

選挙研究における因果推論の研究動向

飯田 健・松林哲也

要旨：本稿は選挙研究における因果推論の方法を概観する。まず因果関係とは何を意味するのかについて日本の社会科学研究における代表的な知見をまとめ、その応用における問題点を指摘する。次に近年主流となりつつある潜在的結果にもとづいた因果的効果の定義を紹介し、その枠組みの中で観察データを用いた回帰分析の問題点を指摘する。次に厳密な因果推論を行うために必要となる方法とその応用例を紹介する。第一に、サーベイ実験やフィールド実験などの実験アプローチ、第二にイベントを利用した自然実験などの疑似実験アプローチ、第三にマッチングを用いた統計学的アプローチの特徴とその応用研究についてまとめる。

はじめに

社会科学研究の主要な目的の一つは変数間の因果関係を明らかにすることにある。選挙研究においてもリサーチクエスションはほとんどの場合、因果関係に関するものである。例えば、なぜ有権者は投票に行ったのか、なぜその候補者に投票したのか、経済状態についての認識は投票にどのような影響を及ぼすのか、教育程度は政治参加にどのような影響を及ぼすのか、党派性はなぜ形成されるのか、選挙制度は有権者の行動にどのような影響を及ぼすのか、など。選挙に関わる現象を理解しさらに理論を構築する上で、変数間の因果関係を解明することは欠かせない。

また因果関係の解明は実際の政治への政策提言を行う上でも役に立つ。例えば選挙競争や政策過程についてある望ましい状態が与えられたとき、それを実現するためにどのように選挙制度を変更すれば良いのか、あるいは選挙における投票率を上げるためには何をすれば良いのか、といった問題に答えるためには、因果的効果の理解とそれにもとづく予測が不可欠である。

本稿の目的は選挙研究における因果推論の研究動向を概観することにある。次節ではまず、因果推論の入門書として日本の社会科学教育で頻繁に用いられる高根正昭の『創造の方法学』における因果的効果の議論をまとめ、因果推論がこれまでどのように理解されてきたかを論じる。また高根の説明が選挙研究における因果推論に与えてきた影響について論じ、その限界を指摘する。

次に3節では、因果推論にまつわる従来の議論には限界があることを踏まえ、近年の社会科学研究において主流になりつつある潜在的結果にもとづく因果的効果の定義を紹介する。この過程で、観察データと回帰分析を組み合わせた方法による因果推論にはどのような問題が存在するかを明らかにする。4節では回帰分析を用いて厳密な因果推論を行うために必要となる方法を三つにまとめ、その特徴および応用例を順番に紹介する。第一にサーベイ実験やフィールド実験を用いた実験デザイン、第二に歴史的イベントを利用した自然実験などの疑似実験アプローチ、第三にマッチング手法を用いたその他の統計学的アプローチについて、それぞれ特徴および選挙研究への応用の可能性を論じる。そ

の際にはそれらの方法の背後にある仮定の重要性に注目する。

2. 従来の因果推論の考え方とその批判 —『創造の方法学』を超えて

どのようにして変数間の因果関係を特定するのかについては、経済学、政治学、社会学など社会科学の分野においてこれまで数々の議論がなされてきた。こうした議論はさまざまな点で主張を異にするとはいえ、基本的なコンセンサスも存在する。日本における政治学や社会学の場合、高根正昭・1979年『創造の方法学』（講談社現代新書）は因果推論の方法についてそうした最もオーソドックスな説明を提供しており、ほとんど古典的な地位を占めている⁽¹⁾。われわれの見たところ、この本の影響は選挙研究においても濃厚に見られる。まずはその因果推論の方法についての説明を見てみよう。

高根（1979）は、原因としての独立変数と結果としての従属変数とを結び付ける因果法則を確定する上で三つの条件を挙げている。第一に、「独立変数の変化が、従属変数の変化に先行する」（p.83）ということ。これは、ある変数が別の変数の変化の原因となるためには、最初の変数における変化がまず生じて、次に別の変数における変化が生じなければならないという直観にもとづいており、計量経済学のグレンジャー因果性（Granger 1969）の中心的な定義を構成しているものである。第二に、「両変数間の共変関係」（p.83）が存在すること。これは、独立変数の値が変化すれば、従属変数の値も変化するということを意味し、当然どちらかが一切変化せず一定であれば、共変関係を確かめることは不可能である。このことから、基本的に社会科学の営みは変化によって変化を説明することであり、「変化しないこと」すなわち「なぜ変化しないのか」を説明することはできないということがわかる。

そして第三の条件として、「他の重要な変数が、変化しない」（p.83）というものがある。こ

れはすなわち、従属変数に影響を与えうる他の変数の値を一定にした上で、関心の対象となる独立変数の値を動かし、それによって従属変数の値が変化するかを確認するということである。高根の例に則して考えてみると、ローソクの燃焼を説明する変数として酸素の他、温度、気圧、外気の性質など様々な要因があるとする。実際に酸素がローソクの燃焼の原因であると主張し、その因果法則を確定するためには、温度、気圧、外気の性質を一定にした上で、酸素の供給を止めることによって、ローソクの燃焼が止まるということを示さなければならない。もし酸素の供給のみならず、気圧までも同時に変化させてしまった場合、ローソクの焰が消えたところでどちらの要因によってその結果が引き起こされたのか、因果法則が確定できないのである。

これらの三つの条件を踏まえた上で、高根は実証研究における因果関係の検証方法について論じる。ここで最も問題となるのは第三の条件、つまりいかに他の重要な変数の影響を排除するか、ということである。まず実験室実験の場合、それは無作為化によって実現される。実験の被験者を処置群と統制群とに無作為に振り分け、二つの出来るかぎり等質の集団を作り出す。これらの集団は関心の対象となる独立変数の値だけが異なる。もしその違いによって従属変数の値に違いがもたらされるのであれば、他の変数は一定に保たれているので、その独立変数は従属変数に影響を及ぼしているということが言える。つまり無作為化を行うことで、関心の対象である因果関係に影響を与えうる他の変数の影響を除去するのである。

ただし、こうした実験的手法は必ずしも社会科学の研究関心の全てについて適用可能なものではない。そこで、高根はサーベイデータ（およびアグリゲートデータ）の多変量解析による統計的統制（statistical control）の考え方を紹介する。ここでは、独立変数と従属変数の二変数間の共変関係が、別の重要な変数（統制変数）

の値によってグループ分けした場合にもそれぞれのグループ内で見られるかを検証することで、「偽の関係」の可能性を排除し、二変数間の因果法則を確定しようとしている。高根が例として挙げるのは、1950年代のアメリカにおいて「赤狩り」を推進したマッカーシーに対する国民の態度を従属変数に、政治的寛容を独立変数にした分析である。まず、これら二つの変数間の関係についてみると両者の間には、政治的寛容がマッカーシーに対する態度に時間的に先行しているという前提の下、政治的寛容が高い人ほどマッカーシーに反対するという明らかな共変関係が見られた。しかし、従属変数に影響を与えうる別の重要な変数として教育程度が考えられるかもしれない。そこで教育程度を統制変数として考慮し、学歴ごとにグループ分けした中で政治的寛容とマッカーシーに対する態度の関係を見ると、ほとんどその二変数の関係は認められなくなった。すなわち、これは「教育程度が高いと政治的寛容が増す」、「教育程度が高いとマッカーシーに反対する」という二つの関係が存在することで、一見政治的寛容が高い人ほどマッカーシーに反対するかのように見えているに過ぎないということである。当初観察された政治的寛容とマッカーシーに対する態度の関係は偽の関係だったのである。これは計量経済学の教科書での重回帰分析の説明における統計的統制と同じ考え方であり、重要な統制変数がモデルに含まれない場合、独立変数と従属変数との間の因果関係の検証が正しく行われないということを意味する。

因果推論をめぐる以上のような考え方は、統制された比較(controlled comparison)として、従来の社会科学における因果推論の方法論に登場するものである。例えば、社会科学の方法論の教科書として名高い King, Keohane, and Verba (1994) は、統制変数に関するアドバイスとして、理論モデルに依拠して従属変数の原因と考えられるあらゆる変数を特定し、さらにその変数群の中で独立変数と従属変数と相関関係にあ

る変数を選択しそれらの影響を統制すべきだと述べている (p.174, 邦訳書 p.208)⁽²⁾。また、同じく政治学の方法論の教科書として有名な Van Evera (1997) でも、統制された比較の一つとして、関心の対象となる変数の値が異なる一方、それ以外の従属変数に影響を与えうる変数の値は一定であるようなケースを選択する差異法(method of difference)が紹介されている。その他、Lave and March (1975), Isaak (1975) などのより古い社会科学、政治学方法論の教科書にも同様の記述が見られる。

しかし近年指摘されるように、上記のような統制変数による因果関係の検証がもたらす弊害も存在する。それは、高根のいう第三の条件を気にするあまり、あまりに多くの変数が回帰分析モデルに投入されるという問題である。高根の議論は独立変数以外の第三の変数群の影響を統制する必要性を指摘しているが、何を統制すべきかについて明確な指針を与えてくれない。従属変数に影響を与える可能性をもつ全変数を統制する必要があると論じているようにも読み取れる。また King, Keohane, and Verba (1994) は何を統制すべきかについて指針を示しつつも、統制変数のリストは長大なものになる可能性も示唆している。結果として、統制された比較という考え方はそれ自体統制変数をできるだけたくさんモデルに含めることを勧めるものではないにせよ、現実問題として統計モデルに多くの変数が投入される状況を正当化している。それは実験では無作為化を用いることで、研究者の意識に上らない変数の影響まで統制できるのに対し、実験データを用いない多変量解析では無作為化が不可能なため、とにかく多くの統制変数を含めることで他の変数を統制し少しでも正確に因果関係の検証を行いたいという気持ちの表れなのかもしれない。このような事情を背景に、従属変数に影響を与えらるすべての変数を統制変数として回帰式に投入するということが頻繁に行われるのである。

われわれの見るところ、この問題はとりわけ

サーベイデータから比較的多くの変数が入手可能な日本の選挙研究において深刻である。すなわち、分析結果を示した表には20を超える独立変数が示され、時にそれは2ページにもわたって印刷される。また、同じ変数を従属変数とするモデルがいくつも推定され、統制変数の数や種類により結果がどう変わるか頑健性のチェックが行われる。モデルに投入された変数には理論的根拠が弱いものもある。とにかく、統制変数は多いに越したことはない、多くても問題がない、かのようである。

しかし多くの統制変数がモデルに含まれれば含まれるほど、関心の主たる対象となる独立変数と従属変数の因果関係の検証がより厳密にできるという保証はない。それは一つには、真のモデルというものを経験的に特定することは不可能だからである。つまりどのモデルが正しいかは、経験的には分らないのである。古い世代の社会科学の研究者は、決定係数、対数尤度、推定値の標準誤差などの「当てはまりの良さ」(goodness of fit)の尺度を用いて、モデルを比較することで真のモデルなるもの、あるいはより真のモデルに近いものが分かると考えていた。Achen (2002) はこうした古い世代の研究者を「当てはまり狂」(“fitness buffs”), さらに長い独立変数のリストをもつ回帰式を「ゴミ缶回帰」(“garbage-can regressions”)と呼ぶ。またそもそも、こうした考えは「回帰モデルは正確に特定されている」、すなわち「経験分析で使用されるモデルにおいて特定バイアスや誤差は存在しない」とのガウスの古典的直線回帰モデルの前提の一つに違反している⁽³⁾。

さらに後に示すように、サーベイデータの分析における統制変数の使用によって、実験における無作為化と同じ効果を得られるとの高根の主張には、実は重要な前提が隠されており、この前提が満たされることは非常に稀である。

このような状況を克服すべく、Achen (2002) や、Achen 自身が主唱者の一人である EITM (Empirical Implication of Theoretical Models) 運

動に関係する研究者たちは因果関係を検証するための計量モデルを特定する上での理論、とりわけフォーマルセオリーの利用を勧める⁽⁴⁾。また、コンピュータ上の世界に一つのパラメータの値のみを変化させたパラレルワールドを複数作り出し、それらを比較することで独立変数の従属変数への因果的效果を見ようとするシミュレーションの手法を提唱する研究者もいる⁽⁵⁾。しかし、本稿ではあくまで、選挙研究で主流を占める経験的手法における近年の因果的效果の検証法の発展について概観する。そこでは上で見たような従来からの因果推論についての考え方に依拠しつつも因果的效果の定義についての新しい考え方が見られるようになってきており、それにもとづいて因果的效果を正確に推定するためのさまざまな方法が提案されている。以下、若干重複もあるが、数式を使って詳細に見ていくことで問題の性質を確認し、どのようにして近年の選挙研究において因果推論をめぐる問題が解決されているのか議論する。また回帰分析における統制変数の意味を問い直し、何を統制すべきかを検討する。

3. 潜在的結果にもとづく因果的效果

本節ではまず始めに統計学や計量経済学で主流となりつつある潜在的結果にもとづいた因果的效果の定義を紹介する⁽⁶⁾。次に定義された因果的效果をどのように測定するかを議論する。その際には無作為化を用いた方法と、選挙研究で頻繁に用いられるクロスセクショナルデータと回帰分析を組み合わせた方法を対比することによって、後者の問題点を明らかにする⁽⁷⁾。

まず個人 i を観察単位とした結果変数 Y_i に対する因果的效果を考える。この結果変数に影響を与えと考えられる2項処置変数 $D_i = \{0,1\}$ があるとする⁽⁸⁾。この場合、潜在的結果は以下のように表せる：

$$Y_i = \begin{cases} Y_i^1 & \text{if } D_i = 1 \\ Y_i^0 & \text{if } D_i = 0 \end{cases} \quad (1)$$

(1)において Y_i^1 は処置を受けた場合の個人 i の状態を意味し、 Y_i^0 は処置を受けていない場合の状態を意味する。例として大学教育が政治知識量に影響を与えるという因果関係を想定する。 D_i が大卒以上 (=1) かそれ以外 (=0) で測られる場合、 Y_i^1 はある個人が大学教育を受けた場合の政治知識量を意味し、 Y_i^0 は同一個人が大学教育を受けなかった場合の政治知識量を意味する。さらに(1)は以下のように表すこともできる。

$$\begin{aligned} Y_i &= Y_i^0 + (Y_i^1 - Y_i^0) D_i \\ \delta_i &= Y_i^1 - Y_i^0 \end{aligned} \quad (2)$$

(2)では Y_i に対する D_i の因果的効果 δ_i は $Y_i^1 - Y_i^0$ と表されている。大学で教育を受けることにより政治知識量が変化した場合、その変化量は δ_i と等しくなる。同一個人の結果の比較を行っているので、処置状態以外の性質は全て一定である。よって結果変数の差は処置変数の因果的効果である δ_i によって生み出されたと結論できる。

この因果的効果の定義は洗練されており、また他の定義に比べて様々な長所がある（詳しくは Imbens and Wooldridge 2009）。しかしこの定義にもとづいた因果的効果を実際に測定することはできない。なぜなら、理論上は同一個人に関する二つの潜在的結果、 Y_i^1 および Y_i^0 、を定義することができるとはいえ、現実にはこれら二つの結果を同時に観察することはできないからである。 Y_i^1 が観察された場合、 Y_i^0 は同時に観察されえないし、その逆も同様である。この意味で潜在的結果は相互に排他的である。これはここでの例の場合、ある個人が大学進学と非進学の両方を同時に経験することは不可能であるということの意味する。つまり同一個人について潜在的結果は事実 (factual) と反事実 (counterfactual) に分割され、その両方を観察することはできないのである。よって δ_i を測定することはできない。どちらかの潜在的結果が観察できないことは欠損データの問題と見なすことも

できる⁽⁹⁾。

個人単位での δ_i の観察は不可能であるので、集団単位での因果的効果を考えてみる⁽¹⁰⁾。個人 $i = 1 \cdots N$ を含む母集団における平均因果的効果は以下のように定義できる⁽¹¹⁾：

$$\begin{aligned} E(\delta_i) &= E[Y_i^1 - Y_i^0] \\ &= E[Y_i^1] - E[Y_i^0] \end{aligned} \quad (3)$$

この式で $E[Y_i^1]$ は母集団内の全個人が処置を受けた状態の平均値を意味し、 $E[Y_i^0]$ は全個人が処置を受けていない状態の平均値を意味する。個人単位での定義と同じように、(3)において $E[Y_i^1]$ と $E[Y_i^0]$ の違いは処置の有無による違いのみを反映している。よって平均値の違いは処置変数の因果的効果のみを反映していると言える。ところがこの場合も $E[Y_i^1]$ と $E[Y_i^0]$ のどちらかが反事実となるため両方を同時に観察することは不可能である。例えばある集団が大学に入学し4年後に卒業する、または同じ集団が大学に進学せず4年間過ぐすという状況が同時に起こることはありえない。よって $E[Y_i^1] - E[Y_i^0]$ を直接測定することはできない。

(3)にもとづく集団レベルでの因果的効果を測定するためには、観察可能なデータを使う必要がある。処置変数 D_i の値にもとづいて母集団を2つの集団に分割することができるとする。処置群に属し処置を受けた集団における結果変数の平均値を $E[Y_i^1 | D_i = 1]$ 、統制群に属し処置を受けていない集団における結果変数の平均値を $E[Y_i^0 | D_i = 0]$ と定義する。これらの平均の差を取ることで、平均因果的効果を以下のように定義することができる：

$$E^*(\delta_i) = E[Y_i^1 | D_i = 1] - E[Y_i^0 | D_i = 0] \quad (4)$$

母集団を処置群と統制群に分割できるのなら、この $E^*(\delta_i)$ は観察可能である。

この $E^*(\delta_i)$ と $E(\delta_i)$ が等しくなるにはどのような条件が必要だろうか。 $E^*(\delta_i)$ と $E(\delta_i)$ の差をとると、

$$E^*(\delta_i) - E(\delta_i) = E[Y_i^1 | D_i = 1] - E[Y_i^0 | D_i = 0] - E[Y_i^1 - Y_i^0] \quad (5)$$

となる。さらに(5)の右辺に $E[Y_i^0 | D_i = 1] - E[Y_i^0 | D_i = 1]$ を加えると

$$\begin{aligned} E^*(\delta_i) - E(\delta_i) &= E[Y_i^1 | D_i = 1] - E[Y_i^0 | D_i = 0] \\ &\quad - E[Y_i^1 - Y_i^0] + E[Y_i^0 | D_i = 1] \\ &\quad - E[Y_i^0 | D_i = 1] \\ &= E[Y_i^0 | D_i = 1] - E[Y_i^0 | D_i = 1] \\ &\quad - E[Y_i^1 - Y_i^0 | D_i = 1] - E[Y_i^1 - Y_i^0] \end{aligned} \quad (6)$$

と書き換えることができる。(6)から $E^*(\delta_i)$ と $E(\delta_i)$ が等しくなるには以下の二つの条件が必要となることがわかる。

$$E[Y_i^1 - Y_i^0 | D_i = 1] = E[Y_i^1 - Y_i^0] \quad (7)$$

$$E[Y_i^0 | D_i = 0] = E[Y_i^0 | D_i = 1] \quad (8)$$

それぞれの条件についてその意味を詳しく見てみよう。等式(7)は D_i の因果的効果が母集団の全ての構成要員に対して均一であることを意味している。特に処置群と統制群の間で処置を受けた場合の因果的効果に差が見られないことを指している⁽¹²⁾。本稿では議論の簡略化のためこの仮定は満たされているとみなす。一方、等式(8)は処置群と統制群の間で処置を受けていない場合に平均値の差がないことを意味する。処置の有無にかかわらず処置群と統制群の間で平均値に何らかの差があれば、この仮定は満たされない。

(8)の仮定を満たす最も確実な方法は処置の割り当ての無作為化である。処置の割り当ての無作為化とは、コインを投げて裏表を決めるような完全に作為のないメカニズムによって母集団に含まれる各観察単位を処置群か統制群かに割り当てることを意味する。適切に無作為化が行われた場合、処置の割り当てと潜在的結果とが完全に独立する。結果として処置変数の値から潜在的結果の予測を行うことは不可能となり、(8)の仮定が満たされていると想定することができる。大学教育と政治知識量の例で考えると、高校を卒業する人々について大学に進学するか

どうかを観察対象の個人ごとに研究者が無作為に決定できるとすれば、4年後に大学進学者と非進学者の平均政治知識量を比較することにより、政治知識量に対する大学教育の因果的効果を測定することができる⁽¹³⁾。無作為化が用いられた場合、処置群と統制群の性質は政治知識量を含め全ての点において相似していると想定できる。よって処置群と統制群の平均値の違いは処置変数の因果的効果を反映していると結論できるのである⁽¹⁴⁾。

次に無作為化を用いず、大学進学者と大学非進学者の政治知識量を単純に比較するという状況を考えてみよう。重要なのは大学進学者と非進学者の間にはそもそも進学前から政治知識量に違いがあるという可能性である。例えば大学進学を考える高校生は政治知識量や政治関心度が高く、政治や社会について何かを学びたいと考えるために進学を希望する可能性がある。つまり大学教育を受ける受けないに関わらず、大学進学者のほうが非進学者に比べて進学前から政治知識量が高い可能性があるのである。大学進学という因果的効果が存在しないにもかかわらず、すでに知識量に差が存在する。つまり処置を受けない状態で処置群と統制群の平均値に何らかの差があるので、(8)は満たされない。観察対象の個人の性質にもとづいて処置状態が意図的に選択されることにより結果変数に差異が生じるので、これは計量経済学では選択バイアスの問題と呼ばれている。

では無作為化にもとづかない研究デザインを用いる場合、(8)の仮定を満たし因果的効果を測定することは可能なのだろうか。特に選挙研究における主要な研究デザインであるクロスセクショナルデータを用いた回帰分析は選択バイアスをなくし、因果的効果を測定するためには有効な方法なのだろうか。

上記の問いに答えるため、まず結果変数に対する処置の効果は母集団において一定だという仮定の下で、式(1)を以下のように書き換える (Angrist and Pischke 2009)：

$$Y_i = a + \delta D_i + \eta_i \quad (9)$$

式(9)において $E(Y_i^0) = a = Y_i^0 + \eta_i$ であり、また $\delta = Y_i^1 - Y_i^0$ である。処置変数の状態によって Y_i の期待値を求めると、

$$\begin{aligned} E[Y_i|D_i = 1] &= a + \delta + E[\eta_i|D_i = 1] \\ E[Y_i|D_i = 0] &= a + E[\eta_i|D_i = 0] \end{aligned} \quad (10)$$

となる。この差をとると

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = \delta + E[\eta_i|D_i = 1] - E[\eta_i|D_i = 0] \quad (11)$$

となる。もし式中の $E[\eta_i|D_i = 1] - E[\eta_i|D_i = 0]$ が 0 である場合、因果的効果 δ は正確に観察されることがわかる。さらに

$$E[\eta_i|D_i = 1] - E[\eta_i|D_i = 0] = E[Y_i^0|D_i = 1] - E[Y_i^0|D_i = 0] \quad (12)$$

なので、もし潜在的結果と処置変数が独立しているならば、式(9)における期待値の差は因果的効果を表していると結論づけることができる。式(12)の右辺の $E[Y_i^0|D_i = 1] - E[Y_i^0|D_i = 0]$ が 0 という仮定は(8)と同一であり、選択バイアスがないということを意味する。つまり回帰分析の枠組みでも因果的効果の測定には $E[Y_i^0|D_i = 1] - E[Y_i^0|D_i = 0]$ という仮定を満足させる必要があることがわかる。もし政治知識量の高い人ほど大学教育を求めるならば、大学教育を受けない場合でも $E[Y_i^0|D_i = 1] - E[Y_i^0|D_i = 0]$ の値は正になる。この場合、大学教育の効果は実際よりも過大に測定されてしまう。

さらに $E[\eta_i|D_i = 1] - E[\eta_i|D_i = 0]$ が 0 という条件は、計量経済学における外生性の仮定、 $E(\eta_i|D_i) = 0$ と同じ意味を持つ。もし選択バイアスがなく誤差項と説明変数の間に全く関係がない場合、 D_i の因果的効果は不偏性あるいは一貫性をもって推定されるのである。

実験の場合は選択バイアスが存在しないという仮定は無作為化によって達成されるが、上記のような回帰式を用いた分析ではこの仮定はどのように満たされるのか。最も一般的なのは回

帰式(9)の右辺に統制変数を含める方法である。処置変数と結果変数の両方に影響を与えると考えられる全ての変数ベクトル X_i を特定し、回帰式に投入することで以下の仮定を満たすことができる：

$$(Y_i^0, Y_i^1) \perp D_i | X_i \quad (13)$$

X_i が考慮されているという条件付きで、潜在的結果と処置変数は独立していると仮定するのである。この仮定は条件つき独立性仮定 (Conditional Independence Assumption, CIA) と呼ばれる (Angrist and Pischke 2008)。この仮定の下では因果的効果は

$$E[Y_i|X_i, D_i = 1] - E[Y_i|X_i, D_i = 0] = \delta \quad (14)$$

と表すことができる。つまり選択バイアスはないと想定できる。また回帰式は

$$Y_i = a + \delta D_i + X_i\beta + \eta_i \quad (15)$$

となる。式(15)で β は各統制変数の Y_i に対する統制変数 X_i の係数を表している。

政治知識量の例で上記の回帰式の意味を考えてみる。大学教育が政治知識量に対しどのような効果を与えるのかを調査する場合、実験を行うことは現実的ではない。そこである年に行われた世論調査のデータを用いて、大学教育の因果的効果の検証を行う研究デザインを選択するとしよう。この場合、大学進学は無作為に決定されないため先ほど述べたような選択バイアスが存在する可能性が高い。たとえば大学進学前の政治的関心と親の経済的地位の二変数のみが大学進学と政治知識量に影響を与えるとしよう。この場合これらの変数を統制変数として回帰式に含めることにより、条件つき独立性仮定を満たすことができる。上記の回帰式(15)を推定することにより、大学教育を表す変数 D_i の係数 δ は不偏性あるいは一貫性をもって推定される。

ここで重要なのは統制変数 X_i は何かということである。この X_i は処置変数と結果変数の両方に相関している変数でなければならない。

詳しく言うと、統制変数は処置変数の状態に影響を与え、なおかつ結果変数に影響を与えるような変数である。上記の例の場合、大学進学前の政治関心度は大学進学決定に影響を与え、しかも結果変数とも相関する可能性がある。よって進学前の関心度は統制される必要がある。一方で、結果変数のみに影響を与えるような変数は選択バイアスを生み出さないで、統制変数として回帰式に投入する必要はない。例えば婚姻状態は政治知識量に影響を与える可能性があるが、(多くの人が大学卒業後に結婚するという前提で)婚姻状態が大学進学決定に影響を与えることはない。よって婚姻状態の影響を統制する必要はない。さらに処置変数と結果変数の因果メカニズムの間に位置するような変数は統制変数として回帰式に含めるべきではない(King, Keohane, and Verba 1994)。もしそのような変数を回帰式に含めた場合、平均因果的効果の正確な推定は難しくなる。

回帰式に統制変数を含めて条件つき独立性仮定を担保する研究は多くみられる。ところが特にクロスセクショナルデータを用いる研究デザインの場合、条件つき独立性仮定を満たすのは不可能ではないが非常に難しい。つまり $E[\eta_i | D_i = 1] - E[\eta_i | D_i = 0] \neq 0$ あるいは $E(\eta_i | D_i) \neq 0$ である可能性が高い。結果として除外変数バイアス (omitted variable bias) と呼ばれる問題を生み出し、推定された係数 δ は不偏性あるいは一貫性を持たなくなる。

除外変数バイアスには二つの原因がある。一つ目は処置変数と結果変数に影響を与えると考えられる全ての変数を特定することが困難なことにある。政治知識量の例だと、政治関心度や親の経済的地位以外にも様々な変数が大学進学や政治知識量に影響を与えると想像できる。しかし大学進学や政治知識量に関する理論が完全に構築されていないので、可能性のある全ての変数を特定し、統制変数として回帰式に投入するのは難しいだろう。

二つ目は観察が不可能な変数の存在である。

政治知識量の例だと、個人の知能程度がそれにあたる。知能を測るためには IQ テストなど様々な指標が開発されてきたが、どの指標も知能の全側面の特徴を測ることはできないだろう。よって知能程度は観察不可能だと見なさなければならぬ。知能程度は大学進学にも政治知識量にも影響を与える可能性が高い。知能の高い個人は大学に進学する可能性が高く、また記憶力が高いため政治知識量も大きいだろう。よって知能程度は統制変数として回帰式に投入する必要があるが、観察不可能であるためそれは不可能になる。もしその影響を統制できない場合、知能の効果は誤差項に吸収されるため、大学教育を表す変数 D_i の係数 δ は偏りをもって推定されてしまう。

さらに除外変数バイアスの問題を別の角度から考えてみる。実験デザインでは無作為化を用いることにより、処置の状態を除き潜在的結果を含めて全ての点で近似した性質を持つと思われる二つの集団、つまり処置群と統制群を作り出す。処置の実行後、もし結果の平均値に何らかの差があれば、その差は処置の因果的効果によって生み出されたと結論づける。一方非実験データと回帰分析を組み合わせた研究デザインにおいては、そのような平均値の比較は因果的効果を必ずしも明らかにしてくれない。処置群と統制群の間には処置以外の違いがあるかもしれないからである。もしその違いが観察可能な性質によるものならば、統制変数として回帰式に投入することでその違いを統制することができる。一方でもし処置群と統制群には観察不可能な違いがある場合は、統制というアプローチを用いることはできない。よって δ を推定してもそれが因果的効果を示しているのかはわからないのである。

このように統制変数と回帰分析を組み合わせると因果的効果を測定するのは非常に難しい。この方法は因果的効果の測定には不向きであると指摘する研究者もいる (例えば Imbens and Wooldridge 2009)。選挙研究の多くは因果的効

果を測定することを目指して統計分析を行う。ところが選択バイアスがないという重要な仮定にあまり注意を払われてこなかった。また統制変数を選択する際には、処置変数と結果変数の両方に影響を与える変数の影響を排除する必要があるということが明確に認識されていないことが多い。選択バイアスがないという仮定を満たすことなしには因果的効果を有効に測定しているとは言えず、このような状態は結果として選挙に関する理論を構築する障壁をうみだしているかもしれない⁽¹⁵⁾。次節では選択バイアスがないという仮定を満たし、因果的効果の測定を可能にするために有用な研究デザインを概観する。適切な研究デザインを見つけることが因果的効果の測定の鍵となる (Freedman 1991)。

4. 因果的効果の識別方法

本節では潜在的結果にもとづく因果的効果を測定する上で望ましいと思われる研究デザインを3つのグループに分け、それぞれについてその特徴と主要な方法をまとめる。また選挙研究における各方法の応用例を紹介する。研究デザインのグループ分けは処置変数の無作為化の程度にもとづく。一つ目のグループは研究者自身が処置変数の処置の割り当てを管理し、その無作為化を行う実験デザインである。実験デザインは実験室実験、フィールド実験、そしてサーベイ実験を含む。二つ目のグループは研究者が処置変数の処置の割り当てを管理することができないが、その割り当てが何らかの理由で（少なくとも結果に対して）無作為に近い状態で行われていると想定できる疑似実験デザインである。疑似実験デザインはパネル分析、操作変数法、不連続回帰デザインを含む。三つ目は研究者が処置変数の処置の割り当てを管理することもまた無作為化が行われていると想定することもできないが、処置群と統制群について性質の近似したグループを作る研究デザインである。このデザインはマッチング手法を含む。

4. 1. 実験デザイン

研究者が処置変数の処置の割り当てを管理し調整できるような研究デザインが実験である。処置の割り当ての際には無作為化を用いることで選択バイアスがないという仮定を満たし、平均因果的効果を測定することを試みる。実験を用いた研究は政治学全般で増加傾向にあり、今後重要な役割を果たすべきであると考えられている (Druckman, Green, Kuklinski, and Lupia 2006)。近年までは大学生などを対象として人工的な状況で実施される実験室実験が主要な方法であったが、最近は因果的効果が発生する実際の状況に近い環境で実施されるフィールド実験も増えている (Green and Gerber 2003)。また世論調査において調査方法や設問形式を無作為に回答者に割り当てるサーベイ実験も選挙研究における主要な方法の一つとなっている。ここではフィールド実験とサーベイ実験について日本での研究例を紹介する。

Horiuchi, Imai, Taniguchi (2007) は政治情報量が投票参加に与える影響を調べるために、インターネットを利用した実験を行った。彼らは2004年の参議院選挙における年金制度改革に注目し、年金制度改革について自民党と民主党のマニフェストの情報を持つ有権者ほど投票に参加する確率が高いのかどうかを検証した。その際に標本を無作為に分割し、自民党に関する情報のみに接するグループ、民主党に関する情報のみに接するグループ、両党の情報に接するグループ、そしてそれらの情報に全く接しないグループを作った。そしてこれらのグループについて投票率を比較したのである。無作為化を用いることによって情報の入手をめぐる選択バイアスをなくし、情報量以外の面で性質の近似したグループを作り出した。そしてそれらのグループ間の投票率の比較を行っている。

西澤・栗山 (2009) は、面接世論調査における社会的期待迎合バイアスが政治参加に関する回答にどのような影響を与えるかを調べるためにサーベイ実験を利用した。従来の質問紙を使

って調査員が回答を記録する調査方法では、調査員が介在することにより回答のプライバシーが守られないため社会的期待迎合バイアスが高くなると考えられる。一方でパソコンを利用し回答者が自ら回答を記録する調査方法では、調査員が介在しないためプライバシーが守られるのでバイアスが低くなると考えられる。そこで、この二つの調査方法を無作為に回答者に割り当てることによって、調査方法の違いが生み出すプライバシーの保護の程度が政治参加についての質問に対する回答にどのような影響を与えるかを調べた。調査方法は無作為に割り振られているので、選択バイアスの問題は生じず、調査方法の異なる二つの集団は比較可能であると想定できる。

実験の実施やデータ分析には様々な手法が存在する。今井（2007）は社会科学の実験デザインでは単純無作為化や完全無作為化が用いられることが多いが、ブロッキングやマッチングといった手法を用いれば無作為化を行う前に観察可能な性質について処置群と統制群の違いを減らすことができると指摘している。また有権者を標本として利用する実験では、各個人に処置の状態を強制できないことが多い。処置群に割り振られた個人が処置を受けなかったり、また統制群に割り振られた個人が自らの意思で処置を受けてしまうような状況が発生する。あるいは実験から脱落してしまう人々も存在する。実験データの分析の際にはこのような不遵守や欠損データの問題も考慮に入れる必要があることを今井（2007）は指摘している。上記の Horriuchi, Imai, Taniguchi (2007) はこのような問題を考慮した上で、処置変数の因果的効果を測定している⁽¹⁶⁾。

因果的効果の測定には実験は最も有効な方法であるが、その一方で実験の成果の一般化には注意を要することが多い。内的妥当性のために外的妥当性を犠牲にしていると言われることが多いのである。実験室実験やフィールド実験は実験の観察対象の数や事例に限られてしまうた

め、母集団全体における因果的効果を推定しているのかは必ずしも明らかではない。また多くの場合実験は特殊な状況で行われることが多い。例えばサーベイ実験では一般化可能な標本を用いることが多いが、その一方でサーベイ実験で使われる設問や場面描写は有権者が日常生活で直面するような状況を作り出せているかは疑わしい。Barabas and Jerit (2010) はサーベイ実験の実験プロトコルに近似した現実の事例を利用することによって、サーベイ実験によって推定された因果的効果は現実の事例では必ずしも観察できなかったことを報告している。このように実験デザインを用いる場合やその実証結果を解釈する場面では外的妥当性に注意を払う必要がありそうである。

4. 2. 疑似実験デザイン

実験デザインと疑似実験デザインの違いは研究者が処置変数の処置の割り当てを管理できるかどうかである。疑似実験デザインでは研究者は処置変数の割り当てを決定することはできない。しかし、何らかの理由でその割り当てが（特に結果変数に対して）無作為に近い状況で行われている事例を利用することによって、近似した性質を持つ処置群と統制群を作り出す。疑似実験デザインは自然あるいは政治経済環境が作り出した無作為性を利用するので、自然実験とも呼ばれる。本節では疑似実験デザインを実際に利用する手法として、パネル分析、操作変数法、そして不連続回帰デザインを紹介する⁽¹⁷⁾。

パネル分析の特徴は同一の観察単位内における処置変数の変動を利用することによって、観察可能な変数及び観察不可能な変数の両方の影響を統制することにある。処置変数の変動は完全な無作為化によってもたらされるわけではないが、少なくとも観察単位の行為者にとってはその変動を調整できないような状況を利用する。クロスセクショナル分析との大きな違いは、観察不可能な変数についてもその影響を統

制できることにある。まず以下のような回帰式を考える：

$$Y_i = \alpha + \delta D_i + X_i\beta + W_i\gamma + \eta_i \quad (16)$$

式(16)において D_i は仮説の検証に必要な処置変数, X_i は観察可能な説明変数, そして W_i は観察不可能な説明変数とする。右辺に含まれるこれらの変数は互いに相関しているとする。観察不可能な W_i を無視して(16)を推定した場合, W_i は誤差項に含まれる。結果として $E(\eta_i|D_i) \neq 0$ となり, δ は不偏/一致推定量とならない。観察不可能な変数について選択バイアスが生じると言うこともできる。

この問題の解決のために, 例として同一観察単位について時系列の比較を試みる。 Y_i , D_i , そして X_i が2つの時点 $t = (1, 2)$ において観察可能だとする。またこれらの変数は2つの時点でその値に変動が見られるとする。観察不可能な W_i は二つの時点の間に変動はないとする。この場合, 各時点について回帰式(16)を以下のように書き直せる：

$$Y_{i1} = \alpha + \delta D_{i1} + X_{i1}\beta + W_i\gamma + \eta_{i1} \quad (17)$$

$$Y_{i2} = \alpha + \delta D_{i2} + X_{i2}\beta + W_i\gamma + \eta_{i2} \quad (18)$$

(17)と(18)の差を取ると,

$$\Delta Y_i = \alpha + \delta \Delta D_i + \Delta X_i\beta + \varepsilon_i \quad (19)$$

となる。(19)において $\Delta Y_i = Y_{i1} - Y_{i2}$, $\Delta D_i = D_{i1} - D_{i2}$, $\Delta X_i = X_{i1} - X_{i2}$, $\varepsilon_i = \eta_{i1} - \eta_{i2}$ である。 W_i は変動がないため, 差を取ることで相殺されている。 D_i と相関関係にあり, 二つの時点で変動する性質が全て変数 X_i に含まれるのであれば, $E(\varepsilon_i|D_i) = 0$ となり処置変数の係数 δ は不偏/一致推定量となる。

この方法は同一の観察単位について時系列の比較を行うことにより, 選択バイアスの問題をなくそうとしている。重要な仮定は二つの時点で観察不可能な変数の変動がないことである。もしそのような変動が存在し, それが処置変数の変動と関連するのであれば, この方法は有効

ではない。この仮定を満たす方法の一つは二つの時点の間隔をできるだけ短くすることである。たとえば人間の好みや能力などが観察不可能な変数として問題となっている状況を考えてみる。好みや能力は短期間には変動しないと想定できるので, (19)を推定することによりそれらの変数の影響を排除できるだろう。

上記の説明では比較を二つの時点に限った。一般化を目的としてより長期間にわたりデータを集め, 時点の数を増やすことも可能である。そのような場合は観察不可能な変数に変動がないという仮定を満たすことが難しくなるかもしれない。また差を取るのではなく, $W_i\gamma$ の代わりに各観察単位について固定効果を投入することにより(19)による推定と同じ結果を得ることもできる⁽¹⁸⁾。

このような方法を使った研究例として Horiuchi and Saito (2003) や Gentzrow (2006) をあげることができる。Horiuchi and Saito (2003) は日本の1990年代の選挙制度改革を利用して一票の重みが政策形成にどのような影響を与えるかを検証した。彼らの分析は中選挙区制から小選挙区制に移行することにより一票の重みが増した市町村では, 国から受け取る一人当りの補助金の総額が増加したことを示している。Horiuchi and Saito は(19)のような回帰式を推定することによって同一市町村内データの短期間の時系列比較を行った。この方法により観察可能な変数および観察不可能な変数の影響を統制することに成功したと言える。

Gentzrow (2006) はアメリカの1940年代から1970年にかけてのテレビ放送の開始時期の地域間の違いを利用して, テレビ放送が投票率に与える影響を検証した。テレビ放送の導入の結果, 政治情報の供給が少なくなり, 結果として政治知識量や投票率が低下したことを示している。Gentzrow (2006) の場合も同一地域内のテレビ放送導入前後の投票率の時系列比較を行うことにより, 比較可能な観察単位を作り出している。テレビ放送の導入は必ずしも外生的なシ

ショックでない可能性はあるが、歴史的背景や統制変数を用いて選択バイアスを軽減しようとしている。

時系列の比較だけでなく、同一時点での地域内の処置変数の変動を利用することも可能である。候補者の質や現職であることが得票率にどのような効果を与えるかを推定するために、Hirano and Snyder (2009) はアメリカの州議会選挙における同一選挙区内の同一政党の候補者の属性の違いを利用した。例えば2議員が選出される選挙区で、有権者が2票投じることができる場合を考える。この場合同一政党からの候補者の得票率に何らかの違いが見られた場合、その違いは候補者の属性の違いが原因であると結論できる。というのも同一選挙区内の同一政党の候補者を比較することにより、選挙区の特徴や政党が得票率に与える影響を統制しているからである。さらに同じ対の候補者が同一選挙区で2回の選挙に立候補し、最初の選挙では片方は現職、もう片方は現職ではなく、次の選挙では両候補者とも現職であったとする。この場合、候補者の質は時間を経ても大きく変わらないと想定できることから、1回目の得票率の差と2回目の得票率の差を比較することによって、得票率に対する現職効果を推定することができる。

操作変数法は経済学では古くから使われてきた手法である。まずその手法について簡単にまとめる。(9)のような回帰式を推定しようとするが、選択バイアスの問題が存在すると仮定する。この問題に対処するために、 D_i とは強い関係にあるが η_i とは独立している変数 Z_i があるとする。この Z_i を操作変数として用い、 D_i の係数 δ について以下の式を導ける：

$$\delta_{IV} = \frac{cov(Y_i, Z_i)}{cov(D_i, Z_i)} = \frac{cov(D_i, Z_i)\delta + cov(\eta_i, Z_i)}{cov(D_i, Z_i)} \quad (20)$$

δ_{IV} が一致推定量となるためには二つの仮定が必要である。一つ目の仮定は D_i と Z_i の間に強い相関関係が存在することである。二つ目の仮

定は Z_i と Y_i とが独立していることである。この仮定が満たされた場合 $cov(\eta_i, Z_i) = 0$ となり、 $\delta_{IV} = \delta$ となる。

操作変数法は因果的効果の測定に非常に有効な手法であるが、その適用には大きな問題がある。一致推定量を得るためには上記の二つの仮定を満たすような操作変数が必要だが、実際にそのような変数を見つけるのは難しい。一つ目の仮定については X_i と Z_i の関係の強さをデータを元に確認することで、操作変数が条件を満たしているかを確認することができる。しかし二つ目の仮定については実証的に確認することができない。よって理論的に仮定が正当であることを論じなければならないのだが、多くの場合操作変数と結果変数に何らかの関係がある可能性を完全には否定できないのである。 Z_i が何らかの理由で直接 Y_i に影響を与える可能性や、または Z_i が D_i 以外の経路で Y_i と結びつく可能性がある。

このような問題を受け、近年の研究では操作変数を見つける際に二つの方法を用いることが多くなっている。二つの方法に共通するのは、実験における処置変数のように、結果変数から見た場合に操作変数がまるで無作為に与えられているような事例を利用することである。一つは地域間の政策や歴史的背景の違いを利用する方法である。政策や歴史は多くの場合無作為に作り出された物ではないが、少なくとも結果変数に対しては直接的な影響を及ぼしている可能性がないと主張できることがある。例えばLassen (2005) は政治的情報が投票参加に与える影響を調べるために地域間の分権化の違いを操作変数法として利用している。市（ここではコペンハーゲン市）が分権化されている地域とそうでない地域に無作為に分割されたことに着目し、分権化されている地域に住んでいる有権者は分権に関する情報をより多く持っている想定し、よって分権の是非を問う選挙により高い確率で参加するだろうという仮説を立てた。この場合もし分権化自体が政治情報を与える以外

の経路で投票参加に影響しないと仮定できるなら、分権化は分権に関する政治情報量の正当な操作変数だと論じることができる。

もう一つは自然によって無作為に作り出されていると考えられる操作変数を利用する方法である。経済学では子どもの性別、生年月日、双子の誕生、そして天気などが操作変数として頻繁に用いられている。自然が無作為に作り出すこれらの変数は人間の支配が及ばない。よって操作変数自体が内生性をもつという可能性を否定できる。もし無作為に作られた操作変数が結果変数に直接影響を及ぼさないと仮定できるなら、そのような操作変数は因果推論に大きな手助けとなる。このようなアプローチを Rosenzweig and Wolpin (2000) は “natural natural experiment” と呼んでいる。選挙研究の分野では天気を操作変数として用いた先行研究がある。投票率は選挙の結果やその後の政策形成に大きな影響を与えることが予想されるが、投票率が無作為に決まるわけではないので、その因果的効果を調べるのは難しい。そこで投票率に影響を与えると考えられる天気の状態（降雨や降雪量）を操作変数として利用している研究がある。アメリカにおける研究は Gomez, Hansford, Krause (2007) と Hansford and Gomez (2010)、日本における研究は Horiuchi and Saito (2009) があげられる⁽¹⁹⁾。

不連続回帰デザインは処置の割り当てがあらかじめ設定されていた法則にもとづいて行われるような状況を利用する⁽²⁰⁾。例えば観察対象の個人 $i = 1, \dots, N$ が連続変数 x_i という特徴を持つとする。この x_i の値が前もって設定されていた値 c を超える場合に i は処置群へと割り当てられ、逆に x_i の値が c を上回らない場合は統制群へと割り当てられる。よって x_i の値を知れば、必然的に処置の有無を知ることができる。そして x_i が c に限りなく近い場合、処置群が統制群への割り当てはほとんど無作為に決定されると想定できる。

不連続回帰デザインにおいて処置の無作為な

割り当てがどのような状況で行われるかを理解するために、選挙結果を利用した研究例を紹介する。アメリカの連邦下院選挙は多くの場合民主党と共和党の2大政党間で争われる。ある下院選挙区 $i = 1, \dots, 435$ における民主党の得票率を x_i としよう。選挙における勝者は x_i が $c = 50$ パーセントを上回るかどうかで決定される。もし $x_i > 50$ ならば民主党の候補者がその選挙区の勝者である。逆に $x_i < 50$ ならば共和党の候補者がその選挙区の議席を占めることになる。このように各候補者の得票率によって必然的に選挙区の勝者が決定される。

選挙区の勝者はある状況においては無作為に決定されると想定できる。激戦の選挙区において候補者間の得票率の差が小さい場合、つまり民主党と共和党の候補者の得票率が50パーセントに限りなく近い場合、候補者、政党、そして有権者などのような政治的行為者も選挙結果に決定的な影響を及ぼすことはできない。Lee (2008, 684) は激戦の選挙区においてどちらの候補者が勝利するかは「まるでコイントスのように」決定されると論じている。激戦の選挙区では天気や他の予測不能な要因が得票率に小さな影響を与えることによって、最終的な勝者の決定を左右するのである。例えば投票日当日に雨が降ることによって民主党候補者の得票率が49.9パーセントから50.1パーセントに変化した場合、この予測不能な要因によって民主党候補者は最終的に勝利を収めることになるのである。よって激戦の選挙区ではほぼ無作為に勝者が決定されると言える。

勝者がほぼ無作為に決定されるため、勝者の特徴を除き、激戦選挙区はあらゆる面で近似した性質を持つ。つまり選択バイアスが存在しないはずである。よって各選挙区に対する勝者の特徴の因果的効果を測定することが可能になる。例えば民主党議員と共和党議員は支持母体が異なるため、選挙区に持ち帰る補助金の性質が異なるという仮説を立てる。この仮説の検証には政策分野ごとの補助金配分額について民主

党議員選出選挙区と共和党議員選出選挙区の平均値を比べればよい。激戦の選挙区では候補者の党派性は無作為に決定されるため、党派性を除き他のあらゆる面で選挙区の性質は等しいだろう。よって補助金の平均値の違いは議員の党派性の違いだと主張できる。

選挙結果を利用した不連続回帰デザインは因果推論の有効な道具としてこれまでいくつかの研究で用いられてきた。Lee (2008) は選挙における現職効果を測定するために不連続回帰デザインを用いている。ある選挙区において民主党の議員が現職であるとする。この場合、次の選挙においてこの選挙区における民主党の得票率は現職議席を占めていることによりどれだけの影響を受けるのかというのが問題の関心である。現職という現象そのものが様々な政治経済的要因によって決定されるため選択バイアスの問題が生じる。そこでこの問題を回避するために、選挙結果が無作為に決定されるという不連続回帰デザインを Lee は用いたのである。Lee の分析のよると t の時点における選挙で激戦を勝ち抜いた民主党候補者のいる選挙区では、次の選挙 $t+1$ の時点での民主党得票率が有意に高くなっている⁽²¹⁾。

因果的效果の測定は最も単純な状況では

$$Y_i = a + \beta x_i + \delta D_i + \eta_i \quad (21)$$

という回帰式を推定することにより行える。ここで D_i は処置群か統制群のどちらかに属することを示している。上記の選挙の例では⁽²¹⁾において

$$D_i = \begin{cases} 1 & \text{if } x_i > 50 \\ 0 & \text{if } x_i < 50 \end{cases} \quad (22)$$

と定義されて、 $D_i = 1$ は民主党候補者が勝利した選挙区、 $D_i = 0$ は共和党候補者が勝利した選挙区を意味する。もし x_i と Y_i が非線形な関係にある場合は x_i を $f(x_i)$ とし、他の関数形を用いて非線形関係を考慮する必要がある。また激戦の選挙区のみを分析対象とするのか、あるいは激

戦の選挙区だけでなく 2 候補者の得票率に比較差のある選挙区も分析対象として含めるのか、また観察可能な性質について処置群と統制群に差はないかなど、他に考慮すべき要素もある。Lee and Lemieux (2010) がより詳しい議論を提示している⁽²²⁾。

4. 3. マッチング

実験及び疑似実験デザインと異なり、マッチングは無作為化に依存しない研究デザインである。実験及び疑似実験デザインでは無作為化を用いるか無作為化に近い事例を用いることにより処置変数の状態を除いて性質の近似した集団を作り出すことを試みる。一方マッチングは観察可能な性質について近似した対を人為的に作り出す方法である。例えば性別と年齢のみが処置変数の割り当てに影響を与えるならば、性別と年齢が似通っているが処置状態の異なる対を作り、その対について結果変数の値を比較するのである。もし観察不可能な変数が重要な役割を果たさないと考えられるような状況では、マッチングは因果的效果の測定を可能にするような処置群と統制群を作り出すことができるだろう。一方で観察不可能な変数が大きな意味を持つ場合は、その変数についてマッチングを行うことはできないため、選択バイアスが生じるかもしれない。マッチングを行う際には観察可能な複数の変数を用いて同じ性質を持つ対を作り出す方法や、処置群と統制群に所属する確率を観察可能な変数を用いて計算し、その確率（傾向スコアと呼ばれる）を用いて対を作り出す方法がある。

マッチングを用いた研究例として Kam and Palmer (2008) をあげることができる。アメリカでは教育程度が政治参加に強い影響を与えることが知られているが、先行研究は高校や大学進学をめぐる選択バイアスの影響を考慮していない。そこで Kam and Palmer (2008) は大学進学に関する傾向スコアを求め、それにもとづいて教育程度以外は近似していると想定できる対

を作った。この対について政治参加の程度を比較したところ、Kam and Palmer (2008) は有意な差が見られないと報告している。

このようにマッチングは因果的効果を測定する方法として用いることも可能であるが、今井(2007)はマッチングはむしろ分析前のデータ処置の方法として有効であると論じている。実験で無作為化を用いて処置群と統制群を作り出す前に、マッチングを用いて観察可能な性質について同等な対を作る。そしてそれぞれの対について観察不可能な性質の影響を取り除くために無作為に処置状態を割り当てるのである。

5. まとめ

本稿では、日本において社会科学の方法論の文献として影響力をもつと思われる『創造の方法学』に依拠しつつ、因果的効果を推定する方法に関する従来の考え方をまとめ、その問題点を指摘した。さらにこの問題点を克服できると考えられる方法をいくつかの種類に分けて概観した。因果的効果の推定のためには、関心の対象となる独立変数と従属変数の両方に影響を及ぼしうる変数の影響を統制することが必要である。実験であれば無作為化によってこれが達成されるが、非実験・観察データの場合は別の方法が必要となる。そこで他の変数の影響を排除するために統制変数を回帰式に投入することが行われてきた。しかし、非常に限られた条件の下でしか実験における無作為化と統制変数は同じ効果をもたない。そこで近年、統制変数を使わない研究デザインが数々提案されてきたのである。

本稿で紹介した方法は、これまで回帰分析において数多くの独立変数を含む傾向にある選挙研究のあり方を徐々に変えていくことになるであろう。少なくとも理論的には統制変数のみに頼る方法はもはや多くの場合正統性を持たない。今後統制変数のみにもとづく因果推論が行われ続けるとすれば、それは分野における長年の習慣によるものか、理解しやすさによるもの

に過ぎないだろう。一方で、今後待たれるのが、同じ研究対象について統制変数を用いる従来の方法と実験や疑似実験デザインによる新しい方法とでは、結果がどのような異なるかについての研究である。本稿で紹介したような方法を用いた場合、これまで検証されてきた因果関係についての結論はどの程度変わるのか。実質的な違いがあるのか、それともほとんど違いはないのか。もし後者であれば、いくら理論的により正しくとも、本稿で紹介したような最新の研究は単なる「数学が得意な人の遊び」として、政治学の世界で重要視されることはないかもしれない。その意味でも今後、こうした手法の平易な解説および「大衆化」の努力が待たれる。因果推論の方法を明確に理解した上で用いることは、どのような結果がもたらされるのであれ、選挙に関する理論の構築を進める上で害となることはないはずである。

(1) 実際、この本は多くの日本の大学の社会科学系の方法論の入門科目において参考文献として指定されている。2011年3月現在、Googleで書名と「シラバス」というワードで検索したところ、少なくとも次の大学の授業において『創造の方法学』が参考文献として挙げられていることがわかる。愛媛大学、お茶の水女子大学、学習院大学、九州大学、京都大学、慶應義塾大学、甲南大学、神戸大学、埼玉大学、島根大学、首都大学東京、上智大学、信州大学、成城大学、千葉大学、筑波大学、都留文科大学、同志社大学、常磐大学、徳島大学、長崎大学、日本福祉大学、法政大学、北海道教育大学、山形大学、横浜国立大学、早稲田大学(五十音順)。

(2) King, Keohane, and Verba (1994) は因果的効果について現代的で洗練された定義を紹介している。また、後にも述べるようにこのアドバイスにはいくつかの留保がある。まず、この説明変数の帰結として生じる変数を統制変数に入れてはならないこと。そして、無限に存在する変数群からどの変数を統制変数としてモデルに加えるかを決めるために理論を用いなければならない、ということである。すなわち、できるだけ多くの変数を統制

変数として加えなければならない、とこの本の著者らは言っているわけではない。

- (3) 例えば, Gujarati (1995, p.66) を参照。
- (4) EITM について詳しくは, Granato and Scioli (2004), Granato, Lo, and Wong (2010) を参照。また日本語での紹介については, 飯田 (2009) を参照。
- (5) 例えば, Kollman, Miller, and Page (2003) を参照。また日本語での政治学および選挙研究へのシミュレーションの応用研究についての優れたレビューとして荒井 (2008) がある。
- (6) Rubin Causal Model (Holland 1986) とも呼ばれる。
- (7) 本節の議論は Holland (1986), King, Keohane, and Verba (1994), Winship and Morgan (1999), Morgan and Winship (2007), Angrist and Pischke (2009), Imbens and Wooldridge (2009) を参考にしている。
- (8) 本節では実験デザインにおける用語である処置変数, 処置群, 統制群といった用語を用いて議論を進める。なお本節の議論は処置変数が複数の値を取る場合にも一般化できる。詳しくは Morgen and Winship (2007, 53-7) を参照。
- (9) 同一個人に関する因果的効果を観察できないこの問題は Fundamental Problem of Causal Inference (Holland 1986) と呼ばれる。
- (10) 実際の因果的効果の測定には普通は標本が使われる。この場合は標本誤差も因果的効果の測定に大きな影響を及ぼす。詳しくは Imai, King, and Stuart (2008) を参照。
- (11) この定義の背後にはある観察単位が処置を受けた場合, この処置の効果はこの観察単位に限定され, 他には影響を及ぼさないとする Stable Unit Treatment Value Assumption (SUTVA) が存在する。詳しくは Morgen and Winship (2006, 37-40) を参照。
- (12) $E[Y_i^1 - Y_i^0 | D_i = 1] - E[Y_i^1 - Y_i^0]$ の意味をもう少し詳細に理解するため, この式を以下のように展開する。母集団における処置群と統制群に属する人口の割合を π と $(1 - \pi)$ とする。これらを用いて $E[Y_i^1 - Y_i^0]$ を処置群と統制群それぞれにおける因果的効果に分割する。その際には因果的効果を各群に含まれる人口の割合で重み付けする:

$$\begin{aligned} & E[Y_i^1 - Y_i^0 | D_i = 1] - E[Y_i^1 - Y_i^0] \\ &= E[Y_i^1 - Y_i^0 | D_i = 1] - \pi E[Y_i^1 - Y_i^0 | D_i = 1] - (1 - \pi) E[Y_i^1 - Y_i^0 | D_i = 0] \end{aligned}$$

$$\begin{aligned} &= (1 - \pi) (E[Y_i^1 - Y_i^0 | D_i = 1]) - (1 - \pi) \\ &\quad (E[Y_i^1 - Y_i^0 | D_i = 0]) \\ &= (1 - \pi) \{ (E[Y_i^1 - Y_i^0 | D_i = 1]) - (E[Y_i^1 - Y_i^0 | D_i = 0]) \} \end{aligned}$$

最後の式から $E[Y_i^1 - Y_i^0 | D_i = 1] - E[Y_i^1 - Y_i^0]$ という条件は, 処置群と統制群における因果的効果が均一であることを意味していることがわかる。

- (13) ここでは無作為化のあとで4年間にわたって研究対象の全ての人がこの決定に従うことを前提としている。
- (14) 本節では簡略化のために母集団における因果的効果のみを議論の対象としている。実際の因果的効果の測定には母集団から選ばれた標本が用いられることが多い。その場合は標本平均因果的効果を推定することになるが, その際には選択バイアスの問題に加えて標本選択による誤差も考慮に入れる必要がある。詳しくは今井 (2007) と Imai, King, and Stuart (2008) を参照。
- (15) 経済学では同一の因果関係について実験デザインから得られた結果と非実験デザインから得られた結果が大きく乖離していることを報告している研究もある。詳しくは La-Londe (1986) を参照。
- (16) 近年ではフィールド実験を用いた優れた研究が増加している。選挙における動員の有無やその方法, あるいは選挙広告が投票率に与える影響を調べた研究は過去10年に数多く出版されている (例えば Gerber and Green 2000; Imai 2005; Panagopoulos and Green 2008)。他にも政党帰属意識が候補者評価や投票行動に与える影響を調べた Gerber, Huber, and Washington (2010), 選挙制度が女性の当選率に影響を与えるかを検証した Bhavnani (2009), 政策決定システムが住民の政治満足度にどのような影響を与えるかを調べた Olken (2010) などの研究がある。
- (17) これらの方法を用いず, 処置変数が無作為に近い状態で割り振られていることを利用してクロスセクショナルデータによる回帰式を推定することも可能である。例えば Huber and Arceneaux (2007) を参照。
- (18) 時系列のデータを用いる場合には誤差項同士の相関などの問題にも気をつける必要がある。例えば Bertrand, Duflo, and Mullainathan (2004) を参照。
- (19) 近年頻繁に用いられる上記のような手法に対する批判もある。Rosenzweig and Wol-

- pin (2000) は無作為に作り出された操作変数を用いても因果関係の背後にあるメカニズムを理論的に理解することにつながらないのでとは指摘している（これに対する反論は Angrist and Kruger 2001 を参照）。また操作変数が無作為に生み出されたといっても、操作変数が結果変数に直接影響を与える可能性を完全に排除するわけではないとも論じている。
- (20) 本稿では鋭角的 (sharp) 不連続回帰デザインについてのみ紹介を行う。ファジー不連続回帰デザインについては Lee and Lemieux (2010) などを参照。
- (21) 不連続回帰モデルを応用した研究例として、アメリカにおいて市長の党派性がその市における政策にどのような影響を与えるか研究した Ferreira and Gyourko (2008) やイギリスにおいて議員になることがどのような経済的利益を生み出すかを研究した Eggers and Hainmueller (2009) などがある。
- (22) 他の優れた応用例として議員になることが個人資産にどのような影響を与えるかを調べた Eggers and Hainmueller (2009) や選挙での強さは世襲されるかを調べた Dal Bo, Dal Bo, and Snyder (2009) がある。
- 参考文献（アルファベット順）
- Achen, Christopher H. 2002. "Toward a New Political Methodology: Microfoundations and ART." *Annual Review of Political Science* 5: 423-450.
- Angrist, Joshua D., and Kruger, Alan B. 2001. "Instrumental Variables and the Search for Identification: From Supply to Demand and Natural Experiments." *Journal of Economic Perspectives* 15: 69-85.
- Angrist, Joshua D. and Pischke, Jorn-Stephen. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- 荒井紀一郎. 2008. 「政治学におけるシミュレーション研究」『早稲田政治公法研究』88: 11-21.
- Barabas, Jason, and Jerit, Jennifer. 2010. "Are Survey Experiments Externally Valid?" *American Political Science Review* 104: 226-242.
- Bertrand, Marianne, Duflo, Esther, and Mullainathan, Sendhil. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119: 249-275.
- Bhavnani, Rikhil. R. 2009. "Do Electoral Quotas Work after They Are Withdrawn? Evidence from a Natural Experiment in India." *American Political Science Review* 103: 23-35.
- Dal Bo, Ernesto, Dal Bo, Pedro, and Snyder, James. 2009. "Political Dynasties." *Review of Economic Studies* 76: 115-142.
- Druckman, James N., Green, Donald P., Kuklinski, James H., & Lupia, Arthur. 2006. "The Growth and Development of Experimental Research in Political Science." *American Political Science Review* 100: 627-635.
- Eggers, Andrew C., and Hainmueller, Jens. 2009. "MPs for Sale? Returns to Office in Postwar British Politics." *American Political Science Review* 103: 513-533.
- Ferreira, Fernando, and Gyourko, Joseph. 2008. "Do Political Parties Matter? Evidence from U.S. Cities." *Quarterly Journal of Economics* 124: 399-422.
- Freedman, David A. 1991. "Statistical Models and Shoe Leather." *Sociological Methodology* 21, 291-313.
- Gentzkow, Matthew. 2006. "Television and Voter Turnout." *Quarterly Journal of Economics* 121: 931-972.
- Gerber, Alan S., and Green, Donald P. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review* 94: 653-663.
- Gerber, Alan S., Huber, Gregory A., and Washington, Ebonya. 2010. "Party Affiliation, Partisanship, and Political Beliefs: A Field Experiment." *American Political Science Review* 104: 720-744.
- Gomez, Brad. T., Hansford, Thomas G., and Krause, G. A. 2007. "The Republicans Should Pray for Rain: Weather, Turnout, and Voting in U.S. Presidential Elections." *Journal of Politics* 69: 649-663.
- Granato, Jim and Frank Scioli. 2004. "Puzzles, Proverbs, and Omega Matrices: The Scientific and Social Significance of the Empirical Implications of Theoretical Models (EITM)." *Perspectives on Politics* 2: 313-23.
- Granato, Jim., Melody Lo, and M. C. Sunny Wong. 2010. "A Framework for Unifying For-

- mal and Empirical Analysis.” *American Journal of Political Science* 54: 783-797.
- Granger, Clive W. J. 1969. “Investigating Causal Relations by Econometric Models and Cross-Spectral Models.” *Econometrica* 37: 424-438.
- Green, Donald P., and Gerber, Alan S. 2003. “The Underprovision of Experiments in Political Science.” *The Annals of the American Academy of Political and Social Science* 589: 94-112.
- Gujarati, Damodar N. 1995. *Basic Econometrics*, 3rd ed. New York: McGraw Hill.
- Hansford, Thoas G., and Gomez, Brad. T. 2010. “Estimating the Electoral Effects of Voter Turnout.” *American Political Science Review*, 104: 268-288.
- Hirano, Shigeo, and Snyder, James M. 2009. “Using Multimember District Elections to Estimate the Sources of the Incumbency Advantage.” *American Journal of Political Science* 53: 292-306.
- Holland, Paul W. 1986. “Statistics and Causal Inference.” *Journal of the American Statistical Association*, 81: 945-960.
- Horiuchi, Yusaku, Imai, Kosuke, and Taniguchi, Naoko. 2007. “Designing and Analyzing Randomized Experiments: Application to a Japanese Election Survey Experiment.” *American Journal of Political Science* 51: 669-687.
- Horiuchi, Yusaku, and Saito, Jun. 2003. “Reapportionment and Redistribution: Consequences of Electoral Reform in Japan.” *American Journal of Political Science* 47: 669-682.
- Horiuchi, Yusaku, and Saito, Jun. 2009. “Rain, Election, and Money: The Impact of Voter Turnout on Distributive Policy Outcomes.” Working paper.
- Huber, Gregory A., and Arceneaux, Kevin. 2007. “Identifying the Persuasive Effects of Presidential Advertising.” *American Journal of Political Science* 51: 961-81.
- 飯田健. 2009年. 「投票行動研究における数理モデルと EITM」 山田真裕・飯田健編著『投票行動研究のフロンティア』おうふう, 279-308.
- Imai, Kosuke. 2005. “Do Get - Out - the - Vote Calls Reduce Turnout?: The Importance of Statistical Methods for Field Experiments.” *American Political Science Review* 99: 283-300.
- 今井耕介. 2007年. 「計量政治学における因果的推論」『レヴァイアサン』40: 224-233.
- Imai, Kosuke, King, Gary, and Stuart, Elizabeth A. 2008. “Misunderstandings between Experimentalists and Observationalists about Causal Inference.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171: 481-502.
- Imbens, Guido. W., and Wooldridge, Jeffrey. M. 2009. “Recent Developments in the Econometrics of Program Evaluation.” *Journal of Economic Literature*, 47: 5-86.
- Isaak, Alan C. 1975. *Scope and Methods of Political Science: An Introduction to the Methodology of Political Inquiry*. Homewood, Ill.: Dorsey Press. (喜多靖郎・富岡宣之訳. 1988年. 『政治学方法論序説』晃洋書房.)
- Kam, Cindy D., and Palmer, Carl L. 2008. “Reconsidering the Effects of Education on Political Participation.” *Journal of Politics* 70: 612-31.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press. (真淵勝監訳. 2004年. 『社会科学のリサーチ・デザイン—定性的研究における科学的推論』勁草書房.)
- Kollman, Ken, John H. Miller, Scott E. Page. 2003. *Computational Models in Political Economy*. Cambridge, MA: MIT Press.
- La Londe, Robert. 1986. “Evaluating the Econometric Evaluations of Training Programs.” *American Economic Review* 76: 604-620.
- Lassen, David Dreyer. 2005. “The Effect of Information on Voter Turnout: Evidence from a Natural Experiment.” *American Journal of Political Science* 49: 103-118.
- Lave, Charles A. and James G. March. 1975. *An Introduction to Models in the Social Science*. New York: Harper & Row. (佐藤嘉倫・大澤定順・都築一治訳. 1991年. 『社会科学のためのモデル入門』ハーベスト社.)
- Lee, David S. 2008. “Randomized Experiments from Non-Random Selection in U.S. House Elections.” *Journal of Econometrics* 142: 675-697.
- Lee, David S., and Lemieux, Thomas. 2010. “Regression Discontinuity Designs in Economics.” *Journal of Economic Literature* 48: 281-355.

- Morgan, Stephen. L., & Winship, Christopher. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press.
- 西澤由隆, 栗山浩一. 2010. 「面接調査における Social Desirability Bias—その軽減への full-scale CASI の試み」『レヴァイアサン』46号.
- Olken, Benjamin A. 2010. “Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia.” *American Political Science Review* 104: 243-267.
- Panagopoulos, Costas, and Green, Donald P. 2008. “Field Experiments Testing the Impact of Radio Advertisements on Electoral Competition.” *American Journal of Political Science* 52: 156-168.
- Rosenzweig, Mark R., and Wolpin, Kenneth I. 2000. “Natural ‘Natural Experiments in Economics.” *Journal of Economic Literature* 38: 827-874.
- Sovey, Allison J., and Green, Donald P. 2011. “Instrumental Variables Estimation in Political Science: A Readers’ Guide.” *American Journal of Political Science* 55: 188-200.
- 高根正昭. 1979年. 『創造の方法学』講談社現代新書.
- Van Evera, Stephen. 1997. *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell University Press. (野口和彦・渡辺紫乃. 2009年『政治学のリサーチ・メソッド』勁草書房.)
- Winship, Christopher, and Morgan, Stephen. L. 1999. “The Estimation of Causal Effects from Observational Data.” *Annual Review of Sociology* 25: 659-706.