



操作変数の見つけ方 : 制度の外にも目を向けよ (<特集>計量分析の理論と応用)

山根, 史博
東, 裕三

(Citation)

国民経済雑誌, 209(1):81-95

(Issue Date)

2014-01

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/81008954>

(URL)

<https://hdl.handle.net/20.500.14094/81008954>



操作変数の見つけ方：制度の外にも目を向けよ

山 根 史 博
東 裕 三

国民経済雑誌 第 209 巻 第 1 号 抜刷

平成 26 年 1 月

操作変数の見つけ方：制度の外にも目を向けよ

山 根 史 博
東 裕 三

社会実験の処置効果を推定する上で注意すべきものに自己選択バイアスがある。こうした内生性バイアスを取り除くために考案された手法の1つとして操作変数法があるが、その効能が十分に発揮されるためには、適切な操作変数を見つける必要がある。適切な操作変数を見つけるための王道は、分析したい社会実験に関係する諸制度を良く吟味することとされている。しかし、むしろそれでは適切な操作変数が見つからない場合があり、制度から離れた視点からの操作変数の探索も必要になる。本稿では、操作変数法を実践した先行研究を取り上げながら、そのことを再確認したい。また、具体的なエピソードとして平成の大合併を取り上げ、これが市町村財政の健全性に与えた影響を推定する際に、関連制度を踏まえて考案されている操作変数と我々が制度から離れた視点で考案した操作変数を用い、両者の推定結果を比較した。

キーワード 社会実験、処置効果、自己選択バイアス、操作変数

1 はじめに

実験の処置効果を測る上で最も理想的な状況は、処置を施すグループ（処置群，treatment group）とそうでないグループ（制御群，control group）であらゆる条件が等しい状況である。例えば、実験室では、温度や気圧などの諸条件がほぼ完全に等しい状況を人工的に作り出すことができる。その次に理想的な状況は、処置状態が条件とは無関係に決められた状況である。代表的な例として、施肥の効果を測るために施肥をした農地（処置群）とそうでない農地（制御群）とで作物の収穫量を比較する実験を考えよう（Wooldridge (2006) p. 13 を参照されたい）。ここで注意しなければならないのは処置群の選び方である。仮に耕作条件が有利な農地を処置群に選べば、施肥の効果を過大評価してしまうことになる。一方、処置群を無作為に選べば、処置状態が耕作条件と無関係になるため、処置群と制御群の平均収穫量を比較することで平均処置効果（average treatment effect, ATE）の不偏かつ一致推定量を得ることができる。

しかし、我々の社会において、こうした理想的な実験が行われるケースはあまり多くない。

先程の施肥の例でも、肥料の製造業者が効果を大きく見せようとして、意図的に耕作条件の良い農地を処置群に選ぶ可能性がある。また、そうした意図が無くても、偏った選定が行われることがある。例えば、医者は効果の見込める患者にしか薬の投与や手術を施さないだろう。それらの処置を患者の様態や体質と無関係に施すことは、医者としての本分に反するだけでなく、人道的に問題視されるケースが多いからである。いずれにせよ、処置の期待効果に基づいて処置状態が決められる状況では、単純に処置群と制御群を比較しても ATE を一致推定することはできない。このような理由による真の処置効果からの乖離を自己選択バイアス (self-selection bias) という。

社会実験において自己選択バイアスはありふれた問題であり、これにどう対処するかは計量経済学の重要な課題の1つと言える。対処法は大きく分けて2通りある。1つはマッチングによる対処、すなわち、propensity score に基づいて処置群の観測個体と「似ている」と判断される制御群の個体を選び、両者を比較することである。しかし、類似性の判断方法は複数あり、¹⁾ 同じデータを使っても、それによって推定結果が変わってしまう。また、不可測な変数の類似性までは考慮できない。

もう1つは回帰であり、これには次の3通りのバリエーションがある。

1つ目は、処置状態にとっての欠如変数を制御することである。そうすることで、処置状態が誤差項と無相関になり、自己選択バイアスを取り除くことができる。しかし、この対処法には少なくとも3つの問題がある。第1に、我々は欠如変数を網羅的にリストアップしなければならないが、社会変数間の相互関係はブラックボックスのようなもので、その極めて複雑な構造を完全に把握できていない以上、本来制御すべき欠如変数を見落とすことは往々にして起こり得る。第2に、リストアップした欠如変数のすべてを観測できるとは限らない。第3に、制御した欠如変数が内生変数であった場合、処置効果の推定結果は自己選択とは別の新たな内生性バイアスが生じる。制御する変数が多いほど、この問題に苛まれる可能性が高くなる。

2つ目は、パネルデータを活用して、固定効果と時間固定効果を制御することである。この対処法のメリットは、固定効果・時間固定効果に該当するすべての欠如変数を可測性とは無関係に一網打尽に制御できる点にある。しかし、どちらにも該当しない欠如変数が存在する場合、上記の対処法と同じ問題が残ることになる。

3つ目は、操作変数法を用いることである。この場合、欠如変数を網羅的にリストアップする必要はなく、また、それらの可測性に制約されることもない。その意味で、操作変数法は最も有用な対処法と言ってよいだろう。ただし、この対処法の効能が十分に発揮されるためには、適切な操作変数が存在しなければならない。今、処置効果の推定モデルを次式で特定したとする。

$$y = \alpha + \beta w + \varepsilon$$

y は従属変数, w は処置状態を表すダミー変数, ε は誤差項, α と β は回帰係数であり, β が処置効果に相当する。このとき, 次の2つの性質を満たす変数 z を操作変数という。

$$\text{Cov}(w, z) \neq 0 \wedge \text{Cov}(\varepsilon, z) = 0$$

やや厳密さに欠けるが, Angrist and Pischke (2008) p. 117 はこれらの性質を「操作変数 (z) は処置状態 (w) に影響を与え, また, その経路のみを通じて従属変数 (y) に影響を与える」という直感的な解釈を提示している。

どうすれば, このような変数が見つかるのだろうか? 特に, 2番目の性質が操作変数の肝と言ってよい。操作変数を見つけるための王道的指針は, 分析したい社会実験に関連する制度 (もしくは経済理論) を活用することとされている (Angrist and Pischke (2008) p. 117, Stock and Watson (2010) p. 448)。しかし, それだけで適切な変数が見つかる保証はないため, 別のアプローチが必要になる。例えば, Angrist and Pischke (2008) p. 117 はデータの生成過程そのものを吟味することを提案している。むしろ制度から離れた視点からデータを眺めることで, 欠如変数とは無関係な変数, あるいは Stock and Watson (2010) p. 449 が言及する確率的な外生要因が見つかる可能性がある。

操作変数を見つけるための必勝法はない。したがって, アプローチの仕方は多様であるべきであり, 本稿では, 先行研究における操作変数法の実践例をいくつか取り上げながら, そのことを再確認することにした。また, ここでのエチュードとして平成の大合併を取り上げ, これが市町村財政の健全性に与えた効果を推定する。その際に, 関連制度に基づいて考案されている操作変数と我々が制度から離れた視点で考案した操作変数を用い, 両者の推定結果を比較する。

本稿の構成をまとめる。第2節と第3節では, 操作変数法を実践した先行研究を取り上げながら, 制度に基づく操作変数の探索 (制度アプローチ) が抱える限界を指摘し, その補完的アプローチとして, 制度から離れた視点による探索 (非制度アプローチ) の指針について考察する。第4節では, 両者のアプローチを活用し, 市町村合併が財政の健全性に及ぼした影響を推定する。第5節はまとめである。

2 制度アプローチ

Angrist and Pischke (2008) の chapter 4 では操作変数法のユニークな実践例が数多く紹介されている。その中から制度アプローチに該当するものを取り上げてみよう。

Angrist and Krueger (1991) は就学年数が将来所得にどのような影響を与えるかを推定している。しかし, 就学年数は個々人が将来の期待所得 (その他, 進学にかかる費用や学習意欲など) を念頭に決定するものであるため, 自己選択バイアスが生じる。そこで Angrist

and Krueger (1991) が注目したのが義務教育制度である。アメリカでは、6歳になる年の9月に小学校に入学し、16歳の誕生日に義務教育が終了する。つまり、義務教育の開始は日本と同様に誰でも同じだが、終了時期が誕生日によって異なり、ゆえに義務教育日数も異なる。義務教育日数の少ない早生まれの人は就学年数も少ない傾向があるため、誕生日 (z) と就学年数 (w) には一定の相関関係がある ($\text{Cov}(w, z) \neq 0$)。また、誕生日は、将来所得を規定する先天的な能力や家庭環境などの不可測要因 (ε) とは無関係であるため ($\text{Cov}(\varepsilon, z) = 0$)、適切な操作変数と言えそうである。このように、制度を掘り下げることで、誕生日によって処置状態が変化するという疑似実験の状況を見出したわけである。

このような疑似実験が意図的に制度化されている場合もある。

例えば、Angrist (1990) はベトナム戦争時 (1960~75年) における兵役が生涯所得に与えた影響を推定している。当時は徴兵制度があったが、徴兵召集があった場合でも健康上の問題などを理由に兵役を免れた青年、あるいは、徴兵されなくても自ら志願した青年がいた。免除や志願の理由 (体力、兵役を通じたキャリアアップへの期待など) は所得とも関連するため、処置効果の推定結果にバイアスが生じる。この問題に対する操作変数として、Angrist (1990) は徴兵資格を用いている。当時の徴兵制度は、公平性を期すために、19歳になった青年を対象に誕生日に応じて番号をランダムに割り振り、その番号が打ち切り番号以下であった青年に徴兵資格を与えるという制度であった。それゆえ、徴兵資格 (z) は兵役参加の決定 (w) に明らかな影響を与え ($\text{Cov}(w, z) \neq 0$)、かつランダムに決定されるため、所得の不可測要因 (ε) とは無関係である ($\text{Cov}(\varepsilon, z) = 0$)。

もう1つ興味深い事例として、Angrist (2006) がある。ここでは、家庭内暴力を取り上げ、逮捕という比較的重い措置が暴力の凶悪化を防ぐことに役立つかを検証している。しかし、よほどひどい暴力が行われているからこそ逮捕に至るという逆の因果が存在するため、自己選択バイアスの問題が生じる。そこで、Angrist (2006) が注目したのが Minneapolis Domestic Violence Experiment という実験である。この実験では、家庭内暴力の通報があると、ルーレットのようなものが作動し、暴力の内容とは無関係に逮捕・別居命令・カウンセリングのいずれかの措置をランダムに指示する。ただし、警察官がその指示に従う必要はなく、ひどい暴力が行われていると判断した場合は、指示を無視して逮捕をしてよい。一方、それほどひどい暴力でなくても、「逮捕」という指示があれば、警察官は逮捕に踏み切っていたので、逮捕の指示 (z) は逮捕という意味決定 (w) に影響を与えている ($\text{Cov}(w, z) \neq 0$)。ただし、指示自体はランダムに決まるため、暴力の凶悪さとは無関係である。

注意しなければならないのは、こうした状況設定は常に期待できるものではないということである。そもそも、制度とは物事の因果関係を我々が意図的に構築したものである (経済理論も社会現象の発生メカニズムとして物事の因果関係を仮定している)。したがって、制

度を吟味して分かることは、基本的には因果関係なのであり、このアプローチで操作変数（つまり、その因果の輪に部分的に関係しているが $\text{Cov}(w, z) \neq 0$ ）、それ以外では無関係な要因（ $\text{Cov}(\varepsilon, z) = 0$ ）が見つかるのは、1つ目の事例のように制度自体が特殊か、2つ目と3つ目の事例のように制度内にランダムな要因が組み込まれている場合に限る。

ここで、本稿でエチュードとして取り上げる合併と市町村財政について考察することにした。合併するかどうかの意思決定と財政状況は因果同時の関係にあると考えられる。財政状況の悪化が合併へのインセンティブとして働いた可能性があるからである。この問題への対処として、宮崎（2013）は段階補正の見直しに着目している。この見直しによる地方交付税の削減度合は自治体の人口規模によって異なる。そこで、宮崎（2013）は人口規模を操作変数として利用している。しかし、この操作変数は適切なものとは言い難い。なぜなら、人口によって歳入が変化することは、合併の意思決定に影響を与えるだろうが、それ以前に財政状況に直接的な影響を与えるからである。したがって、人口は操作変数ではなく、合併の意思決定に対する欠如変数と解釈した方がよい。

では、人口を制御するとどうなるだろうか？この場合、自己選択バイアスは軽減されるかもしれないが、別の問題が生じる恐れがある。人口は財政状況に影響を与えるが、一方で、財政状況の良い自治体に人口が流れ、悪い自治体からは流出するという逆の因果が存在する可能性があるため、人口を制御すれば別の内生性バイアス（因果同時バイアス）が発生する恐れがある。財政状況の規定構造分析において人口を考慮することは一般的であるが、このときのパラメータは、人口から財政状況への因果関係ではなく、両者の相関関係を表しているに過ぎない。また、人口を組み込むことでモデルの予測力は高まるかもしれないが、ここではそれを求めている。ここでの目的は処置効果を一致推定することであり、内生性が疑われる変数を制御することはその目的に反する。

3 非制度アプローチ

このように、制度アプローチのみでは適切な操作変数が見つからない恐れがある。それゆえ、多角的な視野からの探索が必要になる。

例えば、Angrist and Evans (1998) は「子沢山の女性は外で働かなくなるか？」を検証している。しかし、労働意欲が低い、あるいは働いても高い所得が見込めない場合、その女性は労働の優先順位を下げ、子供を持つことの相対的な優先順位を上げる可能性があるため、自己選択バイアスが生じるかもしれない。そこで、Angrist and Evans (1998) は分析対象を子供が2人以上いる母親に限定し、子供が3人以上いるなら子沢山（ $w=1$ ）と定義した上で、次の2つの操作変数を利用している。1つは2度目の出産が双子なら1をとるダミー変数である。2度目の出産が双子なら自動的に子供は3人になるため、その母親は子沢山とい

うことになり ($\text{Cov}(w, z) \neq 0$)、また、双子が生まれるかどうかはほぼランダムな事象である ($\text{Cov}(\varepsilon, z) = 0$)。もう1つは、第1子と第2子の性別が同じなら1をとるダミー変数である。例えば、最初の2人がどちらも女の子なら、男の子が欲しくなって3人目を持とうとするかもしれない ($\text{Cov}(w, z) \neq 0$)。また、第1子と第2子の性別が同じという事象はランダムに起こることである ($\text{Cov}(w, z) \neq 0$)。

Stock and Watson (2010) p. 452 は地理情報の活用を提案し、具体例として McClellan *et al.* (1994) を挙げている。McClellan *et al.* (1994) は急性心筋梗塞に対する心臓カテーテルの効果を推定しているが、効果の見込める患者がその治療を受けるため、自己選択バイアスが生じる。そこで、操作変数として活用したのが心臓カテーテルを専門とする病院とそうでない病院の地理情報である。具体的には、患者の自宅から最寄りの心臓カテーテル専門病院までの距離とそうでない最寄り病院までの距離の差を用いており、この差が小さいほど、急性心筋梗塞になった時の搬送先が前者の病院になる確率が高く、ゆえに心臓カテーテルを受ける可能性が高まる ($\text{Cov}(w, z) \neq 0$)。また、住民の健康状態の地理的分布と病院の配置が無関係ならば ($\text{Cov}(\varepsilon, z) = 0$)、この変数は操作変数として適切と言える。

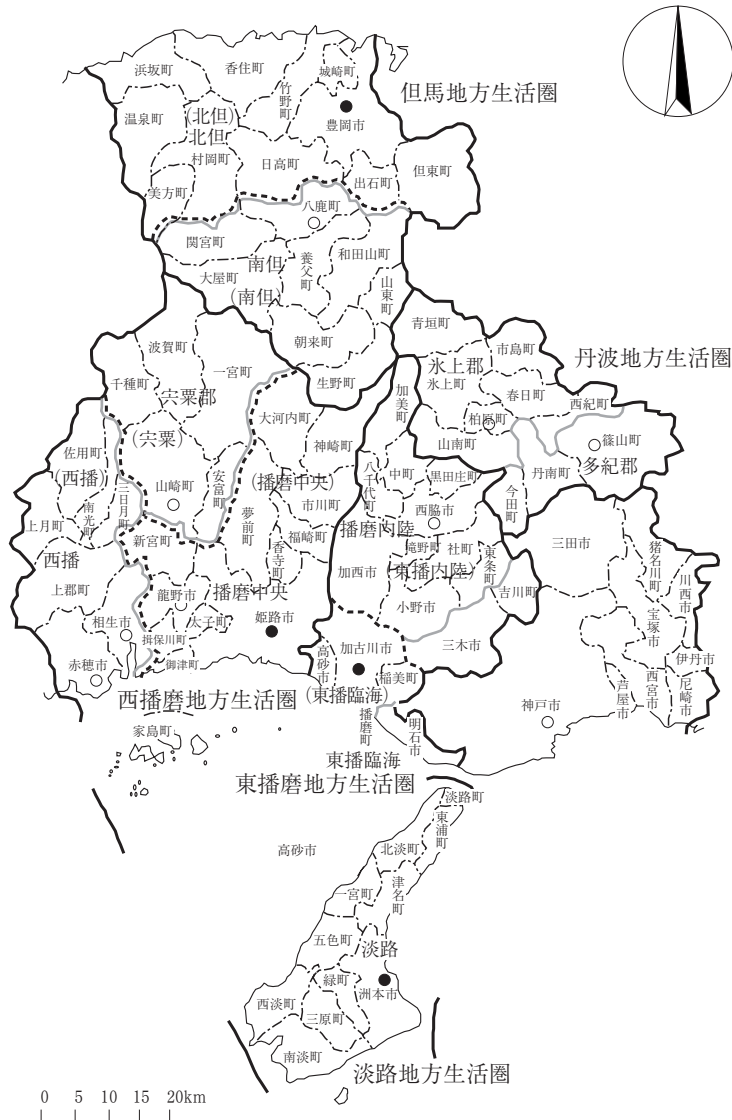
以上のような操作変数は幸運な思いつきによって見つかることがある。また、入念なデータ・マイニングによって見つかる場合もある。例えば、データのシートを見続けるだけで、変数間の漠とした関係が見えてくる。あるいは、観測可能な変数の相関係数を逐一調べ、その数値から操作変数として使えそうな変数を見つけ出し、それが処置状態に影響を与えること、また、その経路を通じてのみ従属変数に影響を与えることをある程度論理的に説明できれば、それが操作変数として適切である可能性が高まる。あるいは、異分野の専門家に相談することで、操作変数の探索に新しい視点を取り入れることができるかもしれない。

ここで、市町村合併に話を戻し、非制度アプローチで我々が考案した操作変数を紹介しよう。基本的に合併は陸続きの隣接自治体同士で行われる。したがって、隣接自治体数が多いほど、合併する確率が高いと考えられる。また、同一の地方生活圏内での合併事例が多い。これは、地方生活圏計画に基づき、従来から圏内で連携して計画的・総合的な公共サービスの整備が図られてきた経緯から、合併に向けた協議・合意形成がしやすい下地が出来上がっていたためと思われる。そこで、我々は新たな操作変数として「同一地方生活圏内の隣接自治体数」を提案したい。

ただし、これが操作変数足り得るか、つまり、 $\text{Cov}(\varepsilon, z) = 0$ が満たされるかについてはもう少し議論が必要である。

第1に、隣接自治体数が多いほど、その自治体は面積が大きい傾向がある。また、林 (2004) によれば面積は財政状況に影響を与える。したがって、隣接自治体数は、合併という経路だけでなく、面積という経路で財政に影響を与えている可能性がある。したがって、

図1 地方生活圏の区分（例：兵庫県）



出典：『地方生活圏要覧 平成5年版』

隣接自治体数を操作変数に用いる場合は、面積を欠如変数としてモデルに組み込んだ方がよいかもしれない。なお、財政状況の良し悪しが自治体の面積に影響を与えるという逆の因果が存在する可能性は、人口の場合に比べて小さいように思われる。人口移動は比較的短期間で行われるが、面積は長期間に渡って固定されているからである。昭和の大合併（1953年から1961年にかけて9,868団体から3,472団体に減少）の際に、当時の財政状況を勘案して合併

し、面積を広げた自治体もあるだろうが、平成の大合併まではそれから40年以上の時間が経過している。したがって、ここでは面積を外生変数として扱う。

第2に、同一地方生活圏内の自治体の財政状況はある程度類似しており、ゆえに操作変数自体が内生変数になっているのではないかと思われるかもしれない。しかし、我々の操作変数は「隣接自治体数」であり、「同一生活地方圏の財政状況」ではないため、その種の内生性問題には苛まれないはずである。

4 エチュード：市町村合併と財政の健全性

ここでは、宮崎（2013）が考案している操作変数「人口」と我々が考案した操作変数「同一地方生活圏内の隣接自治体数」を活用して平成の大合併が市町村財政に及ぼした影響を推定する。なお、宮崎（2013）は住民1人当たり歳出額に与えた影響を推定しているが、ここでは財政の健全性に与えた影響を推定する。

財政の健全性指標として経常収支比率と公債費負担比率を用いる。『市町村別決算状況調』を基に、平成14年度及び平成20年度における両指標の分布を制御群（平成20年度までに合併していない自治体）と処置群（合併した自治体）に分けてまとめたものが表1と表2である。

表1 経常収支比率の変化

	制御群		処置群	
	平成14年度	平成20年度	平成14年度	平成20年度
70%未満	30 (2.5%)	23 (1.9%)	13 (0.6%)	3 (0.5%)
70%以上75%未満	61 (5.0%)	21 (1.7%)	67 (3.3%)	5 (0.9%)
75%以上80%未満	168 (13.9%)	79 (6.5%)	199 (9.9%)	8 (1.4%)
80%以上85%未満	303 (25.0%)	171 (14.1%)	487 (24.3%)	30 (5.3%)
85%以上90%未満	345 (28.5%)	332 (27.4%)	714 (35.7%)	150 (26.5%)
90%以上95%未満	172 (14.2%)	360 (29.8%)	359 (17.9%)	262 (46.2%)
95%以上100%未満	88 (7.3%)	184 (15.2%)	138 (6.9%)	98 (17.3%)
100%以上	43 (3.6%)	40 (3.3%)	25 (1.2%)	11 (1.9%)
計	1,210 (100.0%)	1,210 (100.0%)	2,002 (100.0%)	567 (100.0%)
平均	85.65 %	89.12 %	86.55 %	91.16 %
標準偏差	7.81 %	7.28 %	6.47 %	4.91 %

制御群の経常収支比率は平均して3.47%上昇、処置群は平均して4.61%上昇しており、制御群と処置群の平均値の差は平成14年度時点で0.90%だったものが平成20年度には2.03%に広がっている。一方、公債費負担比率に関しては、制御群は平均して0.12%上昇、処置群は平均して1.36%上昇しており、制御群と処置群の平均値の差は1.21%から2.45%に広がっている。処置群は制御群に比べて健全性が元々低かったこと、また、どちらの群も平成20年度

表2 公債費負担比率の変化

	制御群		処置群	
	平成14年度	平成20年度	平成14年度	平成20年度
10%未満	134 (11.1%)	132 (10.9%)	113 (5.6%)	11 (1.9%)
10%以上15%未満	340 (28.1%)	334 (27.6%)	468 (23.4%)	81 (14.3%)
15%以上20%未満	382 (31.6%)	373 (30.8%)	700 (35.0%)	229 (40.4%)
20%以上25%未満	224 (18.5%)	251 (20.7%)	471 (23.5%)	172 (30.3%)
25%以上30%未満	91 (7.5%)	80 (6.6%)	171 (8.5%)	53 (9.3%)
30%以上	39 (3.2%)	40 (3.3%)	79 (3.9%)	21 (3.7%)
計	1,210 (100.0%)	1,210 (100.0%)	2,002 (100.0%)	567 (100.0%)
平均	17.13 %	17.25 %	18.34 %	19.70 %
標準偏差	6.35 %	6.30 %	5.81 %	5.16 %

にかけて悪化しているが、処置群の悪化は制御群のそれを上回ることが分かる。

以下では、平成14・20年度の2期間パネルデータを用いて合併による処置効果を推定する。なお、観測個体は平成14年度時の3,212団体をベースとし、表3に示すように、合併した自治体の平成20年度時の値には合併後の値を複製して当てはめることにした。

表3 パネルデータの構造（一部抜粋）

	平成14年度				平成20年度		
	自治体名	経常収支比率	公債費負担比率		自治体名	経常収支比率	公債費負担比率
合併した自治体	神戸市	103.3	31.5	→	神戸市	97.3	26.1
	姫路市	75.6	14.2	→	姫路市	84.5	14.6
	家島町	90.5	22.7	→	姫路市	84.5	14.6
	夢前町	81.6	11.0	→	姫路市	84.5	14.6
	香寺町	83.7	11.7	→	姫路市	84.5	14.6
	安富町	79.7	9.7	→	姫路市	84.5	14.6
	尼崎市	100.4	14.4	→	尼崎市	98.6	20.8
	明石市	92.0	18.7	→	明石市	94.9	19.5
	西宮市	99.7	23.5	→	西宮市	98.2	22.0

推定モデルを示す。

$$y_{i,t} = \alpha + \beta w_{i,t} + \text{fixed effects} + \text{time fixed effect} + \varepsilon_{i,t}, \quad i = 1, \dots, 3,212, \quad t = 0, 1.$$

$y_{i,t}$ は経常収支比率もしくは公債費負担比率、 $w_{i,t}$ は合併したか否かを示すダミー変数、 $\varepsilon_{i,t}$ は誤差項であり、下付きの i は自治体、 t は観測時点（平成14年度なら 0，平成20年度なら 1）を表している。なお、任意の i について $w_{i,0} = 0$ であり、平成20年度までに合併した場

合は $w_{i,1}=1$ ，合併していない場合は $w_{i,1}=0$ である。ここでは，この回帰モデルを基にした次式の first-difference モデルを推定する。

$$\Delta y_i = \beta \Delta w_i + \text{time fixed effect} + \Delta \varepsilon_i, i = 1, \dots, 3,212.$$

$\Delta y_i \equiv y_{i,1} - y_{i,0}$, $\Delta w_i \equiv w_{i,1} - w_{i,0}$, $\Delta \varepsilon_i \equiv \varepsilon_{i,1} - \varepsilon_{i,0}$ である。

平成14年度時点における「人口」と「同一地方生活圏内の隣接自治体数」の概要を表4と表5にまとめた。「人口」については，宮崎（2013）に倣い，人口1,000人未満と人口1,000人以上8,000人未満を表す2つのダミー変数を操作変数として用いる。一方，「同一地方生活圏内の隣接自治体数」については，3団体以上4団体以下と5団体以上を表す2つのダミー変数を用いることにした（2団体以下の自治体数は1,080団体，3団体以上4団体以下は1,324団体，5団体以上は808団体でほぼ3等分になる）。

先程述べたとおり，「同一地方生活圏内の隣接自治体数」を操作変数に用いる場合は面積

表4 人口（平成14年度）

	制御群		処置群	
1,000人未満	20	(1.7%)	32	(1.6%)
1,000人以上4,000人未満	137	(11.3%)	322	(16.1%)
4,000人以上8,000人未満	208	(17.2%)	554	(27.7%)
8,000人以上12,000人未満	131	(10.8%)	345	(17.2%)
12,000人以上20,000人未満	179	(14.8%)	311	(15.5%)
20,000人以上30,000人未満	125	(10.3%)	129	(6.4%)
30,000人以上	410	(33.9%)	309	(15.4%)
計	1,210	(100.0%)	2,002	(100.0%)
平均	54,446 人		26,336 人	
標準偏差	178,708 人		75,644 人	

表5 同一地方生活圏内の隣接自治体数（平成14年度）

	制御群		処置群	
0 団体	318	(26.3%)	109	(5.4%)
1 団体	65	(5.4%)	111	(5.5%)
2 団体	165	(13.6%)	312	(15.6%)
3 団体	240	(19.8%)	444	(22.2%)
4 団体	190	(15.7%)	450	(22.5%)
5 団体	122	(10.1%)	289	(14.4%)
6 団体	68	(5.6%)	155	(7.7%)
7 団体以上	42	(3.5%)	132	(6.6%)
計	1,210	(100.0%)	2,002	(100.0%)
平均	2.65 団体		3.63 団体	
標準偏差	2.10 団体		1.90 団体	

表 6 面積（平成14年度）

	制御群	処置群
20 km ² 未満	175 (14.5%)	207 (10.3%)
20 km ² 以上40 km ² 未満	199 (16.4%)	362 (18.1%)
40 km ² 以上80 km ² 未満	268 (22.1%)	547 (27.3%)
80 km ² 以上120 km ² 未満	152 (12.6%)	316 (15.8%)
120 km ² 以上180 km ² 未満	119 (9.8%)	273 (13.6%)
180 km ² 以上240 km ² 未満	96 (7.9%)	119 (5.9%)
240 km ² 以上400 km ² 未満	112 (9.3%)	138 (6.9%)
400 km ² 以上	89 (7.4%)	40 (2.0%)
計	1,210 (100.0%)	2,002 (100.0%)
平均	138.65 km ²	101.66 km ²
標準偏差	176.02 km ²	102.82 km ²

を制御した方がよいかもしれない。平成14年度における「面積」の分布は表 6 に示すとおりである。

推定結果を表 7 と表 8 に示す。(1) 列は操作変数を用いずに first-difference モデルを OLS 推定した結果, (2) 列は「人口」を操作変数に, (3) 列は「同一地方生活圏内の隣接自治体数」を操作変数に 2SLS 推定した結果である。(4) 列は (3) 列の推定において面積を制御した場合の結果である。

始めに 2SLS 推定に関する検定結果を見ていく。まず, 内生性の検定結果から, 合併ダミー (Δw_i) が外生であるという帰無仮説はいずれの推定においても 5 % 有意水準で棄却された。また, 第 1 段階推定に関する F 統計値はいずれの推定においても 10 を超えていた。Stock *et al.* (2002) の基準に基づけば, どの操作変数も Δw_i に弱くない影響を与えていることが支持されている。一方, 過剰識別の検定結果は「同一地方生活圏内の隣接自治体数」の方が操作変数として妥当であることを支持するものであった。(2) 列で「人口」を操作変数に用いた場合, 2 つのダミー変数がどちらも操作変数として妥当であるという帰無仮説は, 経常収支比率の推定において 1 % 水準で棄却されたが (p 値は 0.0005), 公債費負担比率の推定においては 10 % 水準で棄却されなかった (p 値は 0.2489)。これに対し, (3) 列で「同一地方生活圏内の隣接自治体数」を操作変数に用いた場合は, 同様の帰無仮説が経常収支比率の推定と公債費負担比率の推定の双方において 10 % 水準で棄却されなかった。p 値は, 経常収支比率の推定において 0.1370, 公債費負担比率の推定において 0.4451 であり, それぞれ「人口」を操作変数に用いた場合よりも大きかった。(4) 列で面積を制御した場合の p 値はさらに大きかった。

次に, 合併の処置効果の推定結果 (合併ダミーのパラメータ推定値) について考察する。

表7 推定結果 (経常収支比率)

	(1) OLS 推定	(2) 2SLS 推定	(3) 2SLS 推定	(4) 2SLS 推定
First-difference モデルの推定結果				
合併ダミー	1.10 (0.23) ***	-12.40 (2.18) ***	10.63 (1.53) ***	9.85 (1.35) ***
面積 (平成14年度時)	—	—	—	0.0047 (0.0011) ***
固定効果	Yes	Yes	Yes	Yes
時間固定効果	3.47 (0.19) ***	11.89 (1.36) ***	-2.47 (0.98) ***	-2.52 (0.95) ***
修正済み決定係数	0.0066	0.0203	0.0242	0.0245
第1段階の推定結果				
定数項	—	0.717 (0.013) ***	0.493 (0.015) ***	0.545 (0.016) ***
人口ダミー (平成14年度時)	—	—	—	—
人口が1,000人未満	—	-0.102 (0.069)	—	—
人口が1,000人以上8,000人未満	—	-0.153 (0.017) ***	—	—
隣接自治体数ダミー (平成14年度時)	—	—	0.183 (0.020) ***	0.197 (0.020) ***
隣接数が3団体以上4団体以下	—	—	0.220 (0.022) ***	0.247 (0.022) ***
隣接数が5団体以上	—	—	—	-0.0006 (0.0001) ***
面積 (平成14年度時)	—	—	—	—
修正済み決定係数	—	0.0227	0.0372	0.0615
検定結果				
内生性の検定				
Wooldridge's (1995) robust score test	—	78.55 ***	58.94 ***	59.69 ***
Wooldridge's (1995) robust regression-based test	—	81.14 ***	60.85 ***	61.57 ***
Weak instruments の検定				
F test for the first stage estimation	—	40.08 ***	60.46 ***	74.01 ***
過剰識別の検定				
Wooldridge's (1995) robust score test	—	12.18 ***	2.21	1.70
標本数				3,212

注：() 内は不均一分散に頑健な標準誤差. ***p<.01, **p<.05, *p<.10.

表 8 推定結果（公債費負担比率）

	(1) OLS 推定	(2) 2SLS 推定	(3) 2SLS 推定	(4) 2SLS 推定
First-difference モデルの推定結果				
定数項	—	—	—	—
合併ダミー	1.60 (0.17) ***	—	—	—
面積（平成14年度時）	—	—9.90 (1.75) ***	3.65 (0.94) ***	3.63 (0.86) ***
固定効果	Yes	Yes	Yes	0.0002 (0.0007) Yes
時間固定効果	0.12 (0.12)	7.29 (1.10) ***	-1.15 (0.60) *	-1.17 (0.60) *
修正済み決定係数	0.0258	0.0229	0.0048	0.0072
第1段階の推定結果				
定数項	—	0.717 (0.013) ***	0.493 (0.015) ***	0.545 (0.016) ***
人口ダミー（平成14年度時）	—	—	—	—
人口が1,000人未満	—	-0.102 (0.069)	—	—
人口が1,000人以上8,000人未満	—	-0.153 (0.017) ***	—	—
隣接自治体数ダミー（平成14年度時）	—	—	0.183 (0.020) ***	0.197 (0.020) ***
隣接数が3 団体以上 4 団体以下	—	—	0.220 (0.022) ***	0.247 (0.022) ***
隣接数が5 団体以上	—	—	—	-0.0006 (0.0001) ***
面積（平成14年度時）	—	—	—	—
修正済み決定係数	—	0.0227	0.0372	0.0615
検定結果				
内生性の検定				
Wooldridge's (1995) robust score test	—	93.06 ***	4.96 **	6.07 **
Wooldridge's (1995) robust regression-based test	—	96.24 ***	4.97 **	6.09 **
Weak instruments の検定				
F test for the first stage estimation	—	40.08 ***	60.46 ***	74.01 ***
過剰識別の検定				
Wooldridge's (1995) robust score test	—	1.33	0.58	0.56
標本数				3,212

注：（ ） 内は不均一分散に頑健な標準誤差。*** p<.01, ** p<.05, * p<.10.

操作変数を用いずに OLS 推定した場合、経常収支比率と公債費負担比率への処置効果はそれぞれ1.10%、1.60%であり、合併が市町村財政の健全性を悪化させたことを示す結果となった。これに対し、「人口」を操作変数に用いた場合、これらは-12.40%、-9.90%であり、先程の OLS 推定の結果には上方バイアスがかかっていることを示唆している。反対に、「同一地方生活圏内の隣接自治体数」を操作変数に用いた場合は10.63%、3.65%であり、下方バイアスがかかっていることを示唆している。後者の推定結果に関しては、面積を制御しても大きくは変わらなかった。

5 む す び

本稿では、社会実験の処置効果の推定における自己選択バイアスへの対処として操作変数法に着目し、過去の実践例のレビューを通じて、適切な操作変数の見つけ方について考察した。ここでは、操作変数を見つけるためのアプローチを制度アプローチと非制度アプローチに分類し、とりわけ後者のアプローチの重要性を確認した。さらに、平成の大合併が市町村財政の健全性に与える影響を推定し、制度アプローチで考案されている操作変数と我々が非制度アプローチで考案した操作変数を用いて、両者の推定結果の比較を行った。

操作変数を見つけるために、関連する諸制度や経済理論を吟味することは基本中の基本である。しかし、それによって操作変数が見つかる保証はなく、むしろそれらにとらわれ過ぎると適切な操作変数が見つからない恐れがある。一方で、制度を踏まえなくても、ある日突然操作変数が思いつくこともあれば、入念なデータ・マイニングによってそれが見つかることもある。分析したい社会実験の制度や理論を良く知っているほど、それらから逸脱することは困難かもしれないが、その場合は異分野の専門家と議論し、別の視点を取り入れることが大いに役立つかもしれない。実際、第一筆者は市町村合併や市町村財政の専門家ではない。それゆえ、関連する制度を一切考慮せず、単純に合併した自治体同士の地理関係を漠然と考えた結果、「同一地方生活圏内の隣接自治体数」という新しい操作変数を思いついた。推定結果の妥当性については、財政学の専門家の予見や実感との整合性を確認する必要があるが、過剰識別の検定結果は制度アプローチで考案されている「人口」という操作変数よりも妥当であることが示していた。

社会実験の処置効果を一致推定することは、効果的・効率的な政策を実施する上で不可欠である。しかし、社会実験の多くは自己選択を引き起こすのであり、その解決に向け、多様なアプローチで操作変数が探索されることを期待したい。

注

- 1) 最近傍マッチング (nearest-neighbor matching) やカーネル・マッチング (kernel matching),

層化マッチング (stratification matching), 半径マッチング (radius matching) などがある。

2) 段階補正の見直しと地方交付税の削減に関する詳細については宮崎 (2010) を参照されたい。

参 考 文 献

- Angrist, J. D. (1990), “Lifetime Earning and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records,” *American Economic Review*, Vol. 80, No. 5, pp. 313-335.
- Angrist, J. D. (2006), “Instrumental Variables Methods in Experimental Criminological Research: What, Why and How,” *Journal of Experimental Criminology*, Vol. 2, No. 1, pp. 23-44.
- Angrist, J. D., and W. N. Evans (1998), “Children and Their Parents’ Labor Supply: Evidence from Exogenous Variation in Family Size,” *American Economic Review*, Vol. 88, No. 3, pp. 450-477.
- Angrist, J. D., and A. B. Krueger (1991), “Does Compulsory Schooling Attendance Affect Schooling and Earnings?” *The Quarterly Journal of Economics*, Vol. 106, No. 4, pp. 979-1014.
- Angrist, J. D., and J. S. Pischke (2008), *Mostly Harmless Econometrics: An Empiricist’s Companion*, Princeton University.
- McClellan, M., B. J. McNeil, and J. P. Newhouse (1994), “Dose More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality?” *Journal of the American Medical Association*, Vol. 272, No. 11, pp. 859-866.
- Stock, J. H., and M. W. Watson (2010), *Introduction to Econometrics*, 3rd ed., Addison Wesley.
- Stock, J. H., J. H. Wright, and M. Yogo (2002), “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments,” *Journal of Business & Economic Statistics*, Vol. 20, No. 2, pp. 518-529.
- Wooldridge, J. M. (1995), “Score Diagnostics for Linear Models Estimated by Two Stage Least Squares,” *Advances in Econometric and Quantitative Economics: Essay in Honor of Professor C. R. Rao* (ed. by G. S. Maddala, P. C. B. Phillips, and T. N. Srinivasan), pp. 66-87, Willey-Blackwell.
- Wooldridge, J. M. (2006), *Introductory Economics: A Modern Approach*, 4th ed., South-Western.
- 総務省自治税務局 (2002-2008), 『市町村別決算状況調』 (http://www.soumu.go.jp/iken/kessan_jokyo_2.html).
- 地域開発研究所 (1993), 『地方生活圏要覧 平成5年版』 (建設省建設経済局事業調整官監修).
- 林正義 (2004), 「自治体合併の評価」, 『地方分権改革の経済学』 (土居丈朗編), pp. 139-155, 日本評論社.
- 宮崎毅 (2010), 「地方交付税改革が市町村合併に及ぼす影響：段階補正の見直しと地方交付税の削減」, 『日本経済研究』, No. 63, pp. 79-99.
- 宮崎毅 (2013), “Cost Reduction from Municipal Consolidation: Evidence from Japan,” 日本経済学会 2013年度春季大会.