

第2章 科学的手法による評価

1.1 で述べた科学的手法は、実験的 (experimental) 手法と非実験的 (quasi-experimental) 手法に分類される。ここで、非実験的手法とは、実際には同一主体ではないものについて、その客体の個人属性を限りなく近い状態でマッチングさせるといった統計的処置を行いつつ、計量経済学的手法を駆使して、プログラムへの参加グループとプログラムに参加しなかった者からなる比較対照グループについての比較を行うことにより、プログラム本来の効果を抽出するといった、自然科学のような実験が容易ではない社会科学の分野において確立しつつある評価手法のことである。

なお、いずれの場合にも、代表性の欠落や内生性といったセレクションバイアスが発生するため、これらを如何に克服するのかということが問題となる。

2.1 実験的手法による評価

政策プログラムの効果に係る評価を行うにあたっての実験的手法とは、政策プログラムの運営主体が、プログラムへの参加希望者の中から、実際の参加者(トリートメント(treatment)グループ)と非参加者(コントロール(control)グループ)を無作為に選定して、そのグループ間のアウトカムの平均的な差異を見ることにより、プログラムの効果を評価する評価手法である。これは、医薬品の効能を確かめるためのラットの実験や教育心理学の分野における双子の実験等の方法を社会的実験として実現した科学的手法であると言える。

このような実験的手法による政策評価は、アウトカム Y_0, Y_1 が実際に得られるという意味できわめて直接的でわかりやすいものであるということもあってか、アメリカでは、政策評価にあたってよく利用されているところであり、さらに、その評価結果は政策プログラムの運営にあたり重要な役割を果たすものとなっている。ただ、ヨーロッパにおいては、イギリス、スウェーデン等に例がある程度で、実験的手法による政策評価はあまりなじみのある手法とはなっていない。

政策プログラムへの参加希望者のグループから無作為にプログラムへの参加者・非参加者を選定する評価手法である実験的手法により政策プログラムの効果の評価を行うにあたっての基本となる前提は、参加者・非参加者を無作為に選定する方法により得られる政策プログラムに参加することの平均的な効果が、先の「参加者についての参加の効果」に等しくなる、ということである。例えば、

- ・ 属性が X である個人については、全員、 $Y_1 - Y_0 = (\text{ある定数})$ (共通効果モデル)

であるとか、あるいは、さらに一般的に、

- $E(Y_1 - Y_0|X, D = 1) = E(Y_1 - Y_0|X)$
(この場合、「 $D = 1$ 」は、政策プログラムへの参加希望に対応する。)

といった条件が成り立つ場合には、上記の前提は満たされる、すなわち、バイアスは生じないこととなるが、実際にはそういうわけにはいかない。実験的手法によるバイアスとしては、次のようなものがある。

まず、参加意志に関するバイアスで、ランダム化バイアス (randomization bias) と呼ばれるものがある。これは、無作為にプログラム参加者を決定するという方法であるために、人々の参加意思自体に影響を及ぼしてしまうというものである。例えば⁶、通常のプログラム運営では、特定の基準を満たすと思われる者のみに対して参加勧奨が行われるとすると、通常の運営においては参加勧奨を受けられる可能性が高く、参加を希望するであろう者も、無作為に参加が割り当てられるとなると、結果的に参加できなくなる可能性があるために、最初から参加自体を希望しなくなるかもしれない。このとき、無作為割り当てが行われる場合の参加希望者についての潜在能力等の属性は、通常のプログラム運営における参加希望者のものとは異なるであろうことから、推定されるプログラム効果が通常の運営の場合のプログラム効果を示すとは必ずしも言えなくなる。

さらに、政策プログラムへの参加が割り当てられなかった者の中から、外部の何らかのプログラムに参加する者が出てくるという代替バイアス (substitution bias) と呼ばれるバイアスがある。これはコントロールグループに属する非参加者が、対象となるプログラムと代替的な外部のプログラムに参加することにより生じるバイアスである。コントロールグループに属する非参加者が外部のプログラムに参加することにより、そうでなかった場合に比べ賃金等のアウトカムは増加することになる。そうした結果、本来計測されるべき参加者と非参加者のアウトカムの差にバイアスが生じることになる。実際には何らかの外部のプログラムに参加した者は4割程度という話もある。

この他にも実験的手法におけるセレクションバイアスが指摘されている。プログラム修了後の参加者の就業率や賃金の増加等、アウトカムに関する達成目標がプログラムの各運営機関に割り当てられている場合、運営機関担当者は意図的に再就職しやすい者や高賃金を得やすい者をプログラム参加者に割り当てがちになる。こうした行為はクリーミング (creaming) と呼ばれている。無作為抽出を謳っているが、実際には参加者に質の高い者が分配される結果、セレクションバイアスが生じることになる。

実験的手法については、このようにバイアスが生じるといった問題以外にも、そもそも、ひ

⁶ ここでのランダム化バイアスの例については、黒澤 (2001) の説明が非常にわかりやすいので、ほぼそのままの形で引用している (p.9)。

とつのプログラムにおいてその各段階ごとにランダム化を行うことはきわめて大変であるといったことや、先の「参加者についての参加の効果」は評価できるとしてもその他のいろいろな影響を見ることができないといったことその他、以下のような倫理上・費用上の問題点が指摘されている。

まず、プログラムの参加希望者を、人為的にトリートメントグループとコントロールグループに分類しているが、コントロールグループに割り振られた非参加者は、参加希望を持ちながら参加できないということがある。例えば、アメリカの職業訓練パートナーシップ法（Job Training Partnership Act）では、コントロールグループに割り当てられた者は18ヶ月間当該訓練に参加できないことになっている。いくら正確な評価制度を確立するためとはいえ、プログラムへの参加を希望している者の参加権利を奪っていいのかという意見が多く出されているのも事実である。

また、実験的手法は、多額の費用がかかり、しかもデータ収集者の心理的負担が大きいということがある。特に、データ収集にかかる負担の大きさは膨大なものだと言われており、そのためあってトリートメントグループとコントロールグループの抽出比率が2分の1ずつではなく、より負担の大きなコントロールグループの場合には抽出比率が3分の1、そうでないコントロールグループの場合には3分の2とするなどデータの抽出割合にも影響を及ぼしている。プログラム参加希望者を無作為にトリートメントグループとコントロールグループに分類する際には、さらにプログラム修了後数年間にわたりデータを収集しなければならないなど、非実験的手法に比べてコストがかかるという点が指摘されている。

なお、アメリカにおける最も有名な実験的手法による評価の実例としては、次のようなものがある。ひとつは、Hollister et al. (1984) をはじめとする National Supported Work に関するものである。このプログラムは、長期にわたり AFDC（福祉）を受けている者、元麻薬常習者、元受刑者、そして、経済的に不利な立場にある若者を対象とする非熟練業務の職業体験であるが、9（～18）ヶ月間、National Supported Work に参加することの効果の評価にあたり、National Supported Work への参加者・非参加者を無作為に半々に振り分け、参加しない方に振り分けられた者は、他のプログラムに参加できないようにしたというものである。もうひとつは、Heckman and Smith (1998) をはじめとする National JTPA Study に関するものである。このプログラムは、OJT と教室型訓練（座学）の両方を含む訓練プログラムである。

【参考文献】

黒澤昌子（2001）「職業訓練施策の評価：非実験的および実験的方法による検証のレビュー」
『経済研究』（明治学院大学）No. 120: pp. 1-22.

2.2 非実験的手法による評価

2.2.1 非実験的手法とは

非実験的手法とは、社会的実験を行わないで科学的な分析ができる手法である。実験的データと異なり、プログラム参加者についてのデータが非実験的に収集される場合⁷、プログラム非参加者からなる比較対照グループについてのデータは、調査地域や時期の等しい外部調査から再抽出されなければならない⁸。このようにして収集されたデータを非実験的データと呼び、これを使用して分析を行うことを非実験的手法という。

2.2.1.1 比較対照グループを作成する際の仮定と方法⁹

プログラム参加者との比較を行うためにデータ収集されるプログラム非参加者からなる比較対照グループを作成するには、プログラム参加者と等しい個人属性を持った人を、外部調査から再抽出する。このとき、年齢・性別・人種といった観察可能で、かつプログラム参加意思決定に影響を与えないと考えられる個人属性を基準として再抽出を行う。このような再抽出のことをマッチングと呼ぶ。マッチングの基本的プロセスからも明らかのように、“個人 i がプログラム参加者グループと比較対照グループの両方に存在している”、という暗黙の仮定の下、マッチングは行われる。

マッチングを行う際には、二つの仮定を明示的においている。1つめは、利用可能な説明変数をマッチングの基準として用いることで、プログラム参加者とプログラム非参加者の間の受講意思決定プロセスにおける違いを調整によってなくすることができる、というものである (adjust away differences)。プログラム参加者の収入を Y_1 、プログラム非参加者の収入を Y_0 、マッチングに用いる個人属性を X 、プログラム参加の有無を表す変数を D とし、「 $D = 1$ 」はプログラムへの参加を、「 $D = 0$ 」はプログラムへの非参加を表すものとする、この仮定は、

$$\text{Cov}(Y_0, D|X) = 0$$

$$\text{Cov}(Y_1, D|X) = 0$$

と表すことができる。

⁷ プログラム参加者については、実験的手法を用いなくても、追跡調査という形でのデータ収集を行うことができる。

⁸ アメリカでは、The Current Population Survey (CPS) が用いられることが多い。CPS は労働統計局の依頼により人口調査局が毎月一回行う世帯調査で、年齢・性別・人種等の個人属性別に分類されたアメリカ全体の雇用と失業の実態に関する総合的な情報を提供するための調査である。

⁹ ここでの説明は、Heckman et al. (1999, pp.1950-1954) に拠っている。

2つめの仮定は、

$$0 < \text{Prob}(D = 1|X) < 1$$

というものであって、ある個人属性を持つ者がプログラムを受講する確率はゼロより大きく 1 より小さい、すなわち、ある個人属性を持つ者にはプログラム参加者とプログラム非参加者が必ず混在しているという仮定である。

以上2つの仮定が満たされると、外部調査から作成された比較対照グループの観察可能な Y_0 が、プログラム受講者の反事実である Y_0 と同一となる。同様に、プログラム受講者の観察可能な Y_1 は、プログラム非受講者の反事実である Y_1 と同一になる。それゆえ、 Y_0 と Y_1 についての推定バイアス

$$B_0(X) = E(Y_0 | X, D = 1) - E(Y_0 | X, D = 0)$$

$$B_1(X) = E(Y_1 | X, D = 1) - E(Y_1 | X, D = 0)$$

が両方ともゼロとなり、しかも D と誤差項が無相関となることから、 $E(Y_1 - Y_0 | X, D = 1)$ の代わりに $E(Y_1 - Y_0 | X)$ を用いた通常の回帰分析による計測でも、推定によるバイアスを回避することが理論的に可能となる。しかし、このことがサンプルセレクションバイアスがなくなることを意味しているわけではないことに留意が必要である。

マッチングの際に用いられる代表的な手法として、カリパー・マッチング法 (Caliper matching) とカーネル・マッチング法 (Kernel matching) の2つがある。詳細については、Heckman et al. (1999, pp.1953-1954) を参照されたい。

2.2.1.2 サンプルセレクションバイアスの問題¹⁰

計量分析によって求められる推定値は、標本が母集団からのランダムサンプリングであるとき、正確な推定値となる。しかし、政策評価のためのデータ収集を行う際には、サンプルバイアスが生じてしまうことが多い。そこで、以下では、サンプルバイアスの代表的な2つの例について説明する¹¹。

第1に、サンプルの脱落 (dropping out) の問題が挙げられる。政策評価を正確に行うためには、あるプログラムを受講した人の受講後の調査が不可欠である。しかし、プログラム受講後年数が経つにつれて、彼らを捕捉することは難しくなる。特に、定職に就いていない人は住

¹⁰ ここでの説明は、Heckman et al. (1999, pp.1938-1940) に拠っている。

¹¹ これら以外にも、マッチングの基準変数として誤差項と確率的に相関している変数を用いてマッチングを行うと (例えば、学歴などが考えられる。)、誤差項の確率的構造を変えてしまうという問題がある。この場合、推定値にバイアスが生じてしまう。

所を転々としやすく、追跡調査がさらに困難となる。つまり、追跡調査をしやすい人＝効果のあった人のみが標本として残りやすくなり、サンプルセクションが生じることとなる。

第 2 に、汚染バイアス (contamination bias) が挙げられる。データ収集は、比較対照グループはプログラムに参加していないと仮定して行う。しかし、外部調査から再抽出する際に、実際にはプログラムに参加している人が紛れ込んでしまうことがある。また、サンプリング時にはプログラムに参加していなくても、その後プログラム参加機会を見つけて、自主的に同種のプログラムに参加してしまう可能性もある。つまり、比較対照グループが“完全に”プログラムに参加していない状態を保つことは非常に困難なのである。

また、データ収集の過程で生じるサンプルバイアス以外にも、個人のプログラム参加についての合理的意思決定の過程でセクションバイアスが発生することも多い¹²。

繰り返しになるが、計量分析によって求められる推定値は、データが目的集団を正確に代表しており、かつバイアスのない計測方法を用いたとき、正確な推定値となる。しかし、現実に政策評価を行う際には、収集されたデータが目的集団を正確に代表していない場合も少なくない。また、計測手法に関しては、バイアスを回避する手法の開発が積極的に行われてきた。これら計量手法の研究蓄積の概観を、次節以降で行う。

【参考文献】

Heckman, James J., Robert J. LaLonde and Jefferey A. Smith (1999) "The Economics and Econometrics of Active Labor Market Programs," In Ashenfelter, Orley C. and D. Card (eds), *Handbook of Labor Economics*, Vol. 3, Elsevier Science: pp.1865-2097.

2.2.2 クロスセクションデータでの推定方法¹³

2.2.2.1 Heckman の 2 段階推定法

プログラムに参加することから得られる成果を Y とする。プログラムに参加することから得られる成果とは、賃金の上昇であったり、職に就けるようになることなど、労働市場における経済的条件の向上を意味する。そして、就職するということは、収入を伴う。よって、説明を分かりやすくするために、以下ではプログラム参加による成果 Y を収入と称することもある。

Y は、観察可能な個人属性のベクトル X (例えば、年齢、教育年数、居住地域、性など) と観察不可能な属性 u (潜在能力など) の線形関数で表すことができるとすると、

¹² 例えば、潜在的能力が高く、プログラムに参加に対する期待収益の高い者ほど、プログラムを受講するという意思決定を行いやすくなるだろう。

¹³ 本節の説明は、黒澤(2001) に拠っている。

$$Y_{it} = \beta X_{it} + D_i \alpha_t + u_{it} \quad \dots (1)$$

となる。ここで、 D はプログラム受講者の場合は1、非受講者の場合は0となるダミー変数であり、受講者の賃金上昇の平均的効果は α によって計測される。

プログラム受講の有無 D と潜在能力など観察不可能な属性 u の間には相関が発生する可能性が高い。そのルートとして、①個人によるプログラム受講の選択と、②プログラム運営機関による割り当てプロセスの2つが考えられる。①の例として、潜在能力の高い者ほど自発的に訓練を受講しようする傾向があることが挙げられる。②に関しては、訓練運営機関が政府あるいはさらに上部の管轄機関から、ある業績水準を達成することを義務づけられているとき、訓練運営機関はより高い成果を出してくれる人、すなわち潜在能力の高い人に優先してプログラムを受講させようとする、ということが生じる可能性が考えられる。つまり、個人またはプログラム運営機関の合理的意思決定の結果、サンプルセレクションが生じるのである。そこで、このプログラム受講選択プロセスを次のような単純化したモデルで表す。

$$D_i^* = \gamma Z_i + \varepsilon_i \quad \dots (2)$$

D_i^* はプログラムを受講する可能性を示す潜在変数であり、「 $D_i^* > 0$ 」であれば「 $D_i = 1$ 」、つまりプログラムを受講することを示し、「 $D_i^* \leq 0$ 」であれば「 $D_i = 0$ 」、つまりプログラムを受講しないことを示す。 Z_i は受講確率に影響を与える変数を表す。また誤差項 ε は独立・同一分布 (*i. i. d.*) に従うものとする。また、 Z_i と X_i には共通の変数が含まれる場合もある。

以上のようなモデル設定の下、プログラム受講の有無 D と潜在能力など観察不可能な属性 u の間に相関がある場合、すなわち $E(u_{it} \varepsilon_i) \neq 0$ であるとき、 α の一致推定量を得るための計量的推定方法が数多く開発されてきた。その一つが Heckman の 2 段階推定法 (Heckman (1979)) を応用した分析である (Barnow et al. (1980))。

Barnow et al. (1980) では、 u と ε の結合分布が 2 変数正規分布に従うと仮定し、

- ① (2) 式 ($D_i^* = \gamma Z_i + \varepsilon_i$) をプロビットモデルを用いて推定する、
- ② ①で得られた推定値から、 $E(\varepsilon_{it} | Z_i, D_i)$ の推定値 $E^*(\varepsilon_{it} | Z_i, D_i)$ を計算する、
- ③ $E^*(\varepsilon_{it} | Z_i, D_i)$ を (1) 式 ($Y_{it} = \beta X_{it} + D_i \alpha_t + u_{it}$) の独立変数に追加した

$$Y_{it} = \beta X_{it} + D_i \alpha_t + E^*(\varepsilon_{it} | Z_i, D_i) \delta_t + v_{it} \quad \dots (3)$$

を推定する、

という 2 段階推定法を用いることで、 α_t の一致推定量を得られることを示した。

ただし、 Z と X に共通する変数が多いと、 $E^*(\varepsilon_{it} | Z_i, D_i)$ と X の間に高い多重共線性が生じて α_t の推定値が一致推定量でなくなってしまうことを、この手法の使用上の注意点として心

に留めておくべきである¹⁴。

2.2.2.2 操作変数法

2.2.2.1と同じモデル設定の下、プログラム受講の有無 D と潜在能力など観察不可能な属性 u の間に相関がある場合、操作変数法もサンプルセレクション問題を解決して、 α の一致推定量を求めることを可能とする方法の一つである。

具体的には、まず (2) 式 ($D_i^* = \gamma Z_i + \varepsilon_i$) の Z には含まれるが (1) 式 ($Y_{it} = \beta X_{it} + D_i \alpha_t + u_{it}$) の X には含まれない変数、つまりプログラム受講選択には影響を与えるものの、観察不可能な属性 u とは相関のない変数を探し、これを D の操作変数とする¹⁵。そして、この操作変数を用いて推定することにより、望ましい α の推定値を得ることが可能となるのである (Heckman et al. (1999, pp.1961-3), 黒澤 (2001, p.14))。

このように、理論的には操作変数法をプログラム評価に用いることは可能であるが、現実には誤差項 u と相関のない操作変数を見つけることは容易ではない。アメリカで、操作変数として“受講者の自宅と訓練機関の距離”を用いて、職業訓練の効果を推定した論文があるが、これに対しては批判がある。この批判とは、訓練機関が自宅から遠くても潜在能力が高く、訓練を受けることで将来の収入の大幅な上昇が見込まれる人は、訓練を受講するだろう。逆に、潜在能力の低い人は、訓練機関が自宅から遠いということだけで、受講を取りやめてしまう可能性がある。このように、“受講者の自宅と訓練機関の距離”は潜在能力 (= 誤差項 u) と相関をもっていると考えるのが自然であり、操作変数としてふさわしくないという批判である (Heckman et al. (1999, p.1963))。

以上をまとめると、操作変数法はサンプルセレクション問題を解決する手段として理論的には有効であるが、実際には望ましい操作変数を見つけることは困難である。よって、操作変数法が、政策評価を行うための実用的な手法であるかについては疑問が残ると言えよう。

【参考文献】

黒澤昌子 (2001) 「職業訓練施策の評価：非実験的および実験的方法による検証のレビュー」, 『経済研究』 (明治学院大学), 第 120 号: pp. 1-22.

牧 厚志・宮内 環・浪花貞夫・縄田和満 (1997) 『応用計量経済学 II』, 多賀出版.

Barnow, Burt S., Glein G. Cain and Arthur S. Goldberger (1980) "Issues in the Analysis of Selectivity Bias," In Stromsdorger, Ernst W. and George Farkas (eds), *Evaluation*

¹⁴ Heckman の 2 段階推定法の問題点については、牧他 (1997, p.288) に平易にまとめられている。

¹⁵ 操作変数として選ばれる変数は、①誤差項と相関をもたない、②説明変数と相関を持つ、という 2 つの条件を満たさなければならない。

Studies Review Annual, Volume 5, Beverly Hills, CA.: Sage Publications: pp. 43-59 with corrections. (Reprinted in Ashenfelter, Orley C. and Robert J. LaLonde (eds), *The Economics of Training, Volume 1*, Edward Elgar Publishing Limited, 1996.

Heckman, James J. (1979), "Sample Selection Bias As A Specification Error," *Econometrica*, Vol. 47, No. 1: pp.153-161.

Heckman, James J., Robert J. LaLonde and Jefferey A. Smith (1999) "The Economics and Econometrics of Active Labor Market Programs," In Ashenfelter, Orley C. and D. Card (eds), *Handbook of Labor Economics, Vol. 3*, Elsevier Science: pp.1865-2097.

2.2.3 パネルデータでの推定方法

典型的な経済データには時系列データとクロスセクションデータ（横断面データともいう）の2つのタイプがあるが、この2つの性質を併せ持ったデータを、パネルデータという。例として、労働者の職業訓練受講の有無・収入・勤務先など同じ質問項目について、同一人物に対して複数年にわたって追跡調査したデータが挙げられる。

パネルデータを用いることの利点は、観測できない「個体効果 (individual effect)」を制御できることにある。例えば、年齢・学歴・勤務先など観察可能な個人属性が全く同じ労働者同士であっても、彼らの賃金が全く同じということはない。それは、労働者個々人は観測できない本源的な能力を持っており、この影響が労働者間の違いを生じさせているからと考えられる。2.2.1.2で説明したように、このような労働者間の異質性が誤差項に含まれると、推定値にバイアスが生じてしまう。

観測できない要素が大きな影響を持つ経済現象を分析する際に、観察不可能な要素の影響を除去するための解決策の1つとして知られているのが、パネルデータの使用である。もし複数の観測時点で個体効果に変化がなければ、パネルデータを用いることで、個体効果を消すことができるので、観測不能な変数の存在によるバイアスの発生を避けることができるのである¹⁶。

パネルデータは、時系列データとしての性質も持つため、D と誤差項の相関だけでなく、誤差項の自己相関もバイアス発生の要因となる。次にいくつかの分析手法を示すが、自己回帰モデルは自己相関を解消する手法であり、固定効果モデルは D と誤差項の相関を解消する方法である。また、固定効果モデルは賃金の趨勢的上昇が推計結果に紛れ込む危険があるため、これを改善しようとしたのが Difference-in-differences 推計法である。

2.2.3.1 1階の自己回帰モデル (AR(1) モデル)¹⁷

¹⁶ 松田・伴・美添 (2000, pp.219-221) の説明が平易である。

¹⁷ Heckman et al. (1999, p.1939) を参考にした。

政策評価の推定に用いる式を、

$$Y_{it} = \beta X_{it} + D_i \alpha_t + u_{it}$$

とする。ここで推定モデルが1階の自己回帰モデル (AR(1)) に従うとすると、誤差項 u_{it} は、

$$u_{it} = \rho u_{it-1} + \varepsilon_{it}$$

と表すことができる。 ε_{it} は平均がゼロで、*i. i. d.* (independently and identically distributed) に従う誤差項である。すなわち、平均がゼロ、分散が一定、異時点間の相関がゼロである。推定モデルをこのように定式化すると、誤差項 ε_{it} に自己相関がなくなる。こうすることによって、誤差項の自己相関によって生ずる推計バイアスを回避できる。Heckman and Wolpin (1976) は、連邦契約遵守局 (The Office of Federal Contract Compliance (OFCC)) がマイノリティ労働者の地位向上に役立っているかを評価するのに、このモデルを用いている。また、Cooley et al. (1979) では、MDTA が収入に及ぼす効果について、このモデルで分析している。

2.2.3.2 固定効果モデル (The Fixed Effect Model)¹⁸

もし個人 i についてプログラム受講前後についての少なくとも2時点のパネルデータが存在するのであれば、プログラム受講がプログラムによる収入に及ぼす効果を正確に推定することができる¹⁹。なぜならば、パネルデータを用いて「固定効果モデル」で推定を行うと、プログラム受講の有無を表す変数と誤差項の間の相関関係、すなわち内生性の問題を回避できるからである。この内生性回避のメカニズムについて、以下で説明する。

個人 i の t 時点における収入が、以下の式で表せるものとする。

$$Y_{it} = \beta X_{it} + D_i \alpha_t + u_{it}$$

ここで、 X は観察可能な個人 i の t 期における個人属性を表すベクトル (例えば、年齢、教育年数、居住地域、性など) であり、観察可能な個人に固有な変数と期間 t に固有な変数の両方を含んでいる。 D はプログラムを受講した場合 1、受講しなかった場合 0 となるダミー変数である。そして、 u は個人 i の t 期における誤差項で、個人 i に固有で時間 t の影響は受

¹⁸ ここでの記述は、黒澤 (2001)、浅野・中村 (2000, 第12章) に拠っている。

¹⁹ ここで「正確に推定」とは、一致推定量となることを意味する。一致推定量とは一致性を満たす推定量のことで、真のパラメータとの誤差が一定以内となる確率が1に近づく性質をいう。

けない要因（以下、個人効果） μ_i とその他の攪乱要因 ε_{it} に分解できると仮定する²⁰。よって、 u は、

$$u_{it} = \mu_i + \varepsilon_{it}$$

と表すことができる²¹。ここで、個人効果 μ_i の期待値はゼロで分散一定、別の人の個人効果とは無相関、さらに ε_{it} とも無相関であるという仮定を満たすとすると、

$$\begin{aligned} E(\mu_i) &= 0 \\ \text{Var}(\mu_i) &= \sigma_\mu^2 \\ \text{Cov}(\mu_i, \mu_j) &= 0, \quad i \neq j \\ \text{Cov}(\mu_i, \varepsilon_{it}) &= 0. \end{aligned}$$

と表すことができる。この仮定を満たすモデルを、「固定効果モデル」²²という。そして、その他の攪乱要因 ε_{it} は相互に無相関で、分散が均一であると仮定すると、

$$\begin{aligned} E(\varepsilon_{it}) &= 0 \\ \text{Var}(\varepsilon_{it}) &= \sigma_\varepsilon^2 \\ \text{Cov}(\varepsilon_{it}, \varepsilon_{js}) &= 0, \quad i \neq j, \quad t \neq s. \end{aligned}$$

と表せる。つまり、誤差項 u は、 X や D 、及び個人効果 μ とは無相関ということの意味し、次のように表すことができる。

$$\begin{aligned} \text{Var}(u_{it}) &= \text{Var}(\mu_i + \varepsilon_{it}) = \sigma_\mu^2 + \sigma_\varepsilon^2 \\ \text{Cov}(u_{it}, u_{js}) &= 0, \quad i \neq j \\ \text{Cov}(u_{it}, u_{js}) &= \text{Cov}(\mu_i + \varepsilon_{it}, \mu_i + \varepsilon_{js}) = \sigma_\mu^2, \quad t \neq s. \end{aligned}$$

u と D の相関は個人効果を通じてもたらされると考えられるので、この仮定に基づくと、同じ個人についてのプログラム受講前と受講後の収入の差をとることによって、個人効果を除去できる。つまり、プログラム受講者と非受講者の違いが解消されるのである。これは、

$$E(u_{it} - u_{is} | X_{it} - X_{is}, D_i) = 0, \quad s < k < t.$$

²⁰ 時間 t に依存しない個人 i に固有な要素として、能力が考えられる。

²¹ より一般的なモデルでは、期間 t に固有な要因 λ_t も考慮に入れた上でモデル設定を行う。その場合、 u は、 $u_{it} = \mu_i + \lambda_t + \varepsilon_{it}$ 、と表される。ここでは説明を分かりやすくするために、個人 i に固有な要因とその他の攪乱要因だけをモデルに取り入れて説明することとする。つまり、ここでは期間に固有な要因はないと仮定しているが、期間 t に固有な要因 λ_t を考慮に入れた場合も、全く同じ議論が展開できる。

²² LSDV (Least Squares Dummy Variable) 推定量、Within Estimator (級内推定量)ともいう。固定効果モデルの場合、個人効果は確率変数ではなく、固定した定数として扱われることも多い。

と表すことができる。ここで、 k とはプログラムを受講するかしないかを考え、受講すると決心した場合に受講する時期のことである。であるから、 u の構造についての仮定が正しければ、

$$Y_{it} - Y_{is} = \beta (X_{it} - X_{is}) + D_i \alpha_t + (u_{it} - u_{is})$$

と推定式を置くことで、 D と「 $u_{it} - u_{is}$ 」の間の内生性問題を回避できることになり、 α_t の一致推定量を得ることができる。この方法を用いて政策評価を行った代表的な論文として、Bassi (1984) が挙げられる²³。

2.2.3.3 Difference-in-differences 推定法²⁴

パネルデータが非常に充実していて、多時点にわたってデータ収集がなされている場合、difference-in-differences 推定法(以下、D-in-D 法)を用いて、政策評価を行うことができる。これは、プログラム受講後とプログラム受講前の収入の違いと、プログラム受講前のある 2 時点間の収入の違いを比較することで、プログラムの効果を調べる方法である²⁵。つまり、プログラムを受講してからの収入の増分が、プログラムを受講しなかったときに経験した収入の増分を上回っていれば、プログラム受講の効果があったということになる。

個人 i について、 $t, s, s-1$ という 3 時点のデータがあるとする。ここで、 s はプログラム受講前の時点を指し、 k はプログラム受講時点を指す。 $t > k > s > s-1$ 、という関係にあるものとする。収入を表す式を以下の式で表すことができるものとする。

$$Y_{ir} = \beta X_{ir} + D_i \alpha_r + u_{ir}, \quad r = s-1, s, k, t.$$

誤差項 u_{ir} には、固定効果モデルと同じ仮定を置く。このようにモデル設定をすると、 u_{ir} が個人効果以外にも 時間 r とともに変動する部分を含むという仮定に基づいた場合の α_t の一致推定量を得ることができる²⁶。なぜならば、「 $u_{it} - u_{is}$ 」と $(t-s)(u_{is} - u_{is-1})$ との差が X や D と無相関になるからである。すなわち、

²³ アメリカの職業訓練プログラムの CETA が訓練受講者の収入に及ぼす効果を計測している。

²⁴ ここでの記述は、黒澤 (2001)、Ashenfelter and Card (1985)、Heckman et al. (1999, pp.1894-1896) に拠っている。

²⁵ プログラム参加者の t 期と s 期の収入の差と ($t < s$, s 期はプログラム受講後の時点を指す。)非参加者の t 期と s 期の収入の差の違いを用いてプログラムの効果を推定する方法を D-in-D 法と定義することもある。(Wooldridge (2001, pp.129-130))

²⁶ u_{it} が個人効果以外にも、時間の経過を通じて変動する状況としては、プログラム参加者の参加確率や限界生産性が労働市場の需給状況に影響を受ける場合などが考えられる。

$$E[\{(u_{it} - u_{is}) - (t - s)(u_{is} - u_{is-1})\} | (X_{it} - X_{is}), (X_{is} - X_{is-1}), D_i] = 0$$

となることから、

$$(Y_{it} - Y_{is}) - (t - s)(Y_{is} - Y_{is-1})$$

を被説明変数とするモデルを推定することによって、 α_t の一致推定量を得ることができるのである。この α_t をD-in-D推定値という。D-in-D法による政策評価の代表的な論文として、Ashenfelter and Card (1985)を挙げる事ができる²⁷。

【参考文献】

浅野 哲・中村二郎 (2000) 『計量経済学』, 有斐閣.

黒澤昌子(2001) 「職業訓練施策の評価: 非実験的および実験的方法による検証のレビュー」, 『経済研究』 (明治学院大学), 第120号: pp.1-22.

松田芳郎・伴 金美・美添泰人(2000) 『ミクロ統計の集計解析と技法』, 日本評論社.

Ashenfelter, Orley C. and David Card (1985) "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," *Review of Economics and Statistics*, Vol. 67, No. 3: pp.648-660.

Bassi, Laurie J. (1984) "Estimating the Effect of Training Programs with Non-Random Selection," *Review of Economics and Statistics*, Vol. 66, No. 1: pp.36-43.

Cooley, Thomas F., Timothy W. McGuire and Edward C. Prescott (1979) "Earnings and Employment Dynamics of Manpower Trainees: An Exploratory Econometric Analysis," In Bloch, Farrell E. (eds), *Research in Labor Economics, Supplement 1: Evaluating Manpower Training Programs*, JAI Press Inc.: pp.119-147.

Heckman, James J. and Kenneth I. Wolpin (1976) "Does the Contract Compliance Program Work? An Analysis of Chicago Data," *Industrial and Labor Relations Review*, Vol. 29, No. 4: pp.544-564.

Heckman, James J., Robert J. LaLonde and Jefferey A. Smith (1999) "The Economics and Econometrics of Active Labor Market Programs," In Ashenfelter, Orley C. and D. Card (eds), *Handbook of Labor Economics, Vol. 3*, Elsevier Science: pp.1865-2097.

Wooldridge, Jeffrey M. (2001), *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

²⁷ アメリカの職業訓練プログラムの CETA が訓練受講者の収入に及ぼす効果を計測している。