



RIETI Discussion Paper Series 20-J-033

大都市から地方への移住における社会経済的要因の影響 -Elastic net回帰を用いたポアソン重力モデルによる分析-

荒川 清晟
経済産業研究所

野寄 修平
東京大学



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<https://www.rieti.go.jp/jp/>

大都市から地方への移住における社会経済的要因の影響 -Elastic net 回帰を用いたポアソン重力モデルによる分析-*

荒川 清晟（経済産業研究所）[†]

野寄 修平（東京大学）

要 旨

本研究では、国内の市区町村間人口移動を対象に、大都市から地方への人口移動について分析を行った。本研究の目的は、先行研究において、重力モデル等を使用する際に問題となる多重共線性の問題を解決すること、先行研究において使用されている様々な社会・経済的な変数を可能な限り多く同時に分析し、各変数が大都市から地方への人口移動とどのような関連があるかを明らかにすることである。

その結果、本研究では、以下の2点を示した。1点目は、重力モデルによる分析に正則化回帰の1つである Elastic net 回帰を用いることで、社会経済的変数のような相関が強いため線形回帰モデルには同時に投入できない変数を同時に扱いつつ、変数選択をすることができることである。さらに Elastic net 回帰による変数選択でも、重力モデルの基本変数である人口規模、地域間距離は選択され、回帰係数の絶対値は他の変数に比べて大きかった。2点目は、先行研究において使用されている39の説明変数を全て Elastic net 回帰に投入した際には、21の変数が大都市から地方への人口移動と関連があるとして選択されることである。

今後より多くの変数を用いた分析を行い、パネルデータ解析も含め大都市から地方への移住に影響を与える要因を明らかにすることが求められる。

キーワード：地方創生、地域間人口移動、ポアソン重力モデル、Elastic net 回帰

JEL classification: C18, C21, R23

RIETI ディスカッション・ペーパーは、専門論文の形式でまとめられた研究成果を公開し、活発な議論を喚起することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び（独）経済産業研究所としての見解を示すものではありません。

* 本稿の原案に対して、経済産業研究所ディスカッション・ペーパー検討会の方々から多くの有益なコメントを頂いた。ここに記して、感謝の意を表したい。

[†] kiyoaki-arakawa257@g.ecc.u-tokyo.ac.jp

はじめに

日本は 2008 年をピークに人口減少に転じ（増田, 2014, 1 頁）、2050 年には日本の総人口が 9,708 万人にまで減少すると推計されている（国立社会保障・人口問題研究所, 2012）。この要因の 1 つとして、地方の経済や雇用状況の悪化による、大都市圏への人口移動が挙げられている（増田, 2014, 17-22 頁）。このような中、政府は 2014 年 9 月に「まち・ひと・しごと創生法」を制定し、内閣にまち・ひと・しごと創生本部を設置した。2014 年 12 月には 2060 年に一億人程度の人口を維持するなどの中長期的な展望を示した「まち・ひと・しごと創生長期ビジョン」を策定するとともに、まち・ひと・しごと創生法に基づき、5 か年の目標や施策の基本的方向等をまとめた第 1 期の「まち・ひと・しごと創生総合戦略」を策定した。第 1 期「まち・ひと・しごと創生総合戦略」の中に、「地方への人の流れをつくる」という基本目標があり、この基本目標は第 2 期の「まち・ひと・しごと創生総合戦略」でも「地方とのつながりを築き、地方への新しいひとの流れをつくる」として引き継がれており（閣議決定, 2019）、大都市への人口の一極集中の是正は日本の喫緊の課題である。

人口移動要因に関する研究は、大きく 2 つの方法がとられている。1 つは、人口集団とそれを取り巻く地域条件との関連に着目するもの、もう 1 つは、個人の移動理由に着目するものである（日本人口学会, 2002, 606 頁）。日本国内の人口移動を対象に、前者の方法により人口移動要因を分析した研究は数多く行われており、例えば、松本ら（2002）、伊藤（2002, 2011）、田村ら（2016）、田村（2017）、張ら（2016）、當麻（2016）、荒川（2018）などがある。人口集団とそれを取り巻く地域条件との関連に着目したマクロな分析は、個人単位のミクロな分析に比べてデータを収集しやすく、日本全国を対象とした網羅的な分析が可能である一方で、回帰分析における多重共線性の問題が発生するため、相関の高い複数の変数を同時に扱うことができないという問題がある（田村（2017）、荒川（2018）など）。そのため、先行研究においては、相関の高い変数同士を別々にモデルに投入して分析するなど対策をとっているが、回帰分析における説明変数が少なくなるという問題が生じる。

また、日本国内の人口移動を対象とした論文は、都道府県単位で分析しているものが多く、市区町村単位で分析しているものは少ない。近藤（2019）など人口移動に対する政策の効果を分析している近年の論文では市区町村単位の分析も行われ始めたこと、「まち・ひと・しごと創生総合戦略」が市区町村単位で作成され、市区町村単位で施策が実施されていることを考えると、市区町村単位の分析が求められる。

本研究の目的は以下の 2 点である。1 つは、多重共線性の問題を解決するため、正則化回帰の 1 つである Elastic net 回帰を用いた手法を提案することである。もう 1 つは、先行研究において使用されている様々な社会・経済的な要因のうち、可能な限り多くの変数を使用した上で、それぞれの変数が大都市から地方への市区町村単位の人口移動にどのように関連しているか明らかにすることである。

以下、本論文は次の通りに構成される。第 1 章では、関連する先行研究を整理するとともに、本研究の位置付けを示す。第 2 章では、本研究において使用するデータについて説明

し、第3章で本研究において使用するモデルと分析手法を説明する。第4章では、分析結果を示し、第5章でそれらの結果を踏まえた考察を述べ、第6章において、まとめとして本研究の結論、そして今後の研究課題について述べる。

1. 関連する先行研究

人口移動要因に関する研究は国内外で広く行われており、人口集団とそれを取り巻く地域条件との関連に着目するマクロな手法と、個人の移動理由に着目するミクロな手法がある（日本人口学会, 2002, 606 頁）。これらの方法を用いて、様々な人口移動流を対象とした分析が行われているが、Cadwallader (1996, 334-337 頁) によると、人口移動要因は、主に所得格差、就業機会、教育、行政サービス、アメニティ、年齢に分類される。以下、本研究において対象とするマクロな手法を用いた国内の人口移動要因に関する先行研究を表 1 にまとめる。

【表 1】

所得格差は、大阪市内 24 区内の人口移動を対象としている松本ら (2002) を除いて、先行研究の全てにおいて説明変数として使用されている。また、就業機会は、當麻 (2016)、田村 (2017) において、完全失業率や有効求人倍率が説明変数として使用されている。教育は、田村 (2017) において、教員数、進学率が使用されている。行政サービスは、張 (2016) などにおいて、生活関連社会資本が使用されている。アメニティは、伊藤 (2002)、伊藤 (2011) において、平均気温、積雪日数などが使用されている。年齢については、説明変数とせず、伊藤 (2011)、田村 (2017) などで、年齢層ごとに人口移動量を解析している。

これまでの研究の課題として、相関の強い説明変数を 1 つずつしか扱えていないこと、人口移動を都道府県単位で解析していることの 2 点が挙げられる。相関の強い説明変数の取扱いについて、田村 (2017)、荒川 (2018) などで多重共線性の問題を回避するために相関が強い社会経済的変数を同時に扱えていないという課題があった。また、人口移動の解析単位について、松本ら (2002)、荒川 (2018) を除き、都道府県単位の人口移動を分析しているものが多く、市区町村単位の人口移動を分析しているものは少ない。「まち・ひと・しごと創生総合戦略」が市区町村単位で作成されていることを考えると、市区町村間の分析が求められる。

そのため、本研究では、市区町村間の人口移動を対象に、中でも政令指定都市と東京都特別区部から、大都市圏に属する周辺市町村を除く市町村への人口移動を対象とする。先行研究において使用されている様々な社会・経済的な要因のうち、可能な限り多くの変数を使用した上で、相関の強い説明変数を同時に扱うことのできる Elastic net 回帰を用い、それぞれの変数が大都市から地方への市区町村単位の人口移動にどのように関連しているかを明らかにする。

2. 使用したデータ

I. 移動元と移動先の地域の定義

本研究では、移動元を政令指定都市と東京都特別区部、移動先を大都市圏に属する周辺市町村を除く市町村とする。大都市圏に属する周辺市町村は、e-Stat 政府の統計窓口¹に掲載されている平成 27 年国勢調査（総務省）における大都市圏構成市町村を対象とし、大都市圏周辺市町村の定義は、総務省統計局の定義する大都市圏の「中心市」への 15 歳以上通勤・通学者数の割合が該当市町村の常住人口の 1.5%以上であり、かつ、中心市と接続している市町村である（ただし、中心市への 15 歳以上通勤・通学者数の割合が 1.5%未満の市町村であっても、その周囲が周辺市町村の基準に適合した市町村によって囲まれている場合は、「周辺市町村」としている）（総務省統計局、2013）。周辺市町村への転入は、中心市への通勤など中心市の影響が大きいと考えられるため、移動先の地域から除外した。

II. 使用する変数

表 2 に使用した変数を示す。被説明変数として用いる、移動元の地域 (i) から移動先の地域 (j) への人口移動量を M_{ij} とする。 M_{ij} のデータは、住民基本台帳人口移動報告（総務省）を用いた。

重力モデルの基本となる 3 つの変数について、人口 (P_i, P_j) は、平成 27 年国勢調査の人口を用い、地域間の距離 (d_{ij}) については、国土地理院ホームページから各市区町村役場の緯度経度を取得し、国土地理院測量計算サイト²を用いて距離を算出した。その他については、表 1 の先行研究において使用されていた変数のうち、e-Stat 政府の統計窓口の都道府県・市区町村のすがた（社会・人口統計体系）において取得することができるものを用いた（以下 e-Stat と呼ぶ）。その際に、先行研究において用いられているデータと同じデータがない場合は、筆者らにより適切と思われるデータを用いた³。その結果、39 の説明変数を用いた。

【表 2】

2015 年度の全 1741 市区町村から、人口がゼロの町、移動元から移動先への人口移動数が 10 人未満の市町村の組を除外⁴した結果残った 748 市区町村（移動元 43、移動先 705）、4,072

¹ 政府統計の総合窓口（e-Stat）：<https://www.e-stat.go.jp/>（各府省が公表する統計データを一つにまとめ、統計データの検索をはじめとした、さまざまな機能を備えた政府統計のポータルサイト。）

² 国土地理院測量計算サイト：<https://vldb.gsi.go.jp/sokuchi/surveycalc/main.html>

³ 例えば、先行研究において県民所得の変数が使用されている場合、e-Stat には市民所得のデータが掲載されていないため、課税対象所得を用いた。

⁴ 政府統計では 10 人未満（ゼロを含む）のデータは“****”として実数が公開されていないため除外した。

組の人口移動データを解析した。総移動人口は平均 42 人、中央値 20 人、四分位範囲 13-37 人であり、多くの変数で、平均が中央値よりも大幅に大きい、右に尾を引く分布であった。表 3 に、年齢層、性別の総移動人口を示す。男性が全体の 56.9%、女性が 43.1%と男性がやや多く、いずれの年齢層でも男性が女性よりも多かった。また、40、50 歳代は男性が女性のおよそ 2 倍と性別に大きく偏りがあった。年齢の分布は、20、30 歳代が全体の半数を占め、一方で 60 歳以上は全体の 10%未満であった。

【表 3】

3. 使用するモデルと分析手法

I. 修正重力モデル

人口移動を社会の空間的相互作用として扱うため、修正重力モデルを用いた。修正重力モデルの基礎となっている重力モデルは、都市間の人口移動が古典力学におけるニュートンの万有引力の法則に従うことを利用したもので、Ravenstein (1885) によって人口移動のモデリングに導入された。人口移動量 (M_{ij}) が、移動元、移動先双方の人口規模 (P_i, P_j) と地域間の距離 (d_{ij}) に対して(1)式の関係にあるとし (G, a_1, a_2, e は各変数の寄与を定める定数)、移動元、移動先の人口が多いほど、また地域間の距離が小さいほど、人口移動量が多いと考える。

重力モデルは人口移動を分析する際に人口と距離のみ考慮しているが、これに人口と距離以外の様々な要素を追加したものが(2)式に示す修正重力モデルである (V_{i1}, \dots, V_{in} は移動元の変数, V_{j1}, \dots, V_{jm} は移動先の変数, $b_1, \dots, b_n, c_1, \dots, c_m$ は各変数の寄与を定める定数) (Greenwood and Hunt, 2003)。(2)式は変数に対して非線形であるが、両辺を対数変換することで、誤差項が平均 0、分散 σ^2 の正規分布に従うと仮定した線形モデルとして各変数の人口移動量に対する寄与を推定することができる。

$$M_{ij} = G \frac{P_i^{a_1} P_j^{a_2}}{d_{ij}^e} \quad (1)$$

$$M_{ij} = G \frac{P_i^{a_1} P_j^{a_2} V_{i1}^{b_1} \dots V_{in}^{b_n} V_{j1}^{c_1} \dots V_{jm}^{c_m}}{d_{ij}^e} \quad (2)$$

しかしながら、この推定方法の問題点として、①人口移動量の対数を用いることで、結果として人口移動量推定値の合計が実際の値より過小になってしまう傾向にあること、②人口移動量はカウントデータであり非負の整数であるにもかかわらず、連続分布である対数正規分布を仮定してしまっていることなどがある。これらを改良したものがポアソン重力モデルである (Flowerdew, 1982)。ポアソン重力モデルは人口移動量が離散分布の一つであるポアソン分布に従うと仮定したもので、(3)、(4)式により表現される。(3)式は人口移動量が k は母数 λ_{ij} のポアソン分布に従うものとし、このときの母数 λ_{ij} は(4)式のモデルによって規定される。

$$Pr(M_{ij} = k) = \frac{\exp(-\lambda_{ij}) \cdot \lambda_{ij}^k}{k!} \quad (3)$$

$$\lambda_{ij} = \exp\left(\ln G + a_1 \ln P_i + a_2 \ln P_j - e \ln d_{ij} + \sum_{l=1}^n b_l \ln V_{il} + \sum_{l=1}^m c_l \ln V_{jl}\right) \quad (4)$$

モデルには、移動元の人口 P_i 、移動先の人口 P_j 、移動元と移動先の距離 d_{ij} に加え前述の社会経済的変数 $V_{i1}, \dots, V_{in}, V_{j1}, \dots, V_{jm}$ を投入し、Elastic net 回帰によって人口移動数と関連のある変数を選択した。また、説明変数には値がゼロのものも含まれていたため、すべての値に1を足して対数変換を行なった。

II. Elastic net 回帰

説明変数の選択に正則化回帰の手法の1つである Elastic net 回帰 (Zou and Hastie, 2005) を用いた。Elastic net 回帰は L1 ノルムと L2 ノルムを正則化項として変数選択を行なう手法である。L1 ノルムのみを正則化項とした Lasso 回帰では変数の選択はできるものの、変数間の相関が強いものが存在する場合に推定が不安定になるという問題がある。Elastic net 回帰は L2 ノルムも用いることでこの問題を解決し、相関の強い変数が存在する場合には係数の大きさがほぼ同じになるという特徴がある。社会経済的変数には互いの相関が強いため、多重共線性の問題から線形回帰モデルに同時に投入することができない変数の組が複数存在するが、Elastic net 回帰を用いることで、これらの変数を同時に扱いつつ、変数の選択をすることができる。

Elastic net 回帰は一般に、(5)式に示した目的関数を最小化する $\beta = (\beta_0, \beta_1, \dots, \beta_m)$ を求めることにより、与えられたデータに対する最適なモデルを導出する。(3)式の第2項が正則化項であり、 λ は正則化項の大きさを制御するパラメータ ($\lambda > 0$)、 α は L1 ノルムと L2 ノルムの相対的な寄与を調整するパラメータ ($0 \leq \alpha \leq 1$) である。本研究では、10分割交差検証法によって平均2乗誤差の最小値を与えるときの λ 、 α を用いてモデルを決定した。

$$\frac{1}{2n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^m \left(\frac{1-\alpha}{2} \hat{\beta}_j^2 + \alpha |\hat{\beta}_j| \right) \quad (5)$$

回帰係数の95%信頼区間の算出にはブートストラップ法を採用した。重複を許したりサンプリングによる全データと同数のブートストラップ標本を作成し、Elastic net 回帰による回帰係数の推定を、予備的検討から300回繰り返して各説明変数についてパーセンタイル法により95%信頼区間を算出した。95%信頼区間がゼロを含まない場合に、その説明変数は被説明変数を説明するものとして選択した。

統計解析には R version 3.6.1 を使用し、Elastic net 回帰、ブートストラップ法によるサンプリングにはそれぞれ glmnet version 2.0.16、boot version 1.3.20 を用いた。

4. 結果

使用した変数の基本統計量と変数間の相関（対数変換後の Pearson 積率相関係数）をそれぞれ表 4、表 5 に示す。i_総人口と i_図書館数、i_可住地面積 1 km² 当たり人口密度と i_総面積（北方地域及び竹島を除く）、i_第 1 次産業就業者比率（対就業者）はそれぞれ絶対値で 0.8 以上の強い相関がみられた。

Elastic net 回帰により選択された説明変数の回帰係数の 95%信頼区間を表 6 に示す。すべての変数を同等に扱い、変数の選択に制約を課していなかったが、重力モデルで使用される距離、移動元の総人口、移動先の総人口は、総移動人口と関連のあるものとして選択された。係数の符号も距離は負、総人口は正と、重力モデルで使用されているものと一致した。

【表 4】

【表 5】

【表 6】

社会経済的変数については、移動元の変数において、i_可住地面積 1 km² 当たり人口密度、i_課税対象所得（納税義務者 1 人当たり）、i_幼稚園数、i_中学校数、i_第 1 次産業就業者比率（対就業者）、i_第 2 次産業就業者比率（対就業者）、i_完全失業率は係数の符号が負、i_総面積（北方地域及び竹島を除く）、i_高等学校数、i_図書館数、i_病院数は係数の符号が正であった。また、移動先の変数において、j_課税対象所得（納税義務者 1 人当たり）、j_第 1 次産業就業者比率（対就業者）、j_第 2 次産業就業者比率（対就業者）、j_公民館数は係数の符号が負、j_総面積（北方地域及び竹島を除く）、j_可住地面積割合、j_中学校数は係数の符号が正であった。

5. 考察

本研究は、大都市から地方への移住に関連する要因の探索において、相関の強い説明変数を同時に扱いつつ被説明変数と関連のある変数を抽出するため、ポアソン重力モデルに Elastic net 回帰を適用した変数選択を試みた。重力モデルで使用される距離、移動元の総人口、移動先の総人口は、総移動人口と関連のあるものとして選択され、係数の符号も距離は負、総人口は正と先行研究と一致した。大都市から地方への人口移動において、様々な変数を入れて調整しても、距離と総人口は人口移動と関連することが改めて確認された。

また、社会経済的変数について、移動元の変数である i_可住地面積 1 km² 当たり人口密度、i_課税対象所得（納税義務者 1 人当たり）、i_幼稚園数、i_中学校数、i_第 1 次産業就業者比率（対就業者）、i_第 2 次産業就業者比率（対就業者）、i_完全失業率は係数の符号が負であった。i_可住地面積 1 km² 当たり人口密度に関しては、大都市の中でも、都市化が進み

人口密度が高い地域は強い吸引力を持っているため、人口移動量との間に負の関係があったものと考えられる。 $i_{\text{課税対象所得}}$ （納税義務者 1 人当たり）に関しては、先行研究で報告されている人口移動との負の関係を支持する結果となった。伊藤（2002）によると、54 歳以下の女性、59 歳以下の男性の人口移動は実質個人所得と負の関係があり、60 歳以上の男性、65 歳以上の女性の人口移動において、正の関係があった。本研究では、表 3 に示すとおり、総移動人口に占める 60 歳以上の割合は 7.0%と少なく、そのため 50 歳代以下の年齢層における関係が主となり、全体として負の関係になったものと考えられる。このような年齢、性別によって関係が異なる変数との関連を明らかにするため、今後、年齢、性別ごとに分析する必要がある。 $i_{\text{幼稚園数}}$ 、 $i_{\text{中学校数}}$ に関して、幼稚園や中学校に通う子どもを育てている、もしくは今後子育てを行っていく者が比較的多いと考えられる 20 歳から 49 歳が総移動人口の約 70%を占めており（表 3）、幼稚園や中学校が少ない大都市から移住する者が多いと考えられる。 $i_{\text{第 1 次産業就業者比率}}$ （対就業者）、 $i_{\text{第 2 次産業就業者比率}}$ （対就業者）に関しては、第 1 次産業就業者は農林水産業の生産の性質上、特定の場所においてのみ、生産活動が可能であり、第 2 次産業就業者は工場の場合、一度立地すれば移転が少ないと予測され（當麻、2016）、第 1 次、第 2 次産業就業者は移動が難しいため、第 1 次産業就業者比率、第 2 次産業就業者比率の高い大都市からは移住する者が少ないと考えられる。 $i_{\text{完全失業率}}$ に関しては、独立行政法人労働政策研究・研修機構（2016）によると、大都市出身者の地方移住のきっかけで最も多いものが「自身の転勤等」（26.3%）であり、完全失業率が低いほど、就業している者の割合は多く、転勤による移動を行う可能性が高いため、移住する者が多いと考えられる。

移動元の変数の中でも、 $i_{\text{総面積}}$ （北方地域及び竹島を除く）、 $i_{\text{高等学校数}}$ 、 $i_{\text{図書館数}}$ 、 $i_{\text{病院数}}$ に関しては、係数の符号が正であった。 $i_{\text{総面積}}$ （北方地域及び竹島を除く）は、人口移動と負の関係にあった $i_{\text{可住地面積}}$ 1 km² 当たり人口密度と強い負の相関を認めた（相関係数-0.819）。大都市の中でも、都市化が進んだ人口密度の高い地域は強い吸引力を持っているため、総面積は人口密度とは逆に人口移動量との間に正の関係があったものと考えられる。 $i_{\text{高等学校数}}$ に関しては、表 3 より総移動人口における 10 代の人口移動数の割合が低く、人口移動との直接の関連を考察するのは困難であるが、e-Stat の都道府県単位の大学数と高等学校数のデータから、Pearson の積率相関係数が 0.910 であり、相関が非常に高いことが分かった。大学進学時に比べ、大学卒業時に都市圏から地方圏への人口移動が見られるため（田村、2017）、大学卒業時の人口移動と関係していると考えられ、今後の検証が必要である。また、 $i_{\text{図書館数}}$ は、人口移動との直接の関連を考察するのは困難である。しかし、e-Stat の都道府県単位の大学数と図書館数のデータから、Pearson の積率相関係数が 0.926 であり、相関が非常に高いことが分かった。そのため、 $i_{\text{図書館数}}$ に関しては、本研究ではデータの制限上利用できていない大学数が実際に影響している可能性がある。本研究における 20-29 歳の人口移動数は 34.5%を占めており（表 3）、大学卒業に伴う人口移動数が多いと考えられることから、人口移動数と図書館数とは正の関係になったと考えられ、今

後検証を行っていく必要がある。i_病院数について、荒川（2016）において、効果量はほとんどないが、健康・医療のまちという地域イメージと正の相関があると述べられており、先行研究と一致する結果となった。

移動先の変数である j_課税対象所得（納税義務者 1 人当たり）、j_第 1 次産業就業者比率（対就業者）、j_第 2 次産業就業者比率（対就業者）、j_公民館数は係数の符号が負であった。j_課税対象所得（納税義務者 1 人当たり）に関しては、先行研究の結果と異なる結果であった。この理由として、大都市から地方への人口移動を対象とした場合、所得の高い地域ではなく、より生活コストの低いところへ移住することが多い可能性が考えられる。平成 26 年全国消費実態調査（総務省）における消費支出と平成 26 年の課税対象所得（納税義務者 1 人当たり）の Pearson の積率相関係数は 0.551 であり、このうち食料、住居、被服及び履物、教育、教養娯楽との相関係数はそれぞれ 0.627、0.716、0.719、0.739、0.753 であった。さらに、一般社団法人 移住・交流推進機構（JOIN）（2017）によると、地方移住時の所得に関して「変化なしが好ましい」と答える者の割合が最も多かったが、続く回答として 500 万円～700 万円未満の層では「5%程減少まで許容できる」、700 万円以上の層では「10%程減少まで許容できる」など減少を許容する回答が多く、所得をそれほど重視していないことが考えられる。

j_第 1 次産業就業者比率、j_第 2 次産業就業者比率に関して、前述のとおり、第 1 次、第 2 次産業就業者は、第 3 次産業就業者と比較して移動が困難である。そのため、第 1 次産業就業者比率、第 2 次産業就業者比率の低い地方に移住する者が多いと考えられる。j_公民館数に関しては、張ら（2016）の 1985 年から 90 年の若年層を対象とした際の結果と一致したが、1995 年以降の人口移動の結果とは異なる結果となった。文部科学省生涯学習政策局社会教育課（2017）によると、1985 年以前に設立された公民館は全体の約 6 割であり、1981 年の耐震基準改正前に建築された公民館のうち、耐震性なし及び耐震診断未実施の公民館は 23.9%である。そのため、公民館が地域のマイナスの魅力となってしまう可能性がある。また、公民館の耐震化等が行われていないということは、財政状況が悪いことを反映していることも考えられ、今後の検証が必要である。

移動先の変数である j_総面積（北方地域及び竹島を除く）、j_可住地面積割合、j_中学校数は係数の符号が正であった。j_総面積（北方地域及び竹島を除く）、j_可住地面積割合は、先行研究と同じ結果となり、住む場所が広い場所に移住してくる者が多いと考えられる。また、j_中学校数に関しては、先行研究の結果と一致した。前述したとおり、本研究の対象とした人口移動において、中学生の児童を育てる可能性のある層の割合は多く、教育を重視しているものと考えられる。

本研究の限界及び今後の課題は、以下の 3 点である。1 点目は、説明変数として使用した変数において、e-Stat に掲載されていないデータ（地価、平均気温、年齢等）を用いることができていないため、e-Stat に掲載されていないデータを可能な限り収集した上で分析することが求められる。そうすることで、大都市から地方への人口移動に関係する変数をより正

確に見つけることができる。2点目は、性別・年代ごとの分析を行うことである。前述のとおり、本分析においても、年代によって要因が変化することが考えられる。3点目は、パネルデータ分析を行うことである。そうすることで、大都市から地方への人口移動に影響を与える要因を明らかにすることができる。

6. まとめ

本研究では、日本国内の市区町村間の人口移動について、政令指定都市と東京都特別区部から、大都市圏に属する周辺市町村を除く市町村への人口移動を対象に分析を行い、正則化回帰の1つである Elastic net 回帰を用いた、多重共線性の問題を解決する新たな手法を提案した。また、先行研究において使用されている様々な社会・経済的な要因のうち、可能な限り多くの変数を使用した上で、それぞれの変数が大都市から地方への人口移動にどのように関連しているかを明らかにした。

本研究の学術的意義は、以下の2点である。1点目は、重力モデルによる分析を行う際に、Elastic net 回帰を用いることで、従来社会的変数のような互いの相関の強い変数において、多重共線性の問題から線形回帰モデルに同時に投入することができない変数の組が複数存在していたが、Elastic net 回帰を用いることで、これらの変数を同時に扱いつつ、変数の選択をすることができること、そして Elastic net 回帰を用いた際にも重力モデルの基本変数である人口規模、地域間距離は選択されることを明らかにしたことである。2点目は、先行研究において使用されている変数を、Elastic net 回帰により全て使用した際には、21個の変数が大都市から地方への人口移動と関連があることを明らかにしたことである。

本研究の限界及び今後の課題は、以下の3点である。1点目は、e-Stat に掲載されていないデータを用いた上で分析を行うこと、2点目は、性別・年代ごとに解析を行うこと、3点目は、パネルデータ分析を行うことである。これにより、大都市から地方への人口移動についてより詳細な分析を行うことができる。

参考文献

- Cadwallader, M. (1996), *Urban Geography: An Analytical Approach*, New Jersey, Prentice-hall.
- Flowerdew, R. and Aitkin, M. (1982), "A Method of Fitting the Gravity Model Based on the Poisson Distribution", *Journal of Regional Science*, 22(2), pp. 191-202.
- Greenwood, M. J. and Hunt, G. (2003), "The Early History of Migration Research," *International Regional Science Review*, 26(1), pp. 3-37.
- Ravenstein, E. G. (1885), "The laws of migration," *Journal of the Statistical Society of London*, 48(2), pp. 167-235.
- Zou, H. and Hastie, T. (2005), "Regularization and Variable Selection via the Elastic Net," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 67(2), pp. 301-320.
- 荒川清晟. (2018), 「大都市から地方への移住における地域イメージの影響に関する研究」,

- 『東京大学大学院情報学環紀要 情報学研究』, 95, pp. 57-71.
- 伊藤薫. (2002), 「バブル経済期の男女・年齢別人口移動-1990年国勢調査人口移動集計結果を利用して-」, 『地域学研究』, 33(3), pp. 85-102.
- 伊藤薫. (2011), 「高齢者の長距離人口移動の決定因の変化-1960年国勢調査から2000年-」, 『地域学研究』, 41(1), pp. 179-194.
- 一般社団法人 移住・交流推進機構 (JOIN). (2017), 「若者の移住に関する調査」, at https://www.iju-join.jp/material/files/group/1/JOIN_report_201710.pdf, 参照2020年5月17日.
- 閣議決定. (2019), 「第2期『まち・ひと・しごと創生総合戦略』」, 2019年12月20日決定, at <https://www.kantei.go.jp/jp/singi/sousei/info/pdf/r1-12-20-senryaku.pdf>, 参照2020年5月10日.
- 国土地理院. 「測量計算サイト」, at <http://vldb.gsi.go.jp/sokuchi/surveycalc/surveycalc/bl2stf.html>, 参照2020年5月10日.
- 国立社会保障・人口問題研究所. (2012), 「日本の将来推計人口(平成24年1月推計)」, at <http://www.ipss.go.jp/syoushika/tohkei/newest04/gh2401.pdf>, 参照2020年5月10日.
- 近藤恵介. (2019), 「東京一極集中と地方への移住促進」, 『RIETI Policy Discussion Paper Series』, 19-P-006.
- 総務省統計局. (2013), 「平成25年住宅・土地統計調査用語の解説 << 地域 >>」, at <http://www.stat.go.jp/data/jyutaku/2013/1-1.htm>, 参照2020年5月15日.
- 田村一軌, 坂本博. (2016), 「日本の都道府県間人口移動の世代間比較」, 『AGI Working Paper Series』, 17, pp. 1-11.
- 田村一軌. (2017), 「大学進学にともなう都道府県間人口移動の定量分析-修正重力モデルによる分析-」, 『AGI Working Paper Series』, 3, pp. 1-17.
- 張峻屹, 瀬谷創, 兼重仁, 力石真. (2016), 「都道府県間人口移動の影響要因の経年的分析-空間的文脈依存性をもつ選択モデルに基づく分析-」, 『地域学研究』, 71(3), pp. 118-132.
- 當麻雅章. (2016), 「人口移動要因としての地域アメニティ近接性」, 『大阪大学経済学』, 66(3), pp. 1-23.
- 独立行政法人 労働政策研究・研修機構. (2016), 「UIJ ターンの促進・支援と地方の活性化-若年期の地域移動に関する調査結果-」, 『JILPT 調査シリーズ』, 152, at <https://www.jil.go.jp/institute/research/2016/documents/152.pdf>, 参照2020年5月15日.
- 日本人口学会編. (2002), 『人口大辞典』, 培風館.
- 増田寛也編. (2014), 『地方消滅』, 中央公論新社.
- 松本茂, 松下敬一郎. (2002), 「住宅環境が居住地の選定に与える影響: Poisson Gravity Model によるティプー仮説の検証」, 『關西大學經濟論集』, 51(4), pp. 455-471.
- 文部科学省生涯学習政策局社会教育課. (2017), 「公民館の現状と課題」, at

https://www.soumu.go.jp/main_content/000513104.pdf, 参照 2020 年 5 月 18 日.

表1 日本国内の人口移動の要因分析に関する先行研究

筆者名	年	使用しているモデル	従風変数	独立変数
松本 茂ら	2002	ポアソン重力モデル	大阪市内24区の人口移動	総地積、住宅地比率、借地比率、日照条件、保育所利便性、幼稚園利便性、診療所利便性、中心地ダミー
伊藤 薫	2002	修正重力モデル 流入超過率モデル	都道府県間人口移動	人口、距離、実質個人所得水準、地価水準、平均気温、積雪日数、暖房ダグリーデー、冷房ダグリーデー、日照時間
伊藤 薫	2011	修正重力モデル 流入超過率モデル	高齢者の都道府県間人口移動	人口1人当たり実質個人所得、消費者物価地域差指数、平均気温(平年値)、面積1ha当たり医師数、65歳以上人口千人当たり老人ホーム入所者数、男の出生地帰還指数65歳、男の出生地帰還指数75歳、女の出生地帰還指数65歳、女の出生地帰還指数75歳、男の子供指数35歳、男の子供指数65歳(親子の年齢差35歳)、女の子供指数65歳(親子の年齢差35歳)
田村 一軌ら	2016	修正重力モデル	都道府県間人口移動	人口、距離、隣接県かどうかのダミー変数、東北・関東・中部・近畿・中国・四国・九州ダミー変数、15～24歳人口、65歳以上人口、所得格差、年齢格差
張 岐屹ら	2016	集計型離散選択モデル	都道府県間人口移動	大学・短期大学の数(15～34歳の千人あたりの数)、県民所得(百万円/人)、居住面積(1住宅あたりの量数)、都市公園数(ノ千人)、一般病院・診療所数(ノ千人)、幼稚園保育園数(0～14歳の千人あたりの数)、文教施設ストック(万円/人)、事務所数(ノ千人)、平均労働時間(時間)、第3次産業就業者比率(%)、飲食店数(ノ千人)、老人ホーム数(65歳以上の千人あたりの数)、道路ストック(万円/人)、航空ストック(万円/人)、他県との平均距離(km)
當麻 雅章	2016	集計ロジットモデル	移住率 (都道府県間移動：地域から地域への地域間移住者数と地域における域内移動者数の比を対数化したもの)	一人当たり県民所得、有効求人倍率、事務所数、第三次産業就業者比率、大卒者人口比率、大学数、若年者人口比率、地域間距離、累積移住者数(1954年から1973年までの移住者数の総和)、都道府県総面積に占める国立公園、国定公園、および都道府県立自然公園の割合、年間平均気温、年間降水日数
田村 一軌	2017	修正重力モデル	出身高校の所在地県別大学入学者数	出身高校の所在地県別大学進学者数、大学学部別の所在地別大学進学者数、都道府県庁間の距離、一人当たり県民所得、完全失業率(年平均)、大学卒業生数から進学者数などを除いた数に対する就職者数の比率、新規大学卒業者の初任給額、可住人口密度、民営賃貸住宅家賃、学生一人当たり教員数、教員一人当たり科研費配分額、私立大学の授業料、隣接県かどうかのダミー変数、東北・関東・中部・近畿・中国・四国・九州ダミー変数
荒川 清晟	2018	修正重力モデル	市区町村間の人口移動	人口、距離、納税者1人当たりの課税対象所得、地域イメージ(IT・先端技術のまち、スポーツのまち、デザインやセンスの良いまち、学術・芸術のまち、環境にやさしいまち、観光・レジャーのまち、教育・子育てのまち、健康増進・医療福祉のまち、国際交流のまち、住民参加のまち、生活に便利・快適なまち、地産産業が盛んなまち、農林水産業が盛んなまち、歴史・文化のまち)

表2 使用した変数（被説明変数1変数、説明変数39変数）

変数名	単位	出典
総移動人口	人	住民基本台帳人口移動報告(総務省)
距離	km	国土地理院
総人口	人	国勢調査 人口推計(総務省)
可住地面積1 km ² 当たり人口密度	人	国勢調査 全国都道府県市区町村別面積調
総面積(北方地域及び竹島を除く)	ha	国勢調査 全国都道府県市区町村別面積調
可住地面積割合	%	国勢調査 全国都道府県市区町村別面積調
課税対象所得(納税義務者1人当たり)	千円	市町村税課税状況等の調(総務省)
幼稚園数	園*	学校基本調査(文部科学省)
小学校数	校*	学校基本調査
中学校数	校*	学校基本調査
高等学校数	校*	学校基本調査
第1次産業就業者比率(対就業者)	%	国勢調査
第2次産業就業者比率(対就業者)	%	国勢調査
第3次産業就業者比率(対就業者)	%	国勢調査
完全失業率	%	国勢調査
公民館数	館*	社会教育調査(文部科学省)
図書館数	館*	社会教育調査
一般病院数	施設*	医療施設調査(厚生労働省)
病院数	施設*	医療施設調査
老人福祉施設数	所*	社会福祉施設等調査(厚生労働省)
保育所等数	所*	社会福祉施設等調査

*(変数×総人口/100,000)により総人口10万人当たりに統一した。

総移動人口、距離は1つの移動元、移動先の組につき1つずつ、その他の19変数については移動元、移動先それぞれに1つずつ。

表3 年齢・性別の人口移動量

	男性	女性	全体
総移動人口	97,077 (100)	73,677 (100)	170,754 (100)
0～9歳	9,420 (9.7)	8,983 (12.2)	18,403 (10.8)
10～19歳	5,892 (6.1)	3,651 (5.0)	9,543 (5.6)
20～29歳	32,556 (33.5)	26,325 (35.7)	58,881 (34.5)
30～39歳	21,051 (21.7)	17,621 (23.9)	38,672 (22.6)
40～49歳	13,557 (14.0)	7,369 (10)	20,926 (12.3)
50～59歳	8,244 (8.5)	4,025 (5.5)	12,269 (7.2)
60歳以上	6,326 (6.5)	5,677 (7.7)	12,003 (7.0)
不詳/その他	31 (0.0)	26 (0.0)	57 (0.0)

実数 (割合)で記載。

表4 使用した変数の基本統計量 (n=4,072)

変数名	平均	標準偏差	中央値	最小値	最大値
距離	405	391	240	23	2,607
i_総人口	1,357,553	958,192	1,082,159	58,406	3,724,844
i_可住地面積1 km ² あたり人口密度	9,286	5,464	9,326	1,208	22,380
i_総面積（北方地域及び竹島を除く）	38,247	37,394	32,645	1,011	155,806
i_可住地面積割合	80	26	95	24	100
i_課税対象所得（納税義務者1人あたり）	3,803	923	3,558	2,932	10,232
i_幼稚園数	8	2	7	5	21
i_小学校数	10	2	11	8	19
i_中学校数	6	1	6	4	24
i_高等学校数	3	2	3	1	31
i_第1次産業就業者比率（対就業者）	1	1	0	0	4
i_第2次産業就業者比率（対就業者）	18	4	19	8	34
i_第3次産業就業者比率（対就業者）	72	4	73	60	81
i_完全失業率	4	1	4	2	6
i_公民館数	2	3	0	0	10
i_図書館数	2	1	1	1	10
i_病院数	6	3	5	3	26
i_一般病院数	5	2	5	2	26
i_老人福祉施設数	2	1	2	0	8
i_保育所等数	15	3	15	12	27
j_総人口	166,875	144,596	117,146	832	599,814
j_可住地面積1 km ² あたり人口密度	1,180	1,105	940	17	8,140
j_総面積（北方地域及び竹島を除く）	42,418	32,172	36,397	518	217,761
j_可住地面積割合	47	23	44	6	100
j_課税対象所得（納税義務者1人あたり）	2,846	289	2,835	1,993	7,845
j_幼稚園数	11	7	10	0	126
j_小学校数	22	15	18	7	275
j_中学校数	11	10	9	4	225
j_高等学校数	5	4	5	0	120
j_第1次産業就業者比率（対就業者）	7	6	4	0	50
j_第2次産業就業者比率（対就業者）	24	8	24	5	52
j_第3次産業就業者比率（対就業者）	66	8	66	33	90
j_完全失業率	4	1	4	1	8
j_公民館数	20	39	11	0	642
j_図書館数	4	4	2	0	40
j_病院数	9	5	8	0	68
j_一般病院数	8	5	7	0	57
j_老人福祉施設数	6	5	5	0	114
j_保育所等数	25	10	23	0	95

i、jはそれぞれ移動元、移動先の変数を示す。

表5 Elastic net回帰により選択された変数間の相関係数

変数名	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
A 距離	1.000																				
B 総人口	0.019	1.000																			
C 可住地面積1 km2当たり人口密度	0.134	-0.234	1.000																		
D 総面積 (北方地域及び竹島を除く)	-0.083	0.698	-0.819	1.000																	
E 課税対象所得 (納税義務者1人当たり)	0.089	-0.413	0.585	-0.650	1.000																
F 幼稚園数	-0.025	-0.219	-0.145	0.020	0.247	1.000															
G 中学校数	-0.106	-0.263	-0.448	0.208	-0.052	0.460	1.000														
H 高等学校数	-0.051	-0.295	-0.131	-0.038	0.298	0.519	0.787	1.000													
I 第1次産業就業者比率 (対就業者)	-0.136	0.063	-0.871	0.621	-0.454	0.046	0.430	0.149	1.000												
J 第2次産業就業者比率 (対就業者)	-0.064	0.475	-0.521	0.606	-0.618	-0.042	0.031	-0.264	0.423	1.000											
K 完全失業率	-0.057	0.418	-0.381	0.524	-0.766	-0.074	0.132	-0.043	0.149	0.299	1.000										
L 図書館数	0.008	-0.837	0.324	-0.698	0.485	0.307	0.268	0.318	-0.144	-0.357	-0.493	1.000									
M 病院数	-0.076	0.113	-0.456	0.422	-0.492	0.139	0.466	0.286	0.262	0.099	0.659	-0.248	1.000								
N 総人口	0.322	-0.407	0.178	-0.355	0.234	0.106	0.058	0.073	-0.070	-0.103	-0.279	0.399	-0.202	1.000							
O 総面積 (北方地域及び竹島を除く)	0.075	-0.140	0.043	-0.101	0.058	-0.003	0.015	0.016	-0.038	-0.081	-0.075	0.121	-0.029	0.392	1.000						
P 可住地面積割合	0.134	-0.138	0.073	-0.134	0.081	0.035	-0.011	-0.008	-0.031	-0.047	-0.105	0.133	-0.091	0.291	-0.550	1.000					
Q 課税対象所得 (納税義務者1人当たり)	-0.044	-0.271	0.117	-0.220	0.167	0.055	0.032	0.039	-0.070	-0.090	-0.199	0.267	-0.133	0.489	-0.008	0.304	1.000				
R 中学校数	-0.073	0.237	-0.113	0.223	-0.173	-0.073	-0.024	-0.035	0.020	0.024	0.229	-0.251	0.200	-0.577	0.225	-0.536	-0.454	1.000			
S 第1次産業就業者比率 (対就業者)	-0.234	0.269	-0.125	0.238	-0.179	-0.079	-0.042	-0.053	0.043	0.039	0.210	-0.280	0.172	-0.626	0.160	-0.287	-0.542	0.604	1.000		
T 第2次産業就業者比率 (対就業者)	-0.432	0.084	-0.047	0.067	-0.031	-0.002	-0.004	-0.033	0.050	0.123	-0.023	-0.042	-0.092	-0.006	0.060	-0.008	0.137	-0.136	-0.034	1.000	
U 公民館数	-0.219	0.116	-0.035	0.084	-0.057	-0.007	0.007	0.018	0.021	0.070	0.073	-0.100	0.045	-0.209	0.135	-0.251	-0.215	0.239	0.300	0.156	1.000

表 6 Elastic net 回帰により選択された変数の回帰係数

変数名	回帰係数 (95%信頼区間)	
切片	3.198	3.280
距離	-0.944	-0.819
i_総人口	0.451	0.608
i_可住地面積1 km ² 当たり人口密度	-0.451	-0.110
i_総面積 (北方地域及び竹島を除く)	0.242	0.459
i_課税対象所得 (納税義務者 1 人当たり)	-0.309	-0.090
i_幼稚園数	-0.144	-0.008
i_中学校数	-0.207	-0.054
i_高等学校数	0.018	0.140
i_第1次産業就業者比率 (対就業者)	-0.384	-0.125
i_第2次産業就業者比率 (対就業者)	-0.481	-0.289
i_完全失業率	-0.253	-0.023
i_図書館数	0.022	0.194
i_病院数	0.008	0.738
j_総人口	0.542	0.944
j_総面積 (北方地域及び竹島を除く)	0.097	0.433
j_可住地面積割合	0.001	0.238
j_課税対象所得 (納税義務者 1 人当たり)	-0.209	-0.067
j_中学校数	0.040	0.195
j_第1次産業就業者比率 (対就業者)	-0.412	-0.195
j_第2次産業就業者比率 (対就業者)	-0.529	-0.383
j_公民館数	-0.149	-0.055

i、jはそれぞれ移動元、移動先の変数を示す。

付表 全変数の回帰係数

変数名	回帰係数 (95%信頼区間)	
切片	3.198	3.280
距離	-0.944	-0.819
i_総人口	0.451	0.608
i_可住地面積1 km ² あたり人口密度	-0.451	-0.110
i_総面積 (北方地域及び竹島を除く)	0.242	0.459
i_可住地面積割合	-0.012	0.075
i_課税対象所得 (納税義務者1人当たり)	-0.309	-0.090
i_幼稚園数	-0.144	-0.008
i_小学校数	-0.038	0.112
i_中学校数	-0.207	-0.054
i_高等学校数	0.018	0.140
i_第1次産業就業者比率 (対就業者)	-0.384	-0.125
i_第2次産業就業者比率 (対就業者)	-0.481	-0.289
i_第3次産業就業者比率 (対就業者)	-0.016	0.091
i_完全失業率	-0.253	-0.023
i_公民館数	-0.056	0.090
i_図書館数	0.022	0.194
i_病院数	0.008	0.738
i_一般病院数	-0.681	0.010
i_老人福祉施設数	-0.056	0.008
i_保育所等数	-0.087	0.043
j_総人口	0.542	0.944
j_可住地面積1 km ² あたり人口密度	0.000	0.273
j_総面積 (北方地域及び竹島を除く)	0.097	0.433
j_可住地面積割合	0.001	0.238
j_課税対象所得 (納税義務者1人当たり)	-0.209	-0.067
j_幼稚園数	-0.016	0.059
j_小学校数	-0.031	0.127
j_中学校数	0.040	0.195
j_高等学校数	-0.041	0.028
j_第1次産業就業者比率 (対就業者)	-0.412	-0.195
j_第2次産業就業者比率 (対就業者)	-0.529	-0.383
j_第3次産業就業者比率 (対就業者)	-0.148	0.020
j_完全失業率	-0.019	0.073
j_公民館数	-0.149	-0.055
j_図書館数	-0.042	0.046
j_病院数	-0.189	0.044
j_一般病院数	-0.050	0.195
j_老人福祉施設数	-0.090	0.003
j_保育所等数	-0.037	0.068

i、jはそれぞれ移動元、移動先の変数を示す。