

一般化 δ 区間モデルによる準拠集団理論の定式化とその検証

要旨

本稿の目的は、社会学における古典的理論の一つである「準拠集団理論」の単純な数理モデルを提案し、大規模データによって理論モデルの妥当性を検証することである。そのために、先行研究で定式化された δ 区間モデルを拡張し、自分と年齢が近い他者を準拠集団として選択して、その他者の平均所得を比較の基準として認識するという一般モデルを提唱する。拡張した δ 区間モデルを「2015年階層と社会意識全国調査(第1回SSP調査)」データを用いて検証したところ、モデルに基づいて計算した準拠集団平均所得は、「同世代の人の平均的な年収」についての人々の回答を、適切に予測することが判明した。さらに、準拠集団平均所得が増加するほど、主観的幸福感や収入満足度が低下するというインプリケーションが経験的にも正しいことが分かった。

キーワード: 準拠集団, 相対的剥奪, δ 区間モデル

A Formalization of Theory of Reference Group by Generalized δ Interval Model and Its Empirical Test

Abstract

The purposes of this paper is to propose a simple mathematical model of reference group and to examine empirical validity of the model. In the previous study, δ interval model defines the distance between ego and reference group only based on income condition. We generalize the model by assuming that people choose their reference group from those who are in same generation and recognize their average income as a focal point for comparison. To examine the generalized δ interval model, we use nationwide data of the SSP 2015 survey. As a result, we found that the reference group's mean income based on the model can predict the respondents' subjective inference of average income of same generation. Moreover, we found the implication, as reference group mean income increases, subjective well-being and satisfaction of income decreases, is empirically true.

Key words: reference group, relative deprivation, δ interval model, SSP 2015

拡張 δ 区間モデルによる準拠集団理論の定式化とその検証

1 準拠集団と相対的剥奪の理論と実証

個人が、集団内での自分の地位に満足するか、不満を感じるかどうかは、比較の対象となる他者の状態に依存する。この説得的なアイデアは、社会学における古典的理論の一つである「準拠集団理論 (Merton 1957=1961)」ないし「相対的剥奪理論 (Stouffer et. al. 1949)」として、斯界で広く知られている¹⁾。同様の考え方は、経済学においては「相対所得仮説 (Duesenberry 1949; Easterlin 1974)」と呼ばれ、人々の所得満足度や生活満足度や主観的幸福感を説明する理論として、しばしば活用されている (Clark and Oswald 1996; McBride 2001; Ferrer-i-Carbonell 2005; Luttmer 2005; Ball and Chernova 2008)。

相対的剥奪は数学的定式化の進んだ概念の一つであり、経済学者イツハキによる指数の開発は、その先駆的な例である。イツハキは、社会心理学者ランシマンによる相対的剥奪概念の定義 (Runciman 1966) をふまえ、所得比較に基づく社会的な相対的剥奪指数を考案した。彼は社会レベルでの相対的剥奪指数を平均所得で基準化した値が、社会全体のジニ係数と等しいことを証明した (Yitzhaki 1979)。イツハキの相対的剥奪指数は、その後多くの発展的理論研究を生み出した (Hey and Lambert 1980; Kakwani 1984; Berrebi and Silber 1985; Chakravarty et al. 1995; Ebert and Moyes 2000)。

さらに、生活・収入満足度 (D'Ambrosio and Frick 2004; 石田 2011, 2014)、健康指標 (Eibner and Evans 2005; Kondo et al. 2008; Stewart, 2006; Subramanyam et al. 2009; Oshio et al. 2011)、階層帰属意識 (Maeda and Ishida 2013) を説明する変数としても、相対的剥奪指数は国内外の多くの実証研究で利用されている。

一連の実証研究の流れの中で重視されている課題の一つは、準拠集団選択メカニズムの特定である。準拠集団がどのような基準で選ばれるのかという問題は、マートンが理論を提唱して以来、解が定まらぬ経験的難問である (Merton 1957=1961; Eibner and Evans 2005; Ferrer-i-Carbonell 2005; 前田 2014)。これまでに幸福感や満足度に関する実証研究では、年齢・学歴・居住地域・職業・国全体など種々の基準を用いて準拠集団を操作的に定義してきた (Ferrer-i-Carbonell 2005; Kondo et al. 2008; Oshio et al 2011; 石田 2011)。

しかしながら準拠集団そのものは観測不可能なため、データに適合的な準拠集団を研究者が事後的に仮定すると、理論が反証可能性を失う可能性がある (Merton 1957=1961: 212-229)。よって準拠集団理論が科学的説明として機能するためには、準拠集団選択メカニズムを事前に明示しなければならない²⁾。

1.1 先行研究の問題点

主観的幸福感・階層帰属意識・所得満足度に対する準拠集団平均所得の影響の推定を試みた研究は既に複数存在する。これらの研究では、標本を年齢次元で分割し、各個人は自分を含む年齢集団から準拠集団を選択すると仮定している (Ferrer-i-Carbonell 2005; Eibner and Evans 2005; Kondo et al. 2008; Subramanyam et al. 2009; 石田 2011; Medata and Ishida 2013)³⁾。

しかし、このような仮定と変数の操作化には若干の問題がある。例えば研究者が設定した年齢カテゴリが、(21-30), (31-40), (41-50) である場合、「35 才の人は年上も年下も選択できる」のに「31(40) 才の人は年下 (上) を選べない」という不自然な仮定をおかざるをえない。

一般に年齢と年収は正の相関を持つから、この仮定の下ではカテゴリが下端の人は平均年収を高め、上端の人は低めに見積もるバイアスが生じてしまう。このような意図せざる上方比較・下方比較バイアスの発生は、選択プロセスの表現方法の不備を示している。離散変数の次元の場合は上記の方法で問題はないが、収入や年齢のような連続変数上で距離の近さを表現するためには若干の工夫が必要である。

また、これまでの実証研究では応答変数が主観的幸福感や収入満足度等であり、準拠集団収入の間接的な測定値を用いていない。したがって理論モデルの厳密な意味での検証は、不十分である。

そこで本稿は以下の 2 つの問題、準拠集団選択メカニズムの定式化は、まだ不十分であり、それゆえ多くの実証研究で変数の操作化を誤っている、従来の実証研究では、「準拠集団の所得」を(間接的にさえ)測定していないため、モデルの検証用データとしては誤差が大きい、を解決するために新しい理論モデルを提唱し統計データによってその妥当性を検証する。

新モデルを提案するために、準拠集団選択プロセスを明示した先行研究の一つとして、 δ 区間モデル (浜田 2001) に注目する。 δ 区間モデルは、収入満足度を説明するために、連続変数 (所得) 上の本人位置を中心に準拠集団範囲を自然な近さで表現した数理モデルである。他者選択プロセスを積分 (汎関数期待値) で表現する点がイツハキの相対的剥奪指数と共通しており、剥奪指数の特殊例と位置づけることができる (石田 2011)。

ただし δ 区間モデル (浜田 2001) も相対的剥奪指数 (Yitzhaki 1979) も、準拠集団選択の基準となる次元が収入に限定されている点に限界があり、「年齢」という別の次元が選択基準として重要であるという知見がモデルに組み込まれていない。そこで本稿では δ 区間モデルを 2 次元に拡張して、年齢次元を含むより包括的な準拠集団選択プロセスを定式化する。

2 拡張 δ 区間モデルの定式化

2.1 1 次元 δ 区間モデルの概要

はじめに、ベースとなる δ 区間モデルの概要を確認する。モデルの基本的な考え方は、「人々は自分の客観的な所得分布上の近隣から準拠集団を選択し、それとの比較によって相対的な満足/不

満足度を決定する (浜田 2001)」というものである。準拠集団は自分と経済的地位に近い人からなり、遠く離れた人は比較対象として選ばれない。ここでいう「近い」「遠い」は、次に定義する δ 区間に他者が含まれるかどうかによって区別される。

定義 1 (δ 区間). 所得額 y の分布上の近隣を閉区間

$$[y - \delta, y + \delta] = \{z | y - \delta \leq z \leq y + \delta\} \quad \text{ただし } \delta \geq 0$$

で表し、これを y の δ 区間と呼ぶ。 δ は y から見た準拠集団の所得数直線上の距離を表している⁴⁾。

モデルを構成する基本仮定は次の通りである。

- 所得分布は対数正規分布に従う確率変数 Y で近似できる。確率変数 Y の確率密度関数を $f_Y(t)$ で表す⁵⁾。
- y で確率変数 Y の実現値を表す。所得 y を中心として、 $\pm\delta$ の範囲に含まれる所得の割合は、確率密度関数 $f_Y(t)$ を $y - \delta$ から $y + \delta$ の範囲で積分した値に等しい。

$$P(y - \delta \leq Y \leq y + \delta) = \int_{y-\delta}^{y+\delta} f_Y(t) dt. \quad (1)$$

- 所得値が y である個人は、 $[y - \delta, y + \delta]$ の範囲内から準拠集団を選択する。
- 所得 y である個人が所得 $z \in [y - \delta, y + \delta]$ である他者 (準拠集団) を選択する条件付確率密度は、確率密度関数 $f_Y(z)$ を定積分 (1) で除して基準化した関数により定まる⁶⁾。すなわち、

$$\frac{f_Y(z)}{\int_{y-\delta}^{y+\delta} f_Y(t) dt} \quad \text{ただし } z \in [y - \delta, y + \delta] \quad (2)$$

である。この関数 (2) は閉区間 $[y - \delta, y + \delta]$ で z に関して積分すれば 1 になる。

- 関数 (2) を用いた条件付き期待値は、準拠集団の平均所得 $E[Y | y - \delta \leq Y \leq y + \delta]$ である⁷⁾。

$$E[Y | y - \delta \leq Y \leq y + \delta] = \int_{y-\delta}^{y+\delta} z \frac{f_Y(z)}{\int_{y-\delta}^{y+\delta} f_Y(t) dt} dz$$

- 準拠集団平均所得を $y^* = E[Y | y - \delta \leq Y \leq y + \delta]$ で表す。ある個人の収入に対する満足度は、自分の所得 y と準拠集団平均所得 y^* の関数として決まる: $u(y, y^*)$

以上の基本仮定が《 δ 区間モデル》の要諦であった (浜田 2001)。

2.2 2次元に拡張された《 δ 区間モデル》

本節では1次元モデルに基づき、2次元に拡張した δ 区間モデルを定式化する。基本仮定は以下の通りである。

- 二次元の確率変数 (X, Y) を考え、その同時確率密度関数を $f(x, y)$ とおく。 X が年齢の分布を、 Y が所得の分布を表している。
- (x, y) で確率変数 (X, Y) のそれぞれの実現値を表す。
- 年齢が x である個人は、年齢が $[x - \delta, x + \delta]$ の範囲内にある他者を準拠集団として選択する。
- 準拠集団の平均所得は X に対する Y の条件付き期待値、 $E[Y|x - \delta \leq X \leq x + \delta]$ で与えられる。
- 収入に対する満足度は、自分の所得 y と準拠集団平均所得 $y^* = E[Y|x - \delta \leq X \leq x + \delta]$ の関数として決まる: $u(y, y^*)$ 。

2次元に拡張された δ 区間モデルは、1次元の δ 区間モデルの自然な拡張であり、基本仮定に大きな違いはない。ただし個人の特徴を表す分布として所得だけでなく年齢も考慮しているため、準拠集団の平均所得が単なる Y の条件付き期待値ではなく、 X で条件付けられた Y の期待値となる点に注意しなくてはならない。すなわち準拠集団平均所得は、連続確率変数の条件付き期待値である (以下、 δ index と呼ぶ)。

命題 (準拠集団所得平均 (δ index))。年齢の範囲 $[x - \delta, x + \delta]$ から準拠集団を選択した場合の準拠集団平均所得は、次の条件付き期待値

$$E[Y|x_1 \leq X \leq x_2] = \int_{-\infty}^{\infty} y \left(\int_{x_1}^{x_2} \frac{f(x, y)}{\int_{x_1}^{x_2} f_X(x) dx} dx \right) dy \quad (3)$$

で与えられる。ただし $f_X(x)$ は X の確率密度関数 (周辺分布) である。

証明。確率変数の組 (X, Y) の同時確率密度 $f(x, y)$ について、周辺分布としての X の確率密度を $f_X(x)$ とおく。

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

また、 $X = x$ に対する Y の条件付き確率密度 $f_{Y|X}(y|x)$ を $f_{Y|X}(y|x) = f(x, y)/f_X(x)$ とおく。すると $X = x$ に対する Y の条件付き期待値 $E(Y|X = x)$ は

$$E(Y|X = x) = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y|x) dy = \int_{-\infty}^{\infty} y \cdot \frac{f(x, y)}{f_X(x)} dy.$$

ゆえに、 $x_1 \leq X \leq x_2$ に対する Y の条件付き期待値 $E[Y|x_1 \leq X \leq x_2]$ は

$$E[Y|x_1 \leq X \leq x_2] = \int_{-\infty}^{\infty} y \left(\int_{x_1}^{x_2} \frac{f(x, y)}{\int_{x_1}^{x_2} f_X(x) dx} dx \right) dy.$$

□

2次元に拡張した δ 区間モデルのイメージを図に示す (図 2.2)。

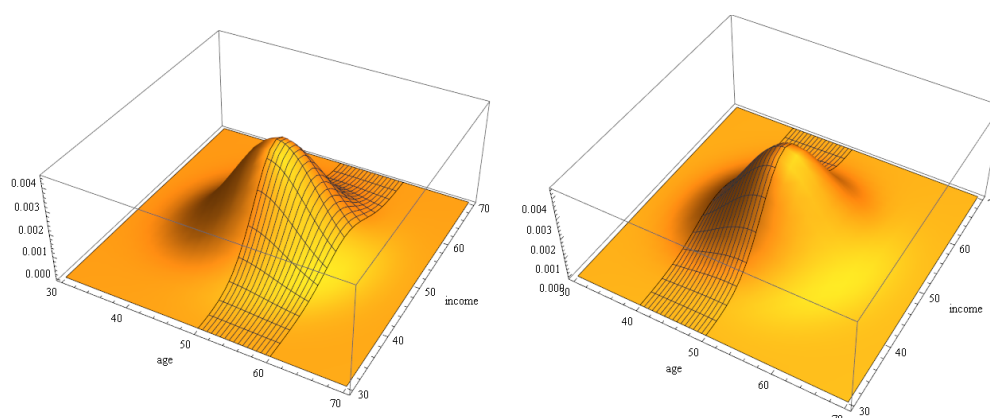


図 1: 2 次元 δ 区間モデルのイメージ。曲面のグラフは所得分布と年齢分布の同時確率密度関数。メッシュで色分けした部分が準拠集団所得（の確率密度）の範囲を表している。

なお、モデルは 3 次元以上にも拡張可能である⁸⁾。

3 モデルの実証

3.1 データの概要

本節では、2 次元 δ 区間モデルの妥当性を調査データによって確認する。使用したデータは「2015 年 階層と社会意識全国調査 (第 1 回 SSP 調査)」である。SSP2015 は、日本国籍を持つ 20 歳以上 64 歳以下の男女を対象に層化 3 段無作為抽出により、抽出された標本に対して、2015 年 1 月～6 月 (実査第 1 期 1 月 24 日～2 月 15 日, 第 2 期 2 月 21 日～3 月 15 日, 第 3 期 5 月 16 日～6 月 21 日) にかけてタブレット端末を用いた個別面接法 (CAPI) で実施された全国調査である (計画標本サイズ 9000, 有効回収 3575(39.7%); 住所不明, 転居, 死亡, 抽出ミスによる対象外を除く回収率は 42.9%)。本論文では年齢に関する準拠集団選択の仮定に基づき、標本を 25 歳から 59 歳までの有職者に限定した。説明変数は準拠集団平均所得 (δ index), 統制変数として年齢, 性別 (男性ダミー), 学歴, 個人収入, 世帯収入, 婚姻状態, 子供有無, 同居家族人数, 職業分類 (EGP6) を用いた。応答変数は同世代の年収予想, 主観的幸福感, 個人収入満足度, 世帯収入満足度である。同世代の年収予想は、「あなたと同年代の人の平均的な年収は、いくらぐらいだと思いますか?」という質問に対する回答カテゴリの階級値 (万円) を用いた。

3.2 統計モデルの選択

個人 i が年齢集団 j に所属している場合の準拠集団平均収入予想値 Y_{ij} は、個人収入 y_{ij} と集団レベルの特性値である準拠集団平均年収 z_j の関数である。そこで、理論的に予測した準拠集団平均収入の誤差を、準拠集団のレベルで生じた誤差と、個人レベルで生じた誤差に区分して、理論値（固定効果）のパラメータ推定誤差を極力小さくするために、集団レベルの誤差を考慮した線形混合モデルを仮定する。グループレベルの固定効果、すなわち拡張 δ 区間モデルに基づく準拠集団平均所得（式 (3)）を明示的に示せば、

$$Y_{ij} = \gamma_{00} + \underbrace{\gamma_{01} \text{個人収入}_{ij}}_{\text{fixed effect}} + \underbrace{\gamma_{10} E[Y|x - \delta \leq X \leq x + \delta]_j}_{\text{fixed effect}} + \underbrace{\beta X}_{\text{統制変数}} + \underbrace{U_j}_{\substack{\text{random effect} \\ \text{集団レベル}}} + \underbrace{R_{ij}}_{\text{個人レベル誤差}}$$

である。ここで、 $E[Y|x - \delta \leq X \leq x + \delta]$ は、単なる年齢別グループ平均 $E[Y|X = j]$ と異なる点に注意する。

3.3 拡張 δ 区間モデルに基づく説明変数の作成方法

前節で仮定した統計モデルのパラメータを推定するために、次のような手順で説明変数（グループレベルの固定効果）をデータから作成する⁹⁾。

1) 本人年齢 x を基準に年齢が $x \pm \delta$ である人々の収入をデータから取り出す（準拠集団の所得分布）。2) 年齢条件で取り出した準拠集団の所得分布から平均値を計算する（準拠集団平均所得の理論値）。3) 準拠集団平均所得の理論値を全ての年齢のレンジで計算する。4) 計算した準拠集団平均所得の理論値をデータ行列に代入する（同じ年齢のユニットは同じ理論値を持つ）。

以上の手順で作成した説明変数が準拠集団平均所得理論値 (δ index) である。準拠集団平均所得理論値は、年齢別集団のグループレベルの固定効果であり、この固定効果によって表現できないグループレベルの誤差はランダム効果で表現する（個人レベルの誤差は個人レベルの誤差項で表す）。統制変数は全て個人レベルで加え、応答変数（回答者が述べた「自分と同じ世代の平均年収（準拠集団平均所得の観測値）」）を理論値（固定効果）に線形混合モデルで回帰する。

以下の分析では $\delta = 5$ と仮定して、本人年齢の ± 5 歳幅に含まれる他者の所得平均値を計算した¹⁰⁾。標本の対象年齢は 20 歳以上 64 歳以下だが、準拠集団の選択に天井効果・フロア効果が生じないように、分析対象を 25 歳から 59 歳に限定した。

3.4 記述統計

以下に年齢別の準拠集団所得分布と、その平均である δ index を比較する (図 2)。

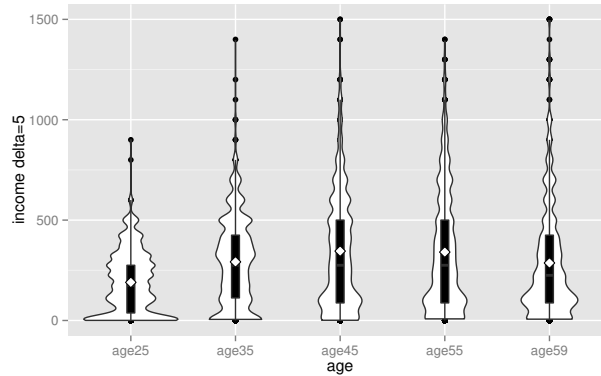


図 2: 年齢 (横軸) と準拠集団所得分布 (縦軸). 本人年齢 ± 5 ($\delta = 5$) 才の所得分布カーネル密度関数と箱ひげ図をプロット. 白い四角は平均 (δ index).

図 2 は, 横軸の本人年齢に対応した $\pm \delta$ 才の他者の所得分布を表示している. 図からは, δ の値は共通でも, 基準年齢が異なれば, 対応する所得分布の形状は異なる様子がうかがえる. このデータは, 年齢に応じて各個人の主観的世界が異なる, という事実を示している. 図内の白い点は, 各所得分布の平均, すなわち δ index である. 単純に増加するのはなく, 50 歳前後をピークに減少する点に注意する. このことは単純に年齢の一次関数として他者所得予想が表現できないことを示唆している.

次に, 応答変数である「同世代の平均的な年収」と δ index の関係を散布図で示す (図 3).

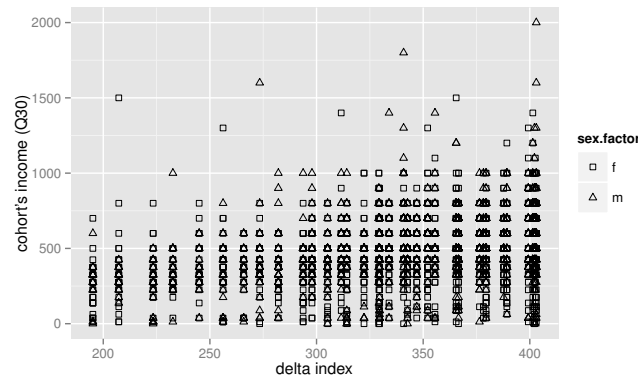


図 3: δ index (横軸) と同世代の人の年収予想 (縦軸) の関係. $\square$ が男性を, $\triangle$ が女性を表す.

図 3 から直感的に分かるように, 性別を考慮してもなお, 同世代平均収入は, δ index だけでは表現できないばらつきを持っている.

それゆえ δ index 以外の統制変数として, 準拠集団平均収入に影響を及ぼす各条件を線形混合モデルの右辺に加える. 以下分析に使用した変数の記述統計量を示す¹¹⁾.

表 1: 記述統計

連続型変数					
	mean	sd	median	min	max
グループレベル					
δ index	335.75	54.12	342.65	195.13	403.04
個人レベル					
同世代所得	445.93	269.13	425	0	4500
年齢	45.27	12.37	46	20	64
個人収入	318.42	349.24	225	0	4500
世帯収入	652.62	464.42	600	0	4500
同居家族人数	3.40	1.42	3	1	10
主観的幸福感	6.78	1.94	7	0	10
離散型変数					
	回答数		%		
性別	男性	1644	0.46		
	女性	1931	0.54		
婚姻状態	結婚	2481	0.69		
	その他	1078	0.30		
	NA	16	0.00		
教育水準	中学校卒業	202	0.06		
	高校卒業	1798	0.50		
	大学卒業	1543	0.43		
	NA	32	0.01		
職業 EGP6	上層ホワイト	878	0.25		
	下層ホワイト	712	0.20		
	自営	243	0.07		
	熟練ブルー	373	0.10		
	非熟練ブルー	533	0.15		
	農業	73	0.02		
	無職	563	0.16		
	学生	88	0.02		
	NA	112	0.03		

ところで、応答変数が「同世代の世帯年収予想」なのだから、拡張 δ 区間モデルに基づいて計算した準拠集団所得期待値が、統計的に有意な効果を持ち、この説明変数を持つ統計モデルがデータにフィットするのは当然ではないか、と読者は疑問を持つかもしれない。

確かに、統計モデルによるデータ解析を、データに対する単なる「あてはめ」と考えた場合には、上記の統計モデル構成は、説明変数と応答変数が概念的に同一であるという意味でほとんどトートロジカルである。しかしわれわれが確かめたいことは、「説明変数が応答変数に因果的な影響を及ぼすか否か」ではなく「理論値（準拠集団所得期待値）で観測値（同世代の世帯年収予想）を予測できるか」である。したがって、観測値の尤度が理論モデルによって高まるならば、われ

われは δ 区間モデルという「数理モデル」の有効性を間接的に主張できるはずだ。

4 分析結果

データから推定した準拠集団平均所得理論値 (δ index) を説明変数として含む計量モデルの分析結果を示す。表 2 の「LME 1」は固定効果として δ index を含む線形混合モデルで、「LME 0」は含まないモデルである。まず、一般化線形モデルではなく、ランダム効果を仮定した線形混合モデルを用いた方がデータをより予測できることを示す。そのために、切片モデル（集団分割だけを仮定した ANOVA）の ICC と、固定効果を含む LME0, LME1 の ICC を比較する。

表 2: ランダム効果の分析結果比較. 応答変数は「同年代の年収予想」

	ANOVA	LME0	LME1
固定効果	Estimate	Estimate	Estimate
切片	433.32 (14.85)	324.85 (26.45)	-8.79 (42.66)
ランダム効果	分散	分散	分散
age (切片)	9005.00 (94.89)	4125.00 (64.23)	29.65 (5.45)
residual	64770.00 (254.50)	48578.00 (220.4)	46236.16 (215.03)
ICC	0.27	0.08	0.00
deviance	48307.39	36157.26	28695.13
AIC	48313.39	36191.26	28731.13

表 2 が示すとおり、切片モデルの ICC は約 0.27 であり、集団間分散は無視できない程度に高い。次に線形混合モデルの固定効果の詳細を表 3 に示す¹²⁾。分析の結果、 δ index を含むモデルの AIC は、含まないモデルより低下しており、理論モデルによる予想の正しさが支持された。

また LME 0 と LME 1 の逸脱度に関する尤度比検定の結果、 δ index を含む LME1 の逸脱度 (28773.4) は、含まない LME0 の逸脱度 (36234.6) と比較して有意 (水準 0.001) に減少していることが分かった。このことはわれわれの理論モデルが経験的に正しいことを間接的に示している。

表 3: 線形混合モデルによる分析結果. () 内は標準誤差

	応答変数:同世代の年収予想		
	GLM	LME0	LME1
	(1)	(2)	(3)
δ index	1.070*** (0.092)		1.071*** (0.093)
個人年収	0.034 (0.023)	0.080*** (0.020)	0.035 (0.023)
世帯年収	0.125*** (0.015)	0.102*** (0.013)	0.125*** (0.015)
男性ダミー	28.531** (11.662)	8.816 (10.357)	28.450** (11.662)
既婚ダミー	37.249*** (11.981)	39.198*** (11.162)	37.046*** (11.988)
世帯人数	-8.091** (3.691)	-3.461 (3.386)	-8.034** (3.692)
定数	-8.135 (42.317)	324.855*** (26.448)	-8.776 (42.655)
Observations	2,120	2,654	2,120
Log Likelihood	-14,387.720	-18,078.570	-14,347.510
AIC	28,807.430	36,191.140	28,731.030
BIC		36,291.160	28,832.890

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

4.1 説明変数としての δ index

前節の分析で、 δ 区間モデルの有効性は間接的に示された。本節では δ 区間モデルから得られるインプリケーションが経験的に正しいかどうかを検証し、数理モデルの有効性をさらに検討する。

先行研究で示された重要なインプリケーションの一つは、「人々の効用は準拠集団平均所得の減少関数である (Ferrer-i-Carbonell 2005; Luttmer 2005; D'Ambrosio and Frick 2007)」だった。簡潔に言い換えれば、本人所得を y , 準拠集団平均所得を y^* とおくと、 $\partial u(y, y^*) / \partial y^* < 0$ である。

このインプリケーションが正しいとするならば、 δ index は説明変数として、主観的幸福感・世帯収入満足度・個人収入満足度の予測に役立ち、かつその効果は負であることが予想できる。

以下に説明変数と統制変数は前節の分析と同じ変数を使い、応答変数が主観的幸福感・世帯収入満足度・個人収入満足度である場合の一般化線形モデル（誤差分布は正規分布）を分析して δ index を説明変数に含む条件と含まない条件を比較する¹³⁾。

用いた統計モデルは誤差関数に正規分布を仮定した恒等リンク関数の GLM である。実質的には OLS 推定と同じ結果だが、AIC を比較するために最尤推定で対数尤度を計算した。なお、前節の分析では線形混合モデルを用いたが、応答変数が主観的幸福感、世帯収入満足度、個人収入満

足度である場合はいずれの場合も年齢集団をクラスターとした ICC が低いため（全て 5% 以下），集団レベルのランダム効果は仮定しない。

理論モデルのインプリケーションが正しければ， δ index を含むモデルは含まないモデルの AIC よりも低く，説明変数（ δ index）には負の効果が観察されるはずである。

分析の結果より，いずれの応答変数に対しても， δ index を含むモデルは AIC が低下していることが分かる。そして尤度比検定の結果，モデルの逸脱度は有意に減少した。個人年収や世帯収入を統制してもなお， δ index が効果を持つ点と逸脱度の統計的に有意な減少は，理論モデルのインプリケーションが正しいことを，さらには拡張 δ 区間モデルの有効性を間接的に指示していると言えるだろう。

5 結論

本稿では浜田 (2001) が提唱した δ 区間モデルを 2 次元に拡張し，年齢基準で準拠集団を選択して他者平均収入を予測するモデルを定式化した。そして理論モデルが経験的に正しいかどうかを SSP2015 全国調査データを用いて検証した。

その結果，われわれの定式化した拡張 δ 区間モデルは，1 次元モデルよりも，準拠集団選択過程をより適切に表現しており，データとも適合的であることが判明した。所得に関する準拠集団は，従来のモデルで考えられていたように，所得階層上の近隣から選択されるというよりは，まず年齢という次元の近さで選択され，その局所的な集団の平均が比較の基準として認識されることを分析結果は示唆している。つまり年齢の近さに基づいて準拠集団を選択した結果，準拠集団の所得レンジが縮小するのではないかと考えられる。

δ index は条件付き期待値であり，年齢と所得を含むデータセット上で簡単に計算できる¹⁴⁾。さらに研究者が分析したいデータセットの中に収入データが含まれていなくても，年齢変数さえあれば外部のデータセット上で計算した δ index を外挿できるという意味で，適用範囲の広い index であると言える。

本稿で提案した単純な数理モデルは，準拠集団理論の定式化と実証のさらなる進展に貢献することが期待できるだろう。

表 4: 主観的幸福感・個人収入満足度・世帯収入満足度を応答変数にした分析。()内は標準誤差.

	応答変数:					
	主観的幸福感			個人収入満足度		
	(1)	(2)	(3)	(4)	(5)	(6)
δ index		-0.006*** (0.001)		-0.001*** (0.0003)		-0.002*** (0.0003)
個人年収	0.0001 (0.0002)	0.0001 (0.0002)	0.001*** (0.0001)	0.001*** (0.0001)	-0.0001* (0.0001)	-0.0002*** (0.0001)
世帯年収	0.0003*** (0.0001)	0.0005*** (0.0001)	0.0002*** (0.00005)	0.0002*** (0.0001)	0.001*** (0.00004)	0.001*** (0.00005)
男性ダミー	-0.479*** (0.085)	-0.422*** (0.097)	-0.144*** (0.036)	-0.151*** (0.040)	-0.109*** (0.033)	-0.052 (0.037)
既婚ダミー	0.981*** (0.089)	1.115*** (0.100)	0.172*** (0.037)	0.256*** (0.041)	0.138*** (0.035)	0.172*** (0.038)
世帯人数	0.002 (0.027)	-0.008 (0.031)	-0.036*** (0.012)	-0.035*** (0.013)	-0.026** (0.011)	-0.033*** (0.012)
定数	5.511*** (0.204)	7.123*** (0.367)	1.740*** (0.094)	2.273*** (0.170)	2.044*** (0.079)	2.564*** (0.147)
観測数	2,685	2,137	2,394	1,904	2,681	2,132
大数尤度	-5,458.133	-4,316.410	-2,712.665	-2,105.995	-2,944.631	-2,247.365
AIC.	10,950.270	8,668.819	5,459.330	4,247.991	5,923.262	4,530.731
BIC.	11,050.490	8,770.828	5,557.602	4,347.921	6,023.459	4,632.698

*p<0.1; **p<0.05; ***p<0.01

[注]

- ¹⁾ 動物行動学者ドゥ・ヴァールらの実験によれば、比較に基づく認識の相対性はヒトに固有の性質ではなくサルにも備わっており、不公平な扱いを受けたサルは不満を表すという (de Waal 2009=2010). またシブリーらはカニクイザルを用いた実験で、集団内で下位のサルは上位のサルより、健康指標が悪化することを示した (Shively et al. 1997). このことは精神的・肉体的健康に相対的剥奪が関連することを示唆している (Eibner and Evans 2005).
- ²⁾ 近年、準拠集団を回答者に直接尋ねる質問項目を含む調査が実施されている (例えば European Social Survey 2006; German Socio-Economic Panel Survey 2008, pretest). 具体的には《所得を他者と比較する場合に、その他者をどのカテゴリから選ぶか》を、回答者に尋ねている. 準拠集団のカテゴリとして、よく使われる選択肢としては「近所の人」「友人」「家族」「同じ年齢」「同じ性別」「同じ職業」「同僚」「年取 (が近い人)」などがある (Wolbring et al 2013; Clark, Snik, and Yamada 2013; 飯田 2009,2011; 前田 2014). 結果に多少の違いはあるものの、回答に共通して多く見られた基準は「年取」や「年齢」や「仕事」が近い他者だった. 直接測定最大のメリットは、研究者が準拠集団選択基準について恣意的な仮定を立てなくてもよいことである. ただし重視される基準が分かっても、その基準にしたがって、どういう範囲から他者が選択されるかまでは分からない点に留意しなくてはならない.
- ³⁾ 参照した各研究では年齢だけでなく、居住地域・人種・教育レベル・性別等も組み合わせて準拠集団の範囲を検討している.
- ⁴⁾ δ は distance の d を表している. 本稿の《 δ 区間》とは、解析学・微積分学の文脈では《 δ 近傍》と呼ばれるものである. 数理社会学上 δ 区間という名称が一般的なので、ここでもそう呼ぶ.
- ⁵⁾ 具体的に書けば

$$f_Y(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{t} \exp - \frac{(\log t - \mu)^2}{2\sigma^2}$$

である. ここで μ と σ は分布のパラメータで、確率変数 Y を対数変換した $X = \log Y$ の平均と分散である.

- ⁶⁾ 連続確率変数なので、(2) は z を選択する確率ではない. Y の実現値である 1 点 z が選ばれる確率は常に 0 である.
- ⁷⁾ 条件付き期待値を考える場合には、 Y の確率密度関数は、 $Y < y - \delta, y + \delta < Y$ の範囲で 0 と考える. よって

$$\int_{-\infty}^{\infty} z \cdot \frac{f_Y(z)}{\int_{y-\delta}^{y+\delta} f_Y(t) dt} dz = \int_{-\infty}^{y-\delta} z \cdot 0 dz + \int_{y-\delta}^{y+\delta} z \frac{f_Y(z)}{\int_{y-\delta}^{y+\delta} f_Y(t) dt} dz + \int_{y+\delta}^{\infty} z \cdot 0 dz = \int_{y-\delta}^{y+\delta} z \frac{f_Y(z)}{\int_{y-\delta}^{y+\delta} f_Y(t) dt} dz.$$

- ⁸⁾ 例えば 3 次元の場合、所得を表す確率変数を Y 、その他の次元 (例えば年齢と学歴) を確率変数 X_1, X_2 とおく. 2 次元の δ 区間を $D \subset R^2$ とおけば $D = [d_1, d_2] \times [d_3, d_4]$ という 2 次元閉区間は δ 区間である. $f_{X_1 X_2}(x_1, x_2)$ で (X_1, X_2) の同時確率密度関数 (周辺分布) を表し、添え字の無い $f(x_1, x_2, y)$ で (X_1, X_2, Y) の同時確率密度関数を表せば

$$f_{X_1 X_2}(x_1, x_2) = \int_{-\infty}^{\infty} f(x_1, x_2, y) dy$$

という関係が成立する. このとき、準拠集団の平均所得は

$$\begin{aligned} E[Y|d_1 \leq X_1 \leq d_2, d_3 \leq X_2 \leq d_4] \\ = \int_{-\infty}^{\infty} y \left(\int_{d_3}^{d_4} \int_{d_1}^{d_2} \frac{f(x_1, x_2, y)}{\int_{d_3}^{d_4} \int_{d_1}^{d_2} f_{X_1 X_2}(x_1, x_2) dx_1 dx_2} dx_1 dx_2 \right) dy \end{aligned}$$

である. 以上の考え方をさらに拡張すると、 n 次元 δ 区間モデルが定義できる. ただし実証分析に対応させるためには、3 次元程度が経験的な上限ではないかと推測される. なお δ 区間に収入次元そのものを加えることも可能である. 例えば 3 次元の場合 $D = [d_1, d_2] \times [d_3, d_4] \times [d_5, d_6] \subset R^3$ として

$$\begin{aligned} E[Y|(x_1, x_2, y) \in D] &= E[Y|d_1 \leq X_1 \leq d_2, d_3 \leq X_2 \leq d_4, d_5 \leq Y \leq d_6] \\ &= \int_{-\infty}^{\infty} y \left(\iiint_D \frac{f(x_1, x_2, y)}{\iiint_D f(x_1, x_2, y) dx_1 dx_2 dy} dx_1 dx_2 dy \right) dy. \end{aligned}$$

- ⁹⁾ パラメトリックに準拠集団平均所得を計算する方法もあるが、本稿では年齢別準拠集団を連続的に変化させるため、データに依拠して経験的に推定する.
- ¹⁰⁾ δ の値として何を仮定すべきかは、理論的かつ経験的な問題である. δ が小さすぎると、準拠集団別平均所得を計算する際に標本が小さくなるせいで誤差が大きくなる. また、 δ が大きすぎると標本全体のサイズが減少するために、全パラメータの推定誤差が大きくなる. したがって本稿では小さ過ぎず大き過ぎず、かつ一般に人が同年代という言葉で思

い浮かべる範囲として無理のない数値として、 ± 5 を仮定した。なお $\delta = 3$ で推定した場合も（誤差は変化するが）結果の傾向は大きくは変化しなかった。

¹¹⁾ 欠損値と EGP6 分類の無職・学生は分析から除外した。

¹²⁾ 統制変数の一部である教育水準と職業カテゴリの係数および有意性検定の表示は省略した。教育水準や職業カテゴリは、 δ 区間モデルの次元として統計モデルの説明変数に変換しなければ、理論的に意味が無い。この分析における統制変数は、観測できない個人レベルの誤差をなるべく抑えるために用いたので、係数についてのアドホックな解釈は一切しない。

¹³⁾ 変数間の相関を確認したところ、個人収入と世帯収入の相関はやや高い (0.59) が、分析に支障をきたす程度ではないので、統制変数として両方を用いた。

¹⁴⁾ ただし計算にはプログラミング可能な統計解析パッケージが必要である。

[付記] この研究は SSP プロジェクト (<http://ssp.hus.osaka-u.ac.jp/>) の一環として行われたものである。SSP2015 データの使用にあたっては SSP プロジェクトの許可を得た。

[文献]

- Ball, Richard and Kateryna Chernova, 2008, “Absolute income, relative income, and happiness,” *Social Indicators Research*, 88(3): 497-529.
- Bartoszynski, Robert and Magdalena Niewiadomska-bugaj, 2008, *Probability and Statistical Inference, Second Edition*, A John Wiley & Sons.
- Berrebi, Z. M. and Jacques Silber, 1985, “Income inequality indices and deprivation: a generalization,” *Quarterly Journal of Economics*, 100(3): 807-10.
- Bossert, W. and Conchita D’Ambrosio, 2006, “Reference groups and individual deprivation,” *Economics Letters*, 90(3): 421-426.
- Chakravarty, Satya R., Chattopadhyay, N. and Majumder, A. 1995, “Income inequality and relative deprivation,” *Keio Economic Studies*, 32: 115.
- Clark, Andrew E., and Andrew J. Oswald, 1996, “Satisfaction and comparison income,” *Journal of Public Economics*, 61: 359-81.
- D’Ambrosio, Conchita, and Joachim R. Frick, 2007, “Income satisfaction and relative deprivation: an empirical link,” *Social Indicators Research*, 81(3): 497-519.
- de Waal, Frans, 2009, *The Age of Empathy: Nature’s Lessons for a Kinder Society*, New York, NY: Harmony Books.(=2010, 柴田裕之訳『共感の時代へ』紀伊國屋書店.)
- Duesenberry, J.S. 1949, *Income, Saving and the Theory of Consumer Behavior*, Cambridge, MA: Harvard Univ. Press,
- Eibner, Christine and William N. Evans, 2005, “Relative deprivation, poor health habits, and mortality,” *Journal of Human Resources*, 40(3): 591-620
- Ebert, Udo. and Patrick Moyes, 2000, “An axiomatic characterization of Yitzhaki’s index of individual deprivation.” *Economics Letters*, 68(3): 263-70.

- Ferrer-i-Carbonell, Ada, 2005, "Income and well-being: an empirical analysis of the comparison income effect," *Journal of Public Economics*, 89(5-6): 997-1019.
- 浜田宏, 2001, 「経済的地位の自己評価と準拠集団 —— δ 区間モデルによる定式化」『社会学評論』206号(第52巻・第2号): 283-99.
- Hey, John. D. and Peter J. Lambert, 1980, "Relative deprivation and the Gini coefficient: comment," *Quarterly Journal of Economics*, 95(3): 567-573.
- 石田淳, 2011, 「相対的剥奪と準拠集団の計量モデル」『理論と方法』26(2): 371-88.
- 石田淳, 2015, 『相対的剥奪の社会学——不平等と意識のパラドックス』東京大学出版会.
- 飯田義郎, 2009, 「相対所得における他者とは誰か——アンケート調査から」『京都産業大学論集. 社会科学系列』26:131-56.
- 飯田義郎, 2011, 「相対所得におけるリファレンスグループの選択とその動機」『京都産業大学論集. 社会科学系列』28:1-23.
- Kakwani, Nanak 1984, "The relative deprivation curve and its applications," *Journal of Business & Economic Statistics*, 2(4): 384-94.
- Kondo, Naoki, Ichiro Kawachi, S.V. Subramanian, Yasuhisa Takeda, Zentaro Yamagata, 2008, "Do social comparisons explain the association between income inequality and health?: relative deprivation and perceived health among male and female Japanese individuals," *Social Science & Medicine*, 67(6): 982-7.
- Luttmer, Erzo, 2005, "Neighbors as negatives: relative earnings and well-being," *Quarterly Journal of Economics*, 120(3): 923-1002.
- 前田豊, 2014, 「比較対象選択と所得イメージ」『理論と方法』29(1): 37-58.
- Maeda, Yutaka, and Atsushi Ishida, 2013, "Income comparison as a determining mechanism of class identification: a quantitative and simulation study using Japanese survey data," *International Journal of Japanese Sociology*, 22: 143-59.
- McBride, Michael, 2001, "Relative-income effects on subjective well-being in the cross-section," *Journal of Economic Behavior & Organization*, 45: 251-278.
- Merton, Robert K. 1957, *Social Theory and Social Structure*, New York: The Free Press. (=1961, 森東吾・森好夫・金沢実・中島竜太郎訳『社会理論と社会構造』みすず書房.)
- Oshio, Takashi, Kayo Nozaki, and Miki Kobayashi, 2011, "Relative income and happiness in Asia: evidence from nationwide surveys in China, Japan, and Korea," *Social Indicators Research*, 104: 351 - 67.
- Runciman, Walter, G. 1966, *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-century England*, London: Routledge & K. Paul.
- Shively, Carol A., Kathy Laber-Laird, and Raymond, F. Anton, 1997, Behavior and physiology of social stress and depression in female cynomolgus monkeys, *Biological Psychiatry*, 41(8):

871-882,

- Stewart, Quincy T. 2006, “Reinvigorating relative deprivation: a new measure for a classic concept.” *Social Science Research*, 35(3): 779 - 802.
- Stouffer, Samuel A., Edward A. Suchman, Leland C. DeVinney, Shirley A. Star, and Robin M. Williams, Jr. 1949, *Studies in Social Psychology in World War II: The American Soldier. Vol. 1: Adjustment During Army Life*, Princeton: Princeton University Press.
- Subramanyam, Malavika, Ichiro Kawachi, Lisa Berkman, and S. V. Subramanian, 2009, “Relative deprivation in income and self-rated health in the United States.” *Social Science & Medicine*, 69(3): 327 - 34.
- Yitzhaki, Shlomo, 1979, “Relative deprivation and the Gini coefficient,” *Quarterly Journal of Economics*, 93(2): 321-324.