察する。非 IIA 型モデルの特徴的なパラメータは、言うまでもなく分散比であるが、これが有意に推定されないケースは一般的に、「選択肢間の類似性を考慮しても意味がない」と考

えることができよう。しかし、通勤鉄道経路選択モデルでの結果のように、類似性を考慮しなければならないにもかかわらず、あまりに類似性の高い選択肢群が混在することにより、分散比が有意に推定されなくなる例も存在する。推定に悪影響を及ぼすという観点から、このような選択肢は事前に 1 つの選択肢に統合されるべきである。一方、表-5 で整理したように、サンプルの類似の度合いを設定すれば分散比を有意に推定することも可能であるが、このような推定に意味があるケースは限られている。以上を鑑みれば、サンプルの平均的な重複度合いを示す何らかの指標を導入して、その値に応じて必要な分散比の t 値が決定される、といった方法で推定結果の妥当性を判断する必要が生じてくるものと考えている。

**謝辞**：本研究で利用した鉄道経路データの作成はJR企画開発にご協力を頂いた。記してここに謝意を表する。

## 付録 代替経路の作成方法

本研究では代替経路の設定は以下の方法で行っている。

- (1)選択経路の発地・着地ゾーンからアクセス・イグレス可能な駅を設定する。
- (2)これらの駅を路線や方角に応じて適宜統合し、「駅グループ」を構成する。
- (3)駅グループ間の他のサンプルによる選択実績を代替経路とする。例えば、発地駅グループ、着地駅グループに駅がそれぞれ2つ含まれていれば、 $2 \times 2 = 4$  経路を代替経路に含める。

## 参考文献

- 1) Hausman, J. A. and Wise, D. A.: A conditional probit model for qualitative choice : discrete decisions recognizing interdependence and heterogeneous preferences, *Econometrica*, Vol.46, No2, pp.403-427, 1978.
- 2) Bunch, D. S.: Estimability in the Multinomial Probit Model, *Transportation Research B*, Vol.25, pp.1-12, 1991.
- 3) McFadden, D. and Train, K.: Mixed MNL models of discrete choice, *working paper of UCB*, 1997.
- 4) 運輸省：東京圏における高速鉄道を中心とする交通網の整備に関する基本計画策定に向けての調査報告書, 2000.
- 5) 兵藤哲朗, 章翔: Mixed Logit モデルの汎用性に着目した特性比較分析, 土木学会論文集, No.660, IV-49, pp.89-99, 2000.
- 6) 福田大輔, 森地茂: 観光目的地選択行動に対する精緻化された個人選択モデルの適用可能性の検討, 土木計画学研究・講演集, No.22(2), pp.655-658, 1999.
- 7) 藤原章正, 杉恵頼寧, 岡村敏之, 神田佑亮: 選択肢間の類似性を考慮した時刻選択モデル, 土木計画学研究・講演集, No.22(1), pp.401-404, 1999.
- 8) Yai, T., Iwakura, S. and Morichi, S.: Multinomial probit model with structured covariance matrix for route choice behavior, *Transportation Research B*, Vol.31, No.3, pp.195 ~207, 1997.
- 9) 清水哲夫, 屋井鉄雄: Mixed Logit Model とプロビットモデルの推定特性に関する比較分析, 土木計画学研究・論文集, No.16, pp.587-590, 1999.
- 10) 屋井鉄雄, 中川隆広, 石塚順一: シミュレーション法により構造化プロビットモデルの推定特性, 土木学会論文集, No.604, IV-41, pp.11-21, 1998.
- 11) Yai, T. and Iwakura, S.: Route choice modeling and investment effects in Metropolitan rail network, *Conference Preprints of 7 International Conference on Travel Behavior*, Vol. 1, pp.363~373, 1994.
- 12) 森地茂, 屋井鉄雄: 非日常的交通への非集計行動モデルと選択肢別標本抽出法の適用性, 土木学会論文集/No.343, pp.161-170, 1984.
- 13) 岡本直久: 観光交通計画のための調査および分析手法に関する研究, 東京工業大学平成8年度博士論文, 1996.
- 14) 森川高行, 竹内博史, 加古裕二郎: 定量的観光魅力度と選択肢集合の不確実性を考慮した観光目的地選択分析, 土木計画学研究・論文集 No.9, 117-124, 1991.
- 15) 吉田朗, 原田昇: 選択肢集合の確率的形成を考慮した集計型目的地選択モデルの研究, 土木学会論文集/No.618, IV-43, pp.1-13, 1999.
- 16) Bolduc, D. and Ben-Akiva, M.: A multinomial probit formulation for large choice set, *6th International Conference of Travel Behavior*, pp.244-258, 1991.
- 17) Yai, T. and Shimizu, T.: Multinomial probit with structured covariance for several choice situations with similar alternative, *Transportation Research Record*, 1645, pp.68-75, 1998.
- 18) 室谷正裕: 新時代の国内観光; 魅力度評価の試み, (財) 運輸政策研究機構, 2000.

(2001.5.7 受付)

## ESTIMATION OF NON-IIA DISCRETE CHOICE MODELS WITH MEDIUM CHOICE SET

Tetsuo YAI, Tetsuo SHIMIZU, Koichi SAKAI and Akiko KOBAYASHI

The objective of the study is to analyze the effect of choice set on the estimation of non-IIA discrete choice models for two choice situation of railway route choice and day-return recreational destination choice model. Firstly, The day-return recreational destination choice model for non-IIA choice model is formulated using the idea of two elements of similarity. Secondary, we succeed to estimate multinomial probit model and mixed logit model for choice set with 10 alternatives. Finally, the effect of choice set on the estimation is deeply considered using actual and imaginary datasets.