

# Jimbun Chiri

Japanese Journal of Human Geography

Vol.64 No.3 2012

## Article

HANAOKA Kazumasa

The Possibility of Using Anonymized Data from Japanese Official Statistics for Implementing Geospatial Analysis at a Small Area Level: Creating a Geographically Disaggregated Synthetic Microdata Set by Using a Spatial Microsimulation..... 1

## Review

A Survey of Geographical Studies in Japan, 2011 ..... 18

## Research Notes

NAKAZAWA Takashi

Factors Contributing to Subcontracted Workers' Vulnerability with Regard to Their Employment and Housing Status: A Case of Visitors to a Local Government Consultation Desk during the Financial Crisis ..... 65

MIHARA Atsumi

Introduction of Advanced Medical Equipment and the Role of Long-distance Travelling Recipients in Non-metropolitan Area: A Case Study of PET Screening Tours in Medical Facilities in Koriyama City, Fukushima Prefecture ..... 84

News ..... 102

## 公的統計「匿名データ」を用いた小地域単位での地理空間分析の可能性

——空間的マイクロシミュレーションによる地理的な合成マイクロデータの生成——

花岡和聖

I はじめに	(1)焼きなまし法の概要
II 研究方法	(2)合成マイクロデータの適合度評価
(1)空間的MSとは	IV 消費特性の地域差
(2)地理的な合成マイクロデータの生成法	(1)草津市内の品目別消費支出額
III 焼きなまし法による合成マイクロデータの生成	(2)草津駅前地区の品目別消費支出額
	V まとめ

キーワード：匿名データ、全国消費実態調査、空間的マイクロシミュレーション、焼きなまし法、合成マイクロデータ、草津市

### I はじめに

社会科学における研究手法は、集計データを用いたマクロな分析から、個票データを用いたミクロな分析へと、近年、大きく転換してきた。地理学も例外ではない。欧米諸国では、1960年代以降、データアーカイブが整備され、多様な調査主体による個票データが公開され、学術研究などに利用されてきた。その中には、政府機関が実施した公的統計も含まれ、アメリカやカナダ、イギリスなどでは国勢調査の個票データが公開されている。こうした公的統計を分析に用いるメリットは、民間調査会社や大学等が実施する調査と比較すると、全国を調査範囲とし、その標本数も数万件を超える点にある。加えて、公的統計は、厳密な標本設計や高い回収率、多様な調査項目数の点でも優れている場合が多く、学術的な利用価値は非常に高い。たとえば、イギリスでは、1990年代後半に、

国勢調査から数%の個票を再抽出した匿名データ(SARs: Samples of Anonymised Records)が公開された。そのデータ提供元のホームページには、SARsを用いた、移民や健康、所得格差、人口移動などに関する研究論文440件以上が紹介され、そのうちの地理学関連の主要雑誌に掲載された論文も多数挙げられる。

日本でも、近年、公的統計の個票データの二次利用が許可されるようになった。我が国の地理学においても、公的統計の個票データに対する関心は高く、国勢調査の個票データを用いた在日外国人研究<sup>4)</sup>や就業構造基本調査を用いた所得格差研究<sup>5)</sup>は、先駆的な研究事例である。ただし、これらの研究では、研究者自身が個別に、データ利用を担当部署に申請する必要があったため、西原が「個票は利用手続きに多大な時間を要する」と指摘するように、個票データ利用自体のハードルが高かった。そうした問題を解決すべく、2009年になって、

統計法が改正され、個人が識別されないよう公的統計の原データに匿名化の措置が施された「匿名データ」が公開された。公的統計のうち「全国消費実態調査」(約4.8万世帯)と「社会生活基本調査」(約27万世帯)、「就業構造基本調査」(約75万世帯)、「住宅・土地統計調査」(約35万世帯)、「労働力調査」(約78万人)、「国民生活基礎調査」(約3.7万世帯)がすでに公開され、独立行政法人統計センターや厚生労働省の窓口を通じて、利用申請手続きが可能である。これによって、公的統計の個票データに、研究者がより容易にアクセスできるようになった。

ただし、匿名データには、匿名化措置として、氏名・住所情報の削除に加えて、全標本のうち8割または1割が再抽出され、世帯人員8人以上の世帯の削除、年齢の階級化、収入のトップコーディング、分類項目の統合などが適用される。また調査項目の一部が表章されない場合もある。加えて、この匿名データを地理空間分析に利用するとすれば、匿名化によって地域区分が大幅に統合されてしまい、実質的に全国単位または都道府県単位での分析に限定されてしまう。具体的には、住宅・土地統計調査の地域区分は「都道府県」、全国消費実態調査と社会生活基本調査、就業構造基本調査のそれは「三大都市圏か否か」、労働力調査と国民生活基礎調査では「全国」のみに統合される。これは地域区分が個人を特定する上で重要な変数であり、匿名化においては一地域内の人口が50万人未満にならないようガイドラインが示されている<sup>8)</sup>。一般的に、匿名化の程度と、統計調査の個票データの公開の程度との関係は、一方を緩くすると他方が強化される関係にある。したがって、匿名データの公開に際して、他の変数と比較して項目数の多い地域区分の開示は、個人が特定されるリスクがあり、ハードルが高い。

個人が特定されるリスクをなくしつつ、分析的な有用性を維持し、公的統計の個票データを広く公開する方法として、最近では、synthetic mi-

crodata という考え方が普及している。これは「模造データ」や「疑似データ」と訳され、原データそのものではないが、その統計的性質を保持したデータである。模造化の方法として、個票をスワッピングする方法や収入や意見などの個人情報に関連する変数の一部もしくはすべてについて、原データの観測値ではなく多変量解析の推定値に置き換える方法がある。実際に、アメリカでは国勢調査や労働力調査の模造データが公開されてきたし、日本でも「レプリカデータ」と呼ばれる公的統計の模造データが試行的に提供された。そして、星野は、詳細な変数を含む模造データの方が、強力な匿名化によって多くの変数が削除・統合された原データよりも、学術研究における統計分析上の利用価値が高いと指摘する。

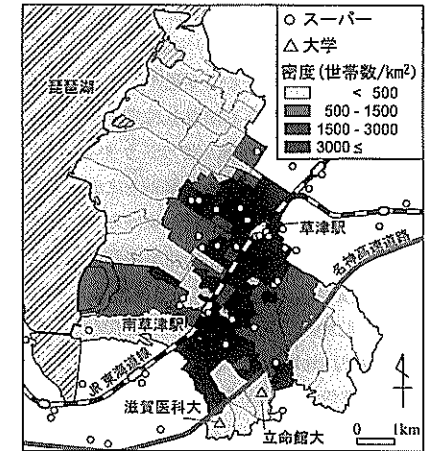
本研究で使用する空間的マイクロシミュレーション法(以下、空間的MS)によるアプローチは、個人情報の匿名化によるデータ公開に主眼を置く統計学の模造データとはその目的を異にするが、実質的には、個票データから、国勢調査小地域集計等の地域統計に整合する「地理的」な模造データ(ここでは「合成マイクロデータ」と呼ぶ)を構築する点で、模造化手法の一種とみなすこともできる。この合成マイクロデータを、原データの代替として利用することで、従来、困難であった公的統計の個票データを用いた小地域単位での地理空間分析が実現可能になる。しかし、日本における空間的MSの研究成果は少なく<sup>12)</sup>、加えて、管見の限り、他の手法で合成マイクロデータを生成する研究事例もない。また合成マイクロデータの生成に使用されるデータも、公的統計の匿名データのように広く公開された大規模な個票データではない。

そこで、本研究の目的は、空間的MSを用いて、公的統計の匿名データから、地理的に詳細な合成マイクロデータを生成し、その精度評価および利用事例を通じて、匿名データを用いた小地域単位での地理空間分析の実現可能性を検討することである。なお、本研究では、全国消費実態調査の

匿名データを用い、消費特性の地域差を推定する。その理由として、個人の生活様式の多様化や所得格差の拡大、都市再開発によって、消費行動の地理的分布も複雑化し、消費実態の把握は、商店街やスーパーマーケット(以下、スーパー)などの小売店舗の立地問題ばかりでなく、中谷がレビューするようにフードデザートや肥満などの食・健康・貧困問題とも密接に関連し、近隣を単位とした地理学的研究において重要な基礎資料となり得ると考えられるからである。

なお、研究対象地域は、滋賀県草津市(第1図)とする。その選定理由は、第一に、立命館大学地理学教室による「草津市消費者購買行動調査」や花岡・クラークによる草津市を対象とした空間的MSに関する既往研究と比較できること、第二に、草津市は、京都や大阪の通勤圏に位置し、JR東海道線の2駅(南草津駅・草津駅)に近接して単身・核家族世帯が入居する多数の中高層マンション、その周辺での戸建て住宅団地、南部での大規模な大学や工場などの開発が進み、1970年代以降、人口・世帯数とも継続して増加してきたことである。平成17年国勢調査によると、草津市の人口は121,159人、世帯数は49,778世帯である。その結果、大型小売店舗が多数進出し、商店街や中小スーパーも含めて店舗間競争が熾烈になっている。加えて、北部の琵琶湖沿岸地域は農業的土地利用が顕著であるが、都市的土地利用とのスプロールの混在も認められる。このような草津市では、消費実態を地理的に詳細に把握する必要が高いことが対象地域の選定理由に挙げられる<sup>16)</sup>。

本研究の構成は以下の通りである。次章で、空間的MSおよび地理的な合成マイクロデータの生成方法について説明する。IIIでは、焼きなまし法を用いて合成マイクロデータを実際に生成し、その適合度評価を行う。IVでは、合成マイクロデータを再集計し、小地域単位での消費実態を地図化することで、地理空間分析での利用例を提示する。Vで、本研究の成果を整理し、匿名データを用



第1図 研究対象地域(滋賀県草津市)  
Figure 1. Study area (Kusatsu City, Japan)  
資料:平成17年国勢調査小地域集計

いた地理空間分析の可能性を考察する。

## II 研究方法

(1)空間的MSとは 空間的MSの端緒は、1950年代後半に経済学者オルコットによって提示された(非空間的な)マイクロシミュレーション<sup>17)</sup>に遡る。この方法は、個人や世帯、企業などの意思決定主体を分析単位とするシミュレーション方法である。そこでは個人に加齢や出生、死亡などの人口推定を行い、その予測をベースにした住宅需要や財政問題、福祉政策の政策評価が進められた<sup>18)</sup>。この考え方は、交通工学にも波及し、自動車一台単位での交通流動の予測モデルが構築された<sup>19)</sup>。

地理学では、経済学でのマイクロシミュレーション研究が政策評価を目的としながらも地域差を軽視しているとの批判にあるように、当初、従来のマイクロシミュレーションに地理的な要素を組み込む形で空間的MSが提示された。したがって、空間的MSとは、大規模な個票データを用いた、個人や世帯などの意思決定単位に基づく地理的なモデリング法である。そして、その分析結

果は、地域政策などの政策評価や意思決定に活用される。これまでにイギリスを中心に、空間的MSを用いて、所得・貧困や犯罪、医療・健康、失業の地域差が推定され議論されてきた。また購買行動の分野では、花岡・クラークや中谷ほか、レーベン、レーベン・リートベルトによる購買需要や買物移動を推定した研究が挙げられる。

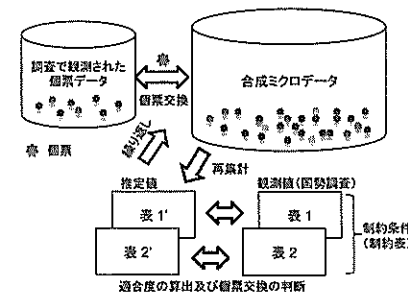
イギリスの既往研究で主に使用されるデータは、イギリス国勢調査の匿名データ SARs の世帯版である。しかし、そこで表章される地域区分は1991年版で13地域のみ、2001年版からは地域区分が一切表章されない。そのため、まず詳細な地理情報を付与した個票データの推定方法が求められた。結果として、空間的MS研究の中で、国勢調査や公的統計の個票データなど複数の統計データを駆使して、地理的に詳細な合成マイクロデータを生成する方法が複数提案され、改良されてきた。第2次節では、空間的MSの既往研究における、そうした地理的な合成マイクロデータの生成法について述べる。

(2) 地理的な合成マイクロデータの生成法 空間的MS研究において、個票データから地理的な合成マイクロデータを生成する手法として、合成復元 (Synthetic reconstruction) 法と組合せ最適化アルゴリズムを用いた手法が提示されてきた。

合成復元法は、第一段階として、IPF (Iterative Proportional Fitting) 法を用いて、複数の集計データの周辺度数から、それらの同時分布を求める。この手法で重要な IPF 法は、標本バイアスの補正や欠損値推定に利用され、レイキング法やRAS法とも呼ばれる。例を挙げると、国勢調査小地域集計では、性別・年齢階級別人口や産業別人口がわかるが、性別・年齢階級別産業別人口は表章されていない。そこで IPF 法を用いて性別・年齢階級と産業の各集計表 (周辺度数) からもっともらしいクロス集計表 (同時分布) を推定する。次に、第二段階として、クロス集計表のセル毎の度数の比率を求め、これを条件付き確率と

する。ここでは男性で年齢階級が20歳代であれば、建設業に就く確率が10%、製造業が30%、サービス業が5%…というように情報が推定される。第三段階では、乱数を発生させ、上の条件付き確率分布に基づき、その地域に居住する一人一人にいずれかの産業を割り当てる。この作業を小地域内の人口数だけ行うことで、合成マイクロデータを生成できる。パーキン・クラークの事例では、合成復元法によって、世帯人数や子の数、子の年齢階級などが逐次的に推定される。同手法のメリットは、集計データのみから合成マイクロデータを一から生成できる点にあるが、非現実的な属性の組合せが生じる可能性もあり、国勢調査との整合性は組合せ最適化アルゴリズムの方が高いとされる。

一方、組合せ最適化アルゴリズムによる合成マイクロデータの生成には、集計データと個票データが必要となる。組合せ最適化アルゴリズムは、計算機科学で発展してきた研究分野であり、山登り法や焼きなまし法、遺伝的アルゴリズム、ニューラルネットワークなどを利用して、効率的かつ高速に最適解を求める手法である。ウィリアムソンらは、この考え方を合成マイクロデータの生成に初めて応用した。この方法では、第2図に示すように、調査で観測された個票データからランダムに抽出を繰り返して合成マイクロデータの初期解を初めに生成する。次に、合成マイクロデータと観測された個票データ間で個票を交換し、合成マイクロデータから国勢調査小地域集計の集計表 (制約条件、制約表) と同一の表を作成し両者を比較する。両者の一致が改善する個票を繰り返し受け入れることで、制約表に一致する個票で構成された合成マイクロデータを求める。この手法のメリットは、個票データを利用することで、調査を通じて観察された実際の人口属性の組合せを利用できる点、世帯や個人などの階層構造を踏まえた推定が可能になる点にある。ただし、推定する合成マイクロデータの規模に応じて、ある程度の標本数があり、制約条件と共通した変数が表章された個票デ



第2図 焼きなまし法による合成マイクロデータ生成

Figure 2. Synthetic microdata creation by simulated annealing

ータが必要な点、計算時間が長いことがデメリットとして挙げられる。

本研究では、匿名データを利用する点および制約条件とする既存統計との整合性が高い点を考慮し、組合せ最適化アルゴリズムによる合成マイクロデータの生成を行う。

### III 焼きなまし法による合成マイクロデータの生成

(1) 焼きなまし法の概要 組合せ最適化アルゴリズムのうち、高熱の金属を冷却する工程を模した焼きなまし法を本研究で用いる。組合せ最適解の探索過程において、局所解に陥る可能性があるが、焼きなまし法では、探索過程の初期段階で制約条件との整合性を低下させるような解を一定割合受け入れることで、局所解からの脱出が図られる。ウィリアムソンらは、複数のアルゴリズムを比較検討した結果、焼きなまし法が、制約条件ともっともよく整合する合成マイクロデータを生成できることを示した。また花岡は、パーソントリップ調査の個票データに焼きなまし法を適用し、標本バイアスを拡大補正させる形で、特殊施設が立地する町丁目を除くと、国勢調査小地域集計と高い一致を示す合成マイクロデータが生成できることを示した。以下、本研究における焼きなまし法を用いた合成マイクロデータの作成手順 (Step 1~3)

を説明する。

Step 1: 国勢調査小地域集計のうち、匿名データと同じ変数で、世帯の消費特性とも関連すると考えられる変数が表章される集計表を、制約表に利用する。そこで、個人レベルの制約表として、①性別年齢5歳階級別人口および②性別産業別人口を、世帯レベルの制約表として、③世帯人員、④住宅の所有関係、⑤住宅の建て方を用いた (第1表)。なお不詳分に関しては、他の分類の度数に応じて比例配分した。また秘匿地域は、合算先の町丁目に統合させた。

Step 2: 2004年全国消費実態調査の匿名データを用意する。全国消費実態調査では草津市は三大都市圏に分類されるため、三大都市圏に区分され、かつ年収票がある個票データ18,364世帯分を使用する。ただし、全国消費実態調査では、施設等の世帯および学生単身世帯が調査対象外である。そこで本研究では施設等の単身世帯は単身世帯の全データ、単身の学生世帯は15-24歳の単身世帯の個票データを割り当てた。

Step 3: 秘匿地域を合算した113町丁目に対して、個別に焼きなまし法を適用し、国勢調査小地域集計と整合する合成マイクロデータを作成する。その適合度指標として、ボアス・ウィリアムソンが提示する以下の指標を用いる。まず Total Absolute Error (TAE) と Overall TAE (OTAE)、1世帯当たりの OTAE を式(1)~(3)に示す。なお、以下に示す適合度指標すべてについて、町丁目別に求めるが、各町丁目を示す添え字は式から省略した。

$$TAE_k = \sum_i |E_{ki} - O_{ki}| \quad (1)$$

$$OTAE = \sum_k TAE_k \quad (2)$$

$$1 \text{世帯当たり } OTAE = OTAE/S \quad (3)$$

ここで、Eは焼きなまし法での推定値、Oは国勢調査から得られる観測値を表す。添え字kは制約表を、iは制約表の各セルを示す。TAEは観

第1表 合成マイクロデータの生成に用いる制約表  
Table 1. Constraint tables for creating synthetic microdata

個人レベル			産業		男女		世帯レベル	
年齢階級	男	女	農業	1	20	世帯人員		
0~4歳	1	19	林業	2	21			
5~9歳	2	20	漁業	3	22	1人	1	
10~14歳	3	21	鉱業	4	23	2人	2	
15~19歳	4	22	建設業	5	24	3人	3	
20~24歳	5	23	製造業	6	25	4人	4	
25~29歳	6	24	電気・ガス・熱供給・水道業	7	26	5人	5	
30~34歳	7	25	情報通信業	8	27	6人	6	
35~39歳	8	26	運輸業	9	28	7人以上	7	
40~44歳	9	27	卸売・小売業	10	29	施設等の世帯	8	
45~49歳	10	28	金融・保険業	11	30	住宅の建て方		
50~54歳	11	29	不動産業	12	31			
55~59歳	12	30	飲食店、宿泊業	13	32	一戸建	1	
60~64歳	13	31	医療、福祉	14	33	長屋建	2	
65~69歳	14	32	教育、学習支援業	15	34	共同住宅(1・2階)	3	
70~74歳	15	33	視覚サービス事業	16	35	共同住宅(3~5階)	4	
75~79歳	16	34	サービス業	17	36	共同住宅(6~10階)	5	
80~84歳	17	35	公務	18	37	共同住宅(11階以上)	6	
85歳以上	18	36	非純業・分類不能	19	38	その他	7	
						間借り	8	
						施設等・住宅以外	6	
						施設等	9	

注) 表中の数字は項目番号を示す。

測値と期待値の差の絶対値の合計であり、OTAEはすべての制約表についてTAEを合計した値である。OTAEは、町丁目の世帯数に応じて増大するので、町丁目の世帯総数Sで除した値として、1世帯当たりOTAEを求める。適合度の目安として、経験的に、この値が1未満であることが望ましい<sup>39)</sup>。

もう一つの指標として、 $\chi^2$ 検定を用いる方法がある<sup>40)</sup>。この方法ではセル数が異なる制約表を比較できるように調整したRelative Sum of Squared Z scores (RSSZ)を用いる(式4・5)。

$$RSSZ_k = \sum_i F_{ki} (E_{ki} - O_{ki})^2 \quad (4)$$

ただし、

$$F_{ki} = \begin{cases} \chi^2_k E_{ki} \left(1 - \frac{E_{ki}}{P_k}\right)^{-1} & \text{if } E_{ki} \neq 0 \\ \frac{1}{\chi^2_k} & \text{if } E_{ki} = 0 \end{cases} \quad (5)$$

ここで、 $P_k$ は制約表kの総度数を示す。 $\chi^2_k$ は、制約表kについて、有意水準5%での $\chi^2$ 値を示す。観測値に応じて $F_{ki}$ を求め、それを観測値と期待値の差の平方値に乘じRSSZを求める。RSSZは合成マイクロデータが制約条件に完全に一致する場合に0となり、両者の差が小さいほど0に近い値をとる。そしてRSSZが1より小さい場合、焼きなまし法の結果と国勢調査の集計値に有意水準5%で有意差がないと判断できる。以下のように、各制約表のRSSZの合計値をOverall RSSZ (ORSSZ)とし、合成マイクロデータの全体的な適合度を評価できる(式6)。

$$ORSSZ = \sum_k RSSZ_k \quad (6)$$

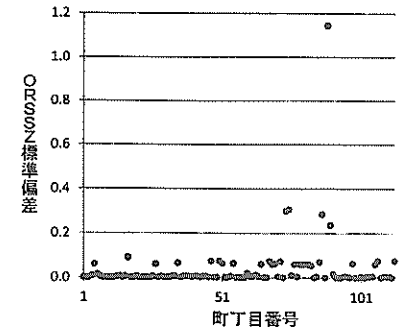
本研究では、適合度指標ORSSZを用いて、匿名データと合成マイクロデータ間で個票の交換を行い、新たな個票の組合せ解を求める。個票の交換を採用するか否かは、以下の基準(式7)で判断する。

$$\text{Prob}(\text{accept}) = \begin{cases} 1 & \text{if } \Delta C < 0 \\ e^{-\frac{\Delta C}{aT}} & \text{otherwise} \end{cases} \quad (7)$$

$$\Delta C = C' - C$$

Prob(accept)は交換した個票の採用確率を、Tは温度パラメータを示す。この温度パラメータTには、個票を採用した回数N回ごとに温度減少率aを適用する。Cは個票の交換前の適合度、C'は交換後の適合度、 $\Delta C$ はその差分を示す。この基準により、適合度が改善する個票の交換は必ず採用し、適合度が改善しない場合でも、温度に応じて個票の交換を採用するかを判断する。適合度が改善しない個票の交換を採用する確率は、探索の初期段階で高く、個票の交換回数が増えるに従い温度が下がり、その確率が徐々に低下する。最終的に、最大回数Mもしくは適合度が基準値L以下になれば処理を終了する。この一連の作業を、町丁目ごとに5回ずつ繰り返す、ORSSZが最も低い値を示した合成マイクロデータを採用した<sup>41)</sup>。なお、各パラメータの初期値は、ウィリアムソンを参考に、 $T=97$ ,  $a=0.95$ ,  $L=0$ ,  $N=970$ 回、 $M=500$ 万回とした。上記の処理は、焼きなまし法のソフトウェアCOで行った<sup>42)</sup>。

(2)合成マイクロデータの適合度評価 町丁目別に得られた5回分の合成マイクロデータについてORSSZの標準偏差を求め、それらを散布図にプロットした(第3図)。この図からは、大半の町丁目目で、標準偏差が0.1未満を示し、高くても0.3前後であることから、試行回によって適合度が大きく変動する可能性は小さいことがわかる。ORSSZの平均値は1.10であり、大半の町丁目



第3図 町丁目別ORSSZの標準偏差(5回試行)  
Figure 3. Standard deviation of ORSSZ by cho-cho-moku (five trials)

関して、高い適合度で安定した合成マイクロデータの生成が可能であった。ただし、外れ値も含めて、草津市南部の大学周辺の町丁目目で標準偏差が高い傾向にある。さらに、これらの地域では、ORSSZの町丁目別平均値も相対的に高い。つまり、適合度が悪い町丁目目で、試行回数ごとの適合度により大きい変動がみられた。

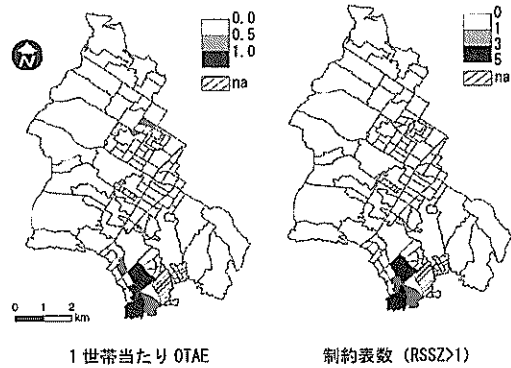
次に、5回の試行のうちORSSZが最小値を示した合成マイクロデータについて、町丁目別の適合度指標を集計した(第2表)。まずOTAEの平均値は39.63を示し、制約条件とした5つの統計表との度数ベースでの乖離は、統計表のセル数を踏まえると十分に小さい。1世帯当たりOTAEの平均値も0.10を示し、目安となる1世帯当たりOTAEが1以下を示す町丁目数は113町目中110であった。ORSSZの平均値は1.06を示し、制約表のRSSZを個別にみても平均値が1よりも十分に小さい。すべての制約表についてRSSZが1未満を示す町丁目数は107であった。したがって、ほとんどの町丁目目で、国勢調査小地域集計と合成マイクロデータに有意差が確認されず、両者は高い精度で整合していると言える。ただし、項目数が多い性別年齢階級および性別産業別の統計表に関しては、TAEとRSSZともに平均値と標準偏

第2表 適合度指標 (最良結果)

Table 2. Goodness-of-fit measures (best results)

	平均値	標準偏差	最小値	最大値
OTAE	39.63	136.75	0.00	1184.00
1世帯当たり OTAE	0.10	0.25	0.00	2.05
ORSSZ	1.06	6.72	0.00	61.66
制約表数 (RSSZ >1)	0.13	0.68	0.00	5.00
TAE (性別年齢5歳階級)	19.92	66.75	0.00	614.00
TAE (性別産業)	12.65	42.19	0.00	346.00
TAE (世帯人員)	2.23	9.47	0.00	78.00
TAE (住宅の所有関係)	1.34	7.09	0.00	54.00
TAE (住宅の建て方)	3.50	15.98	0.00	102.00
RSSZ (性別年齢5歳階級)	0.37	2.37	0.00	23.10
RSSZ (性別産業)	0.47	3.78	0.00	39.70
RSSZ (世帯人員)	0.02	0.15	0.00	1.40
RSSZ (住宅の所有関係)	0.09	0.80	0.00	8.50
RSSZ (住宅の建て方)	0.09	0.50	0.00	3.50
世帯数	447.24	542.27	12.00	4486.00

町丁目数・113



第4図 適合度指標 (町丁目)

Figure 4. Goodness-of-fit measures (cho-cho-moku)

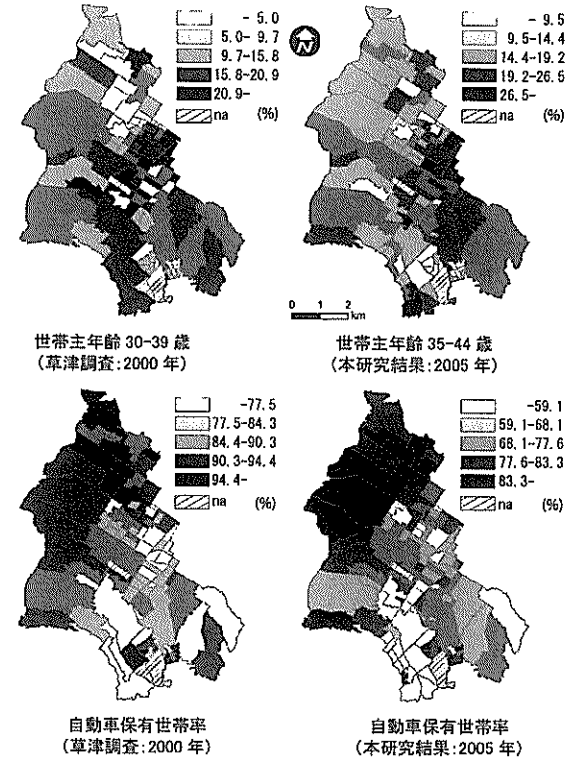
注) 斜線部は隠匿地域

差が若干高く、項目数の多い統計表ほど当てはまりが悪い傾向にある。

さらに、どのような地域で適合度指標が相対的に高いかを把握するため、1世帯当たり OTAE および RSSZ が1より大きい制約表数を地図に示す(第4図)。その結果をみると、草津市南部の大学周辺の町丁目でのみ適合度が極端に悪い。花岡やボアス・ウィリアムソンの研究では、極端に偏った人口構成を示す地区では、焼きなまし法の当

てはまりが悪いことが報告される。草津市のこれらの町丁目でも同様に、医科大学や医療施設、工場、それらの寮・社宅などが立地し、就業者の産業や住宅の建て方の構成に極端な偏りが見られた。この点は、第4図で示した、試行ごとの適合度に大きい変動がみられた要因とも考えられる。

ここで、立命館大学地理学教室が2000年に実施した購買行動に関する悉皆調査「草津市消費者購買行動調査」(以下、草津調査)の結果から、本研



第5図 草津調査結果との比較 (町丁目)

Figure 5. Comparisons to the Kusatsu Consumer Survey (cho-cho-moku)

注) 階級区分は五分位

究結果とも共通して得られる変数として、世帯主年齢が30歳代の世帯割合と自動車保有世帯割合を取り上げ、制約に含まれていない変数について、草津調査と本研究結果を比較する。

まず、世帯主年齢が30歳代の世帯割合を比較する。対象年次が5年異なるため、本研究の結果については、世帯主年齢35-44歳を対象に集計を行った。世帯主年齢は制約表にはないが、性別年齢5歳階級別人口が制約表に含まれているため、両者に正の相関関係が期待される。第5図をみると、鉄道駅周辺から草津市南部にかけての範囲および大学周辺で高い割合がみられ、全体的な空間的分

布パターンはおおよそ類似する(相関係数=0.562)。その町丁目別割合の平均は、草津調査で13.6%に対して、本研究結果で16.7%と若干高いのみであった。次に、自動車保有率の分布では、空間的分布の一致が確認できる(相関係数=0.795)。保有率は、琵琶湖沿岸地域と立命館大学周辺の住宅団地で高い。その平均は、草津調査で85.0%、本研究結果で71.5%であった。

以上のように、制約にはない変数についても、本研究の合成マイクロデータの適合度は高いと考えられる。こうした結果が得られた理由として、焼きなまし法で用いた制約表の数と種類が重要であ

る。制約表にはない変数の適合度は、制約表との関連性に依存する。花岡・クラークによる町丁目別世帯人員別世帯数と草津市全域での世帯主年齢別世帯数を制約表に用いて推定された合成マイクロデータに対して、本研究においては、住宅や性別年齢階級別人口など世帯主年齢を推定する上で必要な制約表をより多く考慮できたことで、世帯主年齢割合や自動車保有割合の適合度の改善に結び付いたと考えられる。ただし、平均割合の差や2変数の相関係数に関しても、草津調査における町丁目間での回収率のばらつきや世帯主年齢が低いほど回収率が低い問題を考慮すると、本研究結果は妥当な範囲にあると考えられる。

以上、全国消費実態調査の匿名データは、標本バイアスが小さく<sup>46)</sup>、国勢調査とも共通する変数が複数表章され、約1.8万世帯もの個票データが含まれるため、結果として、詳細な制約表に整合する個票データの組合せを見つけ出すことができる<sup>49)</sup>。そのことが、一部の例外的な地域を除いて、本研究の合成マイクロデータが、制約表および非制約表ともに適合度が高くなった要因であると考えられる。大規模な個票データほど焼きなまし法の計算に時間がかかるデメリットを除けば、本研究のように、合成マイクロデータの生成に匿名データを用いるメリットは推定精度の点からも大きいと言える。

IV 消費特性の地域差

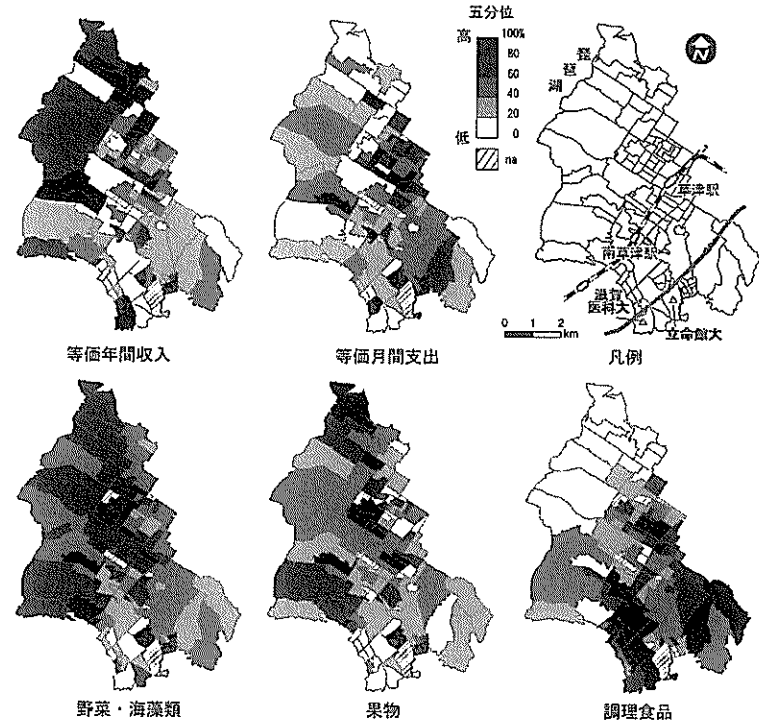
(1)草津市内の品目別消費支出額 空間的 MS の合成マイクロデータを町丁目別に再集計することで、全国消費実態調査に表章される多様な変数の地図化を試みる。ここでは、商店街やスーパーを対象とした商圏分析を第一に想定し、等価年間収入と等価月間支出、1か月当たりの品目別消費支出額のうち野菜・海藻類と果物、調理食品を例に1人当たり消費支出額を求めた(第6図)。

まず、等価年間収入に関しては、草津駅周辺及び琵琶湖岸、草津市南部に位置する立命館大学東

側や南側の住宅団地で相対的に高い傾向にある。一方、等価月間支出は、必ずしも等価年間収入と強い正の相関関係にあるわけではない。草津駅や大学の周辺地域では等価収入・支出額ともに高いが、琵琶湖岸沿いの地域は収入に対して支出額は少ない。この背景として、湖岸沿いの地域は相対的に世帯規模が大きく、かつ高齢者も多いが、草津駅前や大学周辺の地域では単身・核家族世帯が中心で、消費も旺盛な世帯主が20歳代から40歳代の世帯が多いことを指摘できる。

次に、野菜・海藻類の消費支出額は、主としてJR東海道線以北で高い傾向にある。特に、草津駅北側から南側にかけての住宅団地が立地する地域で高い。また立命館大学周辺の住宅団地でも高い地区が集中する。果物の消費支出額も、野菜・海藻類と同様の分布傾向にあるが、その地理的分布は新興住宅団地の範囲で特に高い。草津駅周辺や立命館大学周辺の住宅団地では、野菜・海藻類および果物ともに消費支出額が多いが、湖岸沿いの農業的地域では果物への消費支出額が野菜・海藻類の場合よりも低い。調理食品については、他の2品目とは対極的な分布を示す。草津駅前の若い勤労世帯が多い地区や草津市南部の学生が多い住宅街で、調理食品への消費支出額が多い。

以上、全国消費実態調査の合成マイクロデータを用いて等価年間収入および等価月間支出、町丁目別の消費支出額を地図化した。これらは一見して単純な地図ではあるが、1人当たりの消費支出額や等価調整の計算には、消費支出額および収入・支出と世帯人員の変数が同時に必要なため、マイクロデータを用いずには算出できない。このように空間的 MS を用いて小地域単位での地理的な合成マイクロデータを生成することで、商店街やスーパーの商圏分析に不可欠な、小地域単位での消費特性を把握できるようになった。ただし、上記の消費支出額は、焼きなまし法の制約条件に含まれていない変数であるため、制約条件との関連性から推定できることが前提となる。したがって、た



第6図 等価収入・支出および1か月間の1人当たり品目別消費支出額(町丁目)  
Figure 6. Equivalent income and spending, and monthly spending per person by item (cho-cho-moku)  
注) 階級区分はすべて五分位。濃い色ほど金額が大きい。

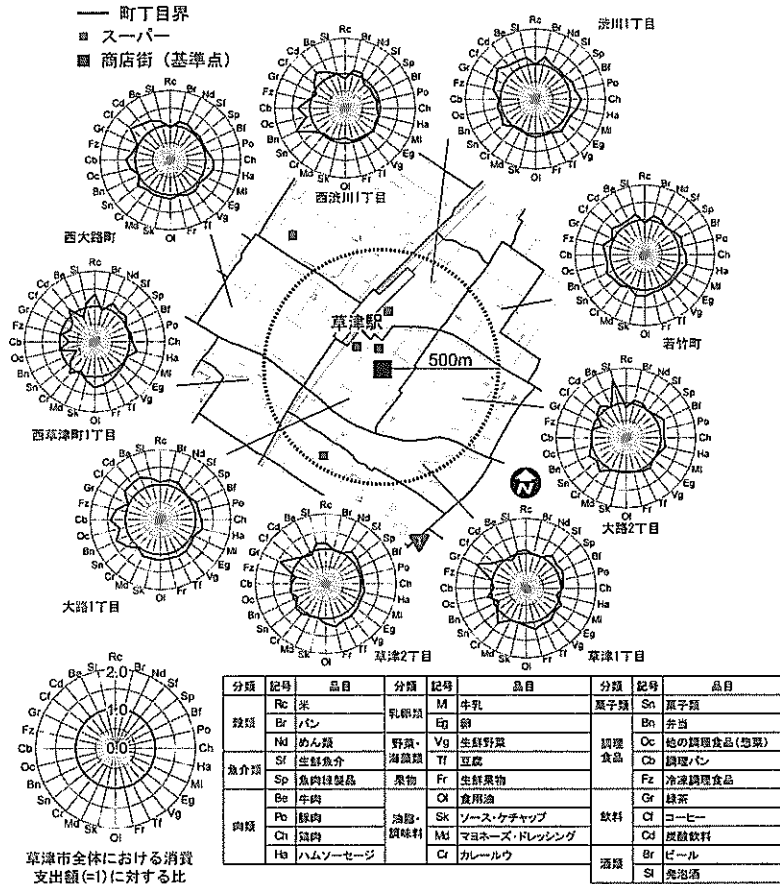
例えば、地域の食文化の影響が大きい品目に関しては、別の資料を用いた補正が必要となる。

(2)草津駅前地区の品目別消費支出額 本節では、商店街や大型小売店舗が多数進出する草津駅前地区を事例に、空間的 MS で推定された合成マイクロデータを用いて、世帯属性および地理的に詳細な消費特性を地図化し、その地理空間分析上の有用性を考察する。

草津駅前には、近鉄百貨店やエルティ932、平和堂、アル・プラザなどの大規模小売店舗が進出する一方で、草津駅周辺の商店街は、市域の人口増加とは相反して、人通りが閑散としており低迷を続けている<sup>51)</sup>。そのため、市街地整備事業の一環

として空き店舗の利活用やインフラ整備などが進められ、「賑わい」の回復が目指されてきた<sup>52)</sup>。こうした都市計画や商店街の活性化策の基礎となる周辺住民の消費特性のうち、外食を含む食料品に対する消費支出を例に、合成マイクロデータを用いた分析を行う。矢野による2000年通行量調査で通行量が最も多い夢大路商店街を基準点とし、徒歩や自転車での買物が見込まれる、基準点から半径500m圏と重複する9つの町丁目を商圏として選定した(第7図)。この9つの町丁目には、約5千世帯が居住する。

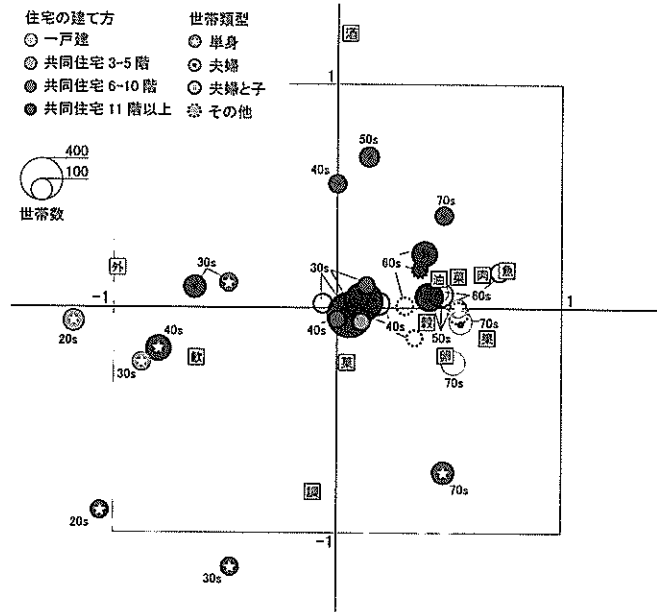
まず、9つの町丁目を対象に、合成マイクロデータを再集計し、生活様式とも強い関連性が想定さ



第7図 草津駅前地区における食料品の消費支出額(町丁目)  
Figure 7. Food spending pattern around Kusatsu Station (cho-cho-moku)

れる世帯主年齢と住宅の建て方、世帯類型の3変数を用いて、消費支出額の品目別構成割合を求め、コレスポネンダンス分析を適用し、商圏内の世帯特性と消費特性との対応関係および世帯数を把握する(第8図)。この図では、各支出品目に近い位置に配置された世帯特性ほど、その品目の構成割合が高いことを意味する。また、支出品目同士または世帯特性同士が近接する場合、それらの性質が類似している。たとえば、魚介類や肉類、穀類は、近接していることから同時に購入される傾向にあ

るが、調理食品や飲料とは離れた位置にあり、一緒に購入されることが相対的に少ないとわかる。草津駅前地区の商圏において、主要な消費者層の<sup>54)</sup>一つは、世帯主年齢が30~40歳代で高層の共同住宅に暮らす夫婦と子から成る世帯(全世帯の約15%を占める)である。こうした世帯の配置を見ると、生鮮食料品に加えて、菓子類も同時に購入する特徴がある。一方、単身世帯(全世帯の約18%を占める)は、調理食品や飲料、外食への支出割合が総じて高い。11階建以上の高層の共同住宅



第8図 世帯特性別にみた消費特性(コレスポネンダンス分析の結果)  
Figure 8. Household characteristics and spending pattern (results of correspondence analysis)

世帯主年齢: 20s: 30歳未満, 30s: 30歳代, 40s: 40歳代, 50s: 50歳代, 60s: 60歳代, 70s: 70歳以上  
 支出品目: 穀: 穀類, 魚: 魚介類, 肉: 肉類, 卵: 卵類, 菜: 野菜海藻類, 果: 果物, 油: 油脂調味料, 菓: 菓子類, 調: 調理食品, 飲: 飲料, 酒: 酒類, 外: 外食  
 注) 世帯主年齢・住宅の建て方・世帯類型別の世帯数が50世帯以上のみを掲載(3427世帯)

に暮らす夫婦世帯(全世帯の約11%を占める)は、比較的、酒類の消費が多い。商店街への主な来街者と考えられる60歳以上の世帯に関しては、70歳代以上の単身世帯で調理食品への支出割合が高いことを除くと、生鮮食料品への支出割合のみが高い傾向にある。

次に、外食を除く食料品に関して、全国消費実態調査の調査品目から購入頻度が高いと考えられる28品目を任意に選択し、1人当たりの消費支出額について草津市全域と町丁目での比を求めグラフ化した(第7図)。これによって、商圏内のどの地区で、どのような品目が消費される傾向にあるかを詳細に把握できる。分析結果をみると、中高層マンションが立地する大路1丁目や渋川1丁目

では、草津市全域と比較して、弁当や総菜、コーヒー、ビール、発泡酒などの支出額が多い。また草津1丁目と2丁目は、生鮮魚介や生鮮野菜、豆腐、生鮮果物への支出が平均よりもやや多く、加えて緑茶への支出額が1.5以上と顕著に高い。西渋川1丁目は、弁当や総菜、調理パン、炭酸飲料の支出額が多い。一方で、西草津1丁目では、米や食用油の支出額が多く、逆に調理食品への支出は平均より少ないことがわかる。以上のように、半径500m程度の商圏内においても、消費特性に多様性が認められた。こうした支出額は、さらに、第8図のように世帯主年齢や住宅の建て方、世帯類型別にも算出が可能であり、地区別ばかりでなく世帯特性とも対応付けた緻密な商圏分析に、空



間的MSの合成マイクロデータを活用できる。

商圏内部の消費特性の地域差の把握は、商店街の中小食料品店による地区を限定した販促・広告活動を効率的に展開する上で、意思決定の重要な判断材料となる。加えて、こうした消費特性が、地域の世帯構成等を踏まえた平均的な消費需要であると解釈するならば、実際に実地調査を行った場合の結果と本推定結果との乖離は、当該地区の小売環境や食生活に地域固有の特性や問題があることを示唆している。

## V まとめ

本研究では、空間的MSを用いて、公的統計の匿名データから地理的に詳細な合成マイクロデータを生成し、消費特性の把握を事例に、匿名データを用いた小地域単位での地理空間分析の実現可能性を検討してきた。本研究で得られた知見は、以下のように整理できる。

第一に、空間的MSを用いることで、全国消費実態調査の匿名データから、国勢調査小地域集計と高い精度で整合する合成マイクロデータを安定して生成できた。小規模なアンケート調査データと比較して、公的統計の匿名データを合成マイクロデータの生成に用いるメリットは、標本バイアスが小さく、標本数が数万件を超え、また国勢調査とも共通する変数が含まれるため、焼きなまし法でより多くの制約条件を考慮できる。そして、計算時間をかければ、それら制約条件にほぼ一致する合成マイクロデータを得られる点にあった。また複数の制約表を考慮できることで、制約として利用していない統計表についても良好な適合度を得た。加えて、公的統計の匿名データは、その申請方法が広く公開されており、多くの研究者らにとって、データにアクセスしやすい点も重要である。

第二に、合成マイクロデータを用いることで、単独のデータだけでは把握できない詳細な消費特性の地域差を明らかにできた。本研究で提示された町丁目単位での品目別消費支出額などの地図は、

一見して単純な地図ではあるが、独自調査を実施する以外には知り得なかった地理情報である。そして草津駅前地区の事例のように、全国消費実態調査の匿名データに表章される多様な変数を用いて合成マイクロデータを再集計することで、世帯属性間や町丁目間で異なる消費特性をも把握できた。こうした事例から、空間的MSを適用することで、匿名データを用いた小地域単位での地理空間分析が実施可能であることが示された。

今後の研究課題として、①他の公的統計の匿名データを用いた合成マイクロデータの生成、②制約条件である国勢調査小地域集計の種類や個票データの標本数、対象地域を変更した合成マイクロデータの適合度比較、③本分析結果の地理学的な応用研究、④小売環境や食文化などの地域性を踏まえた消費実態特性の補正方法の検討が挙げられる。特に③に関連して、本研究で提示した空間的MSの合成マイクロデータを、実際の商圏分析に利用しその有用性を検証する必要がある。さらに、こうしたデータは、食や健康、貧困の地域差の分析<sup>55)</sup>においても、小地域単位で世帯の消費特性や生活様式等を把握できる資料として活用できよう。そうした実践的な研究の積み重ねを通じて、空間的MSによるアプローチの有用性をさらに高めることができると考えられる。

謝辞：全国消費実態調査の匿名データは、神戸大学マイクロデータアーカイブを通じて、独立行政法人統計センターから提供を受けた。(立命館大学・文学部)

## 注

- 1) 村山祐司「GIS—地理情報システムから地理情報科学へ—」(村山祐司・柴崎亮介編「GISの理論」, 2008), 1-16頁。
- 2) <http://www.ccsr.ac.uk/sars/resources/publications/jointpub.html>2012年3月22日閲覧。
- 3) 新統計法から、政府が作成する統計が公的統計と呼ばれるようになった。
- 4) (1)千葉立也・石川義孝・カオリ・リュウ「日本に在住する外国人の国内移動にみられる地域性」(石川義孝編「人口減少と地域—地理学的アプローチ—」京都大学学術出版会, 2007), 197-225頁。(2)石川義孝編「地図でみる日本の外国人」ナカニシヤ出版, 2011。

- 5) 西原純「就業者の収入の地域格差における社会的要因と地域的要因の大きさ—92年就業構造基本調査を用いた分析—」(第24回都市圏研究部会要旨), 人文地理 60-1, 2008, 83-87頁。
- 6) 前掲5) 86頁。
- 7) 公的統計の匿名データが提供されるに至るまでの経緯や関連する法制度、利用手続きについては、以下の著書が詳しい。山口幸三「現代日本の世帯構造と就業形態の変動解析—公的統計のミクロ統計活用序説—」, 日本統計協会, 2011。
- 8) 統計データの二次利用促進に関する研究会「統計データの二次利用促進に関する研究会 報告書」, 2008, <http://www.stat.go.jp/info/kenkyu/2jiryoyu/index-2.htm>2012年2月21日閲覧。
- 9) 星野伸明「公的統計マイクロデータ提供制度の課題」, 日本統計学会誌40-1, 2010, 23-45頁。
- 10) 前掲9)。
- 11) 統計学での「構造データ」と区別するため、著者が既往研究で用いてきた「合成マイクロデータ」を用いる。なお、合成マイクロデータの生成においては、匿名データと他の統計資料との整合性を集計レベルで追求するため、その作業過程や結果を通じて個人・世帯が特定されることはない。
- 12) (1) Hanaoka, K. 'Estimating spatial distributions of earnings at the small area level in Japan: a spatial microsimulation approach', *Journal of City Planning Institute of Japan*, 46, 2011, pp. 142-148. (2) 花岡和聖「焼きなまし法を用いたパーソントリップ調査データの拡大補正法に関する研究—平成12年度京阪神都市圏パーソントリップ調査データを用いて—」, 都市計画論文集41-3, 2006, 91-96頁。(3) Hanaoka, K. and Clarke, G. 'Spatial microsimulation modeling for retail market analysis at the small-area level', *Computers, Environment and Urban Systems*, 31, 2007, pp. 162-187.
- 13) 前掲12(1)と(2)では京阪神都市圏パーソントリップ調査を、前掲12(3)では立命館大学地理学教室が独自に実施したアンケート調査を使用した。
- 14) 中谷友樹「健康と場所—近隣環境と健康格差研究—」, 人文地理63-4, 2011, 360-377頁。
- 15) 前掲12(3)。
- 16) 草津市の小売環境の変容に関しては、矢野桂司編「地域情報システムを用いた地域計画立案支援システム」科学研究費補助金研究成果報告書, 2001に詳しい。
- 17) Orcutt, G. H. 'A new type of socio-economic system', *The Review of Economics and Statistics*, 29, 1957, pp. 116-123.
- 18) 経済学や人口学におけるマイクロシミュレーションのレビューは以下の論文が詳しい。Merz, J. 'Microsimulation: a survey of principles developments and applications', *International Journal of Forecasting*, 7, 1991, pp. 77-104.
- 19) 藤井聡・菊池輝・北村隆一「マイクロシミュレーションによるCO2排出削減に向けた交通施策の検討—京都市の事例—」, 交通工学35-4, 2009, 11-18頁。
- 20) Clarke, G. P. 'Microsimulation: an introduction' (Clarke, G. P., ed. *Microsimulation for urban and regional policy analysis*, Pion, 1996), pp. 1-9.
- 21) 中谷友樹・花岡和聖「ジオシミュレーションと空間的マイクロシミュレーション」(村山祐司・柴崎亮介編「GISの理論」, 2008), 142-160頁。
- 22) (1)前掲12) (1)。(2) Ballas, D., Clarke, G. P., Dorling, D., and Rossiter, D. 'Using SimBritain to model the geographical impact of national government policies', *Geographical Analysis*, 39, 2007, pp. 44-77.
- 23) Kongmuang, C., Clarke, G. P., Evans, A. J., and Jin, J. 'SimCrime: a spatial microsimulation model for the analyzing of crime in Leeds', Working Paper, 06-1, School of Geography, University of Leeds, Leeds, LS29JT, UK, 2006.
- 24) (1) Williamson, P. 'Community care policies for the elderly 1981 and 1991: a microsimulation approach' (Clarke, G. P., ed. *Microsimulation for urban and regional policy analysis*, Pion, 1996), pp. 64-87. (2) Mitchell, R., Dorling, D., and Shaw, M. 'Population production and modelling mortality: an application of geographic information systems in health inequalities research', *Health&Place*, 8, 2002, pp. 15-24. (3) Smith, D. M., Pearce, J. R., and Harland, K. 'Can a deterministic spatial microsimulation model provide reliable small-area estimates of health behaviours? an example of smoking prevalence in New Zealand', *Health&Place*, 17, 2011, pp. 618-624. (4) Tomintz, M. N., Clarke, G. P., and Eigy, J. E. 'The geography of smoking in Leeds: estimating individual smoking rates and the implications for the location of stop smoking services', *Area*, 40-3, 2008, pp. 341-353.
- 25) (1) Ballas, D., and Clarke, G. P. 'GIS and microsimulation for local labour market analysis', *Computers, Environment and Urban Systems*, 24, 2000, pp. 305-330. (2) Ballas, D., and Clarke, G. P. 'Modelling the local impacts of national social policies: a spatial microsimulation approach', *Environment and Planning C: Government and Policy*, 19, 2001, pp. 587-606.
- 26) 前掲12) (3)。
- 27) Nakaya, T., Fotheringham, S., Hanaoka, K., Clarke, G., Ballas, D., and Yano, K. 'Combining microsimulation and spatial interaction models for retail location analysis', *Journal of Geographical Systems*, 9-4, 2007, pp. 345-369.
- 28) Leeuwen, E. S. 'The effects of future retail developments on the local economy: combining micro and macro approaches', *Papers in Regional Science*, 89-4, 2010, pp. 691-710.
- 29) Leeuen, E. S., and Rietveld, P. 'Spatial consumer behavior in small and medium-sized towns', *Regional Studies*, 45-8, pp. 1107-1119.
- 30) イギリスでは、個人情報開示リスクが見直され、2001年版SARsから地域区分が表章されなくなった。代わりに、日本の市区町村程度に相当するLocal authorityが表章されたSmall Area Microdataが作成される。



- 31) Williamson, P., 'Synthetic microdata' (Rees, P., Martin, D., and Williamson, P., eds., *The census data system*, Wiley, 2002), pp. 231-241.
- 32) Birkin, M., and Clarke, M., 'SYNTHESIS: a synthetic spatial information system for urban and regional analysis: methods and examples', *Environment and Planning A*, 20-12, 1988, pp. 1645-1671.
- 33) Huang, Z., and Williamson, P., 'A comparison of synthetic reconstruction and combinatorial optimization approaches to the creation of small-area microdata', 2001, Working Paper2001/2, Department of Geography, University of Liverpool.
- 34) Williamson, P., Birkin, M., and Rees, P., 'The estimation of population microdata using data from small area statistics and samples of anonymised records', *Environment and Planning A*, 30, 1998, pp. 785-816.
- 35) 前掲34)。
- 36) 前掲12) (2)。
- 37) なお住宅の所有関係は「民営の借家」、住宅の建て方は共同住宅を想定し草津市内の共同住宅の階数別割合を基に、いずれかの建て方を単身学生の個票データに無作為に割り当てた。
- 38) Voas, D., and Williamson, P., 'Evaluating goodness-of-fit measures for synthetic microdata', *Geographical & Environmental Modelling*, 5-2, 2001, pp. 177-200.
- 39) 前掲38)。
- 40) (1)前掲33), (2)前掲38), (3) Ryan, J., Maoh, H., and Kanaroglou, P., 'Population synthesis: comparing the major techniques using a small, complete population of firms', Working Paper Series, Centre for Spatial Analysis, McMaster University, 2007.
- 41) Williamson, P., 'CO instruction manual', Working Paper2007/1, Department of Geography, University of Liverpool, 2007.
- 42) ウィリアムソン (Paul, Williamson) の個人ページ <http://pcwww.liv.ac.uk/~william/> (2012年2月20日閲覧) からソフトウェアをダウンロードできる。焼きなまし法の計算時間は Windows PC (CPU: Core-i7 3.4GHz) で1回当たり約3分かかる。
- 43) 前掲12) (2)。
- 44) 前掲38)。
- 45) 草津市南部の笠山7丁目及び8丁目には滋賀医科大学及び医療施設が立地し、「医療、福祉」に従事する就業者割合が8割を超え、住宅の建て方も寮・社宅と考えられる「共同住宅3-5階」でほぼ構成される。
- 46) Voas, D., and Williamson, P., 'An evaluation of the combinatorial optimization approach to the creation of synthetic microdata', *International Journal of Population Geography*, 6, 2000, pp. 349-366.
- 47) 前掲12) (3)。
- 48) 性別や産業別、世帯人員別の人口・世帯割合に関して、国勢調査と全国消費実態調査の匿名データを比較すると、ほとんどの項目で1-2ポイント未満の差しか認められない。
- 49) (1)前掲40) (3), (2)前掲12) (2), (3)前掲33)。
- 50) 等価年間所得および等価月間支出は、年間所得や月間支出をそれぞれ世帯人員の平方根で除した値。全国消費実態調査の年間収入は、公的年金等を含んだ課税前の収入である。月間支出は、税金等の非消費支出を除く支出の総額である。
- 51) 読売新聞滋賀県版「沈む商店街 人口増 県内2位なのに」(2012年2月9日付), [http://www.yomiuri.co.jp/e-japan/shiga/feature/oh1328712553039\\_02/news/20120208-0YT8T01208.htm](http://www.yomiuri.co.jp/e-japan/shiga/feature/oh1328712553039_02/news/20120208-0YT8T01208.htm) 2012年2月21日閲覧。
- 52) (1)草津市「草津市都市計画マスタープラン」, <http://www.city.kusatsu.shiga.jp/www/contents/1222306415602/index.html> 2012年2月21日閲覧。(2)草津市「データブック2008」, <http://www.city.kusatsu.shiga.jp/www/contents/1222859564422/index.html> 2012年2月21日閲覧。
- 53) 矢野桂司「通行量GIS」(矢野桂司編『地域情報システムを用いた地域計画立案支援システム』科学研究費補助金研究成果報告書, 2001), 101-107頁。
- 54) 世帯主年齢・住宅の建て方・世帯類型別の世帯数が50世帯(全体の約1%)以上であった層を主要な消費者層とした。世帯総数は3427世帯である。
- 55) 前掲14)。

### The Possibility of Using Anonymized Data from Japanese Official Statistics for Implementing Geospatial Analysis at a Small Area Level: Creating a Geographically Disaggregated Synthetic Microdata Set by Using a Spatial Microsimulation

HANAOKA Kazumasa

Department of Geography, Ritsumeikan University

In recent times, with growing attention toward data analysis using a microdata set, anonymized data of official statistics were published in Japan; however, these data have limited geographical details so they cannot be used for geospatial analysis at a small area level. Therefore, in order to overcome such problems, we attempted to apply a spatial microsimulation technique to

create geographically disaggregated synthetic microdata from anonymized data from official statistics of the 2004 National Survey of Family Income and Expenditure. Our study area was Kusatsu City, Japan.

As a method for creating a synthetic microdata set, we used a simulated annealing, which is a combinatorial optimization technique. For each *cho-cho-moku* (administrative region for the census), by using a simulated annealing, the samples were repeatedly swapped between anonymized data and synthetic microdata in a manner that they agreed with the five constraint tables created from the small area statistics of the 2005 Japanese Population Census. The output of the synthetic microdata was evaluated based on the Overall Total Absolute Error and the Overall Relative Sum of Squared Z-scores. The results showed that the synthetic microdata almost perfectly matched the constraint census tables, except for a few districts in the southern part of the study area, where skewed population distributions are found. We assumed that the reason for a high rate of accuracy was that the anonymized data contained 1.8 thousand samples of households and the sampling biases were relatively small.

Another merit of using anonymized data from public surveys is that it has a detailed and wide range of variables on demographic and socioeconomic attributes as well as on spending amounts. Therefore, the synthetic microdata were retabulated as per *cho-cho-moku* to visualize the spatial patterns of monthly spending per person by items. In addition, we presented a map of spending patterns for 28 items revealing geographical diversity in consumption. These detailed maps will be used for supporting further analysis relating to food deserts, health, and income inequalities. In conclusion, the use of anonymized data and its geographical disaggregation using a spatial microsimulation allow us to implement geospatial analysis at a small area level using a microdata set.

**Key words:** Anonymized data, National Survey of Family Income and Expenditure, spatial microsimulation, simulated annealing, synthetic microdata, Kusatsu City

報告者：岡田 直（横浜都市発展記念館調査研究員）

報告趣旨：今回の研究報告は、研究機関として、また社会教育機関としての役割をもつ博物館の展示・研究・教育において、歴史的地図資料がどのようにかわるのか、また歴史地理学と博物館のかかわりについて議論を行うものである。関東大震災以降の横浜市の都市形成を主軸とし、近代地図資料を中心とした展示構成をとる横浜都市発展記念館において、その展示の現地説明と、博物館における歴史地理学のかかわり方、またその研究成果の社会的発信のありかたについて研究報告を行う。

連絡先：天野太郎（同志社女子大学） E-mail: amano@ dwc. doshisha. ac. jp

（第44回都市圏研究部会と第128回歴史地理研究部会の会告掲載が遅延したことをお詫びします）

### … 第110回 地理思想研究部会 …

日時：2012年7月28日（土）13：30～17：30

会場：大阪府立大学・中之島サテライト 2階講義室

<http://www.osakafu-u.ac.jp/contribution/lifelong/extension/place/satelite.html>

地下鉄御堂筋線「淀屋橋」駅・京阪電車「淀屋橋」駅下車徒歩約5分

地下鉄堺筋線「北浜」駅下車徒歩約8分

※入り口は中央公会堂側にあります。中之島図書館からは直接行きませんのでご注意ください。

テーマ：ディープサウスからの都市空間論

研究発表：

移動と逃亡の地理学に向けて……………原口 剛（大阪府立大学）

遊歩・逃亡・地図—「通天閣 新・日本資本主義発達史」をめぐって……酒井隆史（大阪府立大学）

連絡先：福田珠己（大阪府立大学） E-mail: tamami@hs.osakafu-u.ac.jp

### … 第129回 歴史地理研究部会 …

日時：2012年9月8日（土）14：00～17：00

会場：神戸市立博物館 地階研修室2（神戸市中央区京町24）

会場へのアクセスは、以下をご参照ください。

<http://www.city.kobe.lg.jp/culture/culture/institution/museum/info/access.html>

テーマ：明治の地籍図の新しい研究視点—近代から現代を考える課題と意義—

報告者：土地台帳と地籍図を用いた近代における災害被災地域の復原……赤石直美（立命館大学・非）  
明治初期大縮尺地図としての地籍図の現代的意味……………藤田裕嗣（神戸大学）

報告趣旨：歴史地理学で盛んに研究利用されてきた明治の地籍図は、近年では様々な学問から関心が寄せられ、学際的な地図資料となってきた。従来は景観変遷の復原に用いられる場面が多かったが、近代以降の変化の分析でも資料の有効性が認められ研究例が増えている。しかし、精緻な復原を行うには方法的課題や資料の性格に対する批判など問題点も多い。また、現代社会とどう結ぶかなど新しい課題も問われている。2つの研究報告を通して、近代以降の研究利用における諸課題や現代的意味について議論を行う。

コメント：水野章二（滋賀県立大学）

連絡先：古関大樹（京都女子大学・非）

inukuro9696@hotmail.com

### … 第3回 政治地理研究部会 第111回 地理思想研究部会 合同研究部会 …

日時：2012年11月3日（土）13：00～17：00

会場：公立学校共済組合奈良宿泊所 春日野荘

（奈良市法蓮町757-2 Tel 0742-22-6021）

近鉄奈良駅またはJR奈良駅より、奈良交通バス西大寺駅行または航空自衛隊行きにて佐保小学校前下車すぐ（下記ウェブサイト参照）。

<http://www.kasugano-so.co.jp/access/>

テーマ：地中海境界都市の政治・文化地理

趣旨：オランダで活躍中の政治・文化地理学者お二人をお招きし、地中海ヨーロッパの周縁に位置するモロッコ・タンジールとイタリア・トリエステの二都市を取り上げ、変動するヨーロッパ空間の「境界」におかれた都市の歴史的・政治的・文化的変容を多面的に議論します（部分通訳付き）。

発表：New Euro-Mediterranean Geographies: Re-mapping Tangiers（地中海ヨーロッパの新しい地理：タンジールの再地図化）……………Luiza Bialasiewicz（University of Amsterdam）  
Trieste: Cultural/Political Border Geographies (of Absence)（トリエステ：文化的・政治的な（不在の）境界地理）……………Claudio Minca（Wageningen University）

連絡先：山崎孝史（大阪府立大学） E-mail: yamataka@lit.osaka-cu.ac.jp

※発表論文は以下のURLからダウンロードいただけます。

Luiza Bialasiewicz <http://www.lit.osaka-cu.ac.jp/user/yamataka/Bialasiewicz20xx.pdf>

Claudio Minca <http://www.lit.osaka-cu.ac.jp/user/yamataka/Minca2009.pdf>

### 初等中等教育におけるGISを活用した授業に係る優良事例表彰の募集（お知らせ）

GISを実践的に活用した授業に取り組んでいる教員等の個人又はグループを対象に、標記の表彰事業が実施されます（募集期間：平成24年7月2日（月）～8月27日（月）必着）。詳細は、[http://www.gisa-japan.org/news/detail\\_1006.html](http://www.gisa-japan.org/news/detail_1006.html)をご覧ください。

### 編集委員

出田 和久（編集理事）	池谷 和信	稲垣 稷	小方 登	貝柄 徹
河原 典史	河本 大地	鳥津 俊之	祖田 亮次	橋 セツ
中谷 友樹	三木 理史	水野 真彦	山崎 健	山近 博義
[英文校閲] Sidney J. Atkins		和田 真理子		

人文地理 第64巻第3号（通巻第375号）  
2012年6月28日発行

編集発行所 ◎ 人文地理学会

〒606-8305 京都市左京区吉田河原町14番地

近畿地方発明センター合同ビル内

電話 075-751-7687

FAX 075-708-5515

振替 01000-6-6611

会員年会費 9,500円

印刷所

〒615-0052 京都市右京区西院清水町156-1

共同印刷工業株式会社

発売所

〒101-0062 東京都千代田区神田駿河台2-10

株式会社古今書院

電話 03-3291-2757

FAX 03-3233-0303

振替 00100-8-35340

定価 2,040円（本体 1,943円）

URL: <http://hgsl.org/>

本誌は直接出版費の一部として、日本学術振興会科学研究費補助金「研究成果公開促進費」の交付を受けております。

本会への会費納入、「入会申込書」の用紙請求、入会申込みは上記の人文地理学会へお願いします。