

飲用乳関連メディア情報による消費者便益への影響と その変化要因に関する定量的研究 —スキャナーパネルデータによるアプローチ—

研究代表者 筑波大学 氏家清和

要旨

本稿では、近年成長が著しいヨーグルト市場を事例に、スキャナーパネルデータならびにヨーグルトに関する新聞記事数を組み合わせてデータセットを作成し、ヨーグルトの市場構造ならびにヨーグルトの需要関数を分析した。市場構造については、アソシエーション分析の一手法である Jaccard 係数による商品ごとの顧客の重複度合いを手掛かりに、クラスター分析等によりアプローチを行った。また、ヨーグルトについての需要関数については、0 消費データが多数みられることから、トービットモデルを需要関数として採用したとともに、階層ベイズの手法により、消費者の選好異質性についてもモデルの中に組み入れて分析した。

分析の結果、小容量型の商品は、様々なメーカーの商品が多数含まれており、競争が激しいことが示唆された。その中で、一部のメーカーは比較的ロイヤルティの高い顧客を獲得して、競合関係からやや距離をとることに成功していることが示された。また、特定保健用食品としての表示も、差別化要因として機能しているようであった。ヨーグルトの需要関数については、階層ベイズモデルとしてトービットモデルを推定した。階層ベイズモデルによる消費者の選好異質性は、定数項、価格係数、記事数係数、トレンド項で存在することが示され、全体的な傾向だけではなく、個人間の差異について注目する価値があることが示された。ただし、個人パラメータと世帯属性との関係性については決定係数が低く、十分解明できなかった。また、記事数の影響についても十分に検出することができなかった。これらの点は今後の課題としたい。

1. 緒言

近年、消費者の健康志向を訴求する食品が多数市場規模を伸ばしている。表示制度についても、特定保健用食品制度に加えて機能性表示食品制度も創設された。特に、ヨーグルトには機能性表示食品制度に届け出がされた商品が多数見受けられ、健康志向食品として一つ代表的なものとして捉えられる（註1）。

本稿では、ヨーグルトについてのスキャナーパネルデータを分析し、商品ごとの購入者の重複程度から、ヨーグルト市場の構造を分析する。またヨーグルトの需要関数を推定し、マスメディアによる情報提供がどのようにヨーグルトの消費に影響を与えるのか検討する。

2. 背景

1) ヨーグルトの概要

日本で初めて本格的なヨーグルトの工場生産が始まったのは1950年のことであり、その当時アメリカでヨーグルトの健康効果に対する関心が高まったことが日本にも伝わり、健康食品として盛んに食されるようになった（齋藤ら、2016¹⁾ : p. 32）。

（註1）消費者庁が公表している機能性表示食品制度の届出情報において、「加工食品（その他）」の分類の中で、ヨーグルト商品は53件あり、これは同分類中の約8%を占めている。

FAO（国連食糧農業機関）とWHO（世界保健機関）によるコーデックス規格において、ヨーグルトは発酵乳の一種として位置づけられており、「サーモフィラス菌とブルガリア菌を使用して乳酸発酵された食品」と定義されている。

日本では、ヨーグルトに関しては上記のような定義づけはなされておらず、同様な食品を乳等省令（註2）によって「発酵乳」として広く定義している。発酵乳とは、「乳又はこれと同等以上の無脂乳固形分を含む乳等を乳酸菌又は酵母で発酵させ、糊状又は液状にしたもの又はこれらを凍結したもの」と定義されている。これは、コーデックス規格とも整合性を保っている。

2) 市販ヨーグルトの草分け

現在日本では、様々なタイプのヨーグルトが販売されており、それらは形状や製造方式の観点から主に5つに分類される（齋藤ら、2016¹⁾ : p.45-47）。

1950年に発売されたヨーグルトは「ハードヨーグルト」と呼ばれるもので、ヨーグルト独特の酸味を抑えるよう砂糖や甘味料など多くの添加物を加え元来のものよりも食べやすく加工されており、当時はおやつ感覚で食されていた。

1969年には、ハードヨーグルトよりも固さを抑え果肉や果汁を入れた「ソフトヨーグルト」が販売された。一般的にフルーツヨーグルトと呼ばれるものはソフトヨーグルトに分類される。1971年には、「プレーンヨーグルト」という伝統的なブルガリアヨーグルトに近いものである。砂糖や甘味料などが入っておらず健康志向や自然志向の強い消費者に徐々に受け入れられるようになった。

1977年には、固形のを液状にした「ドリンクヨーグルト」が発売され、現在ではその手軽さが好まれている傾向にある。1979年には、ヨーグルトを冷凍しアイスクリーム状にした「フローズンヨーグルト」が販売されるようになった。

また、最近では「機能性ヨーグルト」という言葉が徐々に浸透している。主だった定義づけはなされていないが、これは一般的にプロバイオティクス乳酸菌を含むヨーグルトを指している（註3）。

プロバイオティクスとは、1989年にフラーにより定義された言葉で、「腸内フローラ（細菌叢）のバランスを改善することにより、宿主の健康に有益に働く生きた微生物」を指すとされている。つまり、機能性ヨーグルトとは通常のヨーグルトの生理効果に加え、プロバイオティクス乳酸菌によるその他の生理効果が付随したヨーグルトであると捉えられる。プロバイオティクス商品の主な生理効果としては、便秘・下痢の防止、腸内細菌のバランス改善、動脈硬化の改善など多岐にわたる（齋藤ら、2016¹⁾ : p.43）。

3) 日本におけるヨーグルトの市場動向

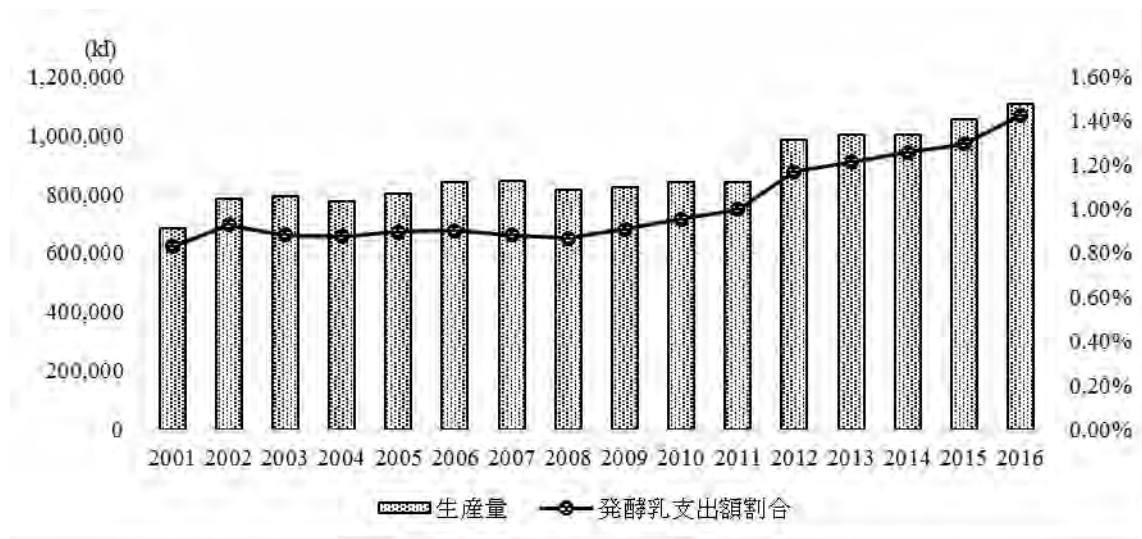
第1図は、2001年から2016年の発酵乳生産量と家計の食料支出における発酵乳支出割合の推移を示しており、両者ともに増加傾向にある。

このようなトレンドの背景には、ヨーグルト市場における新商品の発売による新たな価値提供

（註2）正式名称は、「乳及び乳製品の成分規格等に関する省令」である。

（註3）プロバイオティクス乳酸菌以外にも、プレバイオティクスやシンバイオティクス乳酸菌を使用したものも機能性ヨーグルトとして捉えられている。これらについても齋藤ら（2016）¹⁾が詳しい。

や、特定保健用食品やプロバイオティクス商品の登場による健康イメージの向上、およびその効果や効能が各種マスメディアに取り上げられたこと等に起因していると考えられる(齋藤ら、2016¹⁾: p.43)。



第1図 ヨーグルトの生産量及び家計支出額割合

資料：農林水産省「牛乳乳製品統計調査」(2018年1月)及び総務省「家計調査」(2017年6月)より作成。

3. 目的

以上みたように、ヨーグルトの市場規模は年々増加している。家計調査においてはヨーグルトの支出金額は公表されているものの、購入数量ないしは容量については公表されていない。さらに、ヨーグルトの国内需要に関する分析は広く行われていない現状にある。

そこで、本稿では長期にわたるヨーグルトのスキナーパネルデータを分析し、ヨーグルトの市場構造を分析するとともに、ヨーグルトの消費行動とマスメディアによる情報提供の関係性、ならびに、年齢や性別、家族構成といった世帯属性との関連性について検討する。

4. 分析データ

1) 概要

本稿では、消費者購買履歴データとしてインターネットリサーチ会社である株式会社マクロミルが提供している『QPR』データを用いた。このQPRデータは、スキナーパネルデータの中でもホーム・スキニング・データに分類される(註4)。このQPRデータには、「どのような人が、何を、どれだけ、いくらで」商品を購入したのかの記録が、その商品のJANコード並びにモニターのデモグラフィック属性とともに蓄積されている。

本稿で取り扱うデータは、ヨーグルトのみに関する購買履歴が記録されており、期間は2011年5月から2017年6月までである。なお、モニター数は、より簡潔な解釈を得るために、2012年1月から2016年12月の期間中モニター登録を継続しているモニターの購買履歴データを抽出し、

(註4) スキナーパネルデータの構造については、青木(1993)²⁾が詳しい。また、国内でのスキナーパネルデータを使用した分析手法については、田口(2010)³⁾がコンパクトにまとめている。

さらにその内、世帯年収が不明、家族構成が不明瞭な回答者を取り除いた 11,855 名である。また、商品に関しても一部情報が欠落している商品の購買履歴を取り除いたため 4,141 品目となった。

第 1 表 年別階級別世帯年収の世帯数割合

年	(万円)												
	-200	200-300	300-400	400-500	500-600	600-700	700-800	800-900	900-1000	1000-1200	1200-1500	1500-2000	2000-
2012	6.2%	9.1%	13.2%	14.3%	14.4%	10.3%	9.0%	6.3%	6.3%	5.6%	3.3%	1.2%	0.7%
2013	6.2%	9.0%	13.2%	14.4%	14.7%	10.7%	8.8%	6.3%	6.0%	5.3%	3.2%	1.4%	0.8%
2014	6.1%	9.1%	13.5%	14.3%	14.5%	10.3%	9.2%	6.5%	5.8%	5.4%	3.2%	1.4%	0.8%
2015	6.1%	9.5%	13.2%	14.5%	14.1%	10.0%	9.2%	6.6%	5.8%	5.2%	3.4%	1.3%	0.9%
2016	6.7%	9.2%	13.2%	14.5%	13.3%	10.4%	9.5%	6.5%	5.3%	5.7%	3.2%	1.5%	0.9%

また、ヨーグルトに関する全国的なメディア情報が需要に与える影響をとらえるため、同期間内に朝日新聞に掲載された記事数も分析対象に加えた。具体的には、朝日新聞記事データベース『聞蔵Ⅱビジュアル』にて、「ヨーグルト」というキーワードを含む記事数を入手した。なお、対象とした記事は、すべて全国紙面に記載されているものである。

2) 対象データの記述統計的特徴

本稿の分析に用いる変数についていくつかの単純集計を行った。まず、モニターのデモグラフィック属性について見ていく。

第 1 表は、年別階級別世帯年収の世帯数割合を示している。期間内におけるモニターの世帯年収構造には特別大きな変化は見られない。

第 2 表は、年別家族構成別の世帯数割合を示している。単身世帯、夫婦のみ世帯が増加傾向にあるのに対して、二世帯、三世帯同居世帯は減少傾向にある。これには、いくつかの推測が立つ。まず、本モニターには学生もおり、進学等で家を離れるといった理由で世帯形態が変化している可能性がある。また、比較的青年に近い年齢の構成員がいる場合、同様な観点から夫婦のみ世帯が増加していると考えられ、二世帯、三世帯同居世帯も多少の影響を受けていると示唆される。

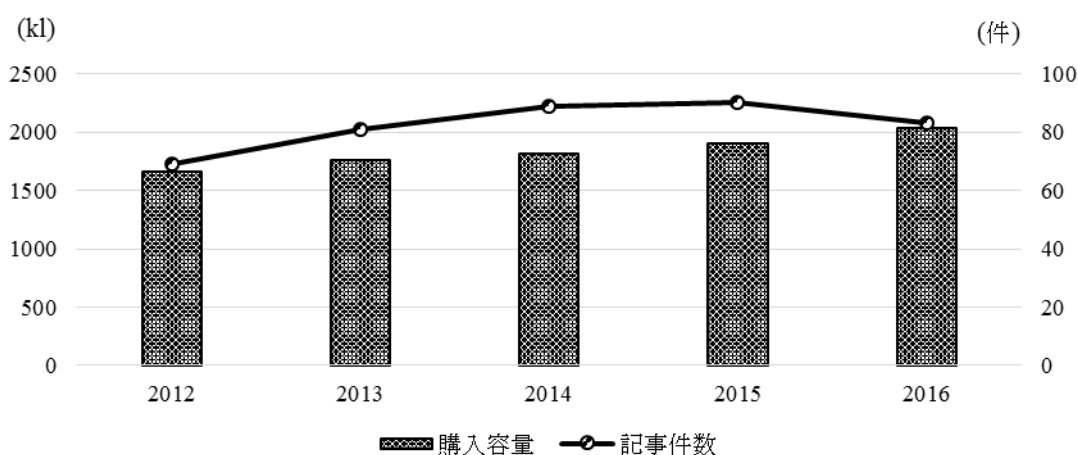
最後に、年別階級別子持ち世帯割合及び年別高齢者あり世帯を第 3 表に示す。幼児あり世帯は減少傾向にあるもののそれ以外の子持ち世帯割合に大きな変化はない。高齢者あり世帯は 5 年間で 8%増加している。

第 2 表 年別家族構成別世帯数割合

年	単身	夫婦のみ	二世帯同居	三世帯同居
2012	11%	21%	59%	9%
2013	11%	22%	58%	9%
2014	11%	23%	57%	8%
2015	12%	23%	57%	8%
2016	12%	25%	56%	7%

第3表 年別階級別子持世帯及び老人世帯数割合

年	幼児	小学生低	小学生高	中学生	高校生	高齢者
2012	14%	8%	8%	9%	10%	36%
2013	13%	9%	8%	8%	9%	38%
2014	12%	9%	8%	8%	9%	40%
2015	11%	9%	8%	8%	9%	42%
2016	10%	8%	8%	8%	8%	44%



第2図 記事数及び購入容量の推移

年別のヨーグルトに関する新聞記事数と対象データにおけるヨーグルト購入容量は第2図のとおりである。購入容量は増加傾向にあることが窺える。記事数は、2016年には減少しているがおおむね増加傾向にあるといえる。

5. ヨーグルトの市場構造についての分析

まず、ヨーグルト市場には様々な商品が存在している。ヨーグルト需要について分析をする前に、まず、商品と購入者の関係からヨーグルト市場にある商品を分類して、市場構造を確認しておきたい。ここでは、直近の市場構造を把握するため、2015年ならびに2016年のデータを分析する。

ヨーグルト市場に存在する各商品について、購入者の重複状況で分類を行う。つまり、同一の購入者に購入されやすいヨーグルト商品は、商品の特徴が似ていると考えられ、逆に購入者がほとんど重複していないヨーグルト商品は、商品の特徴があまり似ていないものであると判断できる。また、購入者の重複関係は商品間の競合関係とも考えることができるだろう。従来、商品間の競合関係は交差価格弾力性により定義されることが多かった。本稿では、データの詳細性を生かして、データマイニング的に競合関係を検討する。この分析により、ヨーグルト市場の中にあるサブ市場の構造を把握することが期待できる。

1) 分析方法

本稿では、第1式のように定義された Jaccard 係数 (cf. 金 (2007) ⁴⁾) を利用して、ヨーグルト商品の近接性を定義し、それらの類似関係を検討した。ヨーグルト商品 ij についての Jaccard 係数 J_{ij} は以下の式で定義される。ただし、 C_k は期間中に商品 k を購入したことがある消費者を元とす

$$J_{ij} = \frac{|C_i \cap C_j|}{|C_i \cup C_j|} \quad (1)$$

る集合である。したがって J_{ij} の分母は商品 i または商品 j を購入したことがある消費者の数であり、分子は商品 i と商品 j をどちらも購入したことがある消費者の数である。商品 i を購入した消費者の集合 C_i と商品 j を購入したことがある消費者の集合 C_j の重複が大きければ Jaccard 係数は 1 に近づき、小さければ 0 に近づく。Jaccard 係数が 1 に近いほど、商品間の類似度は大きいと考えられる。本稿では、 i, j 間の非類似度 d_{ij} を第1式の Jaccard 係数により第2式のように定義する。

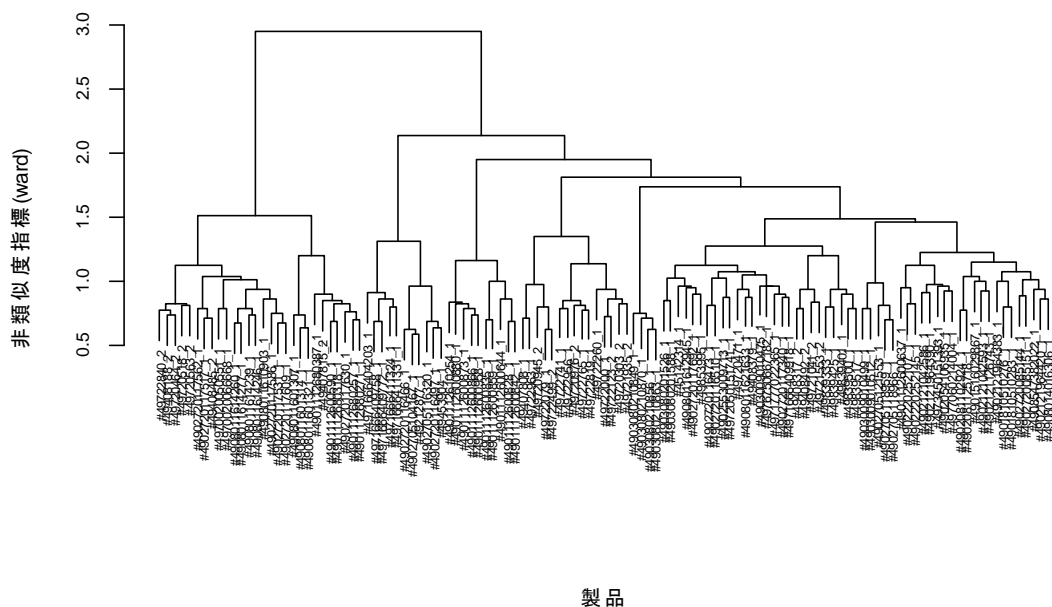
$$d_{ij} = 1 - J_{ij} \quad (2)$$

さらに、非類似度行列 D は d_{ij} を要素とする行列と定義した。

本稿では、この非類似度行列をもとに、階層的クラスター分析を行い、ヨーグルト商品についての商品クラスターと競合関係を整理した。なお、この際に、十分に購入記録がない商品については、他商品との競合関係を評価することはできない。そこで、本稿では、期間中の販売金額合計が 4,000 円以上である 121 商品を対象とした。なお、これらの商品の購入回数の合計は全購入回数の 69.3% となっており、実質的なヨーグルト市場の構造を把握する上では十分であると考えられる。

2) 分析結果

第2式により定義した非類似度行列に基づき、Ward 法により階層的クラスター分析により得られたデンドログラムは第3図に示されている。



第3図 対象商品のデンドログラム

第3図のデンドログラムを基にして、対象商品を9クラスタに分類することとした。分類結果は、第4表に示されている。

第4表 商品クラスタ

商品クラスタ	特徴	商品数	商品例
商品クラスタ1	小容量型	27	A社製特定菌株ドリンクタイプ100g A社製鉄分配合ドリンクタイプ190g B社製低糖180g
商品クラスタ2	小容量複数型	10	A社製70g×3個パック B社製78g×4個パック C社製75g×4個パック
商品クラスタ3	安価大容量型	21	PB無脂肪400g PBビフィズス400g PB脂肪0プロバイオティクス450g
商品クラスタ4	D社製特定ブランド小容量パック型	10	D社特定ブランド75g×4個パック D社特定ブランド佐藤不使用75g×4個パック D社特定ブランド脂肪ゼロ75g×4個パック
商品クラスタ5	健康訴求型大容量型	11	B社特定菌株トクホ450g A社特定菌株トクホ400g A社特定菌株脂肪0400g
商品クラスタ6	B社大容量ドリンクタイプ型	5	B社ドリンクタイプ900g B社カルシウム強化ドリンクタイプ900g B社フルーツ風味ドリンクタイプ900g
商品クラスタ7	小容量複数フルーツ風味型	18	E社アロエ75g×4個パック B社アロエ75g×4個パック B社イチゴ75g×4個パック
商品クラスタ8	E社製小容量ドリンクタイプ型	4	E社ドリンクタイプブルーベリー風味125ml E社ドリンクタイプイチゴ風味125ml E社ドリンクタイプ125ml
商品クラスタ9	B社製小容量型	15	B社特定菌株ドリンクタイプ112ml B社特定菌株ドリンクタイプ低糖・低カロリー112ml

これらの分類は、分析者の先験的基準により分類したものでなく、消費者の購入行動に基づいて導き出されたものである。それぞれのクラスタに含まれる商品は、購入者の重複が比較的多く、ヨーグルト市場のサブ市場を構成し、互いに競合関係にあると考えられる。以下、各商品クラスタごとに特徴を検討しよう。

商品クラスタ1には27商品が含まれており、その多くが100gから200gの小容量の商品である。家庭内で保存されることなく一度に消費される商品が多く含まれている。この中には半固形のタイプのヨーグルトだけでなく、ドリンクタイプのヨーグルトも多く含まれており、スプーンで食べるタイプのヨーグルトといわゆる飲むヨーグルトが同一のサブ市場に位置していると考えられる。

商品クラスタ2には10商品が含まれている。一カップ70gから80gの半固形ヨーグルトが3個から4個でパッケージになった商品がほとんどである。様々なメーカーの商品がこの中には含まれている。

商品クラスタ3には2プライベートブランドの400gから500gの半固形ヨーグルト商品の中

心に 10 商品が含まれている。家庭内で冷蔵してしばらく消費するタイプの商品であり、比較的安価なものが多い。

商品クラスタ 4 には D 社特定ブランドの小容量複数パック商品が 10 商品含まれている。ほかの商品は含まれておらず、D 社特定ブランドがほかの小容量複数パック商品から差別化されて購入されていることが示唆される。

商品クラスタ 5 には健康訴求型大容量商品を中心に 10 商品が含まれている。特定保健用食品として表示が許可されているものが多く含まれており、トクホ商品がほかの安価な大容量商品から差別化されていることが示唆される。2015 年には機能性表示食品制度が導入されているが、今後データが蓄積されれば、トクホとどのように差別化されるか、検討することができるようになるだろう。ただし、現状ではこのクラスタには機能性表示食品はほとんど含まれていない。

商品クラスタ 6 は、B 社と F 社の約 1 リットルの大容量ドリンクタイプの商品が中心に構成されている。一回に消費される小容量型のドリンクタイプは半固形ヨーグルトと同一クラスタに属していたが、大容量型は別のクラスタとなっており、消費者から区別して認識されていると考えられる。

商品クラスタ 7 は、アロエやフルーツなどが混入された小容量複数パック商品が 18 商品含まれている。フルーツ入りの商品は、別のカテゴリとして消費者に認識されているようである。

商品クラスタ 8 には E 社製小容量ドリンクタイプ商品が 4 商品含まれている。E 社のドリンクタイプは商品クラスタ 1 から差別化され、独自の商品として購入されていることが示唆される。

商品クラスタ 9 には、B 社製小容量商品が 15 商品含まれている。B 社のコーポレートブランド訴求が消費者に受け入れられ、商品クラスタ 1 から差別化されているものと考えられる。

前述したとおり、1 回にすべて消費されると考えられる小容量型の商品は、様々なメーカーの商品が多数含まれており、競争が激しいと考えられる。その中で、D 社の小容量特定ブランド商品、E 社ドリンクタイプは、比較的ロイヤルティの高い顧客を獲得して、競合関係からやや距離をとることに成功しているように見られる。また、B 社は大容量ドリンクタイプでも、小容量商品でもクラスタを形成しており、コーポレートブランドの強さが示唆される。また、特定保健用食品としての表示も、差別化要因として機能しているように見られる。

続いて、各消費者について、商品の分類ごとに購入回数を算出し、購入回数比率から、消費者のタイプを分類し、各商品クラスタがどのような消費者に好まれているかを検討した。具体的には 9 種類の商品クラスタについて、各消費者の購入回数比率を計算し、その比率ベクトルを基に、大量の対象でも容易にクラスタリングが可能な k-means 法により分類を行った。なお、消費者の分類を適切に行うために、分析対象となっている 121 商品を 50 回以上購入している 10,067 名の消費者を対象とした。また、本稿では、消費者のクラスタ数については、当該クラスタに含まれる消費者が総数の 5%以上になるように決定し、その結果 6 となった。各消費者クラスタの特徴については第 6 表にまとめている。

第5表 消費者クラス

	小容量型	小容量複数型	安価大容量型	D社製特定ブランド小容量	健康訴求型大容量型	B社大容量ドリンクタイプ	小容量複数フルーツ風味型	E社製小容量ドリンクタイプ	B社製小容量型	特徴	構成比
消費者クラス1	7.5%	3.8%	2.4%	27.9%	12.5%	24.6%	7.0%	7.7%	6.6%	特定メーカー商品購入層	9.37%
消費者クラス2	5.7%	38.3%	4.3%	2.4%	12.3%	2.1%	30.3%	0.6%	3.9%	小容量複数商品	21.57%
消費者クラス3	3.9%	4.0%	5.4%	1.1%	75.6%	2.0%	3.6%	0.4%	4.0%	健康訴求型大容量商品購入層	33.46%
消費者クラス4	68.1%	3.6%	1.6%	2.0%	9.1%	2.0%	3.5%	1.2%	9.0%	小容量商品購入層	10.38%
消費者クラス5	7.4%	2.4%	1.6%	1.3%	8.2%	1.5%	2.7%	0.6%	74.3%	B社小容量商品購入層	19.25%
消費者クラス6	2.9%	6.1%	65.8%	0.9%	16.1%	1.4%	3.2%	0.4%	3.3%	安価大容量商品購入層	5.98%

消費者クラス1はD社製特定ブランドならびにB社製大容量ドリンクタイプの購入が多く、E社製小容量ドリンクタイプの購入回数比率も6クラスで最も大きい。このクラスの消費者は特定のメーカーの商品を好んで購入する傾向が強いと考えられる。女性比率は比較的 low、専業主婦率も低い、低い年収の消費者もやや少ないようである。

消費者クラス2は小容量複数型の商品を多く購入している。比較的年齢が若く、子供の人数も多い。小容量複数型の商品は子供の多い世帯に受け入れられているようである。

消費者クラス3は健康訴求大容量型の商品購入が多い消費者である。消費者クラス2と比較すると、子供の人数は少なく、逆に高齢者の人数が多い。

消費者クラス4は小容量型の商品購入が多い。専業主婦率は最も低く、未婚者比率も高いことから、独身で働いている消費者が多いとみられる。子供の数も少ない。

第6表 消費者クラスタの特徴

	年齢	女性比率	未婚者比率	個人年収 200万円未満比率	世帯年収 400万円未満比率	世帯人数	乳幼児人数	小学校低学年人数	小学校高学年人数	中学生人数	61歳以上人数	専業主婦率
消費者クラスタ1	51.7	57.1%	18.5%	47.5%	23.8%	2.91	0.138	0.08	0.058	0.066	0.887	20.8%
消費者クラスタ2	49.6	65.4%	12.2%	58.4%	24.2%	3.2	0.228	0.139	0.127	0.134	0.71	24.9%
消費者クラスタ3	52.9	62.2%	15.0%	54.8%	27.8%	2.79	0.135	0.072	0.059	0.062	0.907	24.2%
消費者クラスタ4	51.7	53.3%	21.0%	42.4%	27.1%	2.76	0.091	0.04	0.04	0.057	0.846	19.3%
消費者クラスタ5	52.3	56.4%	18.8%	47.5%	25.9%	2.78	0.126	0.069	0.061	0.06	0.909	20.9%
消費者クラスタ6	51.1	65.1%	13.3%	56.8%	27.6%	2.89	0.194	0.088	0.086	0.08	0.713	26.6%
全消費者(10067)	51.8	60.5%	16.0%	52.3%	26.2%	2.89	0.152	0.084	0.074	0.078	0.845	23.0%

消費者クラスタ5はB社製小容量型商品の購入が多い。高齢者人口が最も多く、年齢も高い。高齢世帯にB社製小容量型の顧客が多いように考えられる。

消費者クラスタ6は安価大容量型の商品購入が多いクラスタである。女性比率がやや高く、専業主婦率も高い。

6. ヨーグルト需要関数の分析

本稿では、ヨーグルトの需要に対して、ヨーグルトに関する報道量ならびに価格変数や世帯属性変数、がどのような影響を与えるのかを明らかにすることを目的としている。そこで、ヨーグルトの購入量を被説明変数とする需要関数を推定した。

1) 分析手法

分析対象データには、期間中一度もヨーグルトを購入していない消費者も含めて10,000名を超えるモニターごとの年別の購入履歴が蓄積されている。そのため、モニターによってヨーグルトを購入していない年があり、そのような記録が多数存在している。つまり、被説明変数となる購入容量 Q_{it} が高頻度で0の値をとる。この時、被説明変数を例えばOLSモデルなど通常の方法で分析した場合、バイアスが生じることが知られている。よって、本稿では需要関数として、第3式で示されるようなトービットモデルを採用した。

$$Q_{it} = \begin{cases} Q_{it}^* & (Q_{it}^* > 0) \\ 0 & (Q_{it}^* \leq 0) \end{cases} \quad (3)$$

where $Q_{it}^* = \alpha_i + \beta_{price,i} P_{t,i} + \beta_{a,i} A_t + \beta_t Y_t + \varepsilon_{it}$,
 $\varepsilon_{it} \sim Normal(0, \sigma_\varepsilon)$

ただし、 Q_{it}^* は個人 i 、年次 t の購入容量 Q_{it} の潜在変数を、 P_t は年次平均購入100ml単価、 A_t は年次のヨーグルトに関する記事数、 Y はデータ記録年の西暦下二ケタ、 α_i, β_i はモデルパラメータ、 ε_{it} はモデルの誤差項をそれぞれ示す。

通常、モデルパラメータはサンプルで同一値をとると仮定される。しかし、我々が日常感じる通り、消費者の好みは人によりさまざまである。したがって、サンプルで斉一な選好パラメータを仮定することは、いかにも非現実的であるといえよう。近年、食料消費分析においても、消費者個人の好みの差異、すなわち消費者選好の異質性を考慮したモデルの適用が一般的になっている (cf. 氏家 (2016) ⁵⁾)。選好異質性をモデルに組み込んだ場合、これまでよく用いられているモデルは、選好パラメータが分布を持つと仮定して、パラメータ分布に従って尤度関数の期待値を評価して、最尤推定を行う **Mixed model** である。本稿では、比較的新しい試みとして、階層ベイズアプローチの適用により消費者選好の異質性をモデルに組み入れる。

消費者選好の異質性の分析への階層ベイズモデルの適用については、Rossi ならびに Allenby らの一連の研究において 90 年代から議論されてきた (Rossi and Allenby(1993)⁶⁾、Allenby et al. (1998)⁷⁾、Allenby and Rossi (1998)⁸⁾など)。統計ソフトウェアへのベイズ推定の実装が進んだことで、消費者行動についての実証研究への適用が近年進んでいる。

例えば、Kim and Sugai(2008)⁹⁾は日本のウェブコンテンツに対する消費者選択行動について **Mixed Logit model** に階層ベイズアプローチを適用している。また、Miyawaki et al.(2011)¹⁰⁾は日本の水需要についての **Tobit model** に階層ベイズアプローチを適用している。Volinskiy et al.(2009)¹¹⁾は、キャノーラオイルについての交換型実験オークションデータに対する **Mixed logit model** の推定に階層ベイズアプローチを適用している。また、Osawa et al.(2017)¹²⁾は、小松菜のスキャナーパネルデータについての負の二項分布モデルに対して、階層ベイズアプローチを適用し、原発事故以降の消費者行動の変化を分析している。

従来の **Mixed model** アプローチとの比較については、たとえば、Lagerkvist et al. (2013)¹³⁾が検討している。そこでは、ベストワーストスケーリングによるデータにおける異質性の記述について、**Mixed model** の一つである潜在クラスモデルと比較を行い、階層ベイズモデルの方が選好異質性をより良好にとらえることを主張している。

本稿では、ベイズ統計モデリングソフトウェアである Stan と R を利用して、第 3 式で表されるモデルを階層ベイズモデルと考え、マルコフ連鎖モンテカルロ法 (MCMC) により推定した。

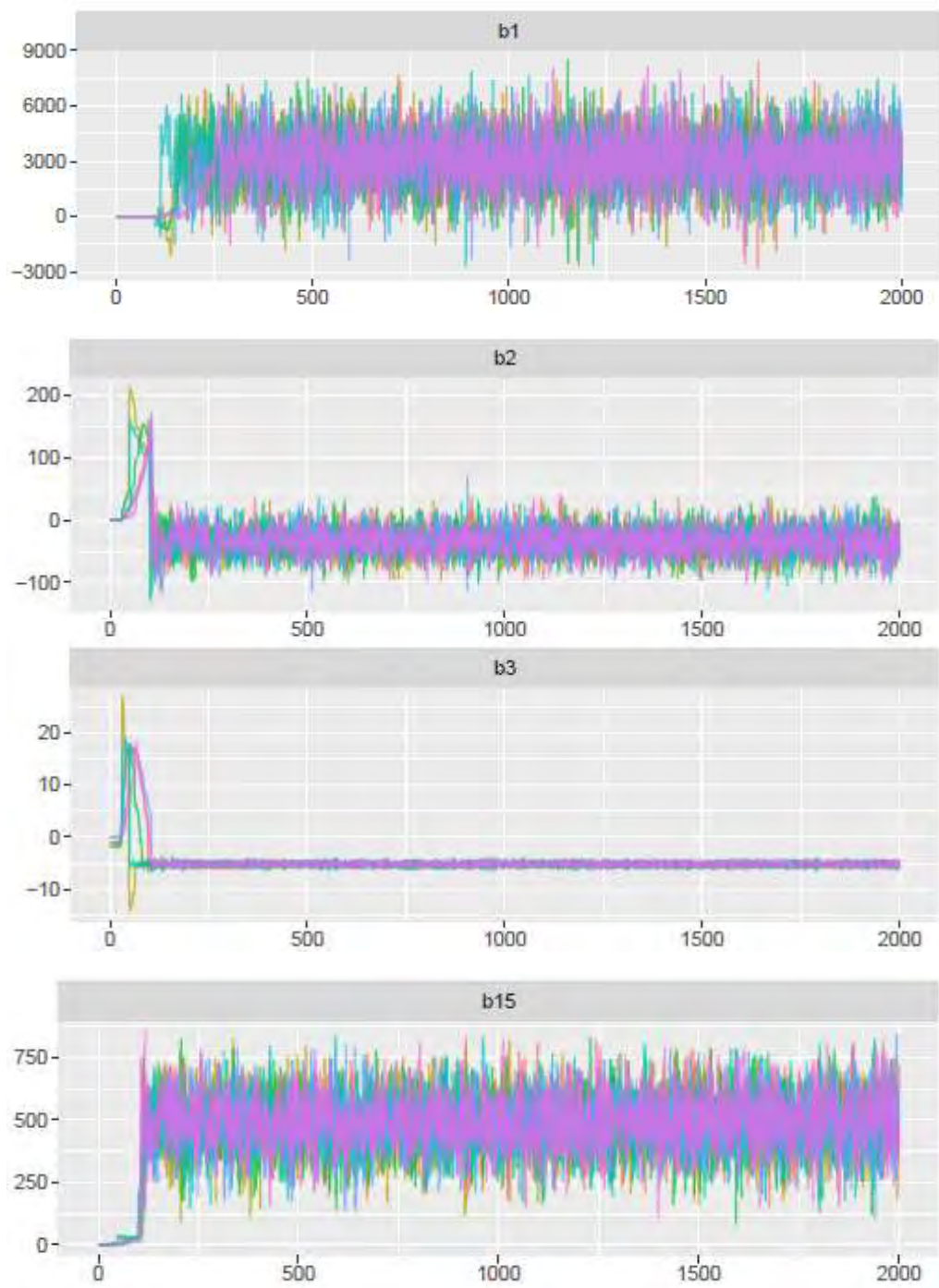
2) 結果

まず階層化せずにベイズモデルを推定し、通常的最尤推定法による結果と比較する。なお、モデルには、シフターとして世帯属性変数を含めている。推定結果は第 6 表に示されている。また、ベイズモデルに関して、定数項ならびに記事件数、価格、ならびに年次トレンド項について、MCMC サンプルの収束状況を示すトレースプロットが第 4 図に示されている。ベイズ推定については、収束状態を示す **Rhat** の値は 1 であり、またトレースプロットからも一定の分布に収束していると判断される。したがって、MCMC によるベイズ推定は良好に行われたと推定される。

最尤推定とベイズ推定の結果を比較すると、推定値はほとんど同一の値となっており、対数尤度も同様の値となっている。最尤推定でもベイズ推定でも同様の結果が得られていることが分かる。

第6表 最尤推定とベイズ推定の結果比較

	最尤推定			ベイズ推定		
	推定値	標準誤差	p 値	MCMCサ ンプル平 均	MCMC サンプル 標準偏差	R hat
定数項	3078	1541	0.046 *	3042.0	1487.7	1
未婚ダミー	-7187	426.9	0.000 ***	-7195.9	419.3	1
乳幼児ダミー	3100	460.6	0.000 ***	3093.5	453.7	1
小学生低学年ダミー	89.68	508.3	0.860	94.6	488.1	1
小学生高学年ダミー	918.8	505.6	0.069 .	915.7	494.4	1
中学生ダミー	2296	512.9	0.000 ***	2294.4	508.5	1
高校生ダミー	1772	507.4	0.000 ***	1770.7	507.5	1
61歳以上ダミー	1660	318.9	0.000 ***	1651.1	318.9	1
世帯人数	126.6	155.6	0.416	126.1	153.3	1
世帯所得400万円未満ダミー	-2821	299.9	0.000 ***	-2816.7	293.8	1
世帯所得1200万円以上ダミー	972.3	293.1	0.001 ***	971.4	288.8	1
年齢	232.8	13.66	0.000 ***	233.0	13.5	1
記事件数	-34.36	21.61	0.112	-34.4	21.5	1
価格(100m l)	-5.222	0.5187	0.000 ***	-5.2	0.4	1
西暦下2桁	487.9	114.7	0.000 ***	490.2	113.0	1
log(sigma)	10.01	0.00432	0.000 ***			
sigma				22171.6	85.81	1
対数尤度	-407146			-407144	2.87	1



第4図 MCMC サンプルのトレースプロット (一部)

残念ながら本稿で注目したい記事数については、有意な結果が得られていない。表示は省略しているが、ベイズ推定の結果でも符号を確定させる結果にはなっていない。したがって、推定結果からは記事文字数の影響を検出することはできなかった。ただし、西暦下2桁の係数がトレンド項となるが、これは有意に推定されており、一年ごとにおよそ500g消費量が増加していることが分かる。

また、世帯属性のシフターの推定結果を見ると、乳幼児がいる消費者は消費量が非常に多くなることが分かる。また世帯所得が400万円未満の世帯では消費量は少なく、逆に世帯年収1200万円以上の場合には消費量が多くなっている。

つづいて、第3式を階層ベイズモデルにより推定した結果について検討する。第7表には個人パラメータ事前分布の平均ならびに標準偏差のMCMCサンプル平均、標準偏差ならびに95%信頼区間、Rhatが示されている。Rhatについては、個人パラメータ推定値も含めて、1.1未満であり、おおむね良好に推定が行われたと評価できる。

第7表 階層ベイズモデル推定結果（事前分布パラメータ）

	MCMCサンプル平均	MCMCサンプル標準偏差	MCMCサンプル95%信頼区間	Rhat
事前分布パラメータ平均				
定数項	11339.4	688.895	[9991 12678]	1.000
価格 (100ml)	-2.92169	0.46912	[-3.868 -2.054]	1.004
記事数	-36.9854	10.3175	[-58.25 -17.25]	0.999
西暦下2桁	681.678	57.0031	[570.6 793.5]	1.000
事前分布パラメータ標準偏差				
定数項	14019.8	339.579	[13313 14671]	1.007
価格 (100ml)	1.91837	0.54195	[0.916 3.015]	1.004
記事数	20.8324	15.2634	[0.865 56.93]	1.087
西暦下2桁	1026.83	24.0473	[978.1 1072]	1.015
sigma	10686.1	46.2068	[10595 10779]	1.000

事前分布パラメータ平均の推定結果を見ると、価格パラメータは負値をとっており理論的要請を満たしており、95%ベイズ信頼区間も狭い。ただし、標準偏差推定値を見ると、一定の大きさがあり、価格をあまり気にしない消費者も存在していることが示唆される。トレンド項については正值となっており、2012年から2016年にかけて、平均的には増加傾向がみられたと判断される。ただし、これについても標準偏差は大きく、消費者によってトレンドのありかたはかなり大きく異なっていることが示唆される。

当初の仮説とは逆の結果が得られたのは記事数であり、パラメータ平均の推定値は負値であり、ベイズ信頼区間も負値域に位置している。結果を素直に読み解けば、記事数が増えるほどヨーグルトの消費量は減少するということになる。ただし、これについても、標準偏差の値は大きく、選好異質性が大きいとみられる。

続いて、階層ベイズモデルにより推定された個々人の選好パラメータと世帯属性との関連性に

について検討する。また、併せて、第 5 節で検討した、消費者クラスタについても、ダミー変数として回帰式に加えた。

第 8 表は個人パラメータを被説明変数とし、各種世帯属性を説明変数とした OLS の推定結果が示されている。ただし自由度修正済み R 二乗の値を見ると、かなり低く、世帯属性が必ずしも個人パラメータの決定要因とはなっていない恐れがある。個人パラメータの推定については、今後の検討課題としたい。

第 8 表 個人パラメータ推定値を世帯属性に OLS 回帰式推定結果

	定数項		価格パラメータ		記事数パラメータ		年次トレンド項	
	推定値	p値	推定値	p値	推定値	p値	推定値	p値
定数項	12613	0.0000 ***	-2.955	0.0000 ***	-36.18	0.0000 ***	1032.9	0.0000 ***
年齢	79.51	0.0001 ***	0.0002	0.5158	0.0233	0.0000 ***	5.9343	0.0001 ***
未婚ダミー	-2333	0.0005 ***	-0.017	0.0428 *	-0.605	0.0010 ***	-152.9	0.0014 **
乳幼児ダミー	799.91	0.2295	0.0015	0.8589	0.48	0.0089 **	90.477	0.0586 .
小学生低学年ダミー	-505.5	0.4945	0.0039	0.6731	-0.153	0.4543	-33.71	0.5262
小学生高学年ダミー	-169.3	0.8164	0.0035	0.7030	0.0225	0.9107	-7.033	0.8933
中学生ダミー	1320	0.0598 .	0.0104	0.2343	0.4334	0.0250 *	75.15	0.1360
高校生ダミー	1134.6	0.1082	0.0064	0.4661	0.3388	0.0818 .	107.5	0.0342 *
61歳以上ダミー	1265.8	0.0044 **	0.0095	0.0876 .	0.4138	0.0007 ***	91.159	0.0043 **
世帯人数	-113.3	0.5939	0.0006	0.8357	-0.073	0.2153	-18.7	0.2206
世帯所得400万円未満ダミー	1569.7	0.0003 ***	0.0161	0.0032 **	-0.435	0.0003 ***	111.46	0.0004 ***
世帯所得1200万円以上ダミー	-13.53	0.9727	-0.015	0.0018 **	0.0542	0.6189	0.9695	0.9728
消費者クラスタ2ダミー	-418.6	0.5143	0.0455	0.0000 ***	-0.501	0.0047 **	-227.6	0.0000 ***
消費者クラスタ3ダミー	1085.1	0.0728 .	0.0271	0.0003 ***	0.0082	0.9606	-76.66	0.0778 .
消費者クラスタ4ダミー	-2999	0.0001 ***	0.0783	0.0000 ***	-1.197	0.0000 ***	-366.7	0.0000 ***
消費者クラスタ5ダミー	-1949	0.0037 **	0.0468	0.0000 ***	-0.925	0.0000 ***	-303.7	0.0000 ***
消費者クラスタ6ダミー	3720.2	0.0000 ***	0.0454	0.0000 ***	0.6881	0.0031 **	69.936	0.2489
自由度修正済みR二乗	0.0463		0.0249		0.052		0.0485	
F統計量		0.0000		0.0000		0.0000		0.0000

まず、定数項についての個人パラメータと世帯属性との関係性について検討する。年齢が高いと定数項が大きくなる傾向がみられる。また消費者クラスタ 4 ならびに 5 の消費者は消費量が少ない傾向がみられるが、消費者クラスタ 6 に該当する消費者は消費量が多いようである。

続いて価格パラメータについて検討する。未婚者において、価格パラメータの値が小さくなり、また、低所得では価格パラメータは大きく、高所得では小さくなりやや奇異な印象がある。ただし、係数の値は小さくなく、影響は軽微であると考えられる。

記事数パラメータについては、年齢が高いほど、正値に振れる傾向がみられ、また高齢者がいる世帯の場合も同様の傾向がある。乳幼児がいる世帯も同様である。

年次トレンドに関しては、高齢者、高校生がいる、世帯所得が低い世帯ほど、増加トレンドが強くなっているとみられる。

7. まとめ

本稿では、増加傾向が著しい乳製品であるヨーグルトを事例に、商品ごとの顧客の重なり具合

を示す Jaccard 係数による市場構造の分析ならびにヨーグルト需要関数の推定を行った。

市場構造の分析では、小容量型の商品は、様々なメーカーの商品が多数含まれており、競争が激しいと考えられる。その中で、D 社の小容量特定ブランド商品、E 社ドリンクタイプは、比較的ロイヤルティの高い顧客を獲得して、競合関係からやや距離をとることに成功しているように見られる。また、B 社は大容量ドリンクタイプでも、小容量商品でもクラスタを形成しており、コーポレートブランドの強さが示唆された。また、特定保健用食品としての表示も、差別化要因として機能しているように見られる。

また、ヨーグルトの需要関数については、階層ベイズモデルとしてトービットモデルを推定した。調査期間全体では増加トレンドがみられ、ヨーグルト市場が拡大していることが示された。一方で、本研究課題の中心的関心であったマスメディアによる需要量への影響については、十分に検出することはできなかった。

階層ベイズモデルによる消費者の選好異質性は、定数項、価格係数、記事件数係数、トレンド項で存在することが示され、全体的な傾向だけではなく、個人間の差異について注目する価値があることが示された。

ただし、個人パラメータと世帯属性との関係性については決定係数が低く、十分解明できなかった。年次集計から月次集計のデータに組み替えて、一人一人の消費行動をより詳細にモデルに反映させることで、推定の品質向上も期待されるだろう。引き続き取り組んでまいりたい。

参考文献

- 1) 齋藤忠夫・伊藤裕之・岩附慧二・吉岡俊満編著 (2016)『ヨーグルトの事典』朝倉出版。
- 2) 青木幸弘 (1993)「スキャナーパネルデータと消費者行動分析」『商學論究』40(3) : 31-59.
- 3) 田口光弘(2010)「POS データを用いた加工食品と生鮮食品の購買行動分析」『フードシステム研究』16(4) : 25-31.
- 4) 金明哲. R によるデータサイエンス: データ解析の基礎から最新手法まで. 森北出版, 2007.
- 5) 氏家清和. “食品表示と消費者行動をめぐる実証的研究の動向.” 農業経済研究= Journal of rural economics 88.2 (2016): 156-171.
- 6) Rossi, Peter E., and Greg M. Allenby. “A Bayesian approach to estimating household parameters.” Journal of Marketing Research (1993): 171-182.
- 7) Allenby, Greg M., Neeraj Arora and James L. Ginter (1998) “On The Heterogeneity of Demand,” Journal of Marketing Research, 35, 384-389.
- 8) Allenby, Greg M., and Peter E. Rossi. “Marketing models of consumer heterogeneity.” Journal of econometrics 89.1-2 (1998): 57-78.
- 9) Kim, Donghun, and Philip Sugai. “Differences in consumer loyalty and willingness to pay for service attributes across digital channels: A study of the Japanese digital content market.” Telecommunications Policy 32.7 (2008): 480-489.
- 10) Miyawaki, Koji, Yasuhiro Omori, and Akira Hibiki. “Panel data analysis of Japanese residential water demand using a discrete/continuous choice approach.” The Japanese

Economic Review 62.3 (2011): 365-386.

- 11) Volinskiy, Dmitriy, et al. "Does Choice Context Affect the Results from Incentive - Compatible Experiments? The Case of Non - GM and Country - of - Origin Premia in Canola Oil." Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 57.2 (2009): 205-221.
- 12) Osawa, Masaki; BenTaieb, Seifeddine; Maruyama, Yuki; Ujiie, Kiyokazu "Changes in Vegetables Consumption after Nuclear Power Plant Accident -a Scanner Panel Data Approach of Japanese Mustard Spinach-", XV EAAE Congress - Towards Sustainable Agri-Food Systems: Balancing between Markets and Society, 2017年8月29日-9月1日
- 13) Lagerkvist, Carl Johan. "Consumer preferences for food labelling attributes: Comparing direct ranking and best-worst scaling for measurement of attribute importance, preference intensity and attribute dominance." Food Quality and Preference 29.2 (2013): 77-88.