

M-308

SPパネルデータにおける回答バイアスの修正

広島大学工学部

正会員 張 峻屹

広島大学大学院国際協力研究科

正会員 杉恵 頼寧

同 上

正会員 藤原 章正

1、はじめに

周知のとおり、個人の交通行動の時間的変化を捉えるにはパネル調査が有効である。しかし、パネル調査は同一個人を何時点も追跡するがゆえに、消耗バイアス、panel fatigue(PF)、panel conditioning(PC)バイアスなど、多くのバイアスが存在する。

パネルデータに存在する各種バイアスを排除するために、まず調査を工夫することが先決であろう。たとえばできるだけ少ない設問数、分かりやすい内容を使って被験者の回答時間を減らすこと、誤解を招く難しい言葉を使わないこと、グラフやコンピュータを活用することなどによって被験者の興味を引きつけることに努める必要がある。もう一つ無視できない重要な要素は適切なインセンティブを被験者に与えることである。

つぎに、調査を工夫しても排除しきれないバイアスに関しては行動モデルを構築する際にそれらを修正することが求められる。この場合、バイアス修正を行う目的は主に2つに分けられる。一つは、母集団代表性の欠如の補正、もう一つはデータの歪みの補正である。

本研究では広島市新交通システムの利用意識に関するSPパネルデータを用いて、non-commitment(NC)、PFとPCバイアスを含む回答バイアス、つまり、回答データの歪みを静的・動的Mass Point(MP)モデルにより修正することを試みる。

2、パネルデータ特有なバイアス

パネルデータ特有なバイアスとして、以下の3つを挙げることができる。

(1) 消耗バイアス：ある特性をもつ被験者が繰り返し設問への回答を拒否するために生じるものである。初期調査からの離脱サンプルによる消耗バイアスに関してはduration modelを利用した修正法、パネルの進行に伴って生じるものに関しては重みづけ法と修正項による方法がそれぞれ提案されている。ただし、重みづけ法では消耗行動と交通行動の独立性を仮定するのに対して、修正項法ではその相関性を仮定する。

(2) panel fatigueバイアス：被験者が複数回の回答を行う時の疲労によって生じるバイアスのことである。

(3) panel conditioningバイアス：過去にとった複数

回の回答行動が被験者の思考回路に影響を与えることにより生じるものである。

上述の(2)と(3)に関する修正方法はまだ十分に提案されていない。

本研究で扱うSPパネルデータの回答バイアスに関しては、冒頭で記述したようにNC、PFとPCバイアスを含む。その修正方法はまだ明らかではない。そこで、本研究では回答バイアスを個人の非観測異質性として扱うことを提案する。その理由は回答バイアスが個人間で異なると考えられるからである。

3、静的・動的Mass Pointモデルの構築

SPパネル調査では同一個人に対して複数時点にわたりそれぞれ複数のSPカードを提示して回答を尋ねる。したがって、回答バイアスに関する個人の非観測異質性は個人間のものと個人内のものに分けられる。前者は異なる個人が同じカードを回答する際に生じるもので、後者は同一個人が異なるカードを回答する際に生じるものである。本研究で前者のみを対象にモデル構築を行う。

MPモデルは個人間の非観測異質性を考慮する方法としてその有効性が実証されてきた。ここではNC、PFとPCバイアスの3者を一括して非観測異質性として扱う静的MPモデル、PCバイアスを前時点の選択結果の導入により修正して残りの2者を非観測異質性として扱う動的MPモデルをそれぞれ定式化する。

(1) 静的Mass Pointモデル

個人*i*の選択肢集合が外生的に与えられる場合、その同時選択確率は以下のように表される。

$$P_i^{\text{static}} = \sum_{k=1}^K \left(\prod_{t=1}^{T_i} \frac{\prod_{j=1}^{J_{it}} [\exp(\beta' x_{ijt} + \xi_{kj})]^{w_{ijt}}}{\sum_{j=1}^{J_{it}} \exp(\beta' x_{ijt} + \xi_{kj})} \right) \rho_k \quad (1)$$

ただし、*i, j, t*はそれぞれ個人、機関、時点を表す。 x_{ijt} 、 β は説明変数とそのパラメータ、 ξ, ρ はMPの位置と重みパラメータ、 w_{ijt} は選択肢グミー変数である。

(2) 動的Mass Pointモデル

ここではPCバイアスを修正するために前時点の選択結果 y_{ijt-1} をモデルのなかに取り入れた動的MPモデルを以下のように構築する。

キーワード：回答バイアス、SP、パネルデータ、Mass Pointモデル

* : 〒739広島県東広島市鏡山1-4-1 TEL 0824-24-7849 FAX 0824-24-7849

** : 同 上 TEL 0824-24-7826 FAX 0824-24-7826

*** : 同 上 TEL 0824-24-7825 FAX 0824-24-7825

表-1 静的・動的Mass Pointモデルの推定結果

説明変数	モデル	静的Mass Pointモデル				動的Mass Pointモデル			
		Card-1	Card-2	Card-3	Card-1+2+3	Card-1	Card-2	Card-3	Card-1+2+3
性別（男1、女0）		0.877**	0.720*	0.578	0.879**	0.907**	0.604	0.584	0.807**
年齢		-0.037**	-0.043**	-0.019	-0.039**	-0.036**	-0.038**	-0.024*	-0.037**
職業（あり1、なし0）		-0.651	0.149	-0.627	-0.229	-0.610	0.216	-0.533	-0.332
主要交通手段（自家用車1、その他0）		0.528*	0.682*	0.406	0.545**	0.465	0.565*	0.437	0.508**
世帯人数		0.262*	0.064	0.161	0.162*	0.247*	0.062	0.176	0.135
乗車時間（分）		0.006	-0.002	-0.010	-0.002	0.005	-0.002**	-0.008	-0.001
コスト（100円）		-0.019	0.018	-0.103*	-0.032	-0.037	0.020	-0.132*	-0.053
待ち時間（分）		-0.088**	-0.078**	-0.089**	-0.084**	-0.095**	-0.075**	-0.089**	-0.089**
アクセス時間（分）		0.004	-0.132**	0.023	-0.020	-0.007	-0.116**	0.021	-0.022
状態依存効果（ γ ）						0.472**	0.348*	0.488**	0.515**
収束Mass Point数 ρ との数		2 5	3 8	3 8	4 11	2 5	2 5	3 8	3 8
初期尤度		-602.04	-566.88	-540.52	-1709.40	-602.04	-566.88	-540.52	-1709.40
最終尤度		-537.49	-480.24	-470.47	-1502.21	-532.35	-480.01	-463.18	-1490.56
自由度調整済み尤度比		0.0591	0.0932	0.0650	0.0980	0.0646	0.1011	0.0754	0.1074
サンプル数		137×4	129×4	123×4	389×4	137×4	129×4	123×4	389×4

(*: 危険率5%で有意; **: 危険率1%で有意)

$$p_i^{\text{dynamic}} = \sum_{k=1}^K \left\{ \frac{\prod_{j=1}^{J_{i1}} [\exp(\beta' x_{ij1} + \xi_{kj})]^{w_{ij1}}}{\sum_{j=1}^{J_{i1}} \exp(\beta' x_{ij1} + \xi_{kj})} \prod_{t=2}^{T_i} \frac{\prod_{j=1}^{J_{it}} [\exp(\beta' x_{ijt} + \gamma y_{ijt-1} + \xi_{kj})]^{w_{ijt}}}{\sum_{j=1}^{J_{it}} \exp(\beta' x_{ijt} + \gamma y_{ijt-1} + \xi_{kj})} \right\} \rho_k \quad (2)$$

ただし、 γ は状態依存パラメータである。

無論、前時点の選択結果以外に、たとえば前時点の選好を表す効用関数を取り入れたり、普段の利用交通手段ダミーを取り込んだりすることも考えられるが、ここではそれらの比較について検討しない。

4、静的・動的Mass Pointモデルの推定と考察

(1) 利用するSPパネルデータの概要

本研究室は広島市新交通システム開業後における通勤通学者の交通機関選択行動の予測を目的に、1987年から88、90、93、94年にかけて計5時点にわたってSPパネル調査を実施した。今回はこのうちサンプリング手法の異なる1988年を除く4時点継続的に得られたパネルデータを利用する。対象交通手段は自家用車、バス、新交通システムである。この4時点において共通して3枚のSPカードを被験者に提示した。各カードの有効サンプル数はそれぞれ137、129、123であった。

(2) 静的・動的Mass Pointモデルの推定

ここではSPカードごとに、そして3枚のカードをプールしたうえでMPモデルをそれぞれ推定した。その結果を表-1に示す。

まず、全体的に動的MPモデルは静的MPモデルより尤度比が高い。そして、状態依存効果パラメータ γ が有意であった。これは前時点の選択結果をモデルに取り入れ、PCバイアスを明確に修正することによりモデルの精度を向上させ、回答バイアスの修正効果が上がったことを意味する。

つぎに、推定パラメータ値がSPカード間に差があるかどうかを χ^2 検定により検証する。その結果、静的MPモデルでは χ^2 値 = $-2\{-1502.21 - (-537.49 - 480.24 - 470.47)\} = 28.02 < \text{限界値} 41.34$ (自由度=14+17+17-20=28、

有意水準95%)、動的MPモデルでは χ^2 値 = $-2\{-1490.56 - (-532.35 - 480.01 - 463.18)\} = 30.04 < \text{限界値} 43.77$ (自由度=15+15+18-18=30、有意水準95%)となった。つまり、推定パラメータ値はカード間で有意な差が見られない。しかし、収束MP数がカードごとに違うため、3枚のカードを単純にプールすることが望ましくない。これに関する議論は紙面上の制限でここでは割愛する。

(3) 静的・動的Mass Pointモデルの現況再現性

静的・動的MPモデルが実際のSP回答をどのぐらい再現できるかを評価するために平均値法により計算した推定分担率と実際分担率との絶対誤差を表-2に示す。

表-2 推定分担率の絶対誤差

モデル	静的pool	動的pool	静的MP	動的MP
Card-1+2+3	0.4712	0.4562	0.4538	0.4500
Card-1	0.4996	0.4667	0.4498	0.4331
Card-2	0.6148	0.6177	0.6028	0.5960
Card-3	0.3656	0.3845	0.3641	0.3806

表-2の中に静的・動的poolモデルはそれぞれ4時点パネルデータをプールし、選択肢固有定数を取り入れたモデルのことで、静的・動的MPモデルはそれぞれ表-1に示されたものである。全体的に動的モデルは静的モデルより、MPモデルはpoolモデルより現況再現性が高いことが分かった。

5、おわりに

本研究ではSPパネルデータに存在する回答バイアスをMPモデルにより修正することを試みた。その成果と今後の課題を以下のようにまとめることができる。

MPモデルは回答バイアスを修正する手法として有効である。そして、前時点の選択結果を用いてPCバイアスを明確に修正することにより回答バイアスの修正効果を向上させることができる。最後にSP調査では提示したカードごとに収束MP数が異なることから、個人内の非観測異質性を考慮したモデルの開発が必要であろう。これに関しては今後の研究課題としたい。