

道路アクセスと市町村内総生産の 因果状態の探索

塚井 誠人¹・杉原 豪²

¹正会員 広島大学大学院准教授 先進理工系科学研究科 (〒739-8527 広島県東広島市鏡山 1-4-1)
E-mail: mtukai@hiroshima-u.ac.jp

²学生会員 広島大学大学院 先進理工系科学研究科 (〒739-8527 広島県東広島市鏡山 1-4-1)
E-mail: m194392@hiroshima-u.ac.jp

経済モデルでは、地域間アクセスの改善が均衡状態の変化を介して生産性の向上につながるものが仮定されているが、DID法以外で実際の経済活動がその仮定を満たすことを検証する手法の開発は十分ではない。本研究では、LiNGAMを用いて、全国道路アクセスと生産活動の因果状態を探索する。国内の全市町村の2時点約3400サンプルを用いたデータから道路アクセス、GRP及び主な交絡変数を作成して検証を行った。その結果、「社会経済変数→アクセス指標・経済パフォーマンス変数」「アクセス指標→経済パフォーマンス変数」という因果構造が得られ、社会基盤の整備と経済成長の間に正の因果関係が存在することを支持した。推計結果の安定性を確認するため部分データセットによる検証を行ったところ、因果構造に多少の違いが現れたが、おおむね上記の結論と合致する結果が得られた。

Key Words : causal relationship, economic growth, confounding factor, socioeconomic variables, economic performance variables

1. はじめに

1990年代以降の社会基盤整備の可否は、費用便益分析に基づいて行われてきた。標準的な費用便益分析の手順は、経済学的な考察に基づく因果構造を前提として、統計学的な仮定を付加して、社会基盤整備後に発生する便益を算出するというものであり、手続き合理的な手法となっている。ただし実際には事業効果の事前予測値と事後観測値の間では、常に乖離が生じる。

平成 28 年 11 月に国土交通省が策定した「ストック効果の最大化に向けて～その具体的戦略の提言～」¹⁾では、社会基盤整備のストック効果の最大化のために、ストック効果に関して得られる知見を事業に有効活用する社会基盤のマネジメントサイクルの確立の必要性が述べられている。しかし予測の基盤となる経済モデルのロジックやモデル構造の妥当性、あるいはその限界が再検討されることは、これまでほとんどなかった。これを実現するには、経済モデルが想定した因果構造の妥当性を、事後的に検証する分析手法が必要である。具体的には、1)経済モデルによる将来便益の算出、2)実証モデルから得られる知見の照査と、3)両モデルの便益・因果構造の比較、という一連のプロセスが必要である。

本研究では、特に上述の2)に関して道路整備のストッ

ク効果に着目し、個別地域・主体での社会基盤の整備と経済成長の因果関係を、それ以外の関連する交絡因子との間の因果構造を含めて明らかにすることを目的とする。人口や域内総生産、ある地域のアクセス性を表すアクセス指標等の6変数について、変数間の因果関係を明らかにするLiNGAMによる推計を行う。推計結果を基に、実証モデルから経済理論が想定する因果構造を再検証する。加えてデータセットを分割した推計を行うことで、LiNGAMの示す性質について考察する。

2. 既往研究と本研究の位置づけ

社会基盤整備の経済効果は、主に建設投資によって発生するフロー効果と、社会基盤供用後にその利用またはその社会基盤が提供するサービスや状態がもたらすストック効果の2種類に分けられる。このうち、費用便益分析においてフロー効果は単純な所得移転にすぎないため便益として計上しないため、問題となるのは後者の効果である。道路整備の費用便益分析では、走行時間短縮便益、走行費用削減便益、および交通事故減少便益の三便益を効果として計上する²⁾。金本³⁾によると、便益の計測手法には、大別して発生ベースと帰着ベースの2種類

がある。発生ベースのアプローチでは、社会基盤整備の便益を、それが発生する状況を想定して計測する。帰着ベースのアプローチは、その便益の発生主体ごとに発生量を計測する。しかし帰着ベースの便益推定はその精度に大きな問題があり、推定者によるモデル形式やパラメータ値の設定による便益推計のばらつきが非常に大きい。このため、実務的に費用便益分析は、発生ベースで行うこととされている。

発生ベースの費用便益分析は、仮説的補償原理によって正当化される。仮説的補償原理は、帰着ベースの総便益が正ならば、ある主体の損失を別の主体が得た便益で補償すれば、社会全体では正の便益が得られるという考え方に基づく。さらに発生ベースの便益と帰着ベースの便益の相等が満たされると前提すれば、発生ベースの便益が正の事業を採択する、という判断になる。

「発生ベースの便益と帰着ベースの便益の相等」は、標準的な経済モデルでは自明な命題である。標準的な経済モデルの想定は、予測期間中に産業の構造変化がないこと、技術革新がないこと、資源有限性がないこと、または独占、寡占といった市場支配がないこと、あるいはサンクコストが投資判断に影響しないという想定の下で導かれる。実際にはこれらが満たされることは稀だが、これを理由として費用便益分析そのものが否定されることはない。たとえ費用便益分析の予測性能に限界があるとしても、依然として手続き的に合理的な事業採択判断の規範を与える手法として、費用便益分析の有用性が高い、と考えられるためである。

近年ストック効果が注目される理由に関して小池²⁾は、仮説的補償原理の仮定から逸脱するような地域開発が実際にあり得る危険性を指摘した。また小池は、複数の社会基盤整備事業を比較、または優先/劣後を判断するのであれば、個別地域・主体に帰着する便益を明示する必要はないことを指摘した。その上で、個別地域・主体への帰着便益は、社会的に望ましい分配に関する計画論を議論する上で重要としている。さらに小池は、帰着便益を明示することによって、その限界や活用方法について技術者や政策決定者が、社会基盤の役割についてより深く議論することを強く推奨している。

費用便益分析の性能を事後的に検証する立場からは、実際に複数の観測データ間の因果関係を明らかにする手法が必要がある。そのような手法は統計的因果推論と呼ばれ、代表的な手法の一つに操作変数法がある。Duranton & Turner⁴⁾は、州間高速道路が 1983 から 2003 年のアメリカの都市の成長に与える影響について構造方程式を定式化して、操作変数法によってモデルを推定した。その結果、道路ストックの 10% の増加は、都市の雇用を 20 年間で約 1.5% 上昇させるという推定結果を得た。

Holt⁵⁾は、1997~2007 年のスペインの製造業に関わる企業を対象に、企業の生産性に対して高速道路へのアクセスが及ぼす影響について操作変数法を用いて分析を行い、高速道路によるアクセス性の向上は集積の効果以上に企業レベルの生産性を向上させたことを明らかにした。さらに、高速道路に近接する郊外地域の企業では生産性が大きく向上する一方で、その周辺地域の企業の生産性は低下するという、高速道路が生産性に及ぼす影響の空間的不均一性を示した。Redding and Turner⁶⁾は、交通基盤整備が経済活動の空間構造に及ぼす効果を、操作変数法などを用いて推定する研究についてレビューを行った。その結果、高速道路や鉄道までの距離が 2 倍になると、人口や雇用が 6-15% 低下すること、高速道路整備は都市人口や製造業の郊外化をもたらすこと、交通インフラの効果は分析の規模や対象都市の経済状況からは影響を受けないこと、および交通インフラの変化の影響は交通機関・産業によって様々であることなどを明らかにした。また織田澤ら⁷⁾は、社会経済統計などの観察データを用いて社会基盤整備評価を行う場合、交絡や逆因果性などの識別が困難なことを指摘した。

本研究は織田澤の指摘に対応可能な因果状態の探索手法によって、道路アクセスに関するストック効果を検証する。具体的には、LiNGAM によって国内の全市町村データから全国の道路アクセスと経済成長の因果関係を明らかにする。この手法に関する杉原の先行研究⁸⁾では、道路アクセスと地域生産額に関する統計的因果探索を試行して、両変数に関して正の因果関係が存在すること明らかにした。ただし同研究が対象としたサンプルは、中国地方の約 100 程度の市町村に留まっていたため、結果の妥当性に関する追検証が必要である。以上より本論では、LiNGAM (統計的因果探索) の推計の安定性の検証を踏まえて、全国の市町村を対象とした分析を試みる。

3. ストック効果の因果探索手法

本研究では、道路アクセス性と経済活動の因果関係を明らかにするため、統計的因果探索の中でも、各変数の同時分関数形に線形性、外生変数の分布に非ガウス分布をそれぞれ仮定し、因果グラフが非巡回で同一の観測変数について閉路構造を持たない LiNGAM (線形非ガウス非巡回モデル Linear Non-Gaussian Acyclic Model) アプローチ⁹⁾を用いる。

LiNGAM の推定原理は、まず変数間の因果的順序を求め、その因果的順序に従う回帰分析をパラメータの疎構造仮定の下で行って、複数の変数間の因果グラフを探索するという 2 段階で構成される。変数の因果的順序とは、

その順序に従って変数を並べ替えると、後の変数が先の変数の原因となることが無いような変数の順序である。因果グラフで表される構造方程式において、説明側（右辺）の変数を祖先、目的変数側（左辺）の変数を子孫と呼ぶ。また非ガウス分布は、「説明変数と残差が独立ならば、推計は正しい因果的順序で行われている」という Darmois-Skitovich の定理の対偶に基づく統計的性質を考慮するために必要な仮定である。

p 種の観測変数 x_1, x_2, \dots, x_p に関する LiNGAM は、式(1)で定式化される。

$$x_i = \sum_{j \neq i} b_{ij}x_j + e_i \quad (i = 1, \dots, p) \quad (1)$$

ここで、 b_{ij} は推定される係数、 e_i は誤差変数を表す。 x_i は、 i 以外の観測変数 $x_j (j = 1, \dots, j \neq i)$ とその重み b_{ij} の加重和に誤差変数 e_i を加えた線形和である。なお、誤差変数 e_i は独立で、非ガウス連続分布に従うと仮定する。行列を使うと、式(1)は、式(2)と表される。

$$\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{e} \quad (2)$$

\mathbf{x} , \mathbf{B} , \mathbf{e} は、それぞれ観測変数 x_i 、係数 b_{ij} 、誤差変数 e_i の行列またはベクトル表記である。 \mathbf{B} のゼロ・非ゼロパターンから、変数間の因果関係を表すグラフを作成する。

LiNGAM の推計は、清水⁹⁾の提案した独立成分分析 (Independent Component Analysis, 以下 ICA) アプローチに則って行う。以下に推計手順の概要を示す。

\mathbf{B} は非巡回の仮定より、対角成分が全て 0 かつ、非ゼロの要素は全て下三角側に限定される下三角行列（以下厳密な下三角行列）となったとき、観測変数は因果的順序に従って並んでいるといえる。この考えを基に、変数の因果的順序を得る。また、使用するアルゴリズムの特性上、変数順序の初期配置によって結果が変わりうる。この問題の理論的な回避は困難なため、以下では便宜的に変数順序の初期配置の候補全て (n 変数の場合、 $n!$ 通り) で推計して、出現頻度が最も多い因果的順序を採用する。

以上から求めた因果的順序を参照して、目的変数よりも因果的順序の早い（その変数の祖先にあたる）全変数を説明変数にとって、線形回帰モデルの推定を行う。ただし、最小二乗法のような通常回帰分析を適用しても、真値が 0 の係数 b_{ij} の推定値が厳密に 0 にはならない。その場合、因果グラフ内に統計的に有意ではないレベルの因果パスが現れてしまう可能性がある。そこで線形回帰モデルの推定と変数選択を同時に実行する Lasso(Least absolute shrinkage and selection operator)によるスパース回帰¹¹⁾

表-1 サンプル数 10000 の結果

因果的順序	出現回数
$x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4 \rightarrow x_5$	120

表-2 サンプル数 1000 の結果

因果的順序	出現回数
$x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4 \rightarrow x_5$	120

表-3 サンプル数 100 の結果

因果的順序	出現回数
$x_4 \rightarrow x_3 \rightarrow x_1 \rightarrow x_2 \rightarrow x_5$	49
$x_3 \rightarrow x_1 \rightarrow x_2 \rightarrow x_4 \rightarrow x_5$	25
$x_3 \rightarrow x_4 \rightarrow x_1 \rightarrow x_2 \rightarrow x_5$	18
$x_3 \rightarrow x_1 \rightarrow x_4 \rightarrow x_2 \rightarrow x_5$	9
$x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4 \rightarrow x_5$	5
⋮	⋮

を行う。これにより、因果グラフ中に現れる本来削除すべき冗長な有向辺を削除する。

以上の手順で LiNGAM の推計を行う。さらに、推計した係数行列 \mathbf{B} のゼロ・非ゼロパターンからは、因果グラフが描ける。なお、本章に示した手順から明らかのように、LiNGAM は純粹に統計的な手法である。このため本分析では、経済学的なインプリケーションを得るための作業仮説を、ストック効果に関連して設定する。その内容は、後述する。

4. 因果的順序の正しさに影響を与える要因

インフラストック効果の推計では、その効果が空間的にスピルオーバー¹²⁾する可能性がある。推計精度を向上させるにはできるだけ広い範囲から得た多くのサンプルを用いるべきである。また、推定される因果的順序の正しさに影響を与える要因として、サンプル数と変数中に現れる極端な外れ値の存在が挙げられる。以上の性質を確認するため、LiNGAM の仮定を全て満たし、因果的順序が明らかな擬似データを作成して、サンプル数を変えて因果的順序の推定を行う。設定した因果的順序と推計された因果的順序を比較して、サンプル数による因果的順序の推定の正しさに関する影響を明らかにする。

非ガウス分布に従う誤差変数 \mathbf{e} と、厳密な下三角行列として係数行列 \mathbf{B} を設定し、因果的順序の明らかな 5 変数の擬似データ \mathbf{X} を生成した。擬似データ \mathbf{X} の因果的順序は、祖先から順に x_1, x_2, x_3, x_4, x_5 と設定した。この擬似データからサンプル数 10000, 1000, 100 のデータセットをランダムに抽出して作成し、5 変数の全順序にあたる 120 通りの変数順序を初期配置として、推定、

表-4 使用データの詳細

年次*	平成 22 年, 平成 27 年
範囲	国内の約 1700 市町村
変数**	就業者数/国勢調査 人口/国勢調査 事業所数/経済センサス 総生産/県民経済計算 地価 (全用途) /一般財団法人 土地情報センター アクセス指標/NITAS

※観測年次が異なるデータは直近のデータとして平成 21 年, 平成 23 年, 平成 26 年, 平成 28 年の値を用いる
 ※※以下はデータの出典を表す

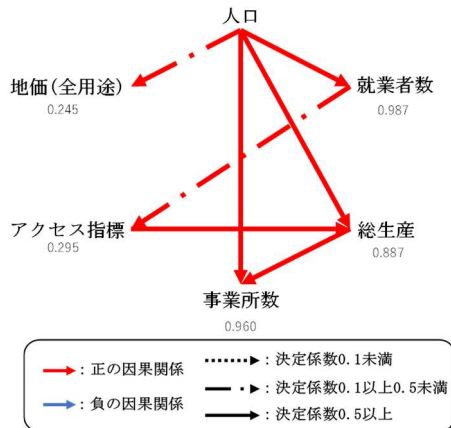


図-1 全データの因果グラフ

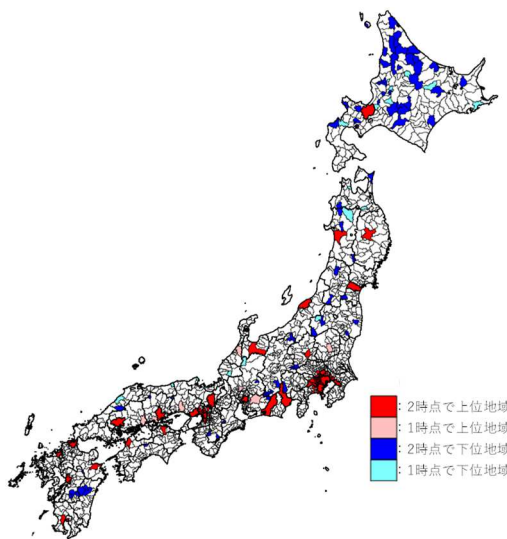


図-2 上位地域, 下位地域の分布

した因果的順序の出現回数を記録した。それらの結果をそれぞれ表-1~表-3 に示す。これらの表より、5 変数の因果的順序の推定には、1000 サンプル以上用意することが望ましいことが分かる。

また統計的因果探索はデータが示す分布特性に基づいて推測する手法であり、極端な外れ値を含む分布を示す変数の影響は無視しえない。なぜなら、尖度を大きく歪めるため結果に大きな影響を与えるからである。そこ

で本研究では、これらの外れ値となる可能性のあるサンプルを上位地域・下位地域と呼び、データから除外した中位地域のサンプルセットを用意して、そのデータセット下での因果探索についても検討する。

4. 全データが示す因果構造

(1) 使用データの概要と作業仮説

サンプル数の議論を踏まえて、本研究では国内の約 1700 市町村の、2 時点分の合計約 3400 サンプルを用いる。推計には表-4 に示す 6 変数を用いる。なお、地域 i のアクセス指標 AC_i は、以下の式(3)で定義する。

$$AC_i = \sum_{j=1}^n \frac{P_j}{t_{ij}^2} \quad (3)$$

ここで、 i : 対象地域 (市町村), j : 都道府県, t_{ij} : i から j への自動車所要時間, n : 地域数, P_j : 地域 j の人口である。また、 t_{ij} は起点を各市町村役場、終点を各県庁とした。なお、アクセス指標算定に用いた NITAS の仕様上、本州と北海道間等の自動車のみの移動が不可能な経路は所要時間が算出されないため、当該区間のアクセシビリティは算出対象外として扱った。つまり北海道の市町村は本州の市町村へのアクセシビリティを計上せず、逆に本州の市町村も北海道の市町村へのアクセシビリティを計上していない。なお計算負荷を軽減するため市町村 i と都道府県 j において異なる解像度を用いたため、都道府県内アクセシビリティは、以下のように算出された点に留意する必要がある。この場合は、県庁所在市でも県庁と市役所の場所が異なるため、両施設間の所要時間が非ゼロとなるため、自都道府県内の市町村と自都道府県のアクセシビリティは計上される。このとき人口 P_j は、あたかも県庁所在地に集中しているかのように扱われる点に注意が必要である。また平成 22 年と平成 27 年では、市町村合併等で市町村数が異なっており、サンプル数は平成 27 年は 1708 市町村、平成 22 年は 1705 市町村となった。

本章では、分析に用いる変数のうち就業者、人口、事業所を地域特性を表す社会経済変数、総生産、地価 (全用途) を経済パフォーマンス変数と分類したうえで、社会経済変数→経済パフォーマンス変数の因果の下で、アクセス指標→社会経済変数、またはアクセス指標→経済パフォーマンス変数の因果が得られる、という作業仮説の検証を行う。

市町村単位の総生産データは、一部の市町村で利用可能なデータが存在しなかった。そこで本研究では、面補

間法を用いて、都道府県単位の県内総生産のデータを市町村単位に分配することで推計した値を分析に用いることとした。この推計は、総生産を目的変数とする Cobb-Douglas 型生産関数に一次同次を仮定して行う。Cobb-Douglas 型生産関数は以下の式(4)で表される。

$$Y = AK^{\alpha_1}L^{\alpha_2} \quad (4)$$

ここで、 Y ：総生産、 A ：全要素生産性、 K ：資本ストック、 L ：労働力、 α_1 ：資本弾力性、 α_2 ：労働弾力性である。本研究では、資本ストックとして、経済センサスから全産業の事業所の売上金額に関する試算値を、労働力として従業者数を使用した。大ゾーン単位のデータ Y_L, K_L, L_L を用いて式(4)の対数をとって回帰分析によって、 $\hat{\alpha}_1, \hat{\alpha}_2, \hat{A}$ を推計し、 $\hat{\alpha}_1, \hat{\alpha}_2, \hat{A}$ を基に、式(5)に基づいて小ゾーン単位の総生産を推計する。

$$\hat{Y}_{S,i} = \frac{\hat{A}K_{S,i}^{\hat{\alpha}_1}L_{S,i}^{\hat{\alpha}_2}}{\sum_{j=1}^{n_R} \hat{A}K_{S,j}^{\hat{\alpha}_1}L_{S,j}^{\hat{\alpha}_2}} Y_{L,R} \quad (5)$$

ここで、 $\hat{Y}_{S,i}$ ：小ゾーン i の総生産の予測値、 $K_{S,i}$ ：小ゾーン i の民間資本ストック、 $L_{S,i}$ ：小ゾーン i の従業者数、 R ：小ゾーン i が所属する大ゾーン、 n_R ：大ゾーン R に所属する小ゾーン数、 $Y_{L,R}$ ：大ゾーン R の総生産の観測値である。なお式(5)は、 $\hat{\alpha}_1, \hat{\alpha}_2, \hat{A}$ から推計される小ゾーン i の予測値の、大ゾーン R の予測値に対する割合に基づいて大ゾーン R の総生産の観測値を分配しており、明らかに体積保存則を満足する。以上で作成した総生産予測値を推計データとして使用し、LiNGAM を適用する。

(2) 全市町村データの道路アクセスのストック効果

推定された因果的順序の中で、出現回数が最多の結果を正しい因果的順序として採用し、有効な因果パスを描くために Lasso を適用する。推定結果に基づいて作成したグラフを、図-1 に示す。同図は因果グラフと呼ばれる。矢印は因果を表しており、特に矢印の出る変数は原因、入る変数は結果を示す。図中の赤矢印は正、青矢印は負の効果を、それぞれ示す。なお、また、変数の下には決定係数を示す。図-1 より「人口→総生産」「人口→地価（全用途）」のパスが見られる。以上の関係は、「社会経済変数→経済パフォーマンス変数」という因果仮説を支持する。アクセス指標の直接的な影響に着目すると、従業者数から正の影響を受け、総生産に正の影響を与えている。前者はアクセス性が従業者という社会経済変数によって定まる内生変数であることを示している。一方後者は、「アクセス指標→経済パフォーマンス変数」という因果構造を示している。

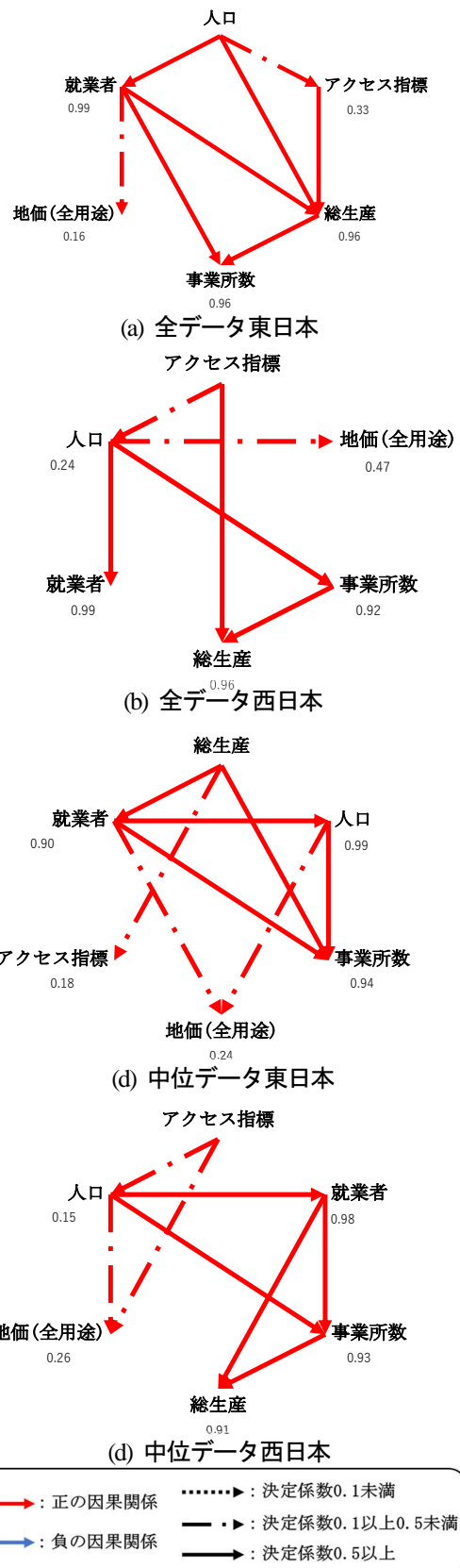


図-3 地理的に分割した因果グラフ

以上より、全データが示す因果構造は、「社会経済変数→アクセス指標・経済パフォーマンス変数」「アクセス指標→経済パフォーマンス変数」であった。また「人

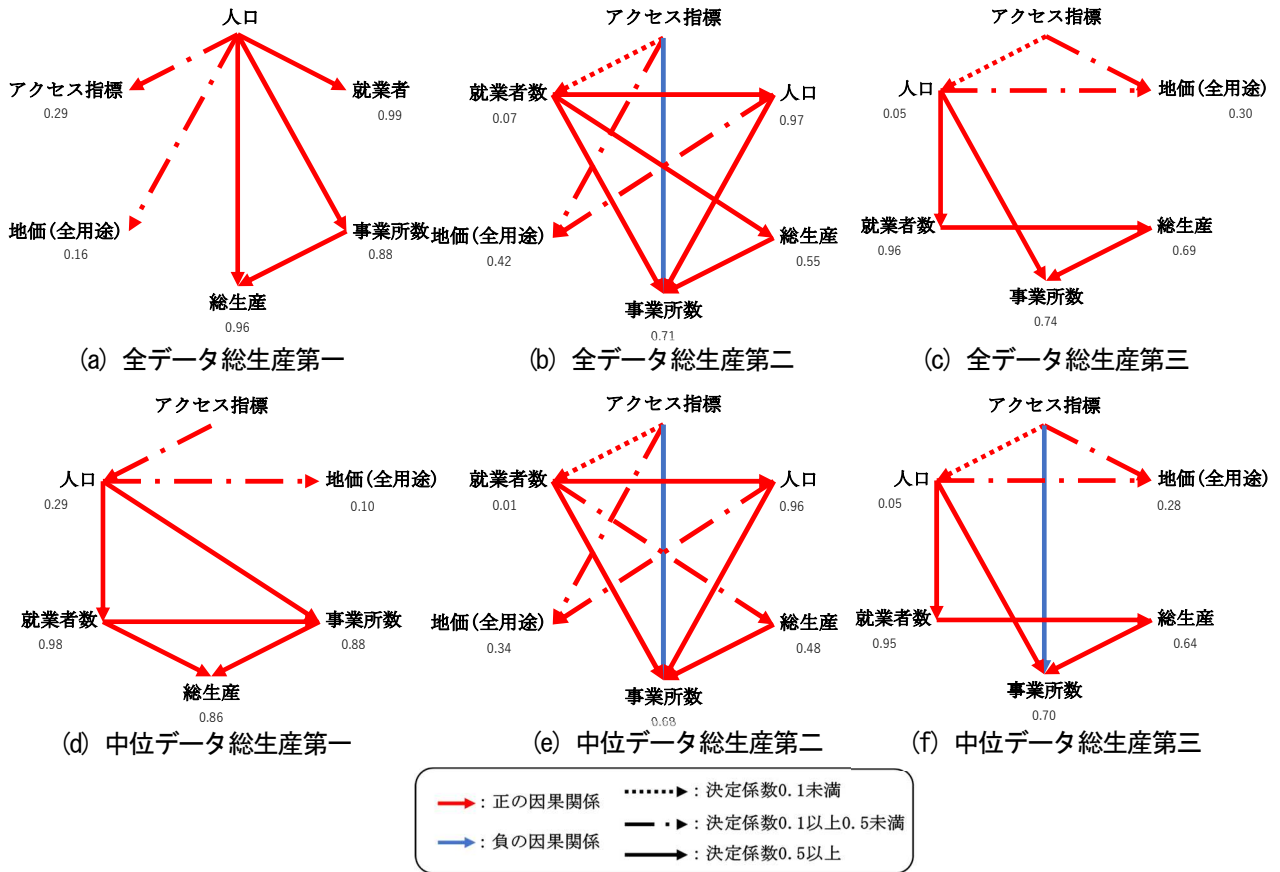


図-4 域内総生産で分割した因果グラフ

「人口・アクセス指標→総生産」となった。すなわち、道
 アクセシビリティが高い地域ほど、地域内総生産が高い傾
 向があることが明らかとなり、道路インフラについて正
 のストック効果が存在することが、支持された。

5. 部分データが示す因果構造

実務的な関心および、本手法の性質を明らかにするた
 めに全国から一部サンプルを抜粋して分析を行い、4.
 全データの示す因果構造との差異の有無を確認する。中
 位データセットと全データそれぞれで、域内総生産で 3
 分割したサンプルセット、および地理的に東西に分割し
 た計 10 種のサンプルセットを作成し、それぞれの因果
 構造を推定して、相互の異同を明らかにする。その上で、
 LiNGAM の運用上の留意点について考察する。

(1) 中位データセットの作成

本節では、まず上位地域、下位地域を定義する。ある
 地域 i の変数 j x_{ij} の、分布中の外れ値の程度指標を
 $Ext(x_{ij})$ と表記し、以下の式(6)で定義する。

$$Ext(x_{ij}) = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (6)$$

ここで、 \bar{x}_j は変数 x_j の平均値、 s_j は変数 x_j の標準偏差で
 ある。 $Ext(x_{ij})$ を全地域 i および全変数 j について計算し、
 各地域 i について、 $U(x_i)$ と $L(x_i)$ を式(7)、式(8)で求める。

$$U(x_i) = \max_j \{Ext(x_{ij})\} \quad (7)$$

$$L(x_i) = \min_j \{Ext(x_{ij})\} \quad (8)$$

各時点の $U(x_i)$ から降順に上位 100 地域、 $L(x_i)$ から昇順
 に下位 100 地域を、それぞれ上位地域、下位地域と定義
 する。上位地域、下位地域の分布を図-2 に示す。図-2
 から、上位地域は政令指定都市を中心に都市部及びその
 周辺に、下位地域は地方部に主に分布し、ほとんどが 2
 時点で共通していることが分かる。以降では上位、下位
 地域を除いたデータを中位データ、元データを全データ
 と呼び、両データからサンプルセットを作成して、推定
 される因果構造の比較を行う。

(2) 地理的に分割したサンプルセットの因果構造

本節では、地理的な基準で東日本・西日本に 2 分割し
 たサンプルセットが示す因果構造の違いを観察する。推

計結果を、全データ東日本・西日本、中位データ東日本・西日本の順に図-3(a), (b), (c), (d)に示す。

西日本の因果グラフ図-3(b), (d)を比較すると、因果的順序や直接の因果関係の有無の違いはあるものの、「アクセス指標→社会経済指標・経済パフォーマンス指標」「社会経済指標→経済パフォーマンス指標」というほぼ同一の因果構造を示している。

東日本の因果グラフ図-3(a), (c)は、最も祖先側にアクセス指標が位置せず、「人口→アクセス指標」等の因果パスが現れている。すなわちアクセス指標が内生変数となっており、西日本とは異なる結果となった。これは、因果構造に空間的異質性が存在することを示唆する。

(3) 域内総生産で分割したサンプルセットの因果構造

本節では、地域の経済発展の水準による因果構造の違いを明らかにするため、総生産を降順に並べて、サンプル数がほぼ等しくなるように3分割する。以下それぞれのサンプルセットを、総生産第一、総生産第二、総生産第三と呼ぶ。なお、総生産第一は、総生産を降順に並べた N サンプルのデータのうち $1 \sim N/3$ 、総生産第二は $N/3 + 1 \sim 2N/3$ 、総生産第三は $2N/3 + 1 \sim N$ のサンプルを含むサンプルセットである。推計結果を、全市町村データ総生産第一、総生産第二、総生産第三、中位データ総生産第一、総生産第二、総生産第三の順に図4(a), (b), (c), (d), (e), (f)に示す。

図-4(b), (c), (d), (e), (f)では、「アクセス指標→社会経済指標・経済パフォーマンス指標」「社会経済指標→経済パフォーマンス指標」という因果構造がうかがえる。また、図-4中のアクセス指標から総生産を含む多くの他の変数への効果は「アクセス指標→人口」「アクセス指標→就業者」のいずれかを通じている。それらの決定係数は、総生産第一で0.3程度、総生産第二、第三で0.1を下回る。これはアクセス性向上の効果が他の変数に与える影響が、特に総生産が中・下位の地域では小さい可能性を示す。

図-4(a)は、上位地域の多くを含むサンプルセットの因果グラフである。「人口→アクセス指標」のパスは、人口の多い市町村が、他の市町村と比較してアクセス性が良い現況を記述している、あるいはこの結果は、人口に対して相対的に道路ストックが不足している地域で道路整備が行われているとの解釈も成立するが、本分析の実からその検証を行うことはできない。いずれにせよこの因果パスから、道路整備が生産活動に関するストック効果をもたらすと結論付けることはできない。

図-4(b), (e), (f)の、総生産第二、第三の因果グラフにおいて、「アクセス指標→事業所数」の間で負の影響がみられる。通常、交通インフラの整備によるアクセ

ス性の向上は、企業の地方進出に伴って雇用を生み、都市部と地方の地域格差を縮める効果が期待される。一方で空間経済学の考え方に基づくと、集積の経済が働くことによって経済活動が都市部に集中する傾向が強まる。これは、一般にストロー効果と呼ばれ、交通インフラの整備が地方経済の衰退をもたらすメカニズムとなることが指摘されている¹³⁾。総生産第二、第三の因果グラフにみられる「アクセス指標→事業所数」の負のパスは、3路のアクセス性の向上によるストロー効果の存在を示すと考えられる。

以上の推計から、部分データの示す因果構造は全データが示す因果構造とは異なる可能性が示唆された。LiNGAMもまた統計手法の一つであり、前サンプルの一部を構成するサブセットに異質性がある場合は、その影響が検出されることがわかった。またこの結果から、社会基盤の整備効果を推計する分析で一部地域のデータしか得られない場合、得られた知見の空間的な移転性は、直ちに支持されず、慎重な検討が必要となった。

6. おわりに

本研究では統計的な手法であるLiNGAMを用いて、道路のアクセス性と市町村総生産の因果関係を、関連する交絡因子との間の因果構造を探索した。この手法の利点は、先験的に因果構造を仮定することなく、交絡因子を含めて因果構造が明らかにできる点である。

まずLiNGAMに用いる変数間の因果的順序に影響を与えるサンプル数について検討したところ、できるだけ多くのサンプルを用いることが望ましいことが分かった。さらにサンプル内のサブセットに(空間的)異質性が存在する可能性を踏まえて考慮すると、道路アクセス性が広域に影響することを踏まえても、地域的な網羅性も高い方が望ましいと考えらる。以上の予備的検討に基づいて、対象地域を国内の約1700市町村全てとして、かつ十分なサンプル数を確保するために2010年と2015年の2時点分のデータを用意して、計約3400サンプルを用いた分析を行った。

全データを用いた推計した因果グラフでは、「人口→総生産」「人口→地価(全用途)」についてそれぞれ正のパスとなり、「社会経済変数→経済パフォーマンス変数」という因果仮説を支持した。また、「就業者→アクセス指標」「アクセス指標→総生産」についてもいずれも正のパスとなった。つまり、アクセス性は内生変数の可能性があることと、「社会経済変数→アクセス指標」「アクセス指標→経済パフォーマンス変数」という因果

構造となっていることが明らかとなった。全体として「社会経済変数→アクセス指標・経済パフォーマンス変数」「アクセス指標→経済パフォーマンス変数」という因果構造が得られた。

次に部分データに関する検討として、域内総生産を降順に並べて全データを 3 分割した部分データから得た因果グラフの多くは、アクセス指標が最も原因側の変数であり、間接的な影響を含めて他の全変数に影響を与える因果構造を示した。一方、総生産が上位の市町村のみで構成したサンプルセットは「人口→アクセス指標」の順で接合する正のパスが得られた。このパスを、人口の多い市町村が他の市町村に対してアクセス性が良いことあらわすと解すると、この因果パスは現況を記述していると理解できる。あるいはこの因果パスをそのような整備が経年的に続いた結果をあらわすと解すると、都市部の人口に対して相対的に道路ストックが不足していることが、現状の道路整備をもたらしているとも理解できる。ただしこの点については追加の検証が必要である。

総生産が中・下位のサンプルセットでは「アクセス指標→事業所数」について負のパスを示し、道路のアクセス性の向上が地方経済の衰退をもたらすストロー効果が存在する可能性を示唆した。地理的な基準で東日本・西日本に分割した部分データでは、両者で異なる因果構造が得られた。これは、東日本と西日本の産業特性の違いによる因果構造に、空間的異質性が存在することを示唆した。推計した因果グラフの中で、アクセス指標が他の変数に負の効果を与える結果は、ストロー効果を表すと考えられる「アクセス指標→事業所数」のみであった。つまり多くのサブセットデータにおいて、社会基盤の整備と経済成長の間に正の因果関係が存在していた。LiNGAM は他の統計的手法と同様に全データと部分データで異なる結果（因果構造）を示す。すなわち一部地域のデータで社会基盤のストック効果を推計すると、全国の傾向と異なる可能性がある。インフラストック効果検証では、分析の目的に応じてサンプル数と空間的な範囲を検討する必要がある。

本研究の課題をまとめる。市町村単位の総生産データの補間について、生産関数アプローチを基に体積保存則を満足する面補間を行ったが、近隣地域で観測されるデータが似通った傾向を示す空間依存性は考慮していない。空間依存性を考慮した面補完法¹⁴⁾により、より精度の高い市町村単位の総生産データ予測値を得ることが望ましい。これにより空間データの分散安定化¹⁵⁾が期待できる。

アクセス指標に関して、本研究では発地と着地で異なる空間解像度を用いることで、結果的に内々所要時間がゼロとなって計測対象から外れる問題を回避した。しかしこの問題は、本来市町村内々の所要時間推計によって

解決することが望ましいため、算出手法を確立する必要がある。

地理的な基準で分割したサンプルセットでの推計では、西日本は上位、下位地域の有無に関わらずほぼ同一の因果構造を示したのに対し、東日本の因果構造に大きな違いが生じた。この違いに関しては、さらにデータの分布特性に関する検討を深めることでその原因を明らかにする必要がある。さらに一連の検討を通じて、LiNGAM の統計的手法としての性質と限界について知見を得る必要がある。

参考文献

- 1) 国土交通省・社会資本整備審議会計画部会専門小委員会: ストック効果の最大化に向けて~その具体戦略の提言~
- 2) 小池淳司: 道路のストック効果は計測可能なのか?, 高速道路と自動車, 第 59 巻, pp.5-8, 2016
- 3) 金本良嗣, 費用便益分析に係る経済学的基本問題, 第 1 章 費用便益分析における効率と公平, pp.5-18, 1999
- 4) Duranton, G., Tumer, M.A.: Urban Growth and Transportation, The Review of Economic Studies, vol.79, Issue 4, pp.1407-1440, 2012
- 5) Holl, A.: Highways and productivity in manufacturing firms, Journal of Urban Economics, vol.93, pp.131-151, 2016
- 6) Redding, S.J. and Tumer, M.A., Transportation Costs and the Spatial Organization of Economic Activity, Handbook of Urban and Regional Economics, vol.5, pp.1339-1398, 2015
- 7) 織田澤利守, 大平悠季: 交通インフラ整備効果の因果推論: 論点整理と展望, 土木学会論文集 D3 (土木計画学), Vol.75, No.5, pp.I-1-I-15, 2019
- 8) 杉原豪, 塚井誠人: 統計的因果探索による社会基盤整備のストック効果の検証, 土木学会論文集 D3 (土木計画学), Vol.75, No.6, pp.I-583-I-589, 2020
- 9) 清水昌平: 機械学習プロフェッショナルシリーズ 統計的因果探索, 講談社, 2017
- 10) J. Eriksson and V. Koivunen : Identifiability, separability, and uniqueness of linear ICA models, IEEE signal Processing letters, 11, pp.601-604, 2004
- 11) 保科架風: Bayesian lasso によるスパース回帰モデリング, 計算機統計学, 第 25 巻・2 号, pp.73-85, 2012
- 12) 塚井誠人, 江尻良, 奥村誠, 小林潔司: 社会資本の生産性とスピルオーバー効果, 土木学会論文集 No.714/IV-57, 53-67, 2002
- 13) 猪原龍介, 中村良平, 森田学: 空間経済学に基づくストロー効果の検証~明石海峡大橋を事例として~, RIETI Discussion Paper Series 15-J-045, 2015
- 14) 津田敏明, 塚井誠人: 小地域補間のための空間集計モデルの適用可能性, 土木学会論文集 D3 (土木計画学), Vol.69, No.5, pp.I-391-I-400, 2013

- 15) 堤盛人, 瀬谷創: 土木計画における応用空間統計学の可能性,
土木学会論文集 D3(土木計画学), Vol.68, No.5, pp.L_1-L_200, 2012

(2021.03.07)

Causal Discovery between the Road Accessibility and Gross Regional Product Using LiNGAM in Whole Japan

Go SUGIHARA and Makoto TSUKAI

In regional economic activities, the stock effect brought by road infrastructure may not always occur as pillowing to the fundamental economic assumptions or models. The purpose of this study is to clarify the causal relationship between road accessibility and regional product using LiNGAM which is an approach to give a causal diagram based on the statistical rationality. The dataset in this study includes all cities, towns, and villages in Japan in 2010 and 2015. Using the dataset, LiNGAM discovered the positive causal structure of "road access indicators to economic performance variables". Through the application of LiNGAM on the subset of original data, it was clarified that the care for data quality such as higher distributional moment or outlines is necessary. Otherwise, the wrong causal structures would be obtained.