労働政策評価の計量経済学

川口　大司
（一橋大学准教授）

労働政策の効果の推定に用いるパラメトリックな計量経済学的な手法を紹介する。非実験データから労働政策の因果的影響を識別するための手法として，観察可能な属性を制御する最小二乗推定法と実験的変動を抽出する操作変数推定法があるがそれぞれを解説する。加えて観察不能な属性あるいはショックを制御する方法として，差の差の推定法，パネル推定法を紹介する。

目　次

Ⅰ　はじめに

独立行政法人・雇用能力開発機構のあり方が現在検討されている。雇用能力開発機構は労働者に公共職業訓練の機会を与えることを主要な業務としているが，そのあり方をめぐってもっとも重要な論点となるべきは，機構の行っている職業訓練が労働市場で評価される能力を労働者に授けているかを数量的に評価し費用便益分析を行うことである。そのためには公共職業訓練が受講者の雇用確率や賃金率をどの程度引き上げているかを推定する作業が必要となる。公共職業訓練評価の貴重な試みとして黒澤（2003）があげられるが，同様の分析が引き続き必要である。

例えばハローワーク来訪者を対象に調査を行い，その１年後に追跡調査して１年の間の公共職業訓練受講の有無と現在の就業状況をきいたデータを作ることが考えられるが，以上のようなデータの存在を前提としてどのように職業訓練の効果を評価すべきかを考えたい。公共職業訓練の受講者と非受講者の就業率や賃金率を比較し，その差を公共職業訓練の効果であると考えるのが一番単純だが，受講者と非受講者はそれぞれに異なったタイプの労働者である可能性が高く，単純比較には問題がありそうだ。

計量経済学はこのような状況において政策効果を抽出するための手法を提供するが，この論文ではその代表的手法を解説する。ここでは政策評価の目的を，ある政策介入（例えば公共職業訓練）が結果（例えば訓練終了後の賃金率）にどれだけの影響を与えたかという処置効果（treatment effect）の推定に限定し，さらに処置効果の大きさが各個人で同じという仮定のもとで行われるパラメトリックな分析手法を紹介する。

政策の結果への効果の推定とは政策から結果への因果的影響の大きさを測定することであるが，因果的影響とは，他の条件を一定としたとき政策 x の変動に反応して結果 y がどれだけ変動するか，その大きさのことである。

非実験データから x から y への因果関係を読み解くためには，①他の条件を一定とするために観察可能な特性を制御したり，観察可能な特性を利用して適切な比較群を作り出す方法と②他の条

件とは無関係に x が変動するような状況を探し出して，そのときに x の変化に対して y がどのような変化をしたのかを調べる方法がある。

他の条件を直接制御しようとする方法として多重回帰法があるが，加えて同一個体の中の x の変化に対して y がどのような変化をするかを調べるパネルデータを用いた固定効果推定法も，他の条件が同一個体の中では通時的には変化しないという仮定の下で，他の条件を直接制御しようとする方法だといえる。

自然科学の諸分野では，変数 y に影響を与えうる他の要因とは独立に変数 x を操作する実験を行い，変数 x から変数 y への因果的影響を測定する。社会科学における実験は難しいケースが多いが，ある歴史事象や制度変化によって結果 y に影響を与えうる他の条件とは無関係に x が動く状況が生まれることがある。この状況を使って因果的影響の推定を行うことを自然実験法と呼ぶ。ある歴史事象や制度変化が x のすべてではないにせよ，その一部を動かすケースもあるが，この一部の x の変動に対して y がどのように反応するかを調べることで x から y への因果的影響を推定する手法のことを操作変数推定法という。

以下ではこれらの手法を順番に解説したい。

Ⅱ　最小二乗推定量

1　母集団における構造モデル

政策分析を行うに当たっては，ある政策 x_1 から結果 y への政策効果を記述すると考えられるモデルを，母集団における x_1 から y への因果関係を描写するものとして設定する。この際に x_1 以外に y に影響をあたえ政策分析を行うに当たって，「他の条件」として一定にとどめたい変数群 $x_2, x_3, ..., x_k$ が存在するとして，x_1 から x_k が結果 y に与える影響は次のような構造で記述できると仮定する。

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_k x_k+u \qquad (1)$$

ただしここで y，x_1，...，x_k ならびに u はすべて確率変数である。最後の項 u は誤差項と呼ばれるもので，観察されない y の決定要因や y の測定誤差といったすべての要因を含む。u は“unobserved”の略である。このとき未知のパラメータ $\beta_j, j=1, ..., k$ を観察可能なデータ $\{y_i, x_{1i}, x_{2i}, ..., x_{ki}\}_{i=1}^{n}$ から推定することが分析の目的となる。特に重要なのは β_1 であり，x_1 が1単位変化したときに y が何単位反応するかという政策効果を示すパラメータである。

たとえば，y が時間当たり賃金率の自然対数値，x_1 が公共職業訓練に参加した場合には1をとり，そうでない場合には0をとる2値変数であるときに β_1 は職業訓練の収益率と呼ばれるものとなる。追加的な説明変数 $x_2, ..., x_k$ には観察可能な時間当たり賃金率決定要因，たとえば教育年数，職業経験年数，勤続年数などを含めることになる。また，誤差項 u にはデータからは観察されない時間当たり賃金率の決定要因で，たとえば一般的には「能力」と漠然と呼ばれているものが含まれる。

先述のようにパラメータ $\beta_j, j=1, ..., k$ の推定が分析目標であるが，これらパラメータを観察可能な確率変数 $\{y, x_1, x_2, ..., x_k\}$ の母集団モーメントの関数として一意に定めることが可能かどうか，すなわちパラメータが識別可能かどうかを議論する必要がある。以上の式のパラメータが識別可能であるための必要条件は $E(u)=0$ と $cov(xj,u)=0, j=1, 2, ...k$ が成立することである[1)]。説明変数ベクトル $x=[x_1 x_2 ... x_k]$ を定義すると，この二つの条件が成立するための十分条件は

$$E(u \mid x)=0 \qquad (2)$$

で，x で条件付けたときの誤差項の条件付期待値が x には依存しないことが識別の条件となる。職業訓練の収益率の例では，訓練を受けた集団と受けなかった集団の平均的な「能力」が異ならないという仮定をおいたときにのみ，データから職業訓練の収益率を識別することができる。ここで，ある説明変数 x_j が動くときに $E(u \mid x)$ が動くとき，変数 x_j は内生であるという。逆に x_j が動いても $E(u \mid x)$ が動かないとき，x_j は外生だという。

構造式(1)と条件付期待値の仮定(2)をあわせることにより，

$$E(y \mid x)=\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_k x_k \qquad (3)$$

が得られる。すなわちxごとのyの条件付期待値がどのように決定されているか構造式はモデル化しているといえる。

条件付期待値の仮定が満たされるかが重要だが，説明変数x_jが外生であるかを見分けるためにはx_jが人によって異なる理由を考えればいい。たとえば，職業訓練を受ける人と受けない人がいる理由を考えると，①職探しに意欲的な人とそうでない人がいる，②家庭の事情で職業訓練を受ける時間が取れる人と取れない人がいる，③住んでいる地域に職業訓練施設がある人とない人がいる，などの理由が思いつくだろう。それでは，これらの要因がyに影響を与えるのはx_jを通じてのみかを考えよう。一つ目の理由である「職探しへの意欲の有無」は職業訓練を受けることを通じてのみ所得に影響を与えるのだろうか。同じ職業訓練受講者の中で職探しの意欲の有無によって所得分布が異なるならば，職探しの意欲の有無は職業訓練受講以外の経路を通じて所得に影響を与えていると考えられる。このとき職探しへの意欲の有無は誤差項と相関するため，結果として職業訓練の受講と誤差項は相関を持つことになる。

2　最小二乗推定量の導出ならびにその特性

パラメータベクトル$\beta=[\beta_1\beta_2\ldots\beta_k]'$を定義すれば，構造式は

$$y=x\beta+u \qquad (4)$$

とかけるが，ここからパラメータベクトルをどのように識別できるかを簡単に導こう。

両辺にxの転置ベクトルであるx'をかけ期待値をとることによって，

$$Ex'y=Ex'x\beta+Ex'u \qquad (5)$$

が得られる。繰り返し期待値の公式により$E(x'u)=EE(x'u \mid x)=Ex'E(u \mid x)=0$なので$E(x'x)$に逆行列が存在すれば，

$$\beta=(Ex'x)^{-1}Ex'y \qquad (6)$$

が得られる。パラメータを母集団のモーメントで示すことで識別可能性が示された。

このサンプル対応である

$$\hat{\beta}=(n^{-1}\sum_{i=1}^{n}x_i'x_i)^{-1}(n^{-1}\sum_{i=1}^{n}x_i'y_i) \qquad (7)$$

が最小二乗推定量である。

この最小二乗推定量は母集団からのランダムサンプル$\{y_i x_i\}_{i=1}^{n}$が得られれば求められるが，$\hat{\beta}$は用いるサンプルごとに異なった値をとる確率変数である。すなわち，母集団において真のパラメータβは一つしか存在しないが，$\hat{\beta}$にはサンプリング変動がある。

上記の最小二乗推定量の表現に$y_i=x_i\beta+u_i$を代入すると，

$$\hat{\beta}=\beta+(n^{-1}\sum_{i=1}^{n}x_i'x_i)^{-1}(n^{-1}\sum_{i=1}^{n}x_i'u_i) \qquad (8)$$

が得られる。以上の表現の第二項目がサンプル変動する部分となる。それではこの確率変数$\hat{\beta}$の期待値はどのような値をとるだろうか。条件付期待値の仮定$E(u \mid x)$が満たされているとすると，

$$\begin{aligned}E(\hat{\beta})&=EE(\hat{\beta} \mid x)\\&=\beta+E[(\sum_{i=1}^{n}x_i'x_i)^{-1}\sum_{i=1}^{n}x_i'E(u_i \mid x)]\\&=\beta \qquad (9)\end{aligned}$$

であるため，期待値は真のパラメータβと等しくなる。この性質のことを不偏性という。条件付期待値に関する仮定$E(u \mid x)=0$が満たされていないと最後の項が0にならないため不偏性が満たされなくなる。最小二乗推定量の期待値と真のパラメータの差のことをバイアスという。仮に誤差項の期待値がxとともに動いてしまうとすると，最小二乗推定量はバイアスを持つ。

以上は有限サンプルでの話であったが，サンプルサイズが無限大となったときに推定量βがどの点に確率収束するかを議論したい。サンプル平均が母集団の期待値に収束すると仮定すると(i.e. $(n^{-1}\sum_{i=1}^{n}x_i'x_i)^{-1}\to E(x'x)^{-1}$と$n^{-1}\sum_{i=1}^{n}x_i'u_i\to E(x'u)$)

$$\hat{\beta}\to\beta+(Ex'x)^{-1}E(x'u). \qquad (10)$$

ここで，条件付期待値の仮定 $E(u \mid x) = 0$ が満たされていれば第二項はゼロとなるため，最小二乗推定量は真のパラメータに確率収束する。この性質のことを一致性という。条件付期待値の仮定が満たされていない場合，最小二乗推定量の確率収束先と真のパラメータの差のことを漸近バイアスという。

3　最小二乗推定量のバイアス

相当数の私立大学が定員割れを起こす中，大学の意義が厳しく問われるようになっているが，大学が果たすひとつの役割は学生に人的資本を蓄積させ彼らの所得獲得能力を上昇させることである。その役割が果たされているかどうかを調べるためには大学で学ぶことが因果関係の意味で賃金率を上昇させているかを調べる必要がある。そこで y が時間当たり賃金率，x_1 が教育年数であるケースを例に取りながら，教育の収益率を一致推定することの難しさを通じて，最小二乗推定量のバイアスについて議論したい。

被説明変数 y を賃金率の自然対数，説明変数 x_1 を教育年数とすれば x_1 の係数である β_1 は教育の収益率を示す。一般に賃金率は職業経験年数に従い上昇していくが，その上昇の仕方が逓減していくことが知られている。このことを捉えるために x_2 に潜在経験年数，x_3 に潜在経験年数の二乗を導入しよう。パラメータが $\beta_2 > 0, \beta_3 < 0$ を満たしていれば，職業経験年数と時間当たり賃金率の自然対数値は右上がりで傾きが逓減していき，潜在経験年数が $-\beta_2/(2\beta_3)$ を超えたところから右下がりの形状を持つようになる。さらに男女間の賃金格差を捉えるために女性ならば1，男性ならば0をとるダミー変数を x_4 として導入しよう。ここで $(\exp(\beta_4)-1)\times 100$ が女性の賃金率が男性の賃金率よりも何%低いかを示す数字となる。

以上の賃金関数の例では誤差項 u には能力など観察不能であるものの賃金率を決定する要因が含まれていると考えられる。x_1 を教育年数としたときに，教育年数が長いほど能力の平均値が高いとすると賃金方程式の各係数はバイアスを持つ。このとき仮に x_2, x_3, x_4 の各説明変数が u と相関を持たないとしても，すべての係数にバイアスが発生しうる。

このバイアスの方向を多重回帰の文脈で明らかにすることは難しいが，定数項以外の説明変数が一つの場合，説明変数と誤差項が正の相関を持つ場合には係数には正のバイアスがかかることを示すことができる。説明変数ベクトルの x_0 を定数項，x_1 を教育年数としたときに教育のリターンの推定量の確率収束先は，

$$\mathrm{plim}\,\hat{\beta} = \beta + cov(x_1, u)/var(x_1) \tag{11}$$

と示すことができる。

この表現から明らかなように漸近バイアスの大きさは説明変数と誤差項の相関が強いほど，説明変数の変動が小さいほど大きくなる。またバイアス項の分母は常に正であるため，説明変数と誤差項の間の相関が正であれば，正のバイアスが，相関が負であれば負のバイアスがかかることがわかる。たとえば誤差項 u のなかに生来の能力のような変数が含まれているとすると，これは教育年数 x_1 と正の相関を持つことが予想される。よって，教育のリターンの推定量 $\hat{\beta}$ は実際の教育のリターン β よりも大きな値に確率収束し，真の教育のリターンを過大推定することになる。これは教育年数が長い者は生来の能力が高いといった要因でも賃金率が高いのであるが，それをすべてあたかも教育の賃金率に与える因果関係であるかのように捉えてしまうために引き起こされているバイアスである。

真の決定因を説明変数に導入しないことによって発生するバイアスのことを省略変数バイアス（omitted variable bias）と呼ぶ。この問題を回避し教育年数が賃金率に与える因果的影響を推定するためには，たとえば各個人のIQテストの成績 x_2 を入手し，y を x_1 と x_2 に多重回帰して，$\hat{\beta}_1$ と $\hat{\beta}_2$ を推定することである。多重回帰を行う主な目的は省略変数バイアスを回避し，真の因果的影響を推定することである。すなわち $(u \mid x_1) \neq 0$ という内生性がある状況で，u のなかに含まれている x_1 との相関をもたらす要因を説明変数に加えることで $E(u \mid x_1, x_2, ..., x_k) = 0$ の成立をもっともらしくさせること，観察可能な変数を制御することによって因果関係の識別を行うということ

が目的である。

説明変数を追加する理由を考えながら説明変数を追加しないと処理後バイアス（post-treatment bias）といった問題が発生してしまう。たとえば，賃金関数の中にあらゆる職種を示す職種ダミーを導入することはできる。しかしながら「弁護士」「配管工」という狭く限定された職種のダミー変数を説明変数に加えることは適切ではない。なぜならば教育年数は職種の選択という経路を通じて賃金率に影響を与えていることが予想されるのに，その重要な経路を遮断した上で教育の収益率を計算すると収益率が過小に推定されることが予想されるからである。この場合は賃金関数の中に職種ダミーを含めずに回帰分析をすることが望ましい。ここでの職種のように教育年数という処理変数によって影響を受ける変数を処理後共変数（post-treatment covariate）と呼ぶ（Rosenbaum 1984）[2]。

もう一つ重要な例としてバイアスの伝播について議論しよう。先と同じく賃金方程式を定数項，教育年数，潜在経験年数の三つの説明変数に回帰するモデルを考える。

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u \tag{12}$$

ここで，y は時間当たり賃金率の自然対数値，x_1 は教育年数，x_2 は潜在経験年数であり，$u = 0$ は成立しているが，$cov(x_1, u) > 0$ で教育年数が内生であるとする。ただし，潜在経験年数と能力が相関しているとは考えにくいので，$cov(x_2, u) = 0$ が成立しているとしよう。このとき $\hat{\beta}_1$ に上方バイアスがかかることは予想されるが，$\hat{\beta}_2$ は一致性を持つだろうか？ 上記の設定の下での OLS 推定量のバイアスは以下のように示すことができる。

$$\begin{aligned} &\mathrm{plim}\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix} - \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} \\ &= \frac{1}{var(x_1)var(x_2) - (cov(x_1, x_2))^2} \\ &\quad \cdot \begin{pmatrix} var(x_2)\cdot cov(x_1, u) \\ -cov(x_1, x_2)\cdot cov(x_1, u) \end{pmatrix} \end{aligned} \tag{13}$$

すなわち x_1 と x_2 に相関があれば $\hat{\beta}_2$ にも漸近バイアスが発生することになる。

ここで x_1 は教育年数，x_2 は潜在経験年数であるが，潜在経験年数が長い年長者ほど世代が古いため教育年数が短いことが予想され，$cov(x_1, x_2) < 0$ であることが予想される。仮にそうであるとすると $\hat{\beta}_2$ には上方バイアスがかかることになる。若年と中高年の賃金を比べた場合，若年は教育年数が長いがこの教育の効果が過大に推定されていて真の平均賃金よりも高い平均賃金を予想してしまう。その一方で中高年は教育が短いがその効果が過大に推定されているので真の平均賃金よりも低い平均賃金を予想してしまう。教育のリターンの推定にかかるバイアスゆえに，若い人は予測される平均賃金が高すぎる一方で，中高年の人は予測される賃金率が低すぎる。この不整合のつじつまを合わせるために潜在経験年数のリターンが過大に推定されるということが起こっているわけである。二つの変数が正の相関を持ち，一方に正のバイアスがかっているとき，もう一方には負のバイアスがかかってつじつまを合わせていると考えれば記憶しやすい。

4　差の差による推定

OLS による多重回帰分析は観察可能な変数を制御することで，内生性の問題に対処しようとするアプローチであった。しかしながら，観察可能な変数を用いて制御できることには限界があり，観察不能な属性が興味のある処置変数と相関する可能性を否定できないことがある。そのような状況で，きわめて同質な人々であるけれども，政策介入の有無だけが異なると想定できる 2 つのグループの被説明変数の動きを比較するアプローチを用いることがあり，差の差の推定法と呼ぶ。

米国では母子家庭の世帯主の就業を促しつつ，所得移転を行う手段として税額所得控除（EITC）の制度が導入されており，実質的には負の所得税として機能している。1986 年にこの制度は拡充されたが，Eissa and Liebman（1996）はその就業促進効果を推定した。EITC 拡充の対象となったのは子供のいる貧困世帯だけだったので，1986 年前後の子のいる独身女性の就業確率の変化をみれば，政策変更の効果を推定できそうだが，1986 年を境として就業率の変化に影響を与えるような景気変動などがあった可能性もある。この可能性

を捉えるために子のいない独身女性の就業率の変化で景気変動の要因を捉えることを彼らは考えた。表1に推定の結果を紹介しているが，子のいる独身女性と子のいない独身女性の就業率への景気変動の影響が同一であれば，差の差の推定量は政策効果を推定する。この結果はEITCの拡充が政策の目標となった子供のいる独身女性の就業確率を上昇させたことを示唆している。

差の差の推定量がゼロとは異ならないという帰無仮説の検定を行うに当たっては，差の差の推定量の正確さを示す標準誤差を計算することが必要となるが，容易な方法が以下の回帰式をOLS推定することである。

$$y = \beta_0 + \beta_1 T + \beta_2 A + \beta_3 TA + u. \tag{14}$$

ここで，条件付期待値ゼロの仮定 $E(u \mid T, A) = 0$ は満たされているとする。表2により示唆されるように，この式をOLS推定することによって得られる $\hat{\beta}_3$ は差の差の推定量となり，$\hat{\beta}_3$ の標準誤差が差の差の推定量の標準誤差ということになる。

しばしば上記の回帰式を推定するに当たりマイクロデータが利用されるが，標準誤差の推定に当たっては注意が必要になる。Eissa and Liebman（1996）の例を見たときに，処置群と比較群に違うショックが加わっていたとしても彼らの結果は得られると思われた読者がいるかもしれない。条件付期待値ゼロ $E(u \mid T, A) = 0$ を仮定しているので，平均的にはその可能性は無視できるのだが，分析サンプルにそのようなショックが加わっている可能性は否定できず，差の差の推定量がたまたま大きな値をとっている可能性がある。その「たまたま」を表現するように標準誤差も大きく推定されれば問題はないのだが，通常のOLS推定を行った際の標準誤差は独立サンプリングを仮定して推定されているため，同一グループの構成員全員が共通のショックを経験した可能性を許す形となっていない。各個人が独立のショックを経験していると仮定するため，ショックの影響は互いに打ち消しあって小さなものであるかのように考え，結果として標準誤差の過小推定がおこる。

一つの解決策として以下のような定式化を考えたい。

$$y_{ijt} = \beta_0 + \beta_1 T_j + \beta_2 A_t + \beta_3 T_j A_t + c_{jt} + u_{ijt} \tag{15}$$

ここで，添え字 i は個人，j はグループ，t は時間を示すものである。ここで，誤差項にグループ全体が時間ごとに異なるショックを経験している可能性を許していることに注意してほしい。$E(c_{jt} \mid T_j, A_j) = 0, E(u_{ijt} \mid T_j, A_j) = 0$ は引き続き仮定する。個人レベルの誤差項は $e_{ijt} = c_{jt} + u_{ijt}$ と定義できるが，同じグループ j から t 時点にサンプルされた観察値同士では誤差項が相関することになる（i.e. $cov(e_{ljt}, e_{mjt}) \neq 0$ for $l \neq m$）。このような相関を許す標準誤差の推定法がclustering robust standard errorである。この問題は経済学の文献ではMoulton（1990）によって指摘された。この水準での標準誤差の補正は少なくとも必要である。

上記の問題に加えてBertrand, Duflo and Mullainathan（2004）は c_{jt} が系列相関を持つ可能性を指摘し，時系列方向に長い繰り返し横断面データ（repeated crosssection data）を使った推定ではDD推定量の標準誤差がさらに過小推定される可能性を指摘した。仮に誤差項が強い系列相関

表1　EITC拡充前後の就業率

	EITC導入前	EITC導入後	差	差の差
子供がいる未婚女性（処置群）	72.9	75.3	2.4	2.4
子供がいない未婚女性（比較群）	95.2	95.2	0.0	

表2　回帰を用いた差の差の推定

	前	後	差	差の差
処置群	$\beta_0+\beta_1$	$\beta_0+\beta_1+\beta_2+\beta_3$	$\beta_2+\beta_3$	β_3
比較群	β_0	$\beta_0+\beta_2$	β_2	

を持つとすると，政策変更の前後で誤差項の平均的な水準は異なった水準でとどまり続ける可能性がある。この場合誤差項の水準が異なった場所にたまたまとどまり続けたことがあたかも政策の効果であるかのように推定されてしまう可能性が出てくる。彼女たちは時系列の相関構造をそのまま持ち込んだ block bootstrap や政策変更の前後の時系列データを前と後の 2 時点に集約することによって，標準誤差の過小推定を補正する方法を提案している。

Ⅲ　操作変数推定量

1　直感的理解とワルド推定量

この章では説明変数の動きのうち外生的な要因によって変動している部分だけを使って，説明変数から被説明変数への因果関係を推定する方法を紹介する。この推定手法は操作変数法と呼ばれるが，その直感的な理解を鮮やかに描き出した Angrist（1990）による例から入ろう。Angrist（1990）は従軍経験が除隊後の労働所得に与える影響を推定しようとした。構造式は

$$y_i = \delta d_i + x_i\beta + u_i, \tag{16}$$

ただし，y_i は除隊 10 年後の年間所得，d_i は従軍経験の有無，x_i は年齢や人種などの賃金決定要因，u_i は誤差項である。従軍経験は直接戦地に送られることで身体的・精神的な障害を負うことによって除隊後の所得を低下させることもあるし，民間労働市場での就業経験を失うことで人的資本の蓄積が遅れ所得の低下を招く可能性もある。この大きさを測ることは兵を徴用することの社会的な機会費用を算出する上でも重要である。

ここで問題となるのは従軍経験の有無に関する内生性である。ある個人が従軍するかどうかは無作為割り当てで決まっているわけではなく，民間での給与が低いものや就業機会が限られているものが志願して従軍していることが多いことが指摘されている。従軍するかどうかを決定する低所得の要因が x_i で捉えられないとすると，x_i で条件付けたうえでも d_i と u_i は負の相関を持ち，OLS 推定量 $\hat{\delta}$ には負のバイアスがかかる可能性がある。もとより民間部門での給与が低いものが従軍する可能性が高いため，あたかも従軍経験そのものが所得を低下させるように見えるバイアスが発生するわけである。

その一方で体力的に屈強なもののみが従軍している可能性もあり，この場合，体力に優れているものはもとより所得が高いという可能性が高いので d_i と u_i は正の相関を持つ。結果として OLS 推定量 $\hat{\delta}$ には正のバイアスがかかることになる。

この潜在的なバイアスに対応するために Angrist（1990）が着目したのはベトナム戦争中の 1970 年から 1971 年にかけて行われたくじ引きによる徴兵である。不足する兵員を補充するためすべての成人男子の誕生日に基づくくじ引きが行われ，くじの結果に応じた徴兵が同期間に行われた。従軍要請が来た場合に 1，こなかった場合に 0 をとるダミー変数 z_i を定義すると，この値は誕生日に基づく無作為割り当てであるため所得方程式の誤差項 u_i とは相関を持たない。よって，従軍要請が割り振られた人々と，割り振られなかった人々の間の平均所得を比較すれば，従軍がどれだけの所得の低下をもたらすかを識別できそうである。

しかしながら，無作為割り当てによる従軍要請の有無は実際に従軍したかどうかと完全に対応するわけではない。従軍要請が来たとしても学校に残ることによって実際には従軍しなかった人々もいたし，従軍要請が来なくても志願して従軍した人々もいたからである。これらの無作為割り当てに従わなかった人々が半数以上の相当数に上っていることが知られておりこの問題は深刻であるが，問題の解決は容易である。従軍要請の有無が従軍確率を有意に変動させたことは事実なので，従軍要請のあった人々となかった人々の間の平均所得の違いを計算し，その違いを従軍要請のあった人々となかった人々の実際の従軍確率の違いで割れば，従軍の有無が除隊後の所得に与えた影響を推定できる。

以上の直感を厳密に取り扱うのが操作変数推定法である。従軍要請の有無を示す無作為割り当ての変数 z_i は操作変数と呼ばれる。この操作変数

は以下の2つの仮定を満たしているものとする。

- ●ワルド推定仮定1　操作変数 z_i は説明変数ベクトル x_i で条件付けた上で構造方程式の誤差項 u_i の期待値を変動させない。あるいは $E(u_i \mid x_i, z_i = 1) = E(u_i \mid x_i, z_i = 0) = 0$ が成立する。これは従軍要請の有無が無作為割り当てされているため，観察不能な所得決定要因の分布の期待値と相関関係を持たないことを意味している。
- ●ワルド推定仮定2　操作変数 z_i は説明変数ベクトル x_i で条件付けた上で潜在的な内生変数 d_i の期待値を変動させる。あるいは $E(d_i \mid x_i, z_i = 1) \neq E(d_i \mid x_i, z_i = 0)$ が成立する。これは無作為割り当てによる従軍要請の有無が実際の従軍確率を変動させることを意味している。

以上の二つの仮定が満たされているときに操作変数推定量を導出することができる。まず x と $z = 1$ のもとでの y の条件付期待値を求めると

$$E(y_i \mid x_i, z_i = 1) = \delta E(d_i \mid x_i, z_i = 1) + x_i\beta + E(u_i \mid x_i, z_i = 1) \qquad (17)$$

が得られる。同様に x と $z = 0$ のもとでの y の条件付期待値は

$$E(y_i \mid x_i, z_i = 0) = \delta E(d_i \mid x_i, z_i = 0) + x_i\beta + E(u_i \mid x_i, z_i = 0) \qquad (18)$$

と求まる。ここで，(17)から(18)を引くとワルド推定仮定1より，

$$E(y_i \mid x_i, z_i = 1) - E(y_i \mid x_i, z_i = 0) = \delta[E(d_i \mid x_i, z_i = 1) - E(d_i \mid x_i, z_i = 0)] \qquad (19)$$

が求まる。ワルド推定仮定2より $E(d_i \mid x_i, z_i = 1) - E(d_i \mid x_i, z_i = 0) \neq 0$ なので

$$\delta = \frac{E(y_i \mid x_i, z_i = 1) - E(y_i \mid x_i, z_i = 0)}{E(d_i \mid x_i, z_i = 1) - E(d_i \mid x_i, z_i = 0)} \qquad (20)$$

となり δ を観察可能な確率変数の母集団モーメントの関数として書くことが可能であること，すなわち δ が識別可能であることが示される。このような形式によって表現される操作変数推定量をワルド推定量と言う。

操作変数推定量の要点は，外生変数 z_i の変動によって内生変数 x_i が動いた分に対応して，被説明変数 y_i がどれだけ動いたかを観察することによって x_i から y_i への因果関係を推定しようとする点にある。ここで重要なのは操作変数が被説明変数に影響を与える経路は x_i を通じてのみであるという点であり，これが操作変数推定仮定1である。

また，外生変数 z_i の変動によってもたらされる内生変数 x_i の変動を使って x_i から y_i への因果関係の推定を行おうとしているので，外生変数 z_i の変動が内生変数 x_i の変動をもたらさない場合には推定は不可能となる。これが操作変数推定仮定2の要点である。

2　ワルド推定量から操作変数推定量へ

ここまでの議論は操作変数が二値変数のケースに限定されていたが，この変数が連続変数であったとしても同様の議論ができることを示したい。構造式が以下のとおり与えられているとしよう。

$$y = x\beta + u \qquad (21)$$

ただし x は $1\times k$ の説明変数ベクトル，β は $k\times 1$ のパラメータベクトルである。ここで x を構成する変数の x_1 から x_{k-1} までは外生だが，x_k は内生であるとしよう。ここで x_1 から x_{k-1} のほかに操作変数 z_1 が存在すると，内生変数の外生変数への線形射影は以下のようにかける。

$$x_k = \delta_0 + \delta_1 x_1 + \ldots + \delta_{k-1} x_{k-1} + \theta_1 z_1 + r_k. \qquad (22)$$

ここで，操作変数は以下の二つの仮定を満たしているとしよう。

- ●操作変数推定仮定1　操作変数 z_1 は外生である。すなわち $cov(z_1, u) = 0$ が成立している。
- ●操作変数推定仮定2　操作変数 z_1 は説明変数ベクトル $x_1, \ldots, x_{k-1}$ で条件付けた上で潜在的な内生変数 x_k の期待値を変動させる。すなわち $\theta_1 \neq 0$ が成立している。

以下の議論を簡潔にするために操作変数ベクトルを以下のように定義する。

$$z \equiv [x_1 x_2 \ldots . x_{k-1} z_1]. \qquad (23)$$

この操作変数ベクトルに含まれる変数のすべてが外生であるため，$Ez'u=0$ が成立する。

構造式に z の転置ベクトルである z' を前からかけて，期待値をとれば，

$$Ez'y = Ez'x\beta + Ez'u \tag{24}$$

が得られるが，最後の項は仮定より 0 である。さらに操作変数推定仮定 2 より，$Ez'x$ の逆行列の存在を示すことができるので，パラメータベクトルを次のように母集団モーメントの関数として書くことができる。

$$\beta = (Ez'x)^{-1}Ez'y. \tag{25}$$

よって，β は識別可能である。

以上の母集団モーメントのサンプル対応をとることによって操作変数推定量を定義することができる。

$$\begin{aligned}\hat{\beta} &= (n^{-1}\sum_{i=1}^{n} z_i'x_i)^{-1}(n^{-1}\sum_{i=1}^{n} z_i'y_i)\\ &= \beta + (\sum_{i=1}^{n} z_i'x_i)^{-1}(\sum_{i=1}^{n} z_i'u_i).\end{aligned} \tag{26}$$

ここで注意すべきことは $E[u \mid z]=0$ がいえたとしても，$E\hat{\beta}=\beta$ は必ずしもいえないことである。これは，

$$\begin{aligned}E[\hat{\beta}] &= EE[\hat{\beta} \mid z]\\ &= \beta + EE[(\sum_{i=1}^{n} z_i'x_i)^{-1}(\sum_{i=1}^{n} z_i'u_i) \mid z_i]\end{aligned} \tag{27}$$

となり，第 2 項がゼロとは限らないことからあきらかである。よって，操作変数推定量は不偏推定量ではない。

その一方で一致推定量であることは

$$\begin{aligned}&\mathrm{plim}\hat{\beta}=\beta\\ &+(\mathrm{plim} n^{-1}\sum_{i=1}^{n} z_i'x_i)^{-1}(\mathrm{plim} n^{-1}\sum_{i=1}^{n} z_i'u_i)\\ &=\beta+(Ez'x)^{-1}Ez'u\\ &=\beta\end{aligned} \tag{28}$$

より示すことができる。しかしながら，操作変数推定量が有限サンプルにおけるバイアスをもつことは欠点として認識されるべきである。

Ⅳ　パネル分析

個人や企業など同一個体 i の情報が t 期間をまたがって入手可能であるようなデータをパネルデータという。ある個人が公共職業訓練に参加したときにどれだけ所得が上昇するかを調べたいとしよう。その公共職業訓練が低所得者を対象としたものだとすると，単純に公共職業訓練に参加した人とそうでない人の所得を比較すると，あたかも公共職業訓練に参加することが所得を引き下げているかのように見えるが，パネルデータを用いた推定を行うことでこの問題をある程度解決することができる。例えば，多数の個人を数年にわたって繰り返し調査しているパネル調査があり，その中で公共職業訓練への参加が記録されているとしよう。個人 i が t 年に公共職業訓練に参加した時の $t-1$ 年から t 年にかけての個人 i の所得の変化を観察することで，各個人の異質性を制御しながら公共職業訓練の効果が推定できそうである。

1　OLS 推定量

以上の状況を構造式で以下のように表す。

$$y_{it} = x_{it}\beta + c_i + u_{it}, \tag{29}$$

ただし y_{it} は個人 i の時点 t における時間当たり賃金率の自然対数値であり，x_{it} は個人 i の時点 t における k 個の変数を含む $1\times k$ の確率変数ベクトル，特に x_{1it} は職業訓練に参加した場合に 1 をとるダミー変数であるとしよう。ベクトル β は k 個のパラメータを含む $k\times 1$ のパラメータベクトルであり，特に β_1 が職業訓練の効果を示すパラメータである。誤差項は c_i+u_{it} として表現されているが，このうち c_i は各個人が持つ固有の観察不能要因で時間とともには変化しないもの，u_{it} は時間とともに変化する観察不能要因である。特に c_i のことを観察されない異質性（unobserved heterogeneity）と呼ぶ。

説明変数ベクトル x_i の線型独立性が満たされていることは前提として，誤差項の条件付期待値に関する仮定 $E(c_i+u_{it} \mid x_{it})=0$ が満たされていれば，OLS 推定量は不偏性と一致性をもつので

効率性を無視すればOLS推定量を用いることに問題はない。また，説明変数x_{it}と誤差項c_i+u_{it}が同時点で相関していなければ一致性があるという意味では，後に触れる変量効果推定よりも頑健である。このようなOLS推定量をプールしたOLS推定量（pooled OLS estimator）と呼ぶ。

もっとも仮にu_{it}の部分に系列相関がなくても，$cov(c_i+u_{it}, c_i+u_{i\tau \neq t})=var(c_i)$となるため誤差項は同一個人内で相関を持つ。よって各個人のレベルでのclusteringを許す標準誤差を計算する必要がある。このようなclusteringを仮定した標準誤差は通常の標準誤差よりも拡大するが，これはたとえば同じ100個の観察値であったとしても，10人の10年分のデータよりも100人の1年分のデータのほうがもっている情報量が多いことを意味している。

2　階差推定量

パネルデータを用いることの真価は誤差項の条件付期待値に関する仮定$E(c_i+u_{it}\mid x_{it})=0$が$x_{it}$と$c_i$の相関ゆえに満たされない時に発揮される。たとえば，恒常的に低所得のもののほうが公共職業訓練を受ける可能性が高いと，$cov(c_i, x_{1it})<0$となる。この相関を無視してOLS推定を行えば推定量$\hat{\beta}_1$には下方バイアスがかかる。

そこで，パネルデータの特性を生かして今期の構造式から前期の構造式を引くことによって各個人の1期間の階差をとれば，

$$y_{it}-y_{it-1}=(x_{it}-x_{it-1})\beta+u_{it}-u_{it-1} \tag{30}$$

が得られる。誤差項にはもはやc_iが含まれていないため，$x_{it}-x_{it-1}$と$u_{it}-u_{it-1}$が相関を持たなければOLS推定量$\hat{\beta}$は一致性を持つ。推定手法によらず係数の解釈は構造式に基づいて行うべきである。

この推定量のことを階差推定量（first difference estimator）というが，ここで着目したいのは，推定に用いられている説明変数の変動は同一個人内での変動だけである点であり，個人間の比較は行われていないことである。このことより二つの点に注意を払う必要がある。

まず第一に，x_{it}に含めることができるのは時間とともに個人の中で変動していく要素である必要がある点である。たとえば人種や性別といった変数は一般的に時間を通じて変動しないため推定に用いることができない。また，公共職業訓練を例にとれば，期間中ずっと訓練に不参加あるいは参加の個人は$x_{1it}-x_{1it-1}$が常にゼロとなるため，推定に必要な説明変数の変動を得るためにはサンプル期間中に公共職業訓練への参加と不参加の状態を推移したものが多数含まれる必要がある。状態を推移したものが少ないと説明変数$x_{1it}-x_{1it-1}$の変動が小さくなり，結果として推定量の標準誤差は大きなものとなってしまう。実際に階差推定を行う前に，興味のある説明変数の個人内の変動が十分に大きなものか，データを確認する必要があるといえる。

第二に個人内での説明変数の変動原因を考えることが重要だ。去年まで公共職業訓練に参加していなかったものが今年は参加しているとすると，それはなぜかを考えなければならない。仮に去年所得に負のショックを経験したものが転職を想定しながら今年から公共職業訓練に通い始めたとすると，去年の所得のショックと今年の公共職業訓練への参加は負の相関を持つことになる（$cov(x_{1it}, u_{it-1})<0$）。結果として階差式の説明変数$x_{1it}-x_{1it-1}$と誤差項$u_{it}-u_{it-1}$は正の相関を持つことになり，公共職業訓練参加の所得への影響は過大推定される。これはAshenfelter（1978）が最初に指摘したことからAshenfelter dipと呼ばれる現象であり，職業訓練に参加していなかったとしても起こった所得平均値への回復をあたかも職業訓練への参加の影響としてカウントしてしまうことから起こる[3]。このようになぜ説明変数が動いているのかを考えることの重要性はパネル分析を行うにあたってもいささかも減じない。

上で指摘したように階差推定量が不偏性，一致性を持つためには異なる時点の説明変数と誤差項が相関していないことが追加的に必要となる。そのための十分条件が厳密な外生性（strict exogeneity）と呼ばれるもので，$E(u_{it}\mid x_{i1}, x_{i2}, \ldots, x_{it}, \ldots, x_{iT})=0$によってあらわされる。すなわちすべての期の説明変数がt期の誤差項の期待値を

説明しないということである。先の公共職業訓練の例では，来期の職業訓練の参加が今期の誤差項の期待値の低さを予想するため，厳密な外生性は満たされていない。今期のショックを観察して経済主体が将来の x を調整するというフィードバックはしばしば経済現象において観察されるが，そのとき厳密な外生性は満たされなくなる。

3　固定効果推定量

今までの議論は観察不能な異質性を階差をとることによって誤差項から消すという作業を行ってきたが，本質的には個人間の説明変数の変動を用いず個人内の変動だけを用いることができれば同様の目的は達成できる。各個人の各時点の説明変数を各個人の説明変数の平均値からの乖離に変換することによっても個人内変動だけを取り出すことは可能であり，そのような操作のことを個体内変換（within transformation）という。構造式の各個人について期間平均を計算すれば，

$$\bar{y}_i = \bar{x}_i\beta + \bar{u}_i \tag{31}$$

が得られる。ただしここで，$\bar{y}_i = (1/T)\sum_{t=1}^{T} y_{it}$, $\bar{x}_i = (1/T)\sum_{t=1}^{T} x_{it}$, $\bar{u}_i = (1/T)\sum_{t=1}^{T} u_{it}$ である。構造式から上記の式を引くことによって，

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)\beta + u_{it} - \bar{u}_i \tag{32}$$

が得られる。この式にはすでに c_i が含まれていないので，説明変数ベクトル（$x_{it} - \bar{x}_i$）の各要素が線形独立性を持ち，厳密な外生性の仮定 $E(u_{it} \mid x_{i1}, x_{i2}, ..., x_{it}, ..., x_{iT}) = 0$ が満たされるならば，この式を OLS 推定して得られる固定効果推定量（Fixed Effects Estimator）は不偏性と一致性を持つ。ここで，説明変数には $\bar{x}_i$ すなわちすべての期の説明変数の平均値が含まれているため，厳密な外生性が満たされていないと説明変数と誤差項の相関が発生する点に注意してほしい。

階差推定量と固定効果推定量はともに観察不能な異質性 c_i と説明変数 x_{it} の相関を許す推定量であった。どちらの推定手法が望ましい推定手法であるかはいくつかの点に依存する。

まず u_{it} の系列相関の構造である。これが $u_{it} = u_{it-1} + e_{it}$, ただし e_{it} は系列相関を持たない，とあらわされる単位根過程であるならば階差式の誤差項は e_{it} となるため系列相関を持たず階差推定量が効率的な推定量となる。その一方で u_{it} がもともと系列相関を持たないならば，個体内変換式の誤差項には系列相関がないため固定効果推定量が効率的な推定量となる。

厳密な外生性が満たされていないときのバイアスの大きさも判断の基準となる。前期のショックに対応して今期の説明変数が選択され $cov(x_{it}, u_{it-1}) \neq 0$ となるため厳密な外生性が破られるケースが多いが，階差推定量の場合この影響はパネルの長さ T が大きくなっても消えない。しかしながら固定効果推定量の場合，問題は $\bar{x}_i$ と u_{it-1} の相関によって引き起こされているため，T が大きくなるにつれ問題は解消していく。

4　変量効果推定量

ここでもう一度説明変数 x_{it} と観察されない異質性 c_i が相関しないケースを考えよう。このケースでは以下の式：

$$y_{it} = x_{it}\beta + c_i + u_{it}, \tag{33}$$

を OLS 推定したプールした OLS 推定量が不偏性と一致性を持つことをすでに述べた。さらに各個人の観察可能な異質性の存在により誤差項が各個人の中で相関を持つため，clustering robust 標準誤差を利用することが必要であることを指摘した。

プールした OLS 推定量は個人間の変動と個人内の変動に同様の重みをかけて推定を行っているが，もしも個人間の変動（σ_y^2）が観察されない異質性の変動（σ_c^2）によってもたらされているとするならば，個人間の変動には x_{it} から y_{it} への因果関係を推定するための情報があまり含まれておらず，むしろ個人内の変動により多くの情報量が含まれていることが示唆される。また，パネルデータの時系列方向が長くなれば（T が大きくなれば）なるほど，個人内変動が持つ情報量が相対的に増える。これらの点を考えると個人間の異質性が大きいときや，時系列方向が長いときには個人内変動の情報に重みをかけた推定を行うことが効率性の観点からは望ましいことがわかる。

以上の直感を推定量の形にしたものが変量効果推定量である。変量効果推定量はプールしたOLS推定量と固定効果推定量の中間に位置する推定量で

$$(y_{it}-\lambda\bar{y}_i)=(x_{it}-\lambda\bar{x}_i)\beta+(v_{it}-\lambda\bar{v}_i), \qquad (34)$$

ただし $v_{it}=c_i+u_{it}$ ならびに $\lambda=1-(\frac{\sigma_u^2}{\sigma_u^2+T\sigma_c^2})^{1/2}$ という式をOLS推定して得られる推定量である[4)]。ここで，$(\sigma_c^2/\sigma_u^2)\to\infty$ のときや $T\to\infty$ のときには固定効果推定量と等しくなることがわかる。これは，観察不能な異質性の相対的な重要性が上がるにつれ，パネルの時系列方向が長くなるにつれ個人内の変動に重みをかけて推定を行うことが効率性の観点より望ましいことを示唆している。一方でパネルが時系列方向に短いとき，観察不能な異質性の重要性が相対的に小さいとき，変量効果推定量はpooled OLS推定量に近づくことがわかる。

ここで注意が必要なのは誤差項に引き続き個人の観察不能な異質性 c_i が含まれるため，説明変数 x_{it} と観察されない異質性 c_i が相関する場合には変動効果推定量はバイアスを持つことである。

V　まとめ

紙幅の関係で政策評価に用いる計量経済学の手法について非常に限定された範囲しか言及できなかったが，政策介入を受けた場合の結果 $y(1)$ と受けなかった場合の結果 $y(0)$ が各個人について存在するという枠組み（counterfactual outcomes model または Rubin causal model と呼ばれる）の中で処置効果の大きさが個人によって異なることを許しつつ，処置効果の分布特性をノンパラメトリックに推定する手法が急速に発展している。この手法について，日本語では黒澤（2005）の，英語では Imbens（2004）のサーベイが非常に優れている。しかし，これらの論文で紹介されている分析手法もその多くが，説明変数 x で条件づけた上での政策処置の外生性を必要としている点は注意が必要である。

また，処置効果の異質性を許すと操作変数推定量の推定するものが，若干の追加的仮定の下で局所平均処置効果（local average treatment effect）となることが Angrist, Imbens and Rubin（1996）などにより指摘され，労働経済学者の中でこの概念は共有知識となりつつある。また事業所の従業員規模によって適用される労働政策が異なるケースなど，ある連続変数に基づいて政策介入が非連続に行われる状況を用いて処置効果を推定しようとする回帰不連続設計（regression discontinuity design）を用いた研究も急速に発展している。これについては Imbens and Lemieux（2008）が識別と推定の双方を手際よくまとめた入門的サーベイである。

この論文では労働政策評価を行うに当たって標準的に用いられるパラメトリックな計量経済学的手法を紹介した。これらの手法を用いれば限定された仮定の下ではあるが，非実験データから政策効果を推定することが可能となる。説明変数群で条件づけた上での政策処置の外生性が満たされることが重要だが，この仮定が満たされる蓋然性が高まるように調査を設計することが肝要である。効果的な政策評価を行うためには計量経済学に対する適切な知識を持つものが調査設計の段階からかかわることが必要だといえる。

説明変数で条件付けた上での外生性が満たされるための実際的な条件は，Heckman, Ichimura and Todd（1997）と Heckman *et. al*（1998）が論じているが，政策処置を受けるか受けないかと，労働市場における結果の両方に影響を与える幅広い変数を条件付けることが重要であるとしている。また，処置群と比較群の同質性を担保するために同一の地域からのサンプリングを行うこと，同一の質問票から取られた変数を用いることの重要性も指摘している。

これらを踏まえて日本において公共職業訓練の効果について非実験的に評価を行おうとするならば，特定の期間のうちに特定のハローワークに職探しに現れた失業者に，学歴，年齢，家族構成，過去数年の就業状態，過去数年の年収，といった項目をきき，それらの人々を１年後に追跡再調査し，１年の間の公共職業訓練への参加の有無を聞くとともに現在の就業状態や年収を聞くといった調査を行うことが有効であろう。いたずらに調査

地点を増やすよりも，都市・地方の代表的な調査地点を選び，それぞれの地点から大量標本を採るという調査設計が公共職業訓練効果の推定のためには重要である。公共職業訓練全体に用いられている予算に比べれば微少な調査費と関連行政機関の協力が得られれば十分に実行可能な調査であり，今すぐにでも実行すべき調査であろう。

また，非実験データから得られる公共職業訓練効果の推定値が正しいものであるかを評価するためには，処置群と比較群の確率的割り振りを伴う社会実験の実行が究極的には必要となる。労働者のためになる安定的な労働政策を実行するためには，科学的な政策評価が欠かせないため，政策関係者の説得を粘り強く続けていくことが必要であろう。

1）さらに $E(x'x)$ に逆行列が存在することが必要条件となるが，これは $1, x_1, x_2, ..., x_k$ が線形独立であれば満たされる。

2）逆に賃金関数に職種ダミーを大量に導入することで面白い分析ができるケースもある。教育年数，潜在経験年数，その二乗項といった通常の説明変数に加えて，女性ダミーを右辺に含むモデルを推定すると，女性ダミーの係数はどこの国のデータを使ってもほぼ負に推定される。しかしながら，細かい分類の職種ダミーを含めると女性ダミーの係数が0に近づき，時には統計的な有意性を失うことがある。これが意味するところは，女性が低賃金の職に多くついているがゆえに男女間賃金格差が発生しているというもので，いったん細かい職種をコントロールすれば男女間賃金格差はなくなるということを意味していている。これは差別がないことを意味しないが，職域分離が男女間賃金格差の大きな原因であることを示唆している。

3）何らかの負のショックがきっかけで処置が行われたときに，仮に処置が行われなかったとしても負のショックからの回復によって正の変化が起こり，その影響を処置効果に含めてしまう問題は「平均値への回帰」として統計学において古典的な問題として知られている。

4）導出については（Wooldridge 2001, p. 286）を参照のこと。

参考文献

Angrist, Joshua (1990) "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *American Economic Review*, Vol. 80, No. 3, pp. 313-36.

Angrist, Joshua, Guido Imbens, and Donald Rubin (1996) "Identification of Causal Effects using Instrumental Variables," *Journal of American Statistical Association*, Vol. 91, No. 434, pp. 444-455.

Ashenfelter, Orley (1978) "Estimating the Effect of Training Programs on Earnings," *Review of Economics and Statistics*, Vol. 60, No. 1, pp. 47-57.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004) "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, Vol. 119, No. 1, pp. 249-275.

Eissa, Nada and Jeffrey B. Liebman (1996) "Labor Supply Response to the Earned Income Tax Credit," *Quarterly Journal of Economics*, Vol. 111, No. 2, pp. 605-637.

Heckman, James, Hidehiko Ichimura, Petra Todd (1997) "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, Vol. 64, No. 4, pp. 605-654.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, Petra Todd (1998) "Characterizing Selection Bias Using Experimental Data," *Econometrica*, Vol. 66, No. 5, pp. 1017-1098.

Imbens, Guido W. (2004) "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review," *The Review of Economics and Statistics*, Vol. 86, No. 1, pp. 4-29.

Imbens, Guido W. and Thomas Lemieux (2008) "Regression Discontinuity Designs: A Guide To Practice," *Journal of Econometrics*, Vol. 142, No. 2, pp. 615-635.

Moulton, Brent R. (1990) "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *The Review of Economics and Statistics*, Vol. 72, No. 2, pp. 334-338.

Rosenbaum, Paul R. (1984) "The Consequences of Adjustment for a Concomitant Variable that has been Affected by the Treatment," *Journal of the Royal Statistical Society, Series A*, Vol. 147, pp. 656-666.

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

黒澤昌子（2003）「公共職業訓練の収入への効果」『日本労働研究雑誌』No. 514，pp. 38-49.

―――（2005）「積極労働政策の評価――レビュー」『フィナンシャル・レビュー』第77号，pp. 197-220.

かわぐち・だいじ　一橋大学大学院経済学研究科准教授。最近の主な論文に「誕生日と学校成績・最終学歴」（森啓明と共著）『日本労働研究雑誌』No. 569，pp. 29-42（2007年12月号）など。労働経済学・応用計量経済学専攻。