

DSSR

Discussion Paper No. J-2

階層ベイズ線形回帰モデルを利用した市町村の支出する除雪費が
各地域に及ぼす経済効果分析

本間 克仁 石垣 司

2017年1月

Data Science and Service Research
Discussion Paper

Center for Data Science and Service Research
Graduate School of Economic and Management
Tohoku University
27-1 Kawauchi, Aobaku
Sendai 980-8576, JAPAN

階層ベイズ線形回帰モデルを利用した市町村の支出する除雪費が各地域に及ぼす経済効果分析

本間 克仁* 石垣 司†

和文概要 除雪事業に支出される除雪費が地域経済にどのような影響を与えるのかについて、階層的な線形回帰モデルを使用しベイズ推定することで考察を行った。推定された影響は、特定の市町村においては負の可能性が高いことが分かり、また地域間における影響力の地域異質性が明らかとなった。

1. はじめに

日本は国土の約 51%が豪雪地帯であり、また当該地には総人口の約 15%が居住している [1]. 恒常的な降積雪は市民生活に大きな影響を与えるため、国や地方自治体など道路管理者は除雪事業を行う。

除雪事業は毎年多額の予算を必要とする。公共サービスのあり方について再検討が求められている中、除雪事業もまたその一つとして考えるべきである。しかしながら、効率化に先立って除雪事業と地域経済の関係を明らかにしなければ、目的に対し適切な施策を行うことはできない。また、除雪事業には全国的に統一的なプロセスが存在しないことから、除雪の出動基準などを筆頭に、市町村を単位とし所与の条件の下で長期間にわたり構築されてきた独自のシステムが存在している [2]. つまり、除雪事業を通じて結ばれる除雪事業と地域経済の関係は、それぞれ地域ごとに異なるものであり、必要となる施策もまた変わってくる可能性がある。しかしながら今日まで、市町村を単位とした除雪に関する詳細なデータを得ることが難しく、市町村単位での除雪費と地域経済の関係の分析は行われていない。

本研究では、道路管理者の支出する除雪事業にかかる経費（以下、除雪費）が各地域に与える経済効果とその異質性を分析し、当該地の行政の意思決定に資する情報を抽出することを目的とする。分析の対象となる道路管理者としては、山形県内の 35 市町村を取り上げる。このような地域異質性の分析を行うため、35 市町村のパネルデータに対して、階層ベイズ線形回帰モデルを適用する。具体的には、市町村内総生産成長率を目的変数、除雪費に関する変数群とマクロ経済指標に関する変数群を説明変数、各市町村の特徴を表す変数群を階層構造における説明変数として、除雪費と市町村単位の地域経済の関係をモデル化する。このモデルを用いた分析を行うことで表 1 の課題を検証し、目的に資する情報を抽出する。

表 1: 検証課題

- | |
|------------------------|
| 1. 除雪費の支出は地域経済に影響を与えるか |
| 2. その影響は市町村ごとに異なるか |

* 米沢市役所総務部 katuhito2981@city.yonezawa.yamagata.jp

† 東北大学大学院経済学研究科准教授 isgk@econ.tohoku.ac.jp

2. 先行研究

2.1. 除雪に関する研究

除雪事業に関する研究は、1. 除雪の効率的運用、2. 除雪事業がもたらす便益の推定の二つに大別できる。前者の研究は、オペレーションズリサーチにおける運搬経路問題の一つとして扱われ、特に除雪車と融雪剤散布車の効率的な運行という観点から注目を集めてきた [3]。また、モントリオール市における都市エリアの分割および効率的な排雪場所への割り当てについての研究 [3] や、数理最適化の観点からエドモントン市における除雪車の運行距離の最小化を検討した研究 [4] もある。このほか、近年に到るオペレーションズリサーチ領域における研究は Gang et al. [4] に詳しい。日本においては、マルチエージェントシステムを使用し米沢市における除雪車の効率的運用を検討した柿崎・松尾 [5] の研究などがある。便益に関する研究については、田辺ら [6] に詳しい。これらでは除雪事業がもたらす便益をそれぞれ定義し経済モデルを構築し計算を行っているが、本研究が対象とする市町村内総生産と除雪費の関係を直接的に統計モデル化した研究は見当たらない。

2.2. 公共事業に関する研究

公共投資と経済成長率の関わりを実証的に扱った研究は、Barro [7] など、数多く存在する。日本においては、1990年代に盛んに研究が行なわれ、近年でも唐木ら [8] など研究の蓄積が進んでいる。しかしながら、ストック効果の欠如と不確実性により、除雪事業は一般的な公共事業とは切り分けて分析を行う必要がある。まず、ストック効果の欠如とは、一般的な公共事業に期待される効果は大きく分けてフロー効果とストック効果の2種類ある [8] が、除雪活動においてはフロー効果のみが期待され、投資がストック化され長期において効果が現れることがない [9] ということの意味する。次に不確実性とは、田辺ら [9] は除雪事業を「不確実性に密接に関連する公共事業」と指摘しているが、恒常的な事業でありながらその年の降雪の様相によって大幅に支出が増減することで、地域経済に負荷をかけやすい事業であることを意味する¹。一般的に公共事業は国単位で実質 GDP を1年目 1.14%、2年目以降もおおむね 1%程度拡大させる [10]。しかしながら、上記2点と前述の地域独自性を踏まえると、除雪事業では異なる結果になることが予想される。

また公共事業に関する研究の対象は、通常国全体もしくは県単位であった。これは、鯉江 [11] が指摘しているとおり、市町村を単位とした詳細なデータを得ることが難しく、市町村ごとの地域計量モデルを構築することは事実上不可能であったことによる。これに対し、鯉江 [11] は以下のような市内総生産を予測するコンパクトな長岡市地域計量経済モデルを提案し、最小二乗推定を行った。

$$YY = \beta_0 + \beta_1 \times EE + \beta_2 \times GDP \quad (1)$$

なお、ここで YY は市内総生産、 EE は従業員数、 GDP は国内総生産であり、 β_0 は切片、 β_1 、 β_2 はそれぞれの説明変数についての回帰係数である。本研究では鯉江 [11] の提案モデルを拡張し、除雪事業を一般的な公共事業と切り分け、直接市町村内総生産への影響を検討していく。

3. データ

本研究では、山形県内 35 市町村を対象に、10 年度分のデータを使用して分析を行う。市町村内総生産成長率を後述のモデルで表現する。目的変数は市町村内総生産成長率であり、

¹山形新聞 2016 年 03 月 24 日などを参照

説明変数に除雪事業に直接関係する変数とマクロ変数を用意する。以下、具体的に使用する変数についての説明を行う。詳細については Appendix A に記載する。

3.1. 目的変数

3.1.1. 市町村内総生産成長率

出典は山形県が作成した市町村民経済計算。データとして存在するのは名目値のみであるが、県内総生産インプリシット・デフレーターを使用して実質化した。総生産は一般的に非定常過程であると想定されているので、対数階差をとって成長率に近似させることで上記の問題に対応する。非定常過程であることが予想されるそのほかの変数に対しても同様の処理を行う。

3.2. 除雪事業に直接関係する説明変数

3.2.1. 除雪費

除雪事業に支出した額。各市町村が作成する歳入歳出決算書などについて情報公開請求を行いデータを得た。山形県において除雪事業が行われるのはおおむね十二月から三月頃であるが、実際除雪費が支払われ地域内に影響を与える期間は、各団体の支払い方によって異なる。たとえば、年度初めに建設業者と委託契約を結び、出納閉鎖期間内の翌年四月あるいは五月に実際の支払いが行われるならば、翌年から経済活動に影響を与えられ一方、月ごとに精算を行うならば、同年内に影響を与え始めるだろう。以上から本研究では、当期に支払われる除雪費と一期前の除雪費に関する影響を分析の対象とする。

3.2.2. 降雪量

各市町村で降った累計降雪量。

表 2: 説明変数

除雪事業に直接関係する説明変数	マクロ変数	階層構造で使用する説明変数
除雪費当期 $x_{i,t}^1$	人口成長率 $c_{i,t}^1$	除雪費平均 z_{1i}
除雪費一期前 $x_{i,t}^2$	国内総生産成長率 $c_{i,t}^2$	降雪量平均 z_{2i}
降雪量 $x_{i,t}^3$	従業者数成長率 $c_{i,t}^3$	一般道路距離 z_{3i}
降雪量一期前 $x_{i,t}^4$	リーマンショックダミー (2009年) $c_{i,t}^4$	市町村面積 z_{4i}
	リーマンショック一期後ダミー $c_{i,t}^5$	
	東日本大震災ダミー (2011年) $c_{i,t}^6$	
	東日本大震災一期後ダミー $c_{i,t}^7$	

3.3. マクロ変数

3.3.1. 人口成長率

出典は山形県山形県の人口と世帯数。原系列は各市町村の人口。非定常過程であると想定されるので、対数階差をとって成長率に変換している。

3.3.2. 国内総生産成長率

出典は内閣府作成の国民経済計算。原系列は日本における国内総生産。非定常過程であると想定されるので、対数階差をとって成長率に変換している。なお、鯉江 [11] は長岡市の地域計量経済モデルを構築するにあたり、国内総生産を外生変数として扱っているが、本研究において扱う山形県内の市町村は経済規模の小さい市町村であることから、同時性があることは想定しにくく同様に扱うことができると考えられる。

3.3.3. 従業者数成長率

出典は経済産業省作成の工業統計調査。原系列は各市町村の製造業従事者数。非定常過程であると想定されるので、対数階差をとって成長率に変換している。

このほか、ダミー変数や階層構造で使用する説明変数を2に整理する。

4. モデル

本研究では階層ベイズ線形回帰モデルを使用する。研究対象となる山形県内35市町村について、データはそれぞれ10年度分しか存在せず、各市町村ごとにモデルを構築し分析を行うにはサンプルサイズが十分とは言い難い。階層ベイズモデルでは、各グループ（市町村）ごとのデータが少ない状況においても、各グループに共通する構造を事前分布として階層的に導入することで各グループのデータ不足を補い、各グループごとのパラメータに関する事後分布を安定的に推定することができる[12]。また、前述の説明変数に加え空間構造を加えることで空間モデルへの拡張を検討する。

4.1. 階層ベイズ線形回帰モデルによる定式化

市町村内総生産成長率を以下のようなモデルで表現する。

$$Y_{i,t} = \beta_{0,i} + x_{i,t}^1 \beta_{1,i} + x_{i,t}^2 \beta_{1,i} + \cdots + x_{i,t}^P \beta_{P,i} + c_{i,t}^1 \lambda_{i,1} + c_{i,t}^2 \lambda_{i,2} + \cdots + c_{i,t}^J \lambda_{J,1} + \varepsilon_{i,t} \quad (2)$$

なお誤差項は正規分布に従い、

$$\varepsilon_{i,t} \sim N(0, \sigma_y^2) \quad (3)$$

とする。ここで $y_{i,t}$ は目的変数である市町村内総生産成長率である。 $x_{i,t}^1, x_{i,t}^2, \dots, x_{i,t}^P$ は除雪事業に直接関係する説明変数のスカラーであり、 i は市町村 ($i = 1, \dots, 35$)、 t は年度 ($t = 1, \dots, 10$)、 P は表2の左列に列挙された変数から選択された数である。 $c_{i,t}^1, c_{i,t}^2, \dots, c_{i,t}^J$ はマクロ変数のスカラーを示しており、 J は表2の中央列に列挙された変数から選択された数である。除雪事業に直接関係する説明変数の $P+1$ 次元回帰係数ベクトルである $\beta_i = (\beta_{0,i}, \beta_{1,i}, \dots, \beta_{P,i})^T$ については誤差項 η_i が多変量正規分布に従うと仮定して以下のように表現する。

$$\beta_i = \delta^T \mathbf{z}_i + \eta_i, \quad \eta_i \sim MVN(\mathbf{0}, \Omega_\beta) \quad (4)$$

とする。 $\mathbf{z}_i = (1, z_{1i}, z_{2i}, \dots, z_{Li})^T$ は階層構造の説明変数ベクトルであり、 L は表2の右列に列挙された変数から選択された数である。また、 $\delta_p = (\delta_{0,p}, \delta_{1,p}, \dots, \delta_{L,p})^T$ は変数 p に対する $L+1$ 次元のベクトルであり、 $\delta = [\delta_0; \delta_1; \dots; \delta_P]$ は $(L+1) \times (P+1)$ の回帰係数行列である。 δ_p については多変量正規分布を仮定し、そのハイパーパラメータである平均ベクトル $\mu = (\mu_0, \mu_1, \dots, \mu_L)^T$ の個々の μ_l ; ($l = 0, 1, \dots, L$) には平均0、分散 σ_μ^2 である独立な正規分布を仮定する。

$$\delta_p \sim MVN(\mu, \Omega_\delta) \quad (5)$$

$$\mu \sim MVN(\mathbf{0}, \sigma_\mu^2 I) \quad (6)$$

一方, $\boldsymbol{\lambda}_i = (\lambda_{1,i}, \lambda_{2,i}, \dots, \lambda_{J,i})^T$ は, $c_{i,t}^1, c_{i,t}^2, \dots, c_{i,t}^J$ の回帰係数ベクトルである. これらのパラメータが β_i と相関をもつと想定するのは合理的ではないため, 本研究ではそれぞれ分散 $\sigma_{\lambda,j}; (j = 1, 2, \dots, J)$ を持ち, 独立に正規分布に従う仮定する.

$$\boldsymbol{\lambda}_i \sim MVN(\mathbf{0}, \sigma_{\lambda,j}^2 I) \quad (7)$$

4.2. 分散に対する事前分布の設定

式 (8) は式 (3) の σ_y , 式 (9) は式 (6) の σ_μ , 式 (10) は式 (7) の $\sigma_{\lambda,j}; (j = 1 \sim L)$ に対する事前分布 (一様分布) を示し, 式 (11) は式 (4) の Ω_β , 式 (12) は式 (5) の Ω_δ の事前分布 (逆ウィッシュャート分布) を示す. なお, 自由度の設定にあたっては, Rossi et al.[13] 参考にしている.

$$\sigma_y \sim uniform(0, 100) \quad (8)$$

$$\sigma_\mu \sim uniform(0, 100) \quad (9)$$

$$\sigma_{\lambda,j} \sim uniform(0, 100) \quad (10)$$

$$\Omega_\beta \sim IW(P + 3, (P + 3)I_P) \quad (11)$$

$$\Omega_\delta \sim IW(L + 3, (L + 3)I_L) \quad (12)$$

4.3. 空間に関する変数

階層ベイズ線形回帰モデルに追加する空間構造の説明を行う. 式 (1) に地域連関を表す変数 ϕ_i を追加し,

$$Y_{i,t} = \beta_{0,i} + x_{i,t}^1 \beta_{1,i} + x_{i,t}^2 \beta_{1,i} + \dots + x_{i,t}^P \beta_{P,i} + c_{i,t}^1 \lambda_{i,1} + c_{i,t}^2 \lambda_{i,2} + \dots + c_{i,t}^J \lambda_{J,1} + \phi_i + \varepsilon_{i,t} \quad (13)$$

とする. ϕ_i に条件つき正規自己回帰モデルを当てはめる. ただし, 市町村 i に近接行列 W を考え, 条件付き分布 $p(\phi_i | \phi_k, k \neq i)$ に正規分布を仮定し, 平均と分散をそれぞれ,

$$E(\phi_i | \phi_k, k \neq i) = \frac{\sum_k W_{ik} \phi_k}{\sum_k W_{ik}} \quad (14)$$

$$Var(\phi_i | \phi_k, k \neq i) = \frac{\tau}{\sum_k W_{ik}} \quad (15)$$

とすると, これはインプロパーな同時分布をもつ Intrinsic Autoregressive モデル (IAR モデル) となる. なお, インプロパーであっても事後分布がプロパーになれば良く, 事前分布として使用可能である. 分散については, Best et al.[14] を参考に $\tau \sim IG(0.1, 0.1)$ に従うとする. 式 (14), (15) における隣接行列 W を作成するにあたっては, Appendix B に記載した表 5 のとおり各市町村の隣接関係を定義する. なお, この隣接関係を定義する際には, 経済的なつながりや地理的な条件を加味しており, 行政区域の隣接関係と厳密に一致するものではない.

5. 実証分析

5.1. 推定方法

事後分布の推定には, マルコフ連鎖モンテカルロ法を使用する. 具体的には, R 言語 (Version 3.2.5), JAGS (Just Another Gibbs Sampler, Version 4.2.0) および WinBUGS (Version 1.4) を使用し推定を行う.

5.2. 計算条件

マルコフ連鎖モンテカルロ法による繰り返し計算を15万回行い、10回に一度サンプルを抽出することで1万5千個のサンプルを得た。このうち最初の5千を初期値の影響が強いサンプルとしてバーンインする。空間構造を導入したモデルについては、事前の分析からより長いバーンインが収束に必要であることが分かったため、繰り返し計算を17万回行い、10回に一度サンプルを抽出して1万7千個のサンプルを得たのち、最初の7千を初期値の影響が強いサンプルとしてバーンインする。

表 3: モデル選択

NO	DIC	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$c^{(1)}$	$c^{(2)}$	$c^{(3)}$	$c^{(4)}$	$c^{(5)}$	$c^{(6)}$	$c^{(7)}$	z_1	z_2	z_3	z_4	ϕ
1	2058	○	○	○			○	○	○		○						
2	2056	○	○	○			○	○	○		○		○	○			
3	2099	○	○	○	○		○	○	○		○		○	○			
4	2055	○	○	○		○	○	○	○		○						
5	2054	○	○	○		○	○	○	○		○		○	○			
6	<u>2051</u>	○	○	○		○	○	○	○	○	○		○	○			
7	2054	○	○	○		○	○	○	○		○	○	○	○			
8	2052	○	○	○		○	○	○	○	○	○	○	○	○			
9	2060	○	○	○		○	○	○	○	○	○		○	○	○		
10	2058	○	○	○		○	○	○	○	○	○		○	○		○	
11	2141	○	○	○		○	○	○	○	○	○		○	○			○
12	2149	○	○	○		○	○	○	○	○	○	○	○	○			○

5.3. モデル選択

市町村内総生産成長率と除雪費の関係を分析するため、第4章に記載済みの階層ベイズ線形回帰モデルの枠組みをもとに、表2の説明変数を組み合わせて複数のモデルを作成する。除雪費に欠損値が存在するが、平均値代入を行い推定を行った。ここでの平均とは、欠損した市町村の除雪費平均に年度平均に対する欠損年度の比を掛けた値を指す。なお、平均値代入は分散を過小に評価することが知られている [15]。この問題に対応するため、解釈の対象とする事後分布の推定を行うにあたっては多重代入法を使用した欠損値補完を行う。しかし多重代入法を使用してマルコフ連鎖モンテカルロ法を実行する場合、非常に大きな計算コストが要求されるため、モデル選択については多重代入法は行わない。モデル選択基準にはDIC (deviance information criterion) を使用する。各モデルで使用された説明変数と、そのモデルより計算されたDICは表3のとおり。変数名は表2に従う。なお、空間モデルについては、階層ベイズ線形回帰モデル (No.1~10) で最もDICが低い二つのモデルに空間構造を導入する。

いずれのモデルにおいてもゲルマン・ルービンの収束診断 (Gelman and Rubin's convergence diagnostic) とトレースから収束が確認できた。表3より、No.6のモデルが選択された。以降はこのモデルをもとに議論を進める。

5.4. 推定結果および考察

一般的に有意性検定では、有意水準5%あるいは10%を設定することで信頼区間を作成する場合が多い。一方、本研究ではパラメータを確率変数と仮定するベイズ推定を行っており、マルコフ連鎖モンテカルロ法の結果から得られた最高確率密度区間 (Highest density interval:HDI) を使用することで、パラメータがある確率で取りうる範囲を直感的に把握す

ることができる。図1に除雪費当期の係数 $\beta_{1,i}$, 図2に除雪費一期前の係数 $\beta_{2,i}$ のHDIを示す。ドットが90%HDI, 直線が80%HDIである。なお解釈にあたっては, 除雪費の単位は百万円であり, 各市町村平均でセンタリングしていることに注意が必要である。

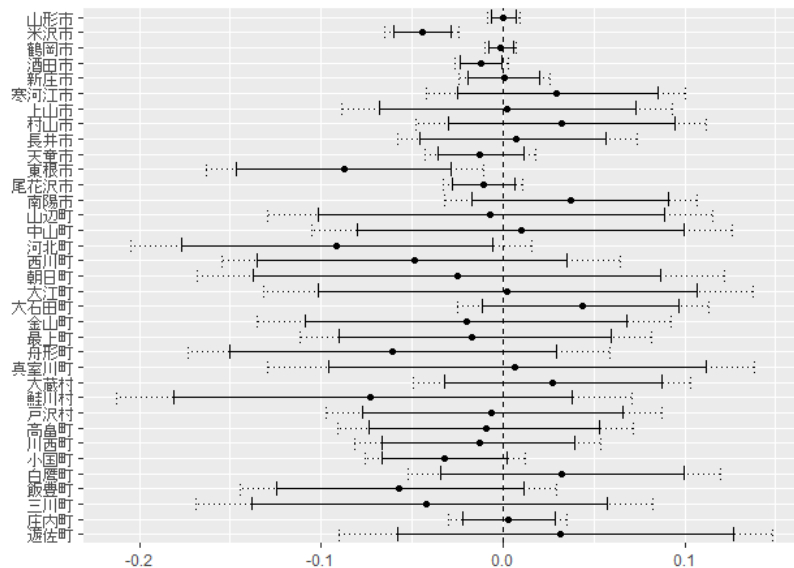


図 1: $\beta_{2,i}$ (除雪費当期)の HDI

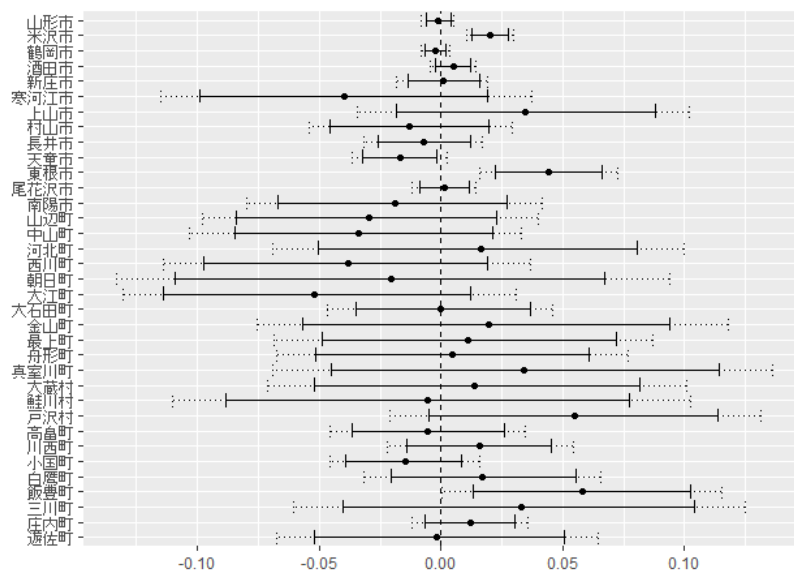


図 2: $\beta_{3,i}$ (除雪費一期前)の HDI

まず除雪費当期の HDI から見ると, 寒河江市, 村山市, 南陽市, 大蔵村, 大石田町では比較的 HDI が 0 を上回る値に偏っているものの, 高い確率で正の影響を与えると判断できる市町村はない。一方, 米沢市, 東根市では 90%以上, 酒田市, 河北町では 80%以上の確

率で負の影響を与えると判断できる。ところが除雪費一期前においては、当期では負の影響を与えると強く判断できる米沢市、東根市の係数が正にふれている。これは前述のとおり、当該市町村の精算方法により次期の出納閉鎖期間（～5月末）までに支出が行われたのか、あるいは地域性によって波及が遅延したものと考えられる。たとえば、冬季間に除雪事業が行われ、除雪車の燃料代など地域内で生産できない製品あるいはサービスへの支出されることによって資金が地域外に流出し、その後遅れて地域に投下された資金が循環し始めるなどの事情が想定できる。

以上のようなタイムラグを加味するために、除雪費当期と除雪費一期前のパラメータを統合した新しいパラメータ $\beta_{con,i}$ を以下のように作成した。

$$\beta_{con,i} = \alpha_i^{-1}(\beta_{1,i} + \beta_{2,i}) \quad (16)$$

なお、ここで α_i^{-1} は正規化項である。このHDIを図3のとおり得た。

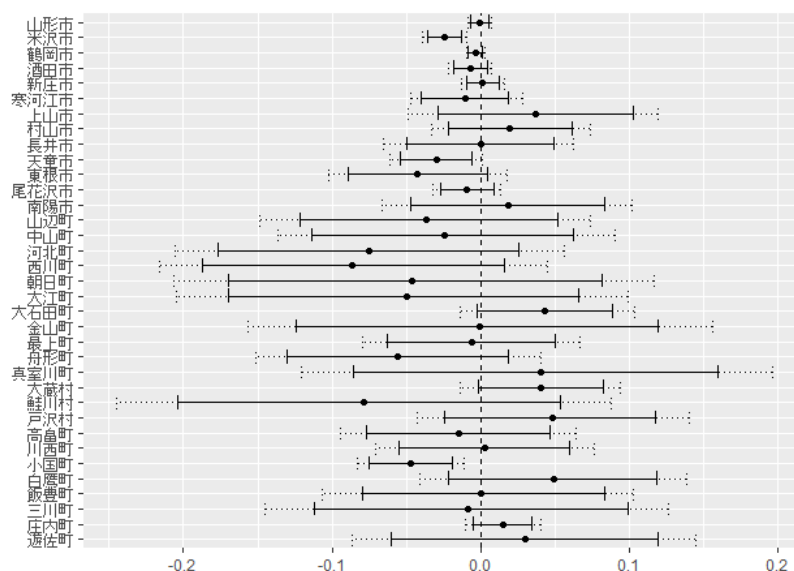


図 3: $\beta_{con,i}$ の HDI

図3において、多くの市町村では0を中心にHDIが広がっている。また全体として、比較的経済規模が大きな自治体についてはHDIが0付近に狭い幅をとる傾向がある。これは地域経済が重層化していることにより、除雪費の大部分が支払われる建設業の動向が地域全体を左右しにくいためと考えられる。一部の市町村においては明確にHDIの偏りが存在している。例えば90%HDIを見ると山形市の下限と米沢市の上限が重なっておらず、市町村ごとに異なる影響力をもつことが明らかである。この傾向は各市町村の80%HDIを確認するとより明瞭に把握できる。また個別に市町村の推定結果をみると、米沢市、小国町が90%以上の確率で、天童市が80%以上の確率で負の影響を与える一方、大蔵村、大石田町など、80%以上ではないものの、正の影響を与える可能性が高い市町村があると判断できる。表4は表1に対する回答をまとめたものである。

表 4: 検証結果

-
1. 除雪費の支出が地域経済に影響を与えると判断できる市町村がある
 2. その影響は市町村ごとに異なる
-

5.5. 推定結果の活用方法

上記結果が示す行政意思決定上の意義は、市町村自治体が除雪事業の再設計を行う際、客観的な検討材料を提供することである。ある地域において今後の人口減少などの社会情勢の変化を踏まえ、除雪サービスの持続可能性をいかに担保するために除雪事業の縮減やインフラの撤退などを検討する必要があるという状況を設定する。このとき問題になるのは住民の理解である。一般的に行政サービスを集約削減していくコンパクトシティ的施策は、計画実施の上で住民やあるいはその代表である議員の反発を招きかねず、実際にこれらはコンパクトシティの推進阻害要因として自治体の政策担当者に認識されている [16]。議論の際、本研究結果のようなエビデンスを土台に据えることは、自治体・住民間のコミュニケーションを円滑なものにすると考えられる。また、地域ごとの特徴を明らかにすることで、各地域に適した効果的な除雪サービスを議論することができる。

6. まとめと今後の展望

本研究では、山形県 35 市町村が支出する除雪費と地域経済の関係を分析した。分析にあたっては、階層ベイズ線形回帰モデルを適用した。結果として、除雪費は地域経済に影響を与えると判断できる市町村があり、またその影響力は地域によって異なることが分かった。

使用できるデータの制限により、本研究では地域異質性の詳細を分析することはできなかったが、今後詳細な市町村ごとのデータ整備・公開が進行していくことによって検討ができるようになる。つまり、本研究で構築したモデルをもとに、 β を目的変数としたモデルに、新しい z_i を説明変数として追加していくことで、 β の変動を理解することが可能である。新しい変数 z_i の例としては、除雪事業の運営を直営方式にするのか民間への委託方式にするかという判断や、消雪・融雪施設が整備された路面距離などが考えられる。特に z_i の中で介入可能なものはその推定結果を確認することで、政策化につなげることができる。

7. Appendix

7.1. Appendix A: データの詳細な説明

7.1.1. 除雪費当期

除雪費に該当する金額が一般的な道路維持にかかる歳出に混入している場合は、担当者に聞き取りを行い、できる限り除雪事業に関する金額を抽出した。単位は百万円であり、各市町村平均でセンタリングを行う。なお、このデータを基に作成した階層構造で使用する説明変数、除雪費平均においても全市町村平均でセンタリングを行っている。除雪事業の定義は各団体により多少異なるので、ここでは道路および施設の除雪事業に要した費用に限定している。なお、「除雪事業」には雪を河川敷などに雪を運搬し廃棄する、いわゆる「排雪」も含まれている。公共事業であることから同時性による内生性についての懸念があるが、除雪事業は災害対応活動の側面を持ち、降雪が多ければ、経済の状況（あるいはそれにとまなう

団体の予算規模)に関わらず除雪事業を行わなければならないという点を踏まえれば、地域の経済状況に左右されることは少ないと考えられる。なお、市町村合併などの理由により、山形市、鶴岡市、酒田市、寒河江市、庄内町、戸沢村において欠損値が存在している。欠損値に対する対応は7.2節に示す。

7.1.2. 降雪量

山形市、米沢市、新庄市、尾花沢市、長井市、小国町、金山町は気象庁の観測データを使用している。各市町村平均でセンタリングを行っている。なお、このデータを基に作成した階層構造で使用する説明変数、降雪量平均においても全市町村平均でセンタリングを行っている。気象庁の観測所が設置されていない市町村については、山形県道路保全課より提供を受けた観測データから主に平地で観測されたものを使用している。なお、山形県道路保全課のデータでは同一の市町村に複数の観測地点が存在するが、平地での観測点に限定し平均をとっている。

7.1.3. 従業者数成長率

鯉江 [11] はこの変数を内生変数として扱っている。つまり、この想定の上でモデルに組み込めば、同時性バイアスの発生が危惧される。しかしながら、工業統計調査における従業者数は製造業従事者の人数であり、製造業において産出される財の消費が産出した地域で行われる量は少ないことから、製造業従事者は地域経済より日本全体、あるいは世界経済の状況に影響を受けると考えるほうが合理的である。以上より、本研究では従業者数を外生変数と仮定する。

7.1.4. 一般道路距離

出典は平成24年度版山形県統計年鑑。

7.1.5. 市町村面積

出典は平成24年度版山形県統計年鑑。

7.2. Appendix B:隣接行列

表 5: 各市町村の隣接関係

市町村名	隣接市町村
山形市	上市市, 天童市, 山辺町, 中山町
米沢市	高畠町, 川西町, 飯豊町
鶴岡市	酒田市, 西川町, 三川町, 庄内町
酒田市	鶴岡市, 真室川町, 鮭川村, 戸沢村, 三川町, 庄内町, 遊佐町
新庄市	金山町, 最上町, 舟形町, 大蔵村, 鮭川村, 戸沢村
寒河江市	村山市, 天童市, 中山町, 河北町, 西川町, 大江町, 大蔵村
上市市	山形市, 南陽市, 高畠町
村山市	寒河江市, 東根市, 尾花沢市, 河北町, 大石田町, 大蔵町
長井市	南陽市, 朝日町, 川西町, 小国町, 白鷹町, 飯豊町
天童市	山形市, 寒河江市, 東根市, 中山町, 河北町
東根市	村山市, 天童市, 尾花沢市, 河北町
尾花沢市	村山市, 東根市, 大石田町, 最上町, 舟形町
南陽市	上市市, 長井市, 高畠町, 川西町, 白鷹町
山辺町	山形市, 中山町, 朝日町, 白鷹町
中山町	山形市, 寒河江市, 天童市, 山辺町, 朝日町, 大江町
河北町	寒河江市, 天童市, 村山市, 東根市
西川町	鶴岡市, 寒河江市, 朝日町, 大江町, 大蔵村
朝日町	長井市, 山辺町, 中山町, 西川町, 大江町, 白鷹町
大江町	寒河江市, 中山町, 西川町, 朝日町
大石田町	村山市, 尾花沢市, 舟形町
金山町	新庄市, 真室川町
最上町	新庄市, 尾花沢市, 舟形町
舟形町	新庄市, 尾花沢市, 大石田町, 最上町, 大蔵村
真室川町	酒田市, 金山町, 鮭川村
大蔵村	新庄市, 寒河江市, 村山市, 西川町, 舟形町, 戸沢村, 庄内町
鮭川村	酒田市, 新庄市, 真室川町, 戸沢村
戸沢村	酒田市, 新庄市, 大蔵村, 鮭川村, 庄内町
高畠町	米沢市, 上市市, 南陽市, 川西町
川西町	米沢市, 長井市, 南陽市, 高畠町, 飯豊町
小国町	長井市, 飯豊町
白鷹町	長井市, 南陽市, 山辺町, 朝日町
飯豊町	米沢市, 長井市, 南陽市, 川西町, 小国町
三川町	鶴岡市, 酒田市
庄内町	鶴岡市, 酒田市, 大蔵村, 戸沢村
遊佐町	酒田市

7.3. Appendix C:欠損値への対応

除雪費について、表6のとおり山形市、鶴岡市、酒田市、寒河江市、庄内町、戸沢村において欠損が生じている。

表 6: 除雪費の欠損値; $n = 17$

市町村名	欠損年度
山形市	2002 ~ 2004
鶴岡市	2002 ~ 2004
酒田市	2002 ~ 2004
寒河江市	2002 ~ 2005
庄内町	2002 ~ 2004
戸沢村	2002

これら欠損の原因は明確であり、市町村合併以前のデータの統合できないこと、除雪に関する歳出が完全に他の土木工事と混在し分離できないこと、データの消失などである。いずれも「市町村」と「年度」という観測値に依存し、除雪費という欠損値に依存しないランダムな欠測 (MAR) に該当する。欠損の割合は約 4%である。

多重代入法を実行するにあたっては、R言語のパッケージmiceを使用する。miceは完全条件付き分布の指定（full conditional specification;FCS）法の一つである、連鎖式による多重代入（multiple imputation by chained equation;MICE）アルゴリズムを実行する。代入ステップにおいて使用する共変量の選択に当たっては、欠損に関連し欠損の起きる変数と相関のある変数が必須であり、また欠損には関連しないが欠損のある変数が存在し、相関のない変数が存在しないことが望ましいとされている[17]。これより、表2に挙げた変数（ただし、ダミー変数と除雪費当期から作成される除雪費一期前、除雪費平均、および降雪量から作られる除雪費一期前、降雪量平均は除く）に加え、各市町村ダミーと各年度ダミーを説明変数として使用したモデルを作成し、事前にLasso回帰とCross-validationを行うことで変数選択を行う。なお、Lasso回帰の結果、欠損した市町村、年度のダミー変数の係数パラメータが0になってしまった場合でも、欠損の関連していると思われ、再度代入モデルに投入する。結果として代入に使用した共変量は表7のとおりであり、

表 7: 代入ステップで使用する共変量

欠損に関連し欠損の起きる変数と相関のある変数	欠損には関連しないが欠損のある変数と相関のある変数
各市町村ダミー（ただし三川町・南陽市・遊佐町を除く） 各年度ダミー（ただし2012年を除く）	市町村道距離 市町村面積 市町村内総生産成長率 人口成長率 従業者数成長率

欠損に関連し欠損の起きる変数と相関のある変数、および欠損には関連しないが欠損のある変数と相関のある変数に分け示している。ただし、ここで使用する共変量は成長率に変換しているため、2002年度分のデータが存在しない。よって実際補完されるのは除雪費のみではなく、市町村内総生産成長率、人口成長率、国内総生産成長率、従業員成長率も同様である。具体的に代入に使用した手法はpredictive mean matching(pmm)である。これにより作成した20個の疑似完全データセットに対してそれぞれ推定を行い結果を統合した。

参考文献

- [1] 国土交通省, 「豪雪地帯対策の推進」, http://www.mlit.go.jp/kokudoseisaku/chisei/crd_chisei_tk_000010.html (2016年12月31日閲覧)
- [2] 山形県, 「除雪出動基準」, <http://www.pref.yamagata.jp/ou/kikakushinko/020024/yuki/yuki/jyosetukijyun.html> (2016年12月31日閲覧)
- [3] A. Labelle, A. Langevin, and J.F. Campbell, “Sector design for snow removal and disposal in urban areas”, *Socio-Economic Planning Sciences*, **36**, pp.183-202, 2002.
- [4] Gang Liu, Yongfeng Ge, Tony Z. Qju, and Hamid R. Soleymani, “Optimization of snow plowing cast and time in an urban environment : A case study for the City of Edmonton”, *TRB 92nd Annual Meeting Compendium of Papers*, pp.667-675, 2014.
- [5] 柿崎敦郎・松尾徳朗, “協調的エージェントに基づく制約付きタスク問題解決”, 情報処理学会研究報告, 2010.
- [6] 田辺慎太郎・原文宏・下僚晃裕・高木秀賞, “除雪の効果に関する考察”, 第十三回寒地技術シンポジウム, pp.644-649, 1997.

- [7] R.J. Barro, “Economic Growth in a Cross Section of Countries”, *The Quarterly Journal of Economics*, **Vol. 106**, No. 2, pp.407-443, 1991.
- [8] 唐木芳博・奥原崇・渡真利諭・朝日ちさと・西畑知明, “社会資本ストックの経済効果に関する研究—都市圏分類による生産力効果と厚生効果—”, 国土交通政策研究, 第 68 号, 2006.
- [9] 田邊慎太郎・林山泰久・原文宏, “除排雪事業が冬期都市環境に及ぼす Option Price の計測”, 環境システム研究, **Vol.27**, pp.45-56, 1999.
- [10] 浜田浩児・堀雅博・花垣貴司・横山瑠璃子・亀田泰佑・岩本 光一郎, “短期日本経済マクロ計量モデル(2015年版)の構造と乗数分析”, 内閣府経済社会総合研究所, 2015.
- [11] 鯉江康正, “長岡市の社会経済の将来像”, 地域研究, 第 10 号, pp.63-70, 2010.
- [12] S. Jackman, “Bayesian Analysis for the Social Sciences”, *Hierarchical Statistical Models*, John Wiley & Sons, pp.302-303, 2009.
- [13] P.E. Rossi, G.M. Allenby, and R. McCulloch, “Bayesian Statistics and Marketing”, *Bayesian Essentials*, John Wiley & Sons, pp.30-31, 2005.
- [14] N.G. Best, R.A. Arnold, A. Thomas, L.A. Waller, and E.M Conlon, “Bayesian Models for Spatially Correlated Diseases and Exposure Data”, *Bayesian Statistics*, **6**, pp.131-156, 1999
- [15] 高井啓二・星野崇宏・野間久史, 『欠測データの統計科学——医学と社会科学への応用——』はじめに：欠測の *Do's and Don'ts* とガイドライン, 岩波書店, pp12-13, 2016.
- [16] 越川知紘・菊池雅彦・谷口守, “コンパクトシティ政策への受容性の変化と障害意識—自治体の都市計画担当者を対象として—”, 土木計画学研究・講演集, 第 54 号, pp.2197-2204, 2010.
- [17] 高井啓二・星野崇宏・野間久史, 『欠測データの統計科学——医学と社会科学への応用——』単一代入と多重代入, 岩波書店, pp127, 2016.