



共同利用・共同研究拠点 Joint Usage/Research Center
東京大学 The University of Tokyo
空間情報科学研究センター
Center for Spatial Information Science

Discussion Paper Series

No. **176**
CSIS Discussion Paper

2022 年 9 月

地域評価に関する基礎的検討—主観的地域満足度を例にして—

西 颯人

（一橋大学 ソーシャル・データサイエンス教育研究推進センター、東京大学 空間
情報科学研究センター）

地域評価に関する基礎的検討

—主観的地域満足度を例にして—*

西 颯人¹

¹ 一橋大学 ソーシャル・データサイエンス教育研究推進センター

2022 年 9 月 22 日

概要

日本の国土は地域によって様々な特色や差異があり、それらの魅力などを評価・比較して順位付けしようという試みは様々な調査団体によって行われている。このような調査結果はしばしばランキングという形で公表され話題となるが、これらのランキングが何をどのように評価しており、どの程度信頼できるのか、また何に活用できるのか、について、十分な議論がなされているとは言えない。そこで本稿では、地域別評価のモデルを通して評価と集計に関する基礎的な検討を行い、地域別評価においては評価主体、つまり「誰にとっての」地域評価なのかが、集計方法においても集計結果の解釈においても重要な要素であることを示す。さらに公表されているランキングに実際に使用されている個票データを用いて、評価主体の個人属性が集計結果に影響する可能性があることを確認した。

1 はじめに

日本の都道府県や市区町村の魅力などを評価し、順位付けしようという試みは従来より行われ、ランキングとして公開された結果は少なからず世間で話題となっている。清水 (2022) ではこうしたランキング指標について、その背景となる課題や理論について幅広く議論を行っている。

現在公表されている地域別のランキングには、例えば以下のようなものが挙げられる。

- 住みたい街ランキング (SUUMO, リクルート)
- 街の住みこち&住みたい街ランキング (いい部屋ネット, 大東建託)
- 地域ブランド調査 (都道府県魅力度ランキング、他) (ブランド総合研究所)

その他、ランキング形式にはなっていないが、「徒歩圏内の施設利用充実度」を評価した Walkability Index (日建設計) も公開され (清水他, 2020)、国土交通省の 3D 都市モデル PLATEAU に提供されている。しかしこれらの公表結果を見ると、ランキングによってその順位が異なっていることがわかる。

そこで公表物の算出方法を見ると、最後に示した Walkability Index を除き、多くのランキングがアンケート調査を行い、それを地域別に集計することで算出されている。アンケート調査は通常は標本調査であるため、調査ごとに統計的な誤差を含むことになるが、それに加えてこれらの調査ではアンケート項目や集計方法が異なるため、調査しようとしている内容自体にも違いがあること

* 執筆にあたり、清水千弘氏、宗健氏には多くの貴重なコメントを頂いた。ここに御礼申し上げます。

になる。

よって各ランキングを見る際には、そのランキングが地域のどんな評価を行おうとしているのかを十分に把握しておく必要があるが、実際に世間でランキング結果が取りざたされる際にはランキング名と順位だけが注目されることもある。もしこれが自治体等の地域政策に影響を与えることがあるとするならば、各ランキングについて「何を評価しようとしているのか」「評価しようとしているものを適切に反映した評価になっているのか」を注意深く検討する必要があると言えるだろう。

そこで本稿では、こうした地域評価について、実際の評価計算を念頭に置いた基礎的な検討を行う。第2章では、評価主体による地域評価のモデルを通して、地域評価をどのように集計できるのかを考察する。第3章では、使用するデータについて説明する。データの分析結果は第4章に整理し、第5章では全体のまとめと今後の研究課題を示す。

2 地域評価のモデル

2.1 個人別モデル

評価主体 i による地域評価（主観的地域満足度など） y_i のモデルを考える。地域 j に住んでいる住民の集合を R_j として、住民の個人属性を α_i 、地域アメニティなどの地域属性を $\mathbf{x}_j$ とすると、Rosen (1974) などで想定される効用関数を参考にして、

$$y_i = u(\mathbf{x}_j; \alpha_i) + \varepsilon_i \quad (1)$$

と定式化できる。ただし $i \in R_j$ であり、 $u(\mathbf{x}_j; \alpha_i)$ は評価関数、 ε_i は誤差項である。この定式化の意味するところは、評価主体による地域評価はその地域で享受できるアメニティ $\mathbf{x}_j$ によって決まるが、それが評価値にどう影響するかを決める関数 u の関数形は、評価主体の個人属性によってそれぞれ決まってくるということである。

このモデルを線形近似すると、以下のようになる。

$$y_i \approx \bar{u}_{i,j} + \frac{\partial u(\mathbf{x}_j; \alpha_i)}{\partial \mathbf{x}}^\top \mathbf{x}_j + \frac{\partial u(\mathbf{x}_j; \alpha_i)}{\partial \alpha}^\top \alpha_i \quad (2)$$

$\bar{u}_{i,j}$ は地域 j に対する評価主体 i の基準満足度である。左辺第2項は地域のアメニティ $\mathbf{x}_j$ の量と質によって決まる部分であり、第3項は評価主体の属性 α によって決まる部分であると解釈できる。

地域別ランキングは多くの場合、地域ごとに評価値を集計してそれを順位付けしたものである。これは、本来評価主体 i ごとに決まっている満足度を、何らかの形で地域 j ごとに割り当て直していることになる。このためには、何らかの評価主体の集合（例えばその地域に住んでいる住民の集合 R_j など）について、個人属性の分布 $p(\alpha)$ を考え、 $p(\alpha)$ に関する評価値 y の期待値を推計することになる。実際のランキング作成では、アンケート調査を行って評価値が集計されている。

このようにして作られたランキングだけを見ると、上位にある地域はあたかも「優れた」地域特性 $\mathbf{x}_j$ を持っているような印象を受けるが、実際には (2) から分かるように、地域特性 $\mathbf{x}_j$ の評価値である係数 $\frac{\partial u(\mathbf{x}_j; \alpha_i)}{\partial \mathbf{x}}$ は個人属性 α_i に本来依存する。すなわち、ある地域 j の属性 $\mathbf{x}_j$ がある住民にとって優れていても、別の評価主体にとっては優れていない可能性があるわけである。

このように、地域評価においては「どこの」地域を評価対象とするかだけではなく、「誰にとっての」評価なのか、という点も重要であることがわかる。これは、例えばその地域の住民と、他の地域の住民と、その地域を訪れる観光客とを考えた場合に、それぞれの個人属性 α の分布 $p(\alpha)$ が異なると想定されるために、集計した結果も異なるためである。

地域 j の評価を考える際に具体的に想定できる評価主体の集合としては、例えば以下の2通りが考えられる。

- いずれかの地域に居住するすべての住民の集合 $\{\dots, R_j, \dots\}$
- 地域 j に居住する住民の集合 R_j

他にも、地域 j に隣接した地域に住む住民の集合や、観光客の集合などが考えられるが、ここでは単純のために上の2つを考える。これらの場合、評価値の集計結果は何を意味することになるだろうか。

まずすべての住民の集合 $\{\dots, R_j, \dots\}$ を想定した場合、例えば日本全国を対象とした調査であれば、(外国籍居住者を含む)国民全体が想定される住民の集合となる。この場合、想定される個人属性の分布 $p(\alpha)$ には地域差がないので、集計値が大きくなるためには多くの国民に評価される(つまり、万人受けする)地域アメニティが多く供給されていればよい。裏を返せば、このような想定に基づく評価が高い地域では、万人受けするアメニティが充実していると予想できる。

一方で、その地域に居住する住民の集合 R_j を考えた場合、住民集合 R_j が高く評価するアメニティが充実していれば、万人受けするアメニティが不十分であっても集計評価値を大きくすることができる。このことから、このような想定に基づく評価が高い地域では、地域が供給するアメニティ x_j と住民属性とのマッチングがうまくいっていると予想できる。住民の居住満足度を評価したい場合には、地域外の住民が高く評価するアメニティが充実している必要性は低いので、こちらの評価主体集合の想定がより適切であると思われる。ただし、こちらの想定では評価主体の分布 $p(\alpha)$ が地域によって異なるために、どこに住んでも充実した生活ができる人が集まっているのか、その地域に住んだことで充実した生活を送ることができているのかを識別するのが難しいという問題が考えられる。

このように、地域別ランキングではその調査対象によって、結果として得られた集計値が持つ意味合いが異なる。各ランキングの題名には「住みたい街」「住みこち」「住み続けたい街」といった表現が付けられているが、こうした表現は背景にあるアンケート調査の対象とそれによる違いを表現しようとしたものであると思われる。一方で、評価主体の違いも統制したランキングや評価指標を求めようとするならば、清水 (2022) が議論するような種々の条件を考えなくてはならない。

2.2 地域別モデル

ここまでの議論により、地域評価においては評価主体である住民の個人属性が重要であることが分かった。このような個人属性を考慮したさらに精緻な地域評価を考えるための基礎的な検討として、以下では地域住民の主観的な地域満足度を評価対象として (2) に追加の仮定を導入することで、さらに単純化した地域別のモデルを考えることにする。

まず住民集合 R_j に属する評価主体集合を考えるとすると、評価する地域 j が十分に小さければ、評価主体が享受する地域属性(アメニティ)は一様に x_j であると見なすことができる。また、住民は自由に居住地選択を行っているとする、地域属性 x_j を高く評価するような個人属性 α を持つ評価主体が地域 j を居住地として選択していることになる。そこで、 R_j に属する住民はおおむ

ね一様な地域評価関数 $u(\mathbf{x}_j; \boldsymbol{\alpha}_i) \approx u_j(\mathbf{x}_j)$ を持っていると仮定する。

このとき、(2) 左辺の第1項第2項は地域 j 内で一様であり、第3項の係数も一様であるから、以下の近似モデルが得られる。

$$\begin{aligned} y_i &\approx c_j + \beta_j^\top \boldsymbol{\alpha}_i \\ \beta_j &\approx \frac{\partial u(\mathbf{x}_j; \boldsymbol{\alpha}_i)}{\partial \boldsymbol{\alpha}} \\ c_j &\approx \bar{u}_{i,j} + \frac{\partial u(\mathbf{x}_j; \boldsymbol{\alpha}_i)}{\partial \mathbf{x}}^\top \mathbf{x}_j \end{aligned} \quad (3)$$

c_j は（地域 j を居住地として選ぶような評価主体による）地域 j の基準評価値で、 β_j は地域 j 内における個人属性 $\boldsymbol{\alpha}_i$ による居住満足度の伸びである。居住満足度の伸びは地域属性 $\mathbf{x}$ の影響を受けるので、 β_j の値に大きな地域差があれば、そこには個人属性に加えて各地域が持つ地域属性の違いも反映されていると予想できる。この近似モデル (3) を地域別に推定することで、個人属性が地域満足度に与える影響を検討する。

3 データと推定方法

第1章に挙げた地域別ランキングに関して、その最新の調査規模を比較すると表1のようになる。なお、いずれの調査も Web アンケートによって行われている。これを見ると、いい部屋ネットによる「街の住みこち&住みたい街ランキング」がサンプルサイズの観点では最大規模となっていることがわかる。そこで本稿では、いい部屋ネットによる調査データを用いて個人属性が地域満足度に与える影響を分析する。

表1 アンケート調査規模

ランキング名	調査団体	調査人数（全国）	調査人数（首都圏）
住みたい街ランキング	SUUMO, リクルート	—	10,000 ¹⁾
街の住みこち&住みたい街ランキング	いい部屋ネット, 大東建託	646,245 ²⁾	207,659 ³⁾
地域ブランド調査	ブランド総合研究所	35,498 ⁴⁾	—

使用するデータは大東建託賃貸未来研究所から提供を受けた「街の住みこち&住みたい街ランキング」に使用されている個票データである。提供を受けたデータのサンプルサイズは 646,502 件であった。

「街の住みこち&住みたい街ランキング」の調査では、「全体として今住んでいる地域について」という設問に対して「大変満足している」「満足している」「どちらでもない」「不満である」「大変不満である」の5段階で回答する項目があり、その地域別平均値が「住みこちランキング」として発表されている。そこでこの項目がアンケート回答者の「主観的地域満足度」を反映していると考え、第2章で示した式 (3) の目的変数とした。

1) 住みたい街ランキング 2022 首都圏 https://bessou.suumo.jp/edit/sumi_machi/ (2022/09/12 閲覧)

2) 街の住みこちランキング 2022 全国版 https://www.eheya.net/sumicoco/ranking/zenkoku/sumicoco_area_pref.html (2022/09/12 閲覧)

3) 街の住みこちランキング 2022 首都圏版 https://www.eheya.net/sumicoco/ranking/shutoken/sumicoco_area.html (2022/09/12 閲覧)

4) 地域ブランド調査 2021 <https://news.tiiki.jp/articles/4693> (2022/09/12 閲覧) ただしこの調査のみ「有効回収数」となっているため、無回答を除外したサンプルサイズが報告されている可能性がある。

表2 「全体として今住んでいる地域について」への回答

回答	割り当て得点	回答数
大変満足している	100	66,153
満足している	75	324,016
どちらでもない	50	187,322
不満である	25	51,418
大変不満である	0	17,593
計		646,502

上記満足度をデータ全体で集計した結果が表2である。「街の住みこち&住みたい街ランキング」ではこの5段階の回答項目を集計上は比例尺度として100点満点の数値を割り当てている。この数値も表中に記載した。集計結果を見ると、「満足している」との回答が最も多く、「大変不満である」との回答が最も少なかった。「大変満足している」と「不満である」を比較しても若干前者の方が多くなっており、全体として満足側に偏った分布となっている。この結果は、その地域の属性を高く評価するような評価主体がその地域を居住地として選択しているという仮定と矛盾しない。

表3では、説明変数として使用した個人属性に関するアンケート回答項目を示した。回答段階が2の項目はダミー変数となっている項目であり、回答段階が5の項目は「そう思う」「どちらかといえばそう思う」「どちらでもない」「どちらかといえばそう思わない」「そうは思わない」の5段階で回答する項目である。5段階の項目に関しては分析の都合上満足度と同様に比例尺度としてみなして、最大が1、最小が0となるように線形のスケーリングを行った。

いい部屋ネットの調査は比較的大規模なものであるが、自治体別や駅周辺別など地域ごとにデータを分割した場合には30件未満の小さなサンプルサイズになってしまう地域も多く、過剰適合への対処が求められる。そこで、Ridge回帰を用いて過剰適合を抑制するとともに、一個抜き交差検証(LOOCV: Leave-One-Out Cross Validation)を行ってRidge回帰の正則化パラメータを選択した(Arlot and Celisse, 2010)。また被説明変数の値域を開区間 $(-\infty, \infty)$ とするために、「大変満足している」「大変不満である」の割り当て得点を90,10に改めたうえで関数 $f(y) = 100 \tanh^{-1}(\frac{y}{50} - 1)$ で変換することによって、0が「どちらでもない」に対応するように変換を行い、被説明変数とした。予測時は逆関数である $f^{-1}(y) = 50(\tanh(\frac{y}{100}) + 1)$ を用いて逆変換することで、予測値を得る。

Ridge回帰ではLOOCVを効率的に計算できる(Hastie, 2020)が、この方法を用いると変数変換後の値に対する予測値に関して、交差検証を行うことになる。そのため、適切な正則化パラメータの値の範囲を調べるために、以下のように2重にLOOCVを行って、正則化パラメータと予測精度の関係を調べた。まずサンプルから1件データを取り除き、残りのデータを用いて(効率的な)LOOCVを行って正則化パラメータを決定する。その後、取り除いたデータに関する予測を行い、予測精度を確認する(このステップでは変数変換前の値に関して予測と観測の二乗誤差を計算した)。この「1件取り除く」プロセスをLOOCVと同様にすべてのデータに関して繰り返し、平均平方誤差(MSE: Mean Squared Error)を計算した。この結果、正則化パラメータが小さく選択された場合に予測精度が大きく低下することが分かったので、正則化パラメータの下限値を10として推定を行った。Ridge回帰の推定にはscikit-learn version 1.0.2⁵⁾を用いた。

5) <https://scikit-learn.org/stable/index.html>

表 3 個人属性項目

回答項目	回答段階
女性ダミー	2
子ども有りダミー	2
大卒ダミー	2
マンションダミー	2
戸建てダミー	2
タワマンダミー	2
地元ダミー	2
大企業勤務ダミー	2
良く飲みに行く	5
自分の才能には自信がある	5
未来は明るい	5
家族関係は良好だ	5
リスクを取らないほうだ	5
子育て支援や図書館の充実など行政サービスの良い街に住みたい	5
自然環境豊かな街に住みたい	5
近所付き合いが煩わしくない街に住みたい	5
老後の生活資金が足りるか不安である	5
コロナをきっかけにして今住んでいる場所が良いと思うようになった。	5
コロナをきっかけにして今住んでいる場所が良くないと思うようになった。	5
コロナで収入が減った	2
仕事には大変満足している	2
社会的地位には大変満足している	2
クルマ系／自分名義のクルマを持っている	2
クルマ系／同居家族がクルマを持っている（一人暮らしの場合はいいえ）	2
ペット系／犬を飼っている	2
ペット系／猫を飼っている	2
ペット系／犬・猫以外のペットを飼っている	2
日常使う交通手段／日常的にクルマを交通手段に使っている	2
日常使う交通手段／日常的に鉄道（地下鉄・モノレール・路面電車を含む）を交通手段に使っている	2

4 結果

2重の LOOCV が可能となる 33 件以上のサンプルサイズが確保できた全国の 1,854 自治体に関して回帰分析を行った。2重の LOOCV によって得られた SSE(Sum of Squared Errors: 誤差二乗和) を決定係数 (R^2) に換算⁶⁾して、サンプルサイズとの関係を示したのが図 1 である。これを見ると、サンプルサイズが少ない場合は精度がばらつく傾向にあり、推定が不安定になっていることがわかる。そこで、安定した推定のためには少なくともサンプルサイズが 200 件程度は必要であると判断し、以降はサンプルサイズ 200 以上となる 192 自治体についての分析結果を示すこととする。200 件未満の自治体を含めた結果は Appendix に示した。

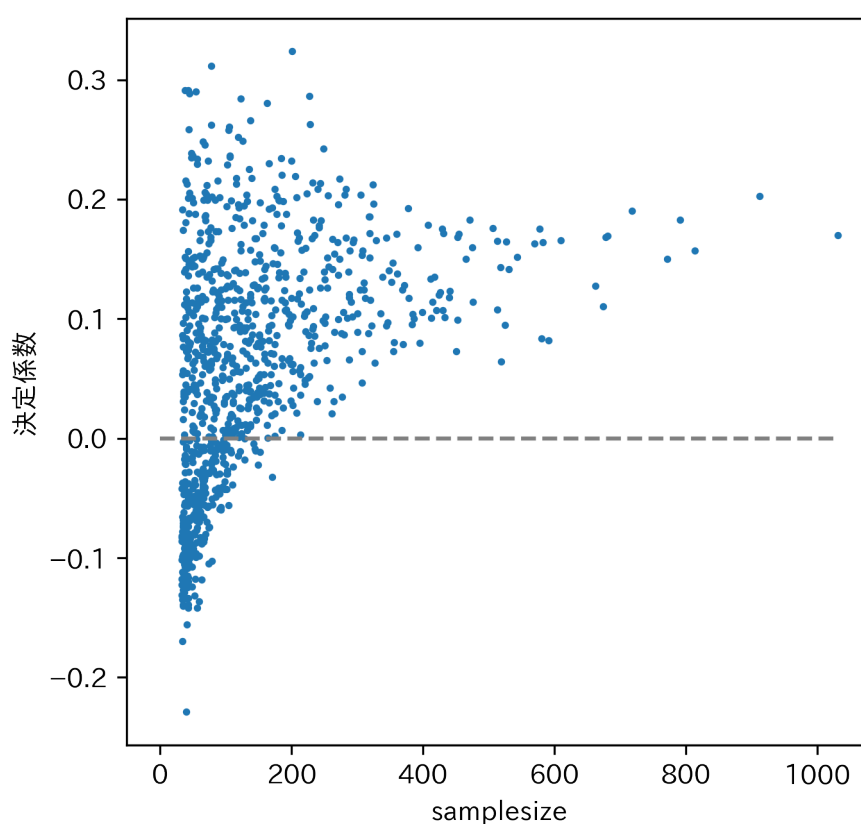


図 1 予測精度とサンプルサイズの関係（自治体別）

図 2 は推定された回帰係数のバイオリンプロットである。係数の分布を見ると、「コロナをきっかけにして今住んでいる場所が良いと思うようになった」「コロナをきっかけにして今住んでいる場所が良くないと思うようになった」の項目が、他の項目と比べてそれぞれ大きな正・負の係数となる傾向にあることがわかる。自宅近辺で過ごす時間が増えることで、居住地に関する知識が増え、それによって地域満足度に対する影響があったものと予想できる。これらの回答項目は、居住

6) 決定係数の定義から、被説明変数の平均からの残差二乗和で SSE を割り、その値を 1 から引くことで決定係数に換算した

地域特性をポジティブに受け取るかという個人的な性質の違いも反映している可能性があるため、実際に各地域に存在する地域特性と比較することで、より詳細な知見が得られると考えられる。現在の居住建物に関する「マンションダミー」「戸建てダミー」「タワマンダミー」の係数の分布をみると、マンション居住者の係数はやや正によっており、一方でタワマン居住者の係数は若干ゼロ近傍になっている地域が多いことがわかる。この結果は、タワマンの少ない地域が多いことを反映している可能性がある。一方でより個人の性格や家庭環境を反映していると思われる「未来は明るい」「家族関係は良好だ」といった項目や、その他の満足度に関連する「仕事には大変満足している」「社会的地位には大変満足している」といった項目に関しては、当てはまるほどに地域満足度も高くなる地域が多い。この要因として、地域満足度に対する回答に個人的な生活の満足度が混入してしまっている可能性が考えられる。よって地域特性から得られる地域満足度を純粋に評価するためには、このような影響を統制する必要性が示唆される。一方で、係数の値にばらつきがあることから、個人の性格や価値観によって評価する地域特性が異なり、それによって係数の地域差が生まれている可能性も考えられるため、地域特性と比較した分析が今後求められる。別途駅別の集計も行ったが、この結果は Appendix に記載した。

第2章で議論したように、居住満足度向上のためには、住民の選好と地域が提供するアメニティとがうまくマッチングされていることが重要となる。このため、推定されたモデルを用いて個人の属性からその人に適した街を予測することで、居住満足度が高まるように街と人とをマッチングすることが考えられる。しかし第2章に記載したように、推定したモデルは住民の選好が地域内で一様であることを仮定しているため、その地域の住民と近い属性を持つ人に対してはある程度の予測ができる可能性があるものの、一方で住民と全く異なる属性・価値観を持つ人の居住満足度に関しては、モデルの仮定を大きくことなるために予測が信頼できないと考えられる。この点を改善するためには、各評価主体の個人属性からその人がどのような選好を持ち、どのような地域特性を評価するのかを追加で予測する必要がある。

5 おわりに

本稿では地域評価の推定に着目して、その基礎的な検討として住民等の評価主体による地域評価のモデルを考察するとともに、既存の調査では比較的大規模な「街の住みこち&住みたい街ランキング」のデータを用いて個人属性と主観的地域満足度の関係を分析した。モデルの考察からは、地域評価においてはその地域を評価する主体が誰になるのかを明確化する必要があり、また評価主体の個人属性によって地域評価が異なってくる可能性が示唆された。この点は、今後の地域評価を改善していくために明確化されるべき点であり、また公表されている地域評価を解釈する上でも留意する必要があると考えられる。例えば、「街の住みこち&住みたい街ランキング」の調査からは多数のランキングが公表されているが、これらはランキングによって異なる評価主体集合を対象として集計されている。具体的には、例えば「街の住みこち 都道府県ランキング<全国版>」では「全国に居住の20歳以上の男女、2019年～2022年の646,245名を対象に集計」とある⁷⁾一方で、「街の住みこち 自治体ランキング<首都圏版>」では「首都圏居住の20歳以上の男女、2019年～2022年合計207,659名を対象に集計」とある⁸⁾。このため、<>内で示された地域の人から見た評価がこれらのランキングでは示されていると解釈できるため、

7) https://www.eheya.net/sumicoco/ranking/zenkoku/sumicoco_area_pref.html(2022/09/22 閲覧)

8) https://www.eheya.net/sumicoco/ranking/shutoken/sumicoco_area.html(2022/09/22 閲覧)

このようなランキングを参考とする際には、集計対象となっている集合がどうなっているのかを確認することで、より正確なランキングの解釈ができると考えられる。

また「街の住みこち&住みたい街ランキング」データの分析からは、個人属性が地域評価に影響を及ぼす影響に地域差がある可能性が示唆された。この結果は、アンケートの地域満足度に関する項目に対して地域とは関係のない満足度も含めて回答してしまっているために発生している可能性もある。このため今後の研究課題として、地域特性データを追加して、個人属性と両方を見た分析を行う必要があると考えられる。

参考文献

- Arlot, Sylvain and Alain Celisse (2010) “A survey of cross-validation procedures for model selection,” *Statistics Surveys*, Vol. 4, pp. 40–79, 1.
- Hastie, Trevor (2020) “Ridge Regularization: an Essential Concept in Data Science,” *Technometrics*, Vol. 62, pp. 426–433, 5.
- Rosen, Sherwin (1974) “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Journal of Political Economy*, Vol. 82, pp. 34–55.
- 清水千弘 (2022) 「地域の魅力の測定方法とその課題 -Walkability Index・再考-」, 『CSIS Discussion Paper』, 第 175 巻.
- 清水千弘・馬場弘樹・川除隆広・松縄暢 (2020) 「Walkability と不動産価値: Walkability Index の開発」, 『CSIS Discussion Paper』, 第 163 巻.

A Appendix

図 3 はサンプルサイズの少ない自治体を含めた回帰係数の分布である。全体的な傾向としては、図 2 と大きな違いがないことがわかる。

また自治体別と同様に、駅エリア別にも推定を行った。なお、駅別推定はサンプルサイズをある程度確保しやすい首都圏・関西圏の駅に限定して行った。図 4 は予測精度とサンプルサイズの関係を図 1 と同様に表示したものである。駅別に地域を分割すると、自治体別と比較して地域あたりのサンプルサイズが小さくなるため、ほとんどの地域でサンプルサイズ 200 未満となり、予測の観点からは良い推定結果を得ることができなかった。このため、精度の良い推定を行うためには、住民の価値観が近い駅エリアを結合するなどの処置が必要になると考えられる。参考に、推定した全駅に関する係数のバイオリンプロットを図 5 に示した。タワマンダミーの係数がゼロ近辺になる傾向が他のバイオリンプロットと比べて強いが、他の係数の分布傾向に自治体別との大きな違いは見受けられない。

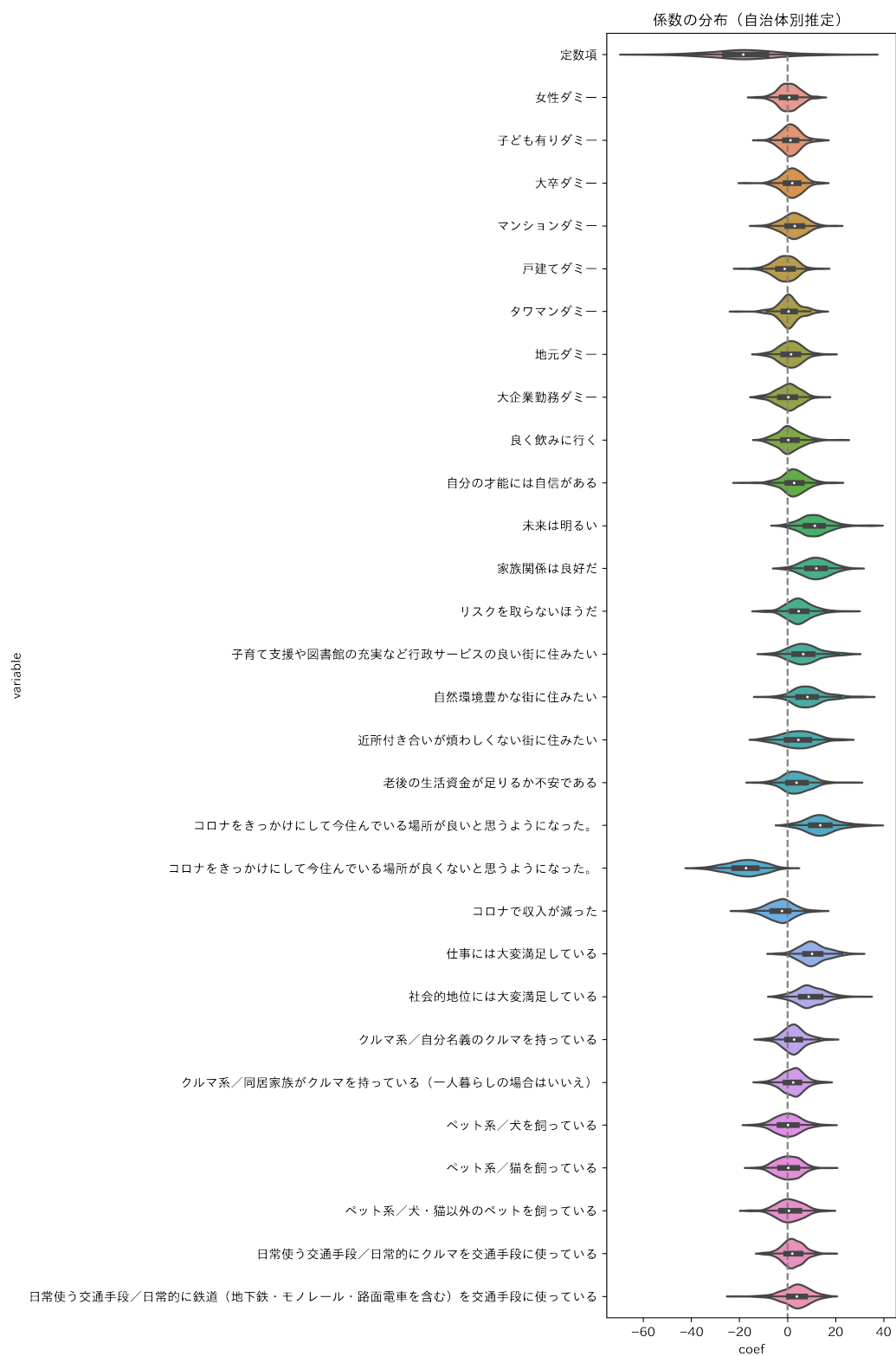


図 2 回帰係数のバイオリンプロット（自治体別）

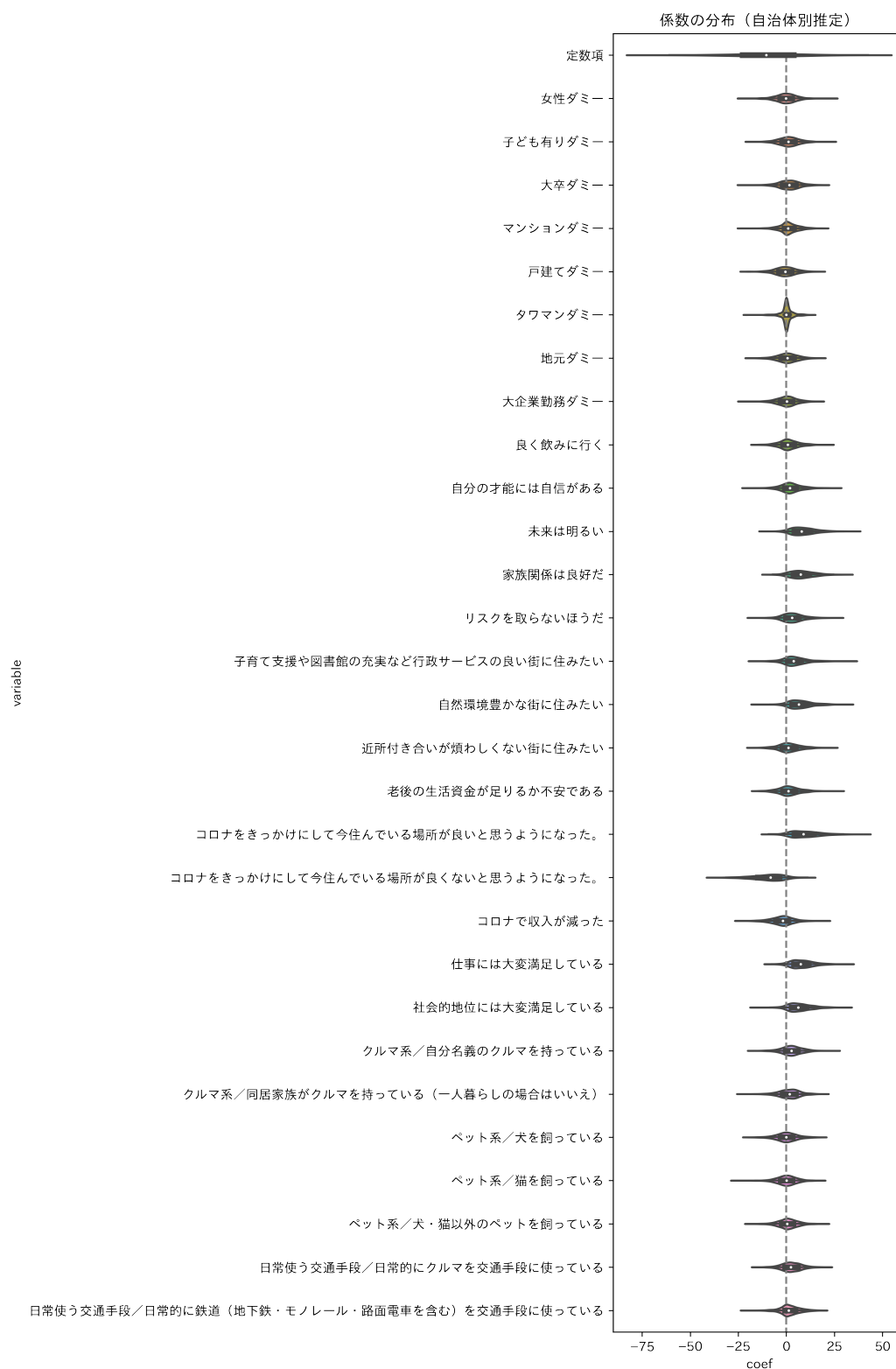


図3 回帰係数のバイオリンプロット（自治体別）（全自治体）

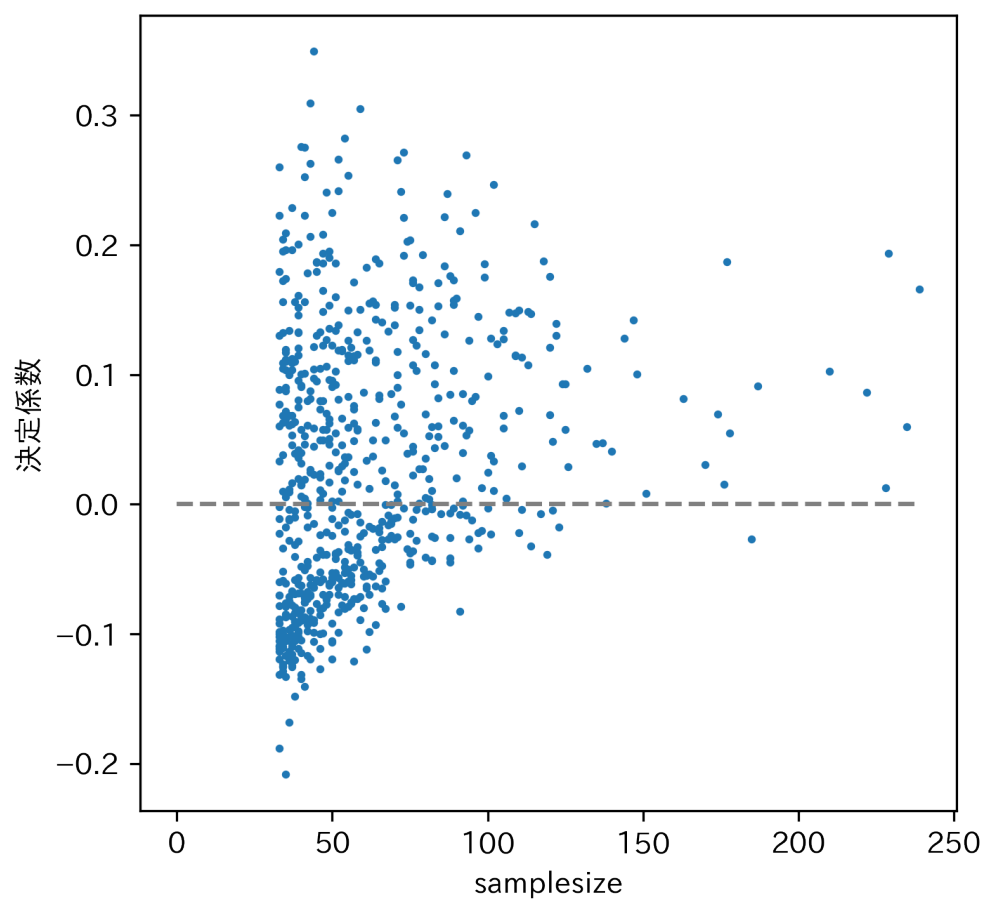


図 4 予測精度とサンプルサイズの関係（駅別）

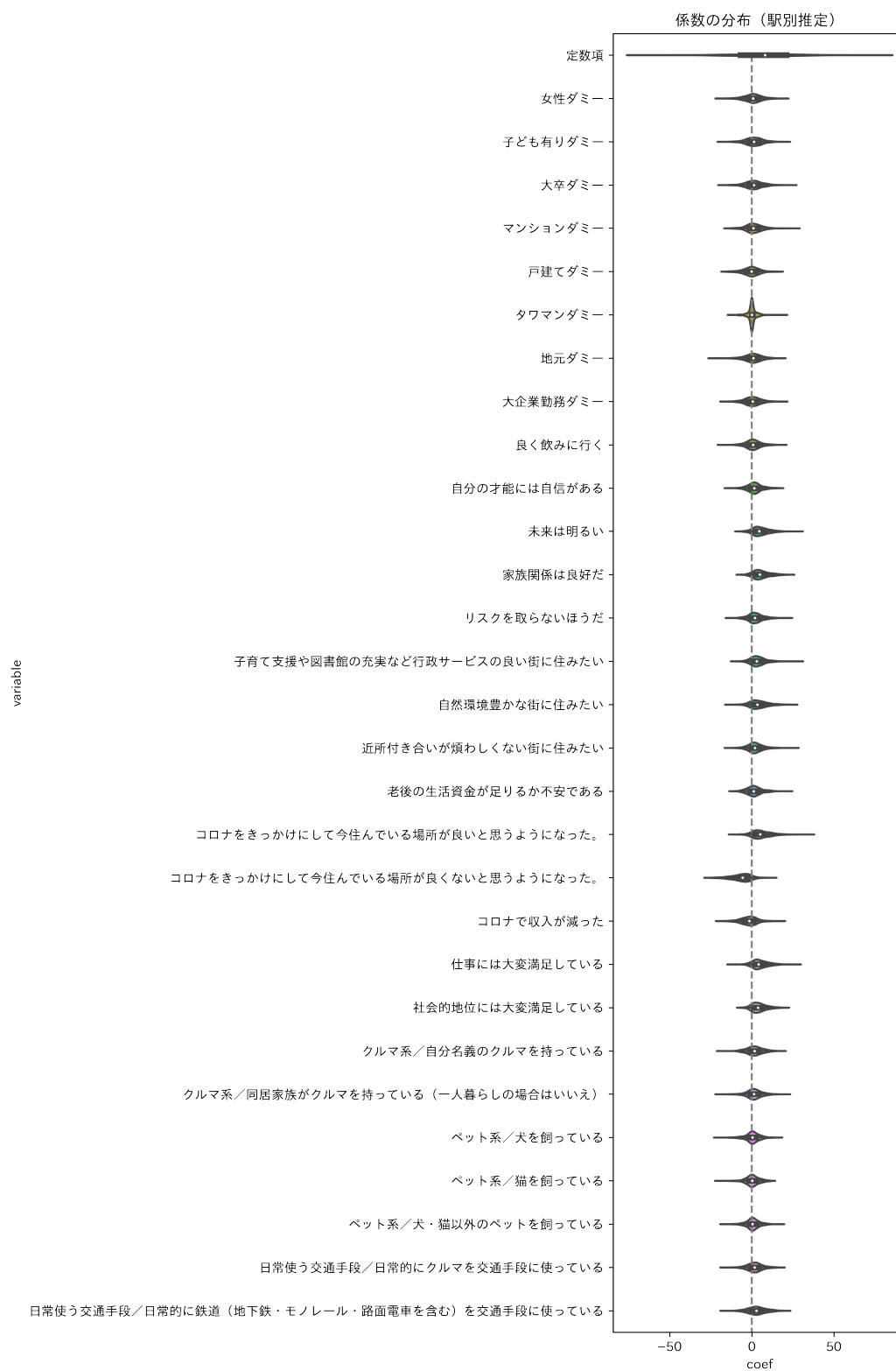


図 5 回帰係数のバイオリンプロット（駅別）