

# マーケティング実務での シングルソースデータの利用

八木 滋

## 1. はじめに

わが国でも日用消費財パッケージ商品のソースマーケティング率が100%に近づき、その詳細性、迅速性、正確性から近年POSをはじめとするスキャナーデータのマーケティング実務での活用が盛んである。本稿ではスキャナーデータのうち消費者単位にデータが整理されたスキャンパネル、なかでも調査世帯のテレビ広告露出状況も同時に調査されているシングルソースのスキャンパネルである「VRホームスキャンシステム<sup>(R)</sup>」のデータがマーケティング実務のどのような場面でのどのように活用されているのか、または活用できるのかを事例データを中心に紹介する。

## 2. VRホームスキャンシステム (VRhomeScan system)

この調査は、(株)ビデオリサーチが1987(昭和62)年4月から本格稼働させた世帯ベースのシングルソースデータ(商品の購買行動とテレビ広告接触状況が同一調査標本で測定され、購入時点あるいは非購入時点の商品プロモーション状況がわかる)調査である。この調査は、パネル(調査対象者)がPOS導入店での商品購買時にIDカードを使用するストア・スキャン方式とは異なり、調査世帯に設置されている入力装置とスキャナーによりすべての店舗での購入商品と付帯情報を記録するホーム・スキャン方式をとっている。POS導入見込みの低い小規模店の多いわが国の商品流通事情のもとでも、この方式によって一般消費世帯での商品購入履歴全体をカバーできることになる。

また、調査世帯には「視聴率調査」と同じ仕様のオンライン・メーターが設置されており、商品購買と同一世帯でのテレビ視聴実態が把握されている。(放送やテレビ

広告業界で広く使用されている「視聴率」調査は、全く別の世帯で行なわれている)

### 2.1 調査の設計

調査エリア	首都圏内の仮想ミニマーケット (半径2km 約5万世帯)
調査世帯	家族人数2人以上で主婦のいる世帯
調査対象者	上記世帯内全員
調査世帯数	1,000世帯
調査内容	(1)ソースマーケティング商品の購買実態 (購入日、購入者、購入店、購入商品、 購入回数、購入金額) (2)テレビ視聴実態
調査方法	VRホームスキャンシステムによる コンピュータオンライン調査
補完データ	テレビCM出稿実態データ 新聞折込広告実態データ インスタプロモーション実態データ 調査地域内の出来事・天候・気温
調査の開始	1987(昭和62)年4月

### 2.2 調査の特徴

この調査のホーム・スキャン方式は、(1)消費世帯の商品購入履歴に購入店の制約を受けず、ブランドチェンジや商品の購入経路分析、新製品のトライアル/リピート分析に威力を発揮する。購買行動と同時にテレビ視聴行動も測定されているので、(2)購買行動とテレビ広告露出との関係を詳しく研究・分析することができる。また、調査エリア内の主要店舗(スーパーマーケット)でのプロモーションや新聞折込広告による商品の販促情報データは整備された購買履歴データと相まって、(3)プロモーション効果の研究(ブランドチェンジ、店舗選択、買い急ぎ/買いため/買い控え etc.)を可能なものにする。

特に、(2)と(3)については従来の集計データ間の時系列相関分析だけでなく、直接「クロス集計」が可能である。検証の難しかった広告効果研究にも新たな切口が用意さ

やぎ しげる (株)ビデオ・リサーチ 消費者分析部

〒104 中央区銀座2-16-7

れることになる。

この調査データの首都圏全域の代表性には議論の余地が残るが、シングルソースデータであることによりマクロなビジネストラッキングデータには難しかったマーケティング戦略仮説の検証に適している。

### 3. データ利用の現状

#### 3.1 マーケットと消費世帯分析

この調査は世帯単位の商品購買データを提供するので当然マーケットの趨勢や消費者（購買家庭）の特徴分析が行なわれる。その場合、次のような観点から眺めてみることが多いようだ。(1)関連する周辺のマーケットとの消費世帯の重複、(2)マーケットの規模・購買サイクル、(3)購入経路(業態)・購入価格分布、(4)競合各社のマーケティング(広告・プロモーション)活動状況、(5)市場構造・製品競合、(6)デモグラフィックス(購買世帯、テレビ広告露出世帯)、(7)購買行動特性によるセグメンテーションと各セグメントの特徴把握、(8)マーケットトレンドなどである。

たとえば、関連する周辺マーケットとの重複の分析を6カ月間のデータを使った「マーガリン」と「バター」をみた場合、両方の商品のいずれかを購入した世帯の49.4%が「マーガリンのみ」、12.2%「バターのみ」、38.4%「両方」となっている。これは「マーガリン」マーケットからみれば消費世帯の重複は43.7%にすぎないが、「バター」マーケットからみれば重複率は大きく75.9%になる。

また、「マヨネーズ」と「ドレッシング」のケースでは同様に、28.8%「マヨネーズのみ」、8.0%「ドレッシングのみ」、63.2%「両方」となっている。「ドレッシング」マーケットからみた消費世帯の重複(88.8%)の方が大きい、いずれにしてもこの2つの製品カテゴリー購入世帯の重複はかなり大きい。この事実と商品の消費目的を考えれば、この2つの商品種類はひとつのマーケットに属しているものという前提でマーケティングを考えるべきであろう。

購入経路や購入価格の分析では、こんな例がある。ある新製品の発売直後はGMS、SMでの購入が多かったが、その後ディスカウンターでのシェアが急増加する現象が発見されたことがある。購入価格分布を調べてみると、やはりこの時期から値崩れが始まっており、犯人はこのディスカウンターであった。このような出来事が意図的に計画されたものか許容される範囲のものであれ

ば事実は事実として報告されるだけで終るが、そうでない場合には報告が早ければ早いほど貴重な警告情報となるだろう。

市場構造に関しては、世帯単位の購買履歴データから図1のような分析が可能である。この事例でのデータ処理はまず購買商品の推移マトリックスを作成する、その結果から製品間相関係数を計算の後クラスター分析をして、そのデンドログラムを解釈している。この「一般歯磨き」マーケットをグラム当り価格を手がかりに解釈すると、低価格品と高価格品マーケットに大きく2分割されているようである。また、高価格品マーケットは「歯石防止」、「塩(歯茎をひきしめる)」、「ヤニとり」、「COOP(ノンフォーム/界面活性剤なし)」、「口臭防止」などの効用によってサブマーケットが形成されているのに対して、低価格品マーケットはブランドネームによってまとまりをみせているが、図中の「?」クラスターは店頭プロモーションが交互に行なわれることが多いことなどからくる特異クラスターと解釈している。市場構造がこのとおりであるとすれば、新たな効用を持った高価格新製品は高価格帯マーケットに新たなサブマーケットを開拓することができることになる。この場合、あくまでその新たな効用が消費者に受け入れられるという前提である。しかし、低価格品マーケットに対してはどのようなマーケティング手段が有効であるのだろうか。頭の痛いところである。この市場構造の分析もただひとつの分析結果や分析手法だけに頼るのは、はなはだ不安である。たとえば、世帯ごとの製品別購入回数データで主成分分析をして、同じような考案が出れば一応信用してもいいのではないだろうか。また、場合によればヘンドリー・モデルやプロデジャー・モデルなどによってその市場構造を検証することも必要であるだろう。

デモグラフィック・プロフィールは、単に「購入」世帯の特徴を製品属性やブランド別、メーカー別に比較するためだけでなく、「広告の露出」世帯との比較をすることによってテレビ広告スケジュールの実績評価にも使用することができる。

図2はあるレベル以上の購入頻度をもつ世帯とある頻度以上のテレビ広告露出世帯のプロフィールを比較したものである。図をみれば、非常に特徴のある消費世帯プロフィールとこれといって特徴のないテレビ広告露出世帯プロフィールが一致していないことがよくわかるだろう。これはテレビ広告の到達効率の悪さを窺わせる結果であり、このような場合テレビ広告キャンペーン全体の

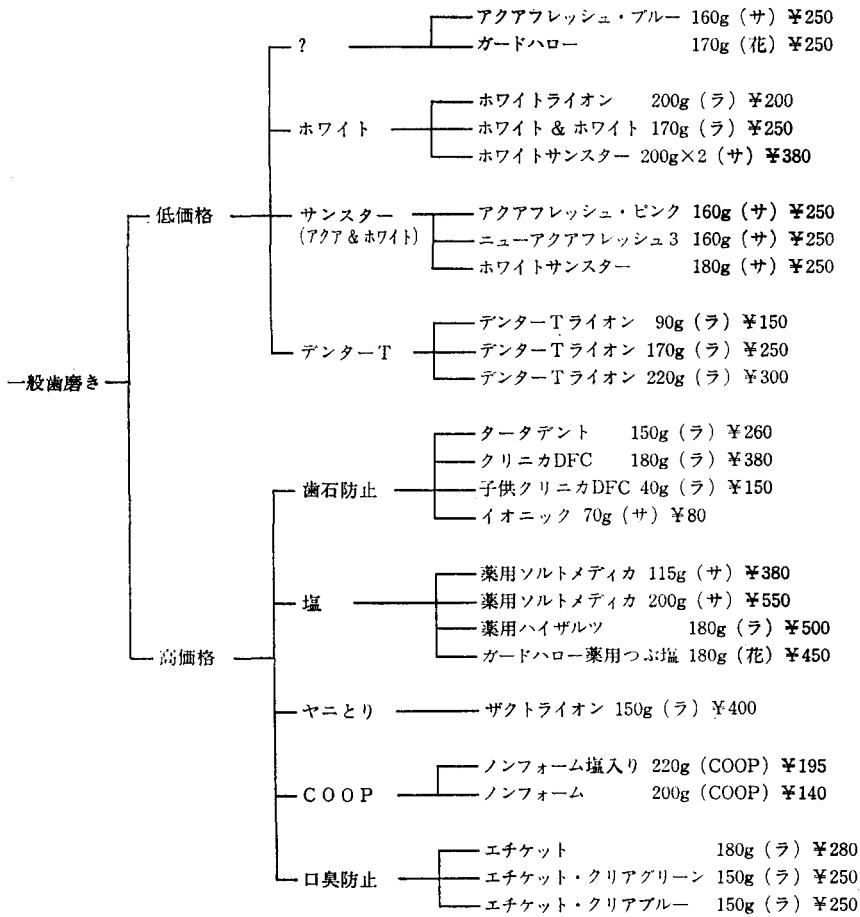
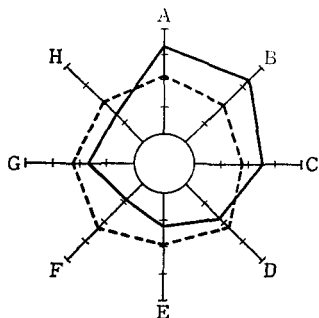


図1 一般歯磨きの市場構造 注：( )内はメーカー名

放送時間を見直す必要がある。

デモグラフィックな消費出帯の属性分析ばかりでなく、過去の購買履歴からの特徴分析も可能である。たとえば、購入頻度や購入量からヘビー/ライトユーザーの分布、購入ブランドの継続状況からロイヤル/スイッチ

ャーの分布、スイッチ状況からさらにバラエティーシーキング度、購入時点の価格やプロモーション状況から価格感度やプロモーション被誘引度、購入インターバルの



凡例

- A 乳幼児のいる世帯
- B 4歳以上の未就学児のいる世帯
- C 小学生のいる世帯
- D 中・高生のいる世帯
- E 高校生以下の子供のいない世帯
- F 19~25歳男性のいる世帯
- G 19~25歳女性のいる世帯
- H 60歳以上の高齢者のいる世帯

- 商品購入率インデックス
- ..... テレビCM露出インデックス

図2 商品購入世帯とテレビCM露出世帯のプロフィール比較 (ある加工食品の場合)

規則性、など購買行動からの多様なセグメンテーションやそのデモグラフィックプロフィール分析が可能である。現在のマーケットやブランドにこれらユーザーグループがどう分布しているのかを知ることは今後のブランドマーケティングを考える上で重要な情報となる。また、新製品が各ユーザーグループにどのように受け入れられているのかを知ることは、その新製品にとっての将来を占う上で重要な鍵になるだろう。

### 3.2 新製品トラッキング

新製品を送り出した後、その市場浸透スピードや購入世帯プロフィール、購入世帯の前購入ブランドをいち早く知ることはそれだけでも日々のブランドマネージメントには有益なことである。(VRホームスキャンは、翌日朝には調査世帯の購買データを分析することが可能である。)そのとき、商品の店頭配荷状況や競合ブランドを含めてテレビ広告やプロモーションの実施状況がわかれば、すでに実施に移されている一連のマーケティング計画になんらかの調整や修正が可能となる。もちろんこのような事例はその後の重要な参考資料として利用されるべきものであることは言うまでもないだろう。

図3は、ある製品の新発売週以降の購入世帯率とテレビ広告露出量、プロモーション状況の推移を示したものである。図からも明らかなようにその時々々の広告やプロモーションが購買行動を促進している様子がみてとれるだろう。

図4は、新たなテレビ広告表現(素材)が購買行動を促進している例である。グロスのテレビCM月別露出量(GRP)は3か月目以降減少傾向が続くが、トライアル購入率は同じパターンではなく微妙に変化している。これをCM表現別月別露出量と対比してみると、それぞれCM表現がその露出量の最も多い月にトライアル購入率を押し上げているのがみてとれるものと思う。テレビCMのオンエア寿命は数週間から数か月のものが多いとみられているが、消費者の購買意欲の喚起のためにもその程度のライフサイクルが妥当なのかなという感じを抱かせる結果である。

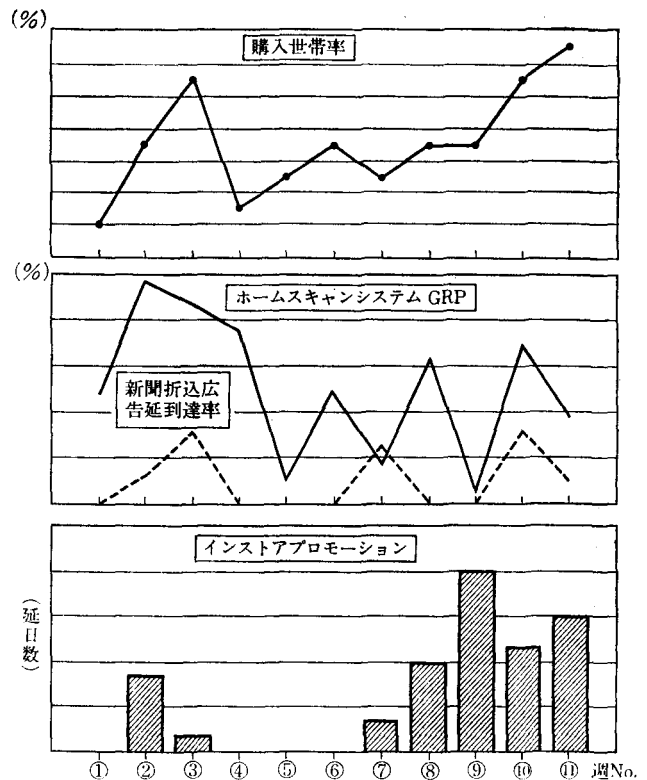


図3 新製品のテレビ広告・プロモーション状況と購入率の推移(ある加工食品の場合)

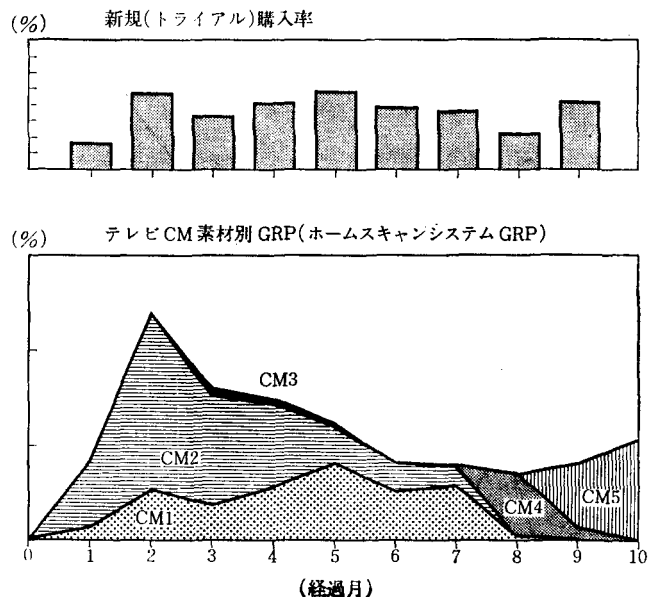


図4 新製品のテレビCM素材別GRPと新規購入の推移(ある日用雑貨品の場合)

### 3.3 広告とプロモーション効果

広告やプロモーションの効果分析をする場合、(1)各種の調査結果を時系列に並べ相互の関係を分析する方法、(2)一定期間のデータをクロス集計する方法、(3)個票データを使った方法の3種類に分類することができるのではないだろうか。

まず、(1)の集計データを用いた時系列回帰分析の例を紹介する。これは、ブランドの購入世帯率の変化を自ブランドと競合ブランドそれぞれのテレビ広告GRP、新聞折込広告の延べ到達率、新聞折込広告の価格、店頭プロモーション延べ日数で説明しようとするものである。

回帰式は、

$$PR_{1t} = \alpha_0 + \beta_1 NI_{1t} + \beta_2 NI_{2t} + \gamma_1 NIP_{1t} + \gamma_2 NIP_{2t} + \delta_1 ISP_{1t} + \delta_2 ISP_{2t}$$

ただし、ここでは

$PR$  ; 購入世帯率

$NI$  ; 新聞折込広告延べ到達率

$NIP$  ; 新聞折込広告掲載価格(ただし、基準容量  
当り換算値)

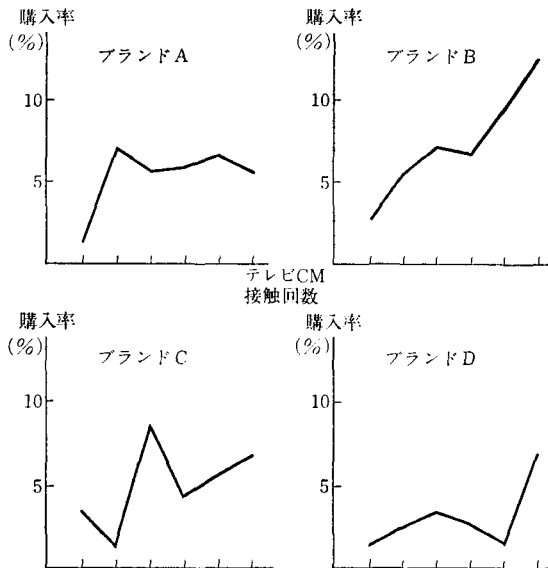


図6 新製品発売直後のテレビCM露出回数区分別購入率(加工食品4ブランドの場合)

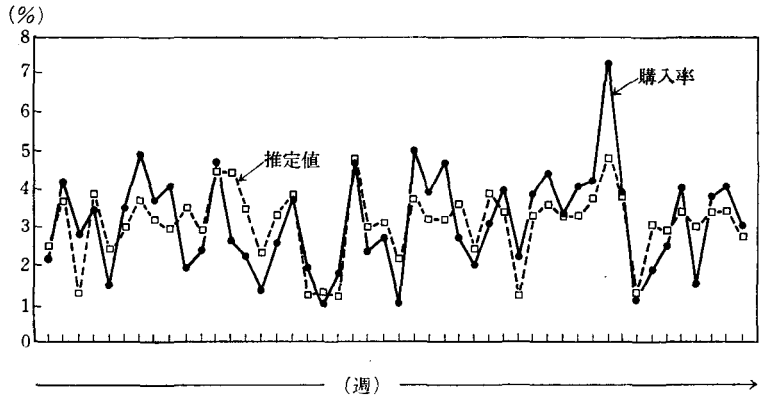


図5 週別購入率の重回帰推定結果(ある加工食品の場合)

$ISP$  ; 店頭プロモーション延べ日数

$t$  ; 週

1, 2 ; ブランドNo. (ブランド1と2は強い競合関係にある)

この回帰式に1987年4月6日(月)以降47週間のデータを適用してみると重相関係数.707(決定係数.500)となり、各時点(週)の観測値と推定値は図5のとおりである。このケースでは競合ブランドを主要な1ブランドに限定していることや年末の特別大売り出し期間(週)についての配慮をしていないことなど今後の改良点がいくつか残されている。また、この手法はマーケットの消費世帯や商品流通事情が同質であり、どの世帯もあるマーケティング刺激には常に同じ反応をすることが仮定されている。

次に、(2)のクロスセクションデータ分析の例である。一般にテレビ広告の購買行動への効果は測定しづらいものであると認識されているが、それは先のような集計データを用いた分析の場合であることが多い。ここで紹介するのは新製品に限定しているが、個票データによるテレビ広告の露出回数と購買行動のクロス集計分析である。この2つの変数間のカロス集計には2つの切口がある。購入経験別CM露出回数の分布または平均とCM露出別回数購入経験の有無である。購入経験の有無の分布は購入率に置き換えることができるが、CM露出回数は世帯毎にみればかなり大きなバラツキをみせるので平均値で結果をみるのは危険である。先の2つの切口のうち前者は適切な方法とは言えないだろう。ここで紹介するのは、CM露出回数をいくつかの「区間」にまとめてそのCM露出回数「区間」別購入率を計算した例である。

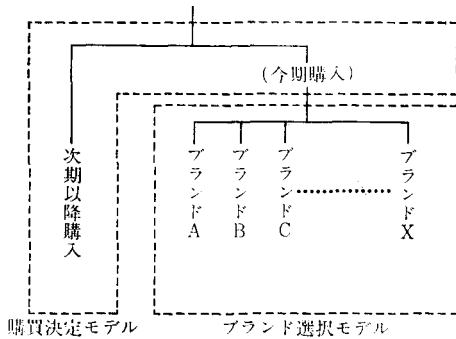


図7 モデルの枠組み

図6は、新製品の発売直後数週間でのテレビCM露出回数区分別商品購入経験率のグラフで、どの図も横軸のCM露出回数区分は同じで右に行くほど露出回数が多いことを示している。

この4商品のケースをみると、少しいびつではあるが各折れ線ともに右上がりテレビCMの露出回数が多くなるほど購入率が高くなっており、テレビCMの購買促進効果が確認される。また、ブランドBの場合はCM露出回数の増加につれて購入率が上昇しており購買促進効果の飽和を感じさせないが、ブランドAの場合はかなり少ないCM露出回数で購入率の横ばいが始まっている。このような違いは製品そのものからくる原因ばかりでなくCM表現要素も関わりが大きいと考えてはいるが、まだ実証的に確認されたわけではない。

最後に(3)の個票データを使った事例である。この事例は片平(1989)「シングルソースデータを用いた広告/プロモーション効果の測定」を基礎にするもので、いわばプロモーションの消費世帯レスポンスモデルである。このモデルは、図7に示したように階層的に組み合された2つのモデルによってできている。

まず、ブランド選択モデルはその週に購買が行なわれたことを条件として次のように規定される。

$$P_{jtk} = \theta_{jk} \exp U_{jtk} / \sum_h \theta_{hk} \exp U_{htk}$$

$$U_{jtk} = \beta_{1k} d_{jtk} + \beta_{2k} c_{jtk} + \sum_{h=1}^{n-1} \gamma_{hk}$$

ただし、

$P_{jtk}$ : 家計  $k$  が  $t$  週にブランド  $j$  を選択する確率

$U_{jtk}$ : 家計  $k$  からみた  $t$  週のブランド  $j$  の魅力度

$d_{jtk}$ : 家計  $k$  が  $t$  週に直面するブランド  $j$  の価格の掛け率 (平常価格を1とする)

$c_{jtk}$ : 家計  $k$  の一番よく行く店で  $t$  週にブランド

$j$  のプロモーションを行なっていれば1、それ以外0をとる0-1変数

$\gamma_{hk}$ : 家計  $k$  のブランド  $h$  に対する固有の効用 (パラメーター)

$\theta_{jk}$ : ブランド  $j$  が家計  $k$  の選択集合に入る時1、それ以外0をとる選択集合規定子 (この事例では使用データ期間の購入ブランド群に含まれるかどうかで判断した)

$\beta_{1k}, \beta_{2k}$ : パラメーター

また、購買決定モデルは次のように規定される。

$$Q_{tk} = \exp V_{tk} / (1 + \exp V_{tk})$$

$$V_{tk} = \alpha_{0k} + \alpha_{1k} \tilde{U}_{tk} + \alpha_{2k} I_{tk}$$

ただし、

$Q_{tk}$ : 家計  $k$  が  $t$  週に当該カテゴリーを購買する確率。本モデルでは各家計が各週ただか1回1単位だけ購買を行なうものと仮定している。

$V_{tk}$ : 家計  $k$  の  $t$  週の購買性向

$\tilde{U}_{tk}$ : ( $= \text{Max. } \hat{U}_{jtk}$ ) 家計  $k$  が  $t$  週に当該カテゴリーの購買から得られるであろう最大効用。ブランド選択モデルで得られる  $\hat{U}_{jtk}$  から計算される。

$I_{tk}$ : 家計  $k$  の  $t$  週の家計内在庫。当該家計の1週間の消費量を1単位としてその家計の購入ヒストリーから「計算」される(計算法は省略)。

$\alpha_{0k}$ : 定数項パラメーターで家計  $k$  の  $\tilde{U}_{tk}, I_{tk}$  では説明できない購買性向の強さを表わしている。

このモデルを適用したデータは、1987年4月から1988年3月の52週間のインスタント・コーヒーである。モデル計算にあたっては消費世帯をその最多購入ブランドによってセグメンテーションし、セグメントごとに最尤法によりパラメータ推定を行なった。ブランド選択モデルの結果は図8と図9に示したが、説明変数が少ないにもかかわらず最終的なブランド選択回数シェアはタイトルの不一致係数( $U$ )が.068とかなりの精度で予測できた。購買決定モデルの結果については紙面の関係で省略させていただく。このモデルとその運用方法にはまだ改良の余地があり、さらに研究を続けたいと考えている。

## 4. おわりに

### 4.1 残された研究テーマ

すでに紹介したようにこのVRホームスキャンデータを使ってこれまでできなかったマーケティングの実証研

究や実務利用が進んでいるが、着手したばかりあるいはこれから着手しようとしている研究テーマがまだいくつかある。たとえば、テレビCMに関するものだけでも、有効フリークエンシー、表現方法による効果の差、秒数による効果の差、番組/スポット種別による効果の差、CMの疲労、複数ブランド・企業広告の効果、長期効果などがある。また、テレビ以外のメディア広告効果も重要なテーマとなっている。

#### 4.2 システム拡張

この調査システムがスタートして2年が経過したが、システムとしてまだ完成されたものではない。今後もデータ処理やアプリケーション研究と平行して、システム自体の改良や拡張が検討されている。現在検討されている主な課題としては、(1)シングルソースとしての完結のために、個人単位の視聴データ測定方法や広告・プロモーションデータの充実、(2)多様な分析次元の準備として消費世帯や商品属性情報の充実・拡大、(3)いくつかの What if に答えるためにサーベイエリアの追加、などがある。

#### 参考文献

片平秀貴：シングルソースデータを用いた広告/プロモーション効果の測定。財団法人 吉田秀雄記念事業財団，昭和63年度助成研究集（第22次）1989年

○は観測された選択ブランド

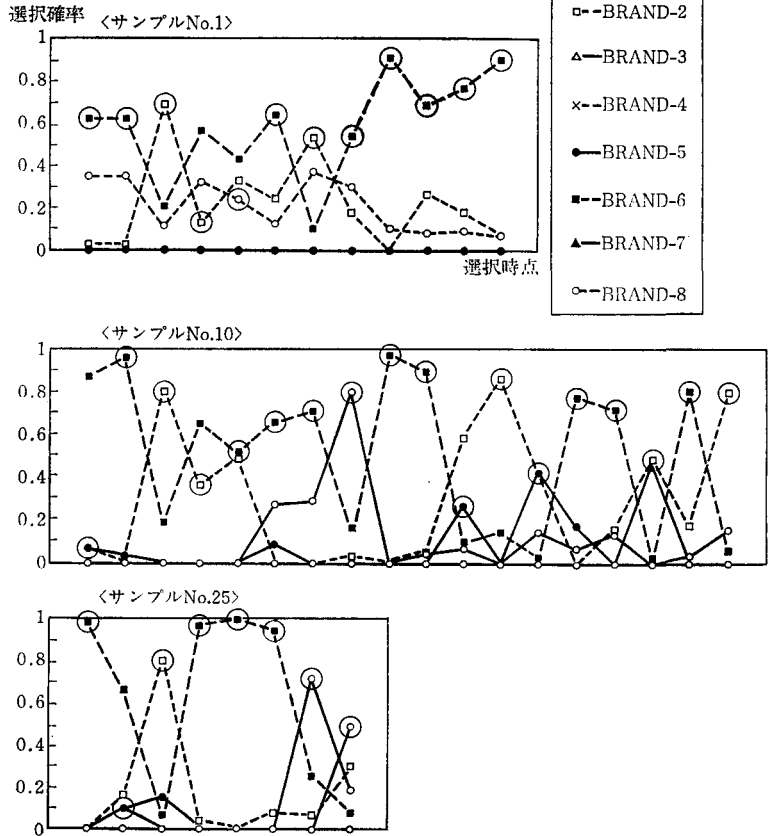


図8 調査世帯の実購入ブランドとモデル推定されたブランド別選択確率

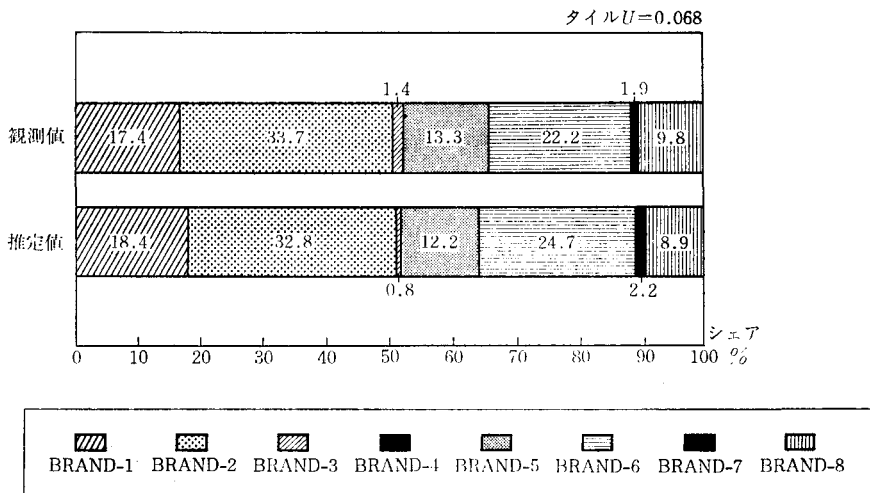


図9 1/2の世帯をホールアウトサンプルとして行なったシミュレート結果