第 7 章：操作変数法 (IV)

Jonathan Roth

数理計量経済学 I
ブラウン大学

モチベーション

- これまでに、観察可能な特徴を条件として、処置が実質的にランダムに割り当てられている（条件付き非交絡性）場合に、平均処置効果を推定する方法を見てきました。

- また、パネルデータがあり、選択バイアスが時間を通じて一定である（並行トレンド）と仮定できる場合に、条件付き非交絡性の仮定を緩和する方法も学びました。

モチベーション

- これまでに、観察可能な特徴を条件として、処置が実質的にランダムに割り当てられている（条件付き非交絡性）場合に、平均処置効果を推定する方法を見てきました。
- また、パネルデータがあり、選択バイアスが時間を通じて一定である（並行トレンド）と仮定できる場合に、条件付き非交絡性の仮定を緩和する方法も学びました。
- しかし、これらの仮定が妥当ではない場合もあります。

モチベーション

- これまでに、観察可能な特徴を条件として、処置が実質的にランダムに割り当てられている（条件付き非交絡性）場合に、平均処置効果を推定する方法を見てきました。
- また、パネルデータがあり、選択バイアスが時間を通じて一定である（並行トレンド）と仮定できる場合に、条件付き非交絡性の仮定を緩和する方法も学びました。
- しかし、これらの仮定が妥当ではない場合もあります。
  - 観察されない交絡変数が存在することを懸念する場合。
  - パネルデータがない場合、あるいは交絡因子が時間とともに変化する（並行トレンドが崩れる）ことを懸念する場合。

モチベーション

- これまでに、観察可能な特徴を条件として、処置が実質的にランダムに割り当てられている（条件付き非交絡性）場合に、平均処置効果を推定する方法を見てきました。
- また、パネルデータがあり、選択バイアスが時間を通じて一定である（並行トレンド）と仮定できる場合に、条件付き非交絡性の仮定を緩和する方法も学びました。
- しかし、これらの仮定が妥当ではない場合もあります。
  - 観察されない交絡変数が存在することを懸念する場合。
  - パネルデータがない場合、あるいは交絡因子が時間とともに変化する（並行トレンドが崩れる）ことを懸念する場合。
- そのような場合、何ができるでしょうか？

「局所的実験」

- 因果推論の黄金律は実験を行うことです。
- しかし、多くの場合、実験の実施は不可能です。

「局所的実験」

- 因果推論の黄金律は実験を行うことです。
- しかし、多くの場合、実験の実施は不可能です。
- 幸いなことに、関心のある処置の「受け入れ（take-up)」に影響を与えるような（自然）実験が存在することがあります。

「局所的実験」

- 因果推論の黄金律は実験を行うことです。
- しかし、多くの場合、実験の実施は不可能です。
- 幸いなことに、関心のある処置の「受け入れ（take-up）」に影響を与えるような（自然）実験が存在することがあります。
- もしそのような実験があれば、（特定の条件の下で）その実験によって処置を受けることになった人々（「コンプライヤー」）における処置の効果を学ぶことができます。

例：Medicaid の効果

- 政策上の重要な問い：Medicaid は健康状態にどのような影響を与えるか？

例：Medicaid の効果

- 政策上の重要な問い：Medicaid は健康状態にどのような影響を与えるか？
- Medicaid の因果的効果を知るための理想的な状況は、誰が加入するかをランダムに決めることです。

例：Medicaid の効果

- 政策上の重要な問い：Medicaid は健康状態にどのような影響を与えるか？
- Medicaid の因果的効果を知るための理想的な状況は、誰が加入するかをランダムに決めることです。
  - 道徳的・予算的な理由から不可能です。

例：Medicaid の効果

- 政策上の重要な問い：Medicaid は健康状態にどのような影響を与えるか？
- Medicaid の因果的効果を知るための理想的な状況は、誰が加入するかをランダムに決めることです。
  - 道徳的・予算的な理由から不可能です。
- しかし、オレゴン健康保険実験（OHIE）があります。これは、所得が連邦貧困線（FPL）の 100%から 138%の間にある人々を対象に、Medicaid の加入資格を無作為に割り当てたものです。

復習：OHIE の背景

In 2008, a group of uninsured low-income adults in Oregon was selected by lottery to be given the chance to apply for Medicaid. This lottery provides an opportunity to gauge the effects of expanding access to public health insurance on the health care use, financial strain, and health of low-income adults using a randomized controlled design. In the year after random assignment, the treatment group selected by the lottery was about 25 percentage points more likely to have insurance than the control group that was not selected. We find that in this first year, the treatment group had substantively and statistically significantly higher health care utilization (including primary and preventive care as well as hospitalizations), lower out-of-pocket medical expenditures and medical debt (including fewer bills sent to collection), and better self-reported physical and mental health than the control group. *JEL* Codes: H51, H75, I1.

抽選の当選 vs Medicaid への加入

- 以前、OHIE を用いて「加入資格の抽選に当選すること」の因果的効果を推定しました。
- しかし、私たちが本当に知りたいのは「Medicaid に加入すること」の因果的効果ではないでしょうか？

抽選の当選 vs Medicaid への加入

- 以前、OHIE を用いて「加入資格の抽選に当選すること」の因果的効果を推定しました。
- しかし、私たちが本当に知りたいのは「Medicaid に加入すること」の因果的効果ではないでしょうか？
- 抽選に当選することと、Medicaid に実際に加入することは同じではありません：

	当選者 (Winners)	落選者 (Losers)
Medicaid 加入経験あり	0.397	0.141

抽選の当選 vs Medicaid への加入

- 以前、OHIE を用いて「加入資格の抽選に当選すること」の因果的効果を推定しました。
- しかし、私たちが本当に知りたいのは「Medicaid に加入すること」の因果的効果ではないでしょうか？
- 抽選に当選することと、Medicaid に実際に加入することは同じではありません：

	当選者 (Winners)	落選者 (Losers)
Medicaid 加入経験あり	0.397	0.141

- 当選しても加入しない人がいます（書類を返送しなかった、加入資格を失ったなど）。
- 落選しても、他の基準で資格を得て加入する人がいます。

Medicaid 加入がうつ病に与える影響

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 抽選の当選が Medicaid 加入に与える推定効果は？

Medicaid 加入がうつ病に与える影響

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 抽選の当選が Medicaid 加入に与える推定効果は？
  $0.397 - 0.141 = 0.256$

Medicaid 加入がうつ病に与える影響

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 抽選の当選が Medicaid 加入に与える推定効果は？
  $0.397 - 0.141 = 0.256$

- 抽選の当選がうつ病に与える推定効果は？

Medicaid 加入がうつ病に与える影響

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 抽選の当選が Medicaid 加入に与える推定効果は？
  $0.397 - 0.141 = 0.256$

- 抽選の当選がうつ病に与える推定効果は？ $0.306 - 0.329 = -0.023$

Medicaid 加入がうつ病に与える影響

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 抽選の当選が Medicaid 加入に与える推定効果は？
  $0.397 - 0.141 = 0.256$

- 抽選の当選がうつ病に与える推定効果は？ $0.306 - 0.329 = -0.023$

- では、「Medicaid への加入」がうつ病に与える効果を推定するにはどうすればよいでしょうか？

Medicaid 加入がうつ病に与える影響

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 抽選の当選が Medicaid 加入に与える推定効果は？
  $0.397 - 0.141 = 0.256$

- 抽選の当選がうつ病に与える推定効果は？ $0.306 - 0.329 = -0.023$

- では、「Medicaid への加入」がうつ病に与える効果を推定するにはどうすればよいでしょうか？

- 自然な推定方法は、うつ病への効果を加入への効果で割ることです：
  $-0.023/0.256 \approx -0.09 \rightarrow$ 加入者 1 人あたり 9 ポイントの減少。

Medicaid 加入がうつ病に与える影響

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 抽選の当選が Medicaid 加入に与える推定効果は？
  $0.397 - 0.141 = 0.256$
- 抽選の当選がうつ病に与える推定効果は？ $0.306 - 0.329 = -0.023$
- では、「Medicaid への加入」がうつ病に与える効果を推定するにはどうすればよいでしょうか？
- 自然な推定方法は、うつ病への効果を加入への効果で割ることです：
  $-0.023/0.256 \approx -0.09 \rightarrow$ 加入者 1 人あたり 9 ポイントの減少。
- これは操作変数法 **(IV)** による推定値と呼ばれます。これはどのような場合に機能するのでしょうか？ また、どのように解釈すべきでしょうか？

局所平均処置効果 (LATE)

- 特定の仮定の下で、操作変数法を用いることで局所平均処置効果 (Local Average Treatment Effect, LATE) を識別できることを示します。

局所平均処置効果 (LATE)

- 特定の仮定の下で、操作変数法を用いることで局所平均処置効果 (Local Average Treatment Effect, LATE) を識別できることを示します。
- これはコンプライヤー (compliers)、すなわち「抽選に当たったからこそ Medicaid に加入した人々」における平均処置効果です。

局所平均処置効果 (LATE)

- 特定の仮定の下で、操作変数法を用いることで局所平均処置効果 (Local Average Treatment Effect, LATE) を識別できることを示します。
- これはコンプライヤー (compliers)、すなわち「抽選に当たったからこそ Medicaid に加入した人々」における平均処置効果です。
- これが「局所的（local)」な効果であるのは、実験によって Medicaid 加入ステータスが変化しなかった人々についての効果は何も教えてくれないからです。

局所平均処置効果 (LATE)

- 特定の仮定の下で、操作変数法を用いることで局所平均処置効果 (Local Average Treatment Effect, LATE) を識別できることを示します。
- これはコンプライヤー (compliers)、すなわち「抽選に当たったからこそ Medicaid に加入した人々」における平均処置効果です。
- これが「局所的（local)」な効果であるのは、実験によって Medicaid 加入ステータスが変化しなかった人々についての効果は何も教えてくれないからです。
- 次に、操作変数法を用いて LATE を識別するために必要な仮定を確認します。

4 つのタイプの人々

集団を以下の 4 つのタイプに分類できます：

- オールウェイズ・テイカー (Always-takers)：抽選の結果に関わらず、必ず Medicaid に加入する人々。

4 つのタイプの人々

集団を以下の 4 つのタイプに分類できます：

- オールウェイズ・テイカー (Always-takers)：抽選の結果に関わらず、必ず Medicaid に加入する人々。
- ネバー・テイカー (Never-takers)：抽選の結果に関わらず、決して Medicaid に加入しない人々。

4 つのタイプの人々

集団を以下の 4 つのタイプに分類できます：

- オールウェイズ・テイカー (Always-takers)：抽選の結果に関わらず、必ず Medicaid に加入する人々。
- ネバー・テイカー (Never-takers)：抽選の結果に関わらず、決して Medicaid に加入しない人々。
- コンプライヤー (Compliers)：抽選に当選した場合にのみ Medicaid に加入する人々。

4 つのタイプの人々

集団を以下の 4 つのタイプに分類できます：

- オールウェイズ・テイカー (Always-takers)：抽選の結果に関わらず、必ず Medicaid に加入する人々。
- ネバー・テイカー (Never-takers)：抽選の結果に関わらず、決して Medicaid に加入しない人々。
- コンプライヤー (Compliers)：抽選に当選した場合にのみ Medicaid に加入する人々。
- ディファイアー (Defiers)：抽選に落選した場合にのみ Medicaid に加入する人々。

4 つのタイプの人々

集団を以下の 4 つのタイプに分類できます：

- オールウェイズ・テイカー (Always-takers)：抽選の結果に関わらず、必ず Medicaid に加入する人々。
- ネバー・テイカー (Never-takers)：抽選の結果に関わらず、決して Medicaid に加入しない人々。
- コンプライヤー (Compliers)：抽選に当選した場合にのみ Medicaid に加入する人々。
- ディファイアー (Defiers)：抽選に落選した場合にのみ Medicaid に加入する人々。
  - ディファイアーは非現実的な存在と考えられ、通常は存在しないと仮定します（単調性の仮定）。

4 つのタイプの人々 – 数式による表現

- D_i を Medicaid に加入したかどうかの指標（処置）とします。
  Z_i を抽選に当選したかどうかの指標（操作変数）とします。

4 つのタイプの人々 – 数式による表現

- D_i を Medicaid に加入したかどうかの指標（処置）とします。
  Z_i を抽選に当選したかどうかの指標（操作変数）とします。
- 潜在的結果と同様に、潜在的処置 $D_i(1)$ と $D_i(0)$ を導入します：
  - $D_i(1)$：$Z_i = 1$（当選）の場合の加入ステータス。
  - $D_i(0)$：$Z_i = 0$（落選）の場合の加入ステータス。

4 つのタイプの人々 – 数式による表現

- D_i を Medicaid に加入したかどうかの指標（処置）とします。
  Z_i を抽選に当選したかどうかの指標（操作変数）とします。

- 潜在的結果と同様に、潜在的処置 $D_i(1)$ と $D_i(0)$ を導入します：
  - $D_i(1)$：$Z_i = 1$（当選）の場合の加入ステータス。
  - $D_i(0)$：$Z_i = 0$（落選）の場合の加入ステータス。

- オールウェイズ・テイカー：$D_i(1) = D_i(0) = 1$

4 つのタイプの人々 – 数式による表現

- D_i を Medicaid に加入したかどうかの指標（処置）とします。
  Z_i を抽選に当選したかどうかの指標（操作変数）とします。
- 潜在的結果と同様に、潜在的処置 $D_i(1)$ と $D_i(0)$ を導入します：
  - $D_i(1)$：$Z_i = 1$（当選）の場合の加入ステータス。
  - $D_i(0)$：$Z_i = 0$（落選）の場合の加入ステータス。
- オールウェイズ・テイカー：$D_i(1) = D_i(0) = 1$
- ネバー・テイカー：$D_i(1) = D_i(0) = 0$

4 つのタイプの人々 – 数式による表現

- D_i を Medicaid に加入したかどうかの指標（処置）とします。
  Z_i を抽選に当選したかどうかの指標（操作変数）とします。

- 潜在的結果と同様に、潜在的処置 $D_i(1)$ と $D_i(0)$ を導入します：
  - $D_i(1)$：$Z_i = 1$（当選）の場合の加入ステータス。
  - $D_i(0)$：$Z_i = 0$（落選）の場合の加入ステータス。

- オールウェイズ・テイカー：$D_i(1) = D_i(0) = 1$

- ネバー・テイカー：$D_i(1) = D_i(0) = 0$

- コンプライヤー：$D_i(1) = 1$ かつ $D_i(0) = 0$

4 つのタイプの人々 – 数式による表現

- D_i を Medicaid に加入したかどうかの指標（処置）とします。
  Z_i を抽選に当選したかどうかの指標（操作変数）とします。
- 潜在的結果と同様に、潜在的処置 $D_i(1)$ と $D_i(0)$ を導入します：
  - $D_i(1)$：$Z_i = 1$（当選）の場合の加入ステータス。
  - $D_i(0)$：$Z_i = 0$（落選）の場合の加入ステータス。
- オールウェイズ・テイカー：$D_i(1) = D_i(0) = 1$
- ネバー・テイカー：$D_i(1) = D_i(0) = 0$
- コンプライヤー：$D_i(1) = 1$ かつ $D_i(0) = 0$
- ディファイアー：$D_i(1) = 0$ かつ $D_i(0) = 1$

主要な仮定

- 独立性 (Independence)：操作変数（例：抽選結果）は実質的にランダムに割り当てられている。$Z_i \perp\!\!\!\perp (Y_i(0), Y_i(1), D_i(0), D_i(1))$

主要な仮定

- 独立性 (Independence)：操作変数（例：抽選結果）は実質的にランダムに割り当てられている。$Z_i \perp\!\!\!\perp (Y_i(0), Y_i(1), D_i(0), D_i(1))$
- 関連性 (Relevance)：操作変数が処置を受ける確率に影響を与える。$P(D_i = 1|Z_i = 1) \neq P(D_i = 1|Z_i = 0)$
  - これはデータで検証可能です（実際に影響していました)。
- 単調性 (Monotonicity)：ディファイアーが存在しない。すべての i について $D_i(1) \geq D_i(0)$。
  - 落選しても加入するような人が、当選したときに加入しなくなるとは考えにくいので、妥当な仮定です。

最後の重要な仮定 — 除外制約！

- 最後の一つであり、しばしば最も議論の的となるのが除外制約 (Exclusion restriction) です。

最後の重要な仮定 — 除外制約！

- 最後の一つであり、しばしば最も議論の的となるのが除外制約 (Exclusion restriction) です。
- 直感的には、操作変数 Z_i（抽選の当選）は、処置（*Medicaid* 加入）を通じてのみアウトカムに影響を与えるという仮定です。

最後の重要な仮定 — 除外制約！

- 最後の一つであり、しばしば最も議論の的となるのが除外制約 (Exclusion restriction) です。
- 直感的には、操作変数 Z_i（抽選の当選）は、処置（*Medicaid* 加入）を通じてのみアウトカムに影響を与えるという仮定です。
- 数学的には、$Y_i = D_i(Z_i)Y_i(1) + (1 - D_i(Z_i))Y_i(0)$ と書けることを意味します。
  - 潜在的結果を $Y_i(d, z)$ と書くなら、$Y_i(d, z) = Y_i(d)$ であるという仮定です。

最後の重要な仮定 — 除外制約！

- 最後の一つであり、しばしば最も議論の的となるのが除外制約 (Exclusion restriction) です。
- 直感的には、操作変数 Z_i（抽選の当選）は、処置（*Medicaid* 加入）を通じてのみアウトカムに影響を与えるという仮定です。
- 数学的には、$Y_i = D_i(Z_i)Y_i(1) + (1 - D_i(Z_i))Y_i(0)$ と書けることを意味します。
  - 潜在的結果を $Y_i(d, z)$ と書くなら、$Y_i(d, z) = Y_i(d)$ であるという仮定です。
- これは、処置ステータスが変化しないオールウェイズ・テイカーやネバー・テイカーにとっては、当選したかどうかがアウトカムに影響を与えないことを意味します。

なぜ除外制約が崩れるのか？

- 操作変数（当選）が、処置ステータスを変えることなくアウトカムに直接影響を与える場合、除外制約は満たされません。

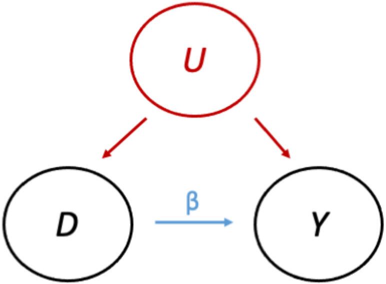
なぜ除外制約が崩れるのか？

- 操作変数（当選）が、処置ステータスを変えることなくアウトカムに直接影響を与える場合、除外制約は満たされません。
- 例えば OHIE において、当選したオールウェイズ・テイカーは当選しなかった場合よりも早く Medicaid に加入できたとします。これは除外制約に違反するでしょうか？

なぜ除外制約が崩れるのか？

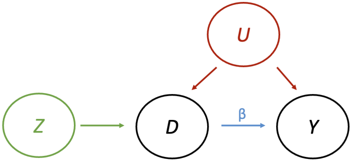
- 操作変数（当選）が、処置ステータスを変えることなくアウトカムに直接影響を与える場合、除外制約は満たされません。
- 例えば OHIE において、当選したオールウェイズ・テイカーは当選しなかった場合よりも早く Medicaid に加入できたとします。これは除外制約に違反するでしょうか？
- はい。Medicaid への加入「時期」が健康に直接影響を与えるのであれば、違反になります。

図解



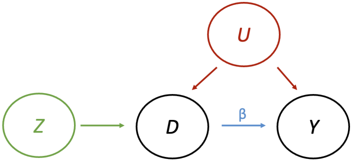
- 非交絡性が満たされないのは、処置 D とアウトカム Y の両方に影響を与える観察されない U が存在する場合です。

図解



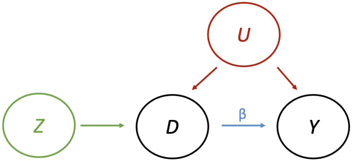
- 非交絡性が満たされないのは、処置 D とアウトカム Y の両方に影響を与える観察されない U が存在する場合です。
- 操作変数 Z があれば、この問題を回避できます。

図解



- 非交絡性が満たされないのは、処置 D とアウトカム Y の両方に影響を与える観察されない U が存在する場合です。
- 操作変数 Z があれば、この問題を回避できます。
- 独立性：Z は未観察因子 U の影響を受けない。

図解



- 非交絡性が満たされないのは、処置 D とアウトカム Y の両方に影響を与える観察されない U が存在する場合です。
- 操作変数 Z があれば、この問題を回避できます。
- 独立性：Z は未観察因子 U の影響を受けない。
- 除外制約：Z は D を通じてのみ Y に影響を与える。

LATE 定理

- 独立性、関連性、除外制約、単調性が成り立つとき：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}=E[Y_i(1)-Y_i(0)|D_i(1)=1,D_i(0)=0]$$

言葉で言えば：

$$\frac{Z\text{ が }Y\text{ に与える効果}}{Z\text{ が }D\text{ に与える効果}}=\text{コンプライヤーにおける平均処置効果}$$

LATE 定理

- 独立性、関連性、除外制約、単調性が成り立つとき：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}=E[Y_i(1)-Y_i(0)|D_i(1)=1,D_i(0)=0]$$

言葉で言えば：

$$\frac{Z \text{ が } Y \text{ に与える効果}}{Z \text{ が } D \text{ に与える効果}}=\text{コンプライヤーにおける平均処置効果}$$

- この成果により、Joshua Angrist と Guido Imbens は 2021 年ノーベル経済学賞を受賞しました！

LATE 定理の直感

- 定理の式：

$$\frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[D_i|Z_i = 1] - E[D_i|Z_i = 0]} = \text{LATE}$$

LATE 定理の直感

- 定理の式：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}=\text{LATE}$$

- 独立性により、分子は Z が Y に与える因果的効果です。

LATE 定理の直感

- 定理の式：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}=\text{LATE}$$

- 独立性により、分子は Z が Y に与える因果的効果です。
- しかし、除外制約によれば、Z の Y への効果は、処置ステータスが変化しない人々（AT, NT）にとっては 0 です。

LATE 定理の直感

- 定理の式：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}=\mathrm{LATE}$$

- 独立性により、分子は Z が Y に与える因果的効果です。
- しかし、除外制約によれば、Z の Y への効果は、処置ステータスが変化しない人々（AT, NT）にとっては 0 です。また単調性によりディファイアーはいません。

LATE 定理の直感

- 定理の式：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]} = \text{LATE}$$

- 独立性により、分子は Z が Y に与える因果的効果です。

- しかし、除外制約によれば、Z の Y への効果は、処置ステータスが変化しない人々（AT, NT）にとっては 0 です。また単調性によりディファイアーはいません。

- したがって、分子は「コンプライヤーにおける平均効果」に「コンプライヤーの割合」を掛けたもの、すなわち $LATE \times P(Complier)$ になります。

LATE 定理の直感

- 定理の式：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}=\text{LATE}$$

- 独立性により、分子は Z が Y に与える因果的効果です。
- しかし、除外制約によれば、Z の Y への効果は、処置ステータスが変化しない人々（AT, NT）にとっては 0 です。また単調性によりディファイアーはいません。
- したがって、分子は「コンプライヤーにおける平均効果」に「コンプライヤーの割合」を掛けたもの、すなわち $LATE \times P(Complier)$ になります。
- 一方、分母は Z が D に与える効果です。これは AT や NT にとっては 0 なので、単にコンプライヤーの割合 $P(Complier)$ に等しくなります。

より形式的な導出

- 集団をオールウェイズ・テイカー (AT)、ネバー・テイカー (NT)、コンプライヤー (C) の 3 つのグループに分けます。

より形式的な導出

- 集団をオールウェイズ・テイカー (AT)、ネバー・テイカー (NT)、コンプライヤー (C) の 3 つのグループに分けます。(ディファイアーはいないと仮定しました)。

- それぞれの割合を $\alpha_{AT} = P(AT), \alpha_{NT} = P(NT), \alpha_C = P(C)$ とします。

より形式的な導出

- 集団をオールウェイズ・テイカー (AT)、ネバー・テイカー (NT)、コンプライヤー (C) の 3 つのグループに分けます。(ディファイアーはいないと仮定しました)。
- それぞれの割合を $\alpha_{AT} = P(AT), \alpha_{NT} = P(NT), \alpha_C = P(C)$ とします。
- Z はランダムに割り当てられている（独立性）ので、
  $P(AT|Z = 1) =$

より形式的な導出

- 集団をオールウェイズ・テイカー (AT)、ネバー・テイカー (NT)、コンプライヤー (C) の 3 つのグループに分けます。(ディファイアーはいないと仮定しました)。

- それぞれの割合を $\alpha_{AT} = P(AT), \alpha_{NT} = P(NT), \alpha_C = P(C)$ とします。

- Z はランダムに割り当てられている（独立性）ので、$P(AT|Z=1) = P(AT) = \alpha_{AT}$ です。

より形式的な導出

- 集団をオールウェイズ・テイカー (AT)、ネバー・テイカー (NT)、コンプライヤー (C) の 3 つのグループに分けます。（ディファイアーはいないと仮定しました）。
- それぞれの割合を $\alpha_{AT} = P(AT), \alpha_{NT} = P(NT), \alpha_C = P(C)$ とします。
- Z はランダムに割り当てられている（独立性）ので、
$P(AT|Z=1) = P(AT) = \alpha_{AT}$ です。
同様に、$P(AT|Z=0) =$

より形式的な導出

- 集団をオールウェイズ・テイカー (AT)、ネバー・テイカー (NT)、コンプライヤー (C) の 3 つのグループに分けます。(ディファイアーはいないと仮定しました)。

- それぞれの割合を $\alpha_{AT} = P(AT), \alpha_{NT} = P(NT), \alpha_C = P(C)$ とします。

- Z はランダムに割り当てられている（独立性）ので、
  $P(AT|Z=1) = P(AT) = \alpha_{AT}$ です。
  同様に、$P(AT|Z=0) = P(AT) = \alpha_{AT}$ です。

より形式的な導出

- 集団をオールウェイズ・テイカー (AT)、ネバー・テイカー (NT)、コンプライヤー (C) の 3 つのグループに分けます。(ディファイアーはいないと仮定しました)。

- それぞれの割合を $\alpha_{AT} = P(AT), \alpha_{NT} = P(NT), \alpha_C = P(C)$ とします。

- Z はランダムに割り当てられている（独立性）ので、
  $P(AT|Z=1) = P(AT) = \alpha_{AT}$ です。
  同様に、$P(AT|Z=0) = P(AT) = \alpha_{AT}$ です。

- 同様の議論で、NT と C の割合も $Z=1$ と $Z=0$ のグループで同じになります。

より形式的な導出

- 次のステップ：分子が $\alpha_C \times LATE$ であることを示します。

より形式的な導出

- 次のステップ：分子が $\alpha_C \times LATE$ であることを示します。

- 反復期待値の法則より：

  $E[Y_i|Z_i = 1] =$

より形式的な導出

- 次のステップ：分子が $\alpha_C \times LATE$ であることを示します。
- 反復期待値の法則より：

$$E[Y_i|Z_i = 1] =$$
$$\alpha_C E[Y_i|Z_i = 1, C] + \alpha_{AT} E[Y_i|Z_i = 1, AT] + \alpha_{NT} E[Y_i|Z_i = 1, NT] =$$

より形式的な導出

- 次のステップ：分子が $\alpha_C \times LATE$ であることを示します。
- 反復期待値の法則より：

$$E[Y_i|Z_i=1]=$$
$$\alpha_C E[Y_i|Z_i=1,C]+\alpha_{AT}E[Y_i|Z_i=1,AT]+\alpha_{NT}E[Y_i|Z_i=1,NT]=$$
$$\alpha_C E[Y_i(1)|Z_i=1,C]+\alpha_{AT}E[Y_i(1)|Z_i=1,AT]+\alpha_{NT}E[Y_i(0)|Z_i=1,NT]=$$

より形式的な導出

- 次のステップ：分子が $\alpha_C \times LATE$ であることを示します。

- 反復期待値の法則より：

$$
\begin{aligned}
&E[Y_i|Z_i=1]= \\
&\alpha_C E[Y_i|Z_i=1,C]+\alpha_{AT}E[Y_i|Z_i=1,AT]+\alpha_{NT}E[Y_i|Z_i=1,NT]= \\
&\alpha_C E[Y_i(1)|Z_i=1,C]+\alpha_{AT}E[Y_i(1)|Z_i=1,AT]+\alpha_{NT}E[Y_i(0)|Z_i=1,NT]= \\
&\alpha_C E[Y_i(1)|C]+\alpha_{AT}E[Y_i(1)|AT]+\alpha_{NT}E[Y_i(0)|NT]
\end{aligned}
$$

最後の行で独立性を用いました。

- 同様に：

$$E[Y_i|Z_i=0]=$$

より形式的な導出

- 次のステップ：分子が $\alpha_C \times LATE$ であることを示します。
- 反復期待値の法則より：

$$E[Y_i|Z_i=1]=$$
$$\alpha_C E[Y_i|Z_i=1,C]+\alpha_{AT}E[Y_i|Z_i=1,AT]+\alpha_{NT}E[Y_i|Z_i=1,NT]=$$
$$\alpha_C E[Y_i(1)|Z_i=1,C]+\alpha_{AT}E[Y_i(1)|Z_i=1,AT]+\alpha_{NT}E[Y_i(0)|Z_i=1,NT]=$$
$$\alpha_C E[Y_i(1)|C]+\alpha_{AT}E[Y_i(1)|AT]+\alpha_{NT}E[Y_i(0)|NT]$$

  最後の行で独立性を用いました。

- 同様に：

$$E[Y_i|Z_i=0]=\alpha_C E[Y_i(0)|C]+\alpha_{AT}E[Y_i(1)|AT]+\alpha_{NT}E[Y_i(0)|NT]$$

- したがって、比の分子は以下のようになります：

$$E[Y_i|Z_i=1]-E[Y_i|Z_i=0]=\alpha_C\underbrace{(E[Y_i(1)|C]-E[Y_i(0)|C])}_{LATE}$$

より形式的な導出

- 分子が $\alpha_C \times LATE$ であることを示しました。
- 最後に、分母が α_C であることを示します。

より形式的な導出

- 分子が $\alpha_C \times LATE$ であることを示しました。
- 最後に、分母が α_C であることを示します。
- 分母は以下の通りです：

$$E[D_i|Z_i = 1] - E[D_i|Z_i = 0] =$$

より形式的な導出

- 分子が $\alpha_C \times LATE$ であることを示しました。
- 最後に、分母が α_C であることを示します。
- 分母は以下の通りです：

$$E[D_i|Z_i = 1] - E[D_i|Z_i = 0] = Pr(AT \text{ または } C|Z_i = 1) - P(AT|Z_i = 0)$$

より形式的な導出

- 分子が $\alpha_C \times LATE$ であることを示しました。
- 最後に、分母が α_C であることを示します。
- 分母は以下の通りです：

$$
\begin{aligned}
E[D_i|Z_i=1]-E[D_i|Z_i=0] &= Pr(AT \text{ または } C|Z_i=1)-P(AT|Z_i=0) \\
&= (\alpha_C+\alpha_{AT})-\alpha_{AT} =
\end{aligned}
$$

より形式的な導出

- 分子が $\alpha_C \times LATE$ であることを示しました。
- 最後に、分母が α_C であることを示します。
- 分母は以下の通りです：

$$\begin{aligned} E[D_i|Z_i=1]-E[D_i|Z_i=0] &= Pr(AT \text{ または } C|Z_i=1)-P(AT|Z_i=0) \\ &= (\alpha_C+\alpha_{AT})-\alpha_{AT}=\alpha_C \end{aligned}$$

より形式的な導出

- 分子が $\alpha_C \times LATE$ であることを示しました。
- 最後に、**分母**が α_C であることを示します。
- 分母は以下の通りです：

$$\begin{aligned} E[D_i|Z_i=1]-E[D_i|Z_i=0] &= Pr(AT \text{ または } C|Z_i=1)-P(AT|Z_i=0) \\ &= (\alpha_C+\alpha_{AT})-\alpha_{AT}=\alpha_C \end{aligned}$$

ゆえに：

$$\frac{\text{分子}}{\text{分母}}=\frac{\alpha_C \times LATE}{\alpha_C}=LATE$$

LATE の推定

- IV の仮定の下で、LATE が母平均の関数として識別されることを示しました：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}=\text{LATE}$$

- これをどのように推定すればよいでしょうか？

LATE の推定

- IV の仮定の下で、LATE が母平均の関数として識別されることを示しました：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]} = \text{LATE}$$

- これをどのように推定すればよいでしょうか？

- 標本平均を代入すればよいのです！

$$\hat{\beta} = \frac{\bar{Y}_{Z=1}-\bar{Y}_{Z=0}}{\bar{D}_{Z=1}-\bar{D}_{Z=0}},$$

ここで、例えば $\bar{Y}_{Z=1}$ は $Z_i=1$ であるユニットの Y_i の標本平均です。

LATE の推定

- IV の仮定の下で、LATE が母平均の関数として識別されることを示しました：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]} = \text{LATE}$$

- これをどのように推定すればよいでしょうか？

- 標本平均を代入すればよいのです！

$$\hat{\beta} = \frac{\bar{Y}_{Z=1}-\bar{Y}_{Z=0}}{\bar{D}_{Z=1}-\bar{D}_{Z=0}},$$

  ここで、例えば $\bar{Y}_{Z=1}$ は $Z_i=1$ であるユニットの Y_i の標本平均です。

- $\hat{\beta}$ は IV 推定量、より正確には二段階最小二乗法 (Two-stage Least Squares, 2SLS) 推定量と呼ばれます（その理由はすぐに明らかになります)。

例：Medicaid

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- Medicaid の例で 2SLS 推定量を計算してみましょう。

例：Medicaid

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- Medicaid の例で 2SLS 推定量を計算してみましょう。

$$\frac{\bar{Y}_{Z=1} - \bar{Y}_{Z=0}}{\bar{D}_{Z=1} - \bar{D}_{Z=0}} =$$

例：Medicaid

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- Medicaid の例で 2SLS 推定量を計算してみましょう。

$$\frac{\bar{Y}_{Z=1} - \bar{Y}_{Z=0}}{\bar{D}_{Z=1} - \bar{D}_{Z=0}} = \frac{0.306 - 0.329}{0.397 - 0.141} =$$

例：Medicaid

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- Medicaid の例で 2SLS 推定量を計算してみましょう。

$$\frac{\bar{Y}_{Z=1} - \bar{Y}_{Z=0}}{\bar{D}_{Z=1} - \bar{D}_{Z=0}} = \frac{0.306 - 0.329}{0.397 - 0.141} = -0.09$$

二段階最小二乗法

- 以下の 2 つの線形回帰を考えます：

$$Y_i = \gamma_0 + Z_i\gamma_1 + \varepsilon_i \tag{1}$$

$$D_i = \pi_0 + Z_i\pi_1 + u_i \tag{2}$$

- OLS 推定値 $\hat{\gamma}_1$ と $\hat{\pi}_1$ は何でしょうか？

二段階最小二乗法

- 以下の 2 つの線形回帰を考えます：

$$Y_i = \gamma_0 + Z_i \gamma_1 + \varepsilon_i \tag{1}$$

$$D_i = \pi_0 + Z_i \pi_1 + u_i \tag{2}$$

- OLS 推定値 $\hat{\gamma}_1$ と $\hat{\pi}_1$ は何でしょうか？

$$\hat{\gamma}_1 = \bar{Y}_{Z=1} - \bar{Y}_{Z=0}$$

二段階最小二乗法

- 以下の 2 つの線形回帰を考えます：

$$Y_i = \gamma_0 + Z_i\gamma_1 + \varepsilon_i \quad (1)$$
$$D_i = \pi_0 + Z_i\pi_1 + u_i \quad (2)$$

- OLS 推定値 $\hat{\gamma}_1$ と $\hat{\pi}_1$ は何でしょうか？

$$\hat{\gamma}_1 = \bar{Y}_{Z=1} - \bar{Y}_{Z=0}$$
$$\hat{\pi}_1 = \bar{D}_{Z=1} - \bar{D}_{Z=0}$$

二段階最小二乗法

- 以下の 2 つの線形回帰を考えます：

$$Y_i = \gamma_0 + Z_i \gamma_1 + \varepsilon_i \tag{1}$$
$$D_i = \pi_0 + Z_i \pi_1 + u_i \tag{2}$$

- OLS 推定値 $\hat{\gamma}_1$ と $\hat{\pi}_1$ は何でしょうか？

$$\hat{\gamma}_1 = \bar{Y}_{Z=1} - \bar{Y}_{Z=0}$$
$$\hat{\pi}_1 = \bar{D}_{Z=1} - \bar{D}_{Z=0}$$

- したがって、2SLS 推定量はこれら 2 つの OLS 係数の比になります：

$$\hat{\beta} = \frac{\bar{Y}_{Z=1} - \bar{Y}_{Z=0}}{\bar{D}_{Z=1} - \bar{D}_{Z=0}} = \frac{\hat{\gamma}_1}{\hat{\pi}_1}$$

二段階最小二乗法

- 以下の 2 つの線形回帰を考えます：

$$
\begin{aligned}
Y_i &= \gamma_0 + Z_i \gamma_1 + \varepsilon_i \quad (1) \\
D_i &= \pi_0 + Z_i \pi_1 + u_i \quad (2)
\end{aligned}
$$

- OLS 推定値 $\hat{\gamma}_1$ と $\hat{\pi}_1$ は何でしょうか？

$$
\begin{aligned}
\hat{\gamma}_1 &= \bar{Y}_{Z=1} - \bar{Y}_{Z=0} \\
\hat{\pi}_1 &= \bar{D}_{Z=1} - \bar{D}_{Z=0}
\end{aligned}
$$

- したがって、2SLS 推定量はこれら 2 つの OLS 係数の比になります：

$$
\hat{\beta} = \frac{\bar{Y}_{Z=1} - \bar{Y}_{Z=0}}{\bar{D}_{Z=1} - \bar{D}_{Z=0}} = \frac{\hat{\gamma}_1}{\hat{\pi}_1}
$$

- 式 (2) は通常「第一段階（first-stage）」、式 (1) は「誘導型（reduced-form）」と呼ばれます。

例：Medicaid

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 以下の「第一段階」回帰の係数は？

$$D_i = \pi_0 + Z_i \pi_1 + u_i$$

例：Medicaid

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 以下の「第一段階」回帰の係数は？

$$D_i = \pi_0 + Z_i \pi_1 + u_i$$

- $\hat{\pi}_1 = 0.397 - 0.141 = 0.256$

例：Medicaid

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 以下の「第一段階」回帰の係数は？

$$D_i = \pi_0 + Z_i \pi_1 + u_i$$

- $\hat{\pi}_1 = 0.397 - 0.141 = 0.256$

- 以下の「誘導型」回帰の係数は？

$$Y_i = \gamma_0 + Z_i \gamma_1 + \varepsilon_i$$

例：Medicaid

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 以下の「第一段階」回帰の係数は？

$$D_i = \pi_0 + Z_i \pi_1 + u_i$$

- $\hat{\pi}_1 = 0.397 - 0.141 = 0.256$
- 以下の「誘導型」回帰の係数は？

$$Y_i = \gamma_0 + Z_i \gamma_1 + \varepsilon_i$$

- $\hat{\gamma}_1 = 0.306 - 0.329 = -0.023$

例：Medicaid

	当選者	落選者
Medicaid 加入経験あり	0.397	0.141
うつ病の症状	0.306	0.329

- 以下の「第一段階」回帰の係数は？

$$D_i = \pi_0 + Z_i \pi_1 + u_i$$

- $\hat{\pi}_1 = 0.397 - 0.141 = 0.256$
- 以下の「誘導型」回帰の係数は？

$$Y_i = \gamma_0 + Z_i \gamma_1 + \varepsilon_i$$

- $\hat{\gamma}_1 = 0.306 - 0.329 = -0.023$
- よって 2SLS 推定量は $\hat{\gamma}_1 / \hat{\pi}_1 = -0.023/0.256 = -0.09$ となります。

コントロール変数がある場合の IV

- しばしば、操作変数が実質的にランダムに割り当てられているという仮定は、観察可能な特徴を条件として初めて妥当になります。
- 実際、OHIE では抽選に当たる確率は家族の人数に依存していました。

コントロール変数がある場合の IV

- しばしば、操作変数が実質的にランダムに割り当てられているという仮定は、観察可能な特徴を条件として初めて妥当になります。
- 実際、OHIE では抽選に当たる確率は家族の人数に依存していました。
- 独立性の仮定を、条件付き独立性 $Z_i \perp\!\!\!\perp (Y_i(1), Y_i(0), D_i(1), D_i(0)) | X_i$ に置き換えたとします。
- 前述と同様の議論により、同じ x の値を持つグループ内で同様の比をとることで、「条件付き **LATE**」を識別できます：

$$\frac{E[Y_i|Z_i = 1, X_i = x] - E[Y_i|Z_i = 0, X_i = x]}{E[D_i|Z_i = 1, X_i = x] - E[D_i|Z_i = 0, X_i = x]} = \underbrace{E[Y_i(1) - Y_i(0)|D_i(1) = 1, D_i(0) = 0, X_i = x]}_{\text{LATE(x)}}$$

条件付き LATE の推定

- 実務において操作変数の条件付き独立性を仮定する場合、コントロール変数を含めた 2SLS の修正版を推定します。
- すなわち、$\hat{\beta}_{2SLS} = \hat{\gamma}_1/\hat{\pi}_1$ を用います。ここで $\hat{\gamma}_1, \hat{\pi}_1$ は以下の OLS 推定値です：

$$Y_i = \gamma_0 + Z_i\gamma_1 + \boldsymbol{X}_i'\boldsymbol{\gamma}_2 + \varepsilon_i \tag{3}$$

$$D_i = \pi_0 + Z_i\pi_1 + \boldsymbol{X}_i'\pi_2 + u_i \tag{4}$$

条件付き LATE の推定

- 実務において操作変数の条件付き独立性を仮定する場合、コントロール変数を含めた 2SLS の修正版を推定します。
- すなわち、$\hat{\beta}_{2SLS} = \hat{\gamma}_1 / \hat{\pi}_1$ を用います。ここで $\hat{\gamma}_1, \hat{\pi}_1$ は以下の OLS 推定値です：

$$Y_i = \gamma_0 + Z_i \gamma_1 + \boldsymbol{X}_i' \boldsymbol{\gamma}_2 + \varepsilon_i \tag{3}$$

$$D_i = \pi_0 + Z_i \pi_1 + \boldsymbol{X}_i' \pi_2 + u_i \tag{4}$$

- $\boldsymbol{X}_i$ がカテゴリー変数（例：家族の人数）のダミー変数のセットである場合、これは各共変量の値における 2SLS 推定値の加重平均を与えます。

条件付き LATE の推定

- 実務において操作変数の条件付き独立性を仮定する場合、コントロール変数を含めた 2SLS の修正版を推定します。

- すなわち、$\hat{\beta}_{2SLS} = \hat{\gamma}_1 / \hat{\pi}_1$ を用います。ここで $\hat{\gamma}_1, \hat{\pi}_1$ は以下の OLS 推定値です：

$$Y_i = \gamma_0 + Z_i \gamma_1 + \boldsymbol{X}_i' \boldsymbol{\gamma}_2 + \varepsilon_i \tag{3}$$

$$D_i = \pi_0 + Z_i \pi_1 + \boldsymbol{X}_i' \pi_2 + u_i \tag{4}$$

- $\boldsymbol{X}_i$ がカテゴリー変数（例：家族の人数）のダミー変数のセットである場合、これは各共変量の値における 2SLS 推定値の加重平均を与えます。

- より一般的には、(4) が第一段階の CEF の良い近似であれば、これは $LATE(x)$ の加重平均の近似的な一致推定量となります。

例：Medicaid

- OHIE では、抽選は家族の人数を条件としてのみランダムでした。

例：Medicaid

- OHIE では、抽選は家族の人数を条件としてのみランダムでした。
- そのため Finkelstein et al. (2012) は、以下の 2SLS を推定しています：

$$Y_i = \gamma_0 + Z_i\gamma_1 + \boldsymbol{X}_i'\boldsymbol{\gamma}_2 + \varepsilon_i \tag{5}$$

$$D_i = \pi_0 + Z_i\pi_1 + \boldsymbol{X}_i'\pi_2 + u_i \tag{6}$$

ここで $\boldsymbol{X}_i$ には家族の人数の固定効果やその他の人口統計学的変数が含まれています。

TABLE IV
Hospital Utilization

	Control mean (1)	ITT (2)	LATE (3)	p-values (4)
Panel A: Extensive margin				
All hospital admissions	0.067	0.0054	0.021	[0.004]
	(0.250)	(0.0019)	(0.0074)	
Admissions through ER	0.048	0.0018	0.0070	[0.265]
	(0.214)	(0.0016)	(0.0062)	
Admissions not through ER	0.029	0.0041	0.016	[0.002]
	(0.167)	(0.0013)	(0.0051)	

TABLE IX
HEALTH

	Control mean (1)	ITT (2)	LATE (3)	*p*-values (4)
Panel A: Administrative data				
Alive	0.992	0.00032	0.0013	[0.638]
	(0.092)	(0.00068)	(0.0027)	
Panel B: Survey data				
Self-reported health good/very good/excellent (not fair or poor)	0.548	0.039	0.133	[<0.0001]
	(0.498)	(0.0076)	(0.026)	{<0.0001}
Self-reported health not poor (fair, good, very good, or excellent)	0.86	0.029	0.099	[<0.0001]
	(0.347)	(0.0051)	(0.018)	{<0.0001}
Health about the same or gotten better over last six months	0.714	0.033	0.113	[<0.0001]
	(0.452)	(0.0067)	(0.023)	{<0.0001}
# of days physical health good, past 30 days*	21.862	0.381	1.317	[0.019]
	(10.384)	(0.162)	(0.563)	{0.018}
# days poor physical or mental health did not impair usual activity, past 30 days*	20.329	0.459	1.585	[0.009]
	(10.939)	(0.175)	(0.606)	{0.015}
# of days mental health good, past 30 days*	18.738	0.603	2.082	[0.001]
	(11.445)	(0.184)	(0.64)	{0.003}
Did not screen positive for depression, last two weeks	0.671	0.023	0.078	[0.001]
	(0.470)	(0.0071)	(0.025)	{0.003}
Standardized treatment effect		0.059	0.203	[<0.0001]
		(0.011)	(0.039)	

Notes. Standard errors in parentheses; per comparison *p*-values in square brackets; family-wise *p*-values in curly brackets. Column (2) reports the coefficient and standard error on *LOTTERY* from estimating equation (1) by OLS. Column (3) reports the coefficient and standard error on *INSURANCE* from estimating equation (3) by IV; for the IV estimates in column (3), the endogenous variable *INSURANCE* is defined as "ever on Medicaid" during our study period and the first stage is given in the first row of Table III. Column (4) reports the per comparison *p*-value and the family-wise *p*-value across the different measures used to create the standardized treatment effect. Standardized treatment effect reports results based on equation (2). All regressions include household size fixed effects and standard errors are clustered on the household. The regressions in Panel A include lottery draw fixed effects, and the dependent variable "alive" is measured from the notification date through September 2009 ($N = 74{,}922$). The regressions in Panel B include survey wave fixed effects and the interaction of survey wave fixed effects with household size fixed effects, and are weighted using the survey weights ($N = 23{,}741$).

*These questions were worded to ask about number of days health "not good" or "impaired"; we switched the sign for consistency with the other measures. See Online Appendix Figure A4 for the exact survey wording.

「擬似実験」的な操作変数

- Medicaid の事例では、操作変数 Z_i は明示的にランダムに割り当てられていました。

「擬似実験」的な操作変数

- Medicaid の事例では、操作変数 Z_i は明示的にランダムに割り当てられていました。
- 他の設定では、直接ランダム化されたわけではないが、潜在的に「ランダムと同様」と見なせるような特異な要因による操作変数を用いることがあります。
- これにより IV を利用できるケースが広がりますが、必要な仮定についてはより厳格な吟味が必要になります！

例：Angrist and Krueger (1991)

- AK (1991) は労働経済学の古典的な問いに取り組んでいます：教育を 1 年追加することの労働市場における収益（賃金への影響）は何か？

例：Angrist and Krueger (1991)

- AK (1991) は労働経済学の古典的な問いに取り組んでいます：教育を 1 年追加することの労働市場における収益（賃金への影響）は何か？
- 単に教育年数が長い人と短い人の賃金を比較するだけではいけないのはなぜでしょうか？

例：Angrist and Krueger (1991)

- AK (1991) は労働経済学の古典的な問いに取り組んでいます：教育を 1 年追加することの労働市場における収益（賃金への影響）は何か？
- 単に教育年数が長い人と短い人の賃金を比較するだけではいけないのはなぜでしょうか？
  - 教育年数はランダムに決まるわけではないからです。教育の選択は、能力や家庭環境などの、賃金に直接影響を与える多くの交絡因子に依存している可能性があります。

例：Angrist and Krueger (1991)

- AK (1991) は労働経済学の古典的な問いに取り組んでいます：教育を 1 年追加することの労働市場における収益（賃金への影響）は何か？

- 単に教育年数が長い人と短い人の賃金を比較するだけではいけないのはなぜでしょうか？
  - 教育年数はランダムに決まるわけではないからです。教育の選択は、能力や家庭環境などの、賃金に直接影響を与える多くの交絡因子に依存している可能性があります。

- IV を用いて教育の収益を推定するには何が必要でしょうか？

例：Angrist and Krueger (1991)

- AK (1991) は労働経済学の古典的な問いに取り組んでいます：教育を 1 年追加することの労働市場における収益（賃金への影響）は何か？
- 単に教育年数が長い人と短い人の賃金を比較するだけではいけないのはなぜでしょうか？
  - 教育年数はランダムに決まるわけではないからです。教育の選択は、能力や家庭環境などの、賃金に直接影響を与える多くの交絡因子に依存している可能性があります。
- IV を用いて教育の収益を推定するには何が必要でしょうか？
- 実質的にランダムに割り当てられ、教育年数を通じてのみ賃金に影響を与えるような操作変数を見つける必要があります。

義務教育法

- 米国のほとんどの州には、16 歳または 17 歳の誕生日まで学校に留まることを義務付ける義務教育法があります。

義務教育法

- 米国のほとんどの州には、16 歳または 17 歳の誕生日まで学校に留まることを義務付ける義務教育法があります。
- AK は、これらの法律により、中退を考えている生徒は、1 年のうちの早い時期に生まれた場合、より少ない教育しか受けられなくなると主張しました。
- つまり、同じ学年で最も年上の子（1 月 1 日生まれ）と最も若い子（12 月 31 日生まれ）を比べると、年上の子は最も若い子よりも 1 年早く合法的に中退できてしまいます。

義務教育法

- 米国のほとんどの州には、16 歳または 17 歳の誕生日まで学校に留まることを義務付ける義務教育法があります。
- AK は、これらの法律により、中退を考えている生徒は、1 年のうちの早い時期に生まれた場合、より少ない教育しか受けられなくなると主張しました。
- つまり、同じ学年で最も年上の子（1 月 1 日生まれ）と最も若い子（12 月 31 日生まれ）を比べると、年上の子は最も若い子よりも 1 年早く合法的に中退できてしまいます。
- AK は、1 年のうちのどの時期に生まれるかは実質的にランダムであると議論し、教育年数の操作変数として「出生四半期」を用いました。

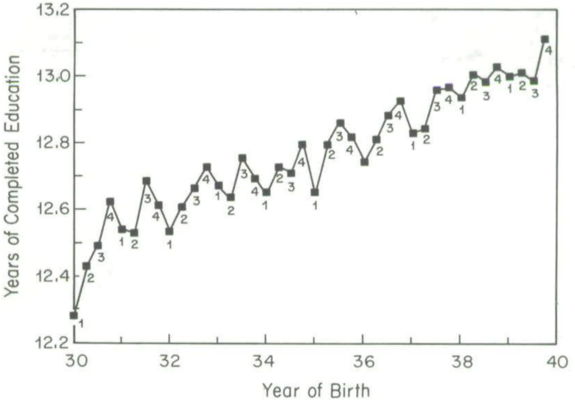


FIGURE I
Years of Education and Season of Birth
1980 Census
Note. Quarter of birth is listed below each observation.

- 実際に AK は、第 1 四半期に生まれた人は、同じ年のそれ以外の四半期に生まれた人よりも平均して教育年数が短いことを見出しました。

PANEL A: WALD ESTIMATES FOR 1970 CENSUS—MEN BORN 192

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year
ln (wkly. wage)	5.1484	5.1574
Education	11.3996	11.5252

PANEL A: WALD ESTIMATES FOR 1970 CENSUS—MEN BORN 192

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year
ln (wkly. wage)	5.1484	5.1574
Education	11.3996	11.5252

- 第一段階の推定値は？

PANEL A: WALD ESTIMATES FOR 1970 CENSUS—MEN BORN 192

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year
ln (wkly. wage)	5.1484	5.1574
Education	11.3996	11.5252

- 第一段階の推定値は？ $11.5252 - 11.3996 = 0.1256$

Panel A: Wald Estimates for 1970 Census—Men Born 192

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year
ln (wkly. wage)	5.1484	5.1574
Education	11.3996	11.5252

- 第一段階の推定値は？ $11.5252 - 11.3996 = 0.1256$
- 誘導型の推定値は？

Panel A: Wald Estimates for 1970 Census—Men Born 192

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year
ln (wkly. wage)	5.1484	5.1574
Education	11.3996	11.5252

- 第一段階の推定値は？ $11.5252 - 11.3996 = 0.1256$
- 誘導型の推定値は？ $5.1574 - 5.1484 = 0.0090$

Panel A: Wald Estimates for 1970 Census—Men Born 192

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year
ln (wkly. wage)	5.1484	5.1574
Education	11.3996	11.5252

- 第一段階の推定値は？ $11.5252 - 11.3996 = 0.1256$
- 誘導型の推定値は？ $5.1574 - 5.1484 = 0.0090$
- 2SLS 推定値は？

PANEL A: WALD ESTIMATES FOR 1970 CENSUS—MEN BORN 192

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year
ln (wkly. wage)	5.1484	5.1574
Education	11.3996	11.5252

- 第一段階の推定値は？ $11.5252 - 11.3996 = 0.1256$
- 誘導型の推定値は？ $5.1574 - 5.1484 = 0.0090$
- 2SLS 推定値は？ $0.0090/0.1256 = 0.0715$

TABLE III

PANEL A: WALD ESTIMATES FOR 1970 CENSUS—MEN BORN 1920–1929[a]

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year	(3) Difference (std. error) (1) − (2)
ln (wkly. wage)	5.1484	5.1574	−0.00898
			(0.00301)
Education	11.3996	11.5252	−0.1256
			(0.0155)
Wald est. of return to education			0.0715
			(0.0219)
OLS return to education[b]			0.0801
			(0.0004)

仮定の評価

- 関連性：

仮定の評価

- 関連性：出生四半期が教育年数と相関している必要がある。

仮定の評価

- 関連性：出生四半期が教育年数と相関している必要がある。
  ✓ これは検証可能であり、実際に相関していました（t 統計量は 12）。

仮定の評価

- 関連性：出生四半期が教育年数と相関している必要がある。
  ✓ これは検証可能であり、実際に相関していました（t 統計量は 12）。
- 独立性：

仮定の評価

- 関連性：出生四半期が教育年数と相関している必要がある。
  ✓ これは検証可能であり、実際に相関していました（t 統計量は 12)。
- 独立性：出生四半期が、賃金の決定要因や義務教育法への反応と独立である必要がある（ $Z_i \perp\!\!\!\perp (Y(.), D(.))$ ）。

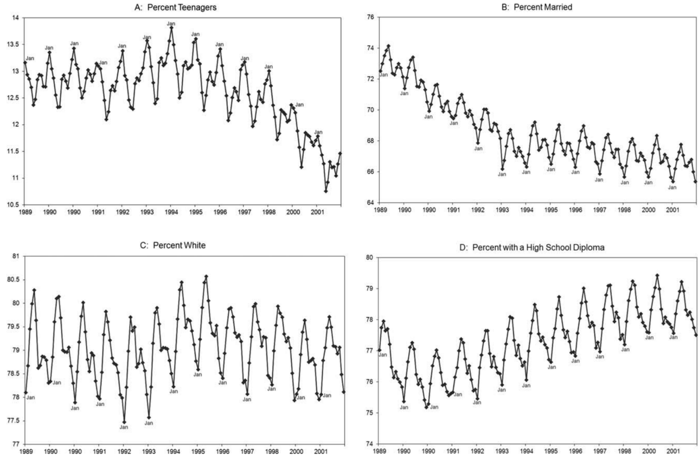
仮定の評価

- 関連性：出生四半期が教育年数と相関している必要がある。
  ✓ これは検証可能であり、実際に相関していました（t 統計量は 12）。
- 独立性：出生四半期が、賃金の決定要因や義務教育法への反応と独立である必要がある（ $Z_i \perp\!\!\!\perp (Y(.), D(.))$ ）。
- 親が子供の誕生時期を正確にコントロールできないのであれば、独立性は妥当に見えます。

仮定の評価

- 関連性：出生四半期が教育年数と相関している必要がある。
  ✓ これは検証可能であり、実際に相関していました（t 統計量は 12）。
- 独立性：出生四半期が、賃金の決定要因や義務教育法への反応と独立である必要がある（ $Z_i \perp\!\!\!\perp (Y(.), D(.))$ ）。
- 親が子供の誕生時期を正確にコントロールできないのであれば、独立性は妥当に見えます。しかし...

FIGURE 1.—MATERNAL CHARACTERISTICS BY MONTH, NATALITY FILES, 1989–2001



- 出生四半期は、いくつかの人口統計学的特徴と相関しています。
- これは独立性の仮定が満たされていない可能性を示唆しています。

仮定の評価...

- 除外制約：出生四半期は教育年数を通じてのみ賃金に影響を与える。

仮定の評価...

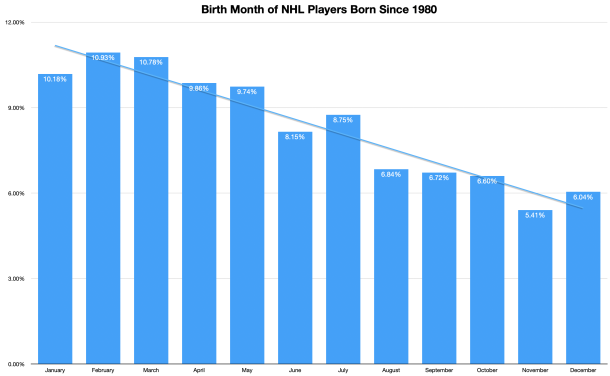
- 除外制約：出生四半期は教育年数を通じてのみ賃金に影響を与える。
- これは、出生四半期によって教育年数が変化しない人々にとっては、いつ生まれても賃金が同じであることを意味します。

仮定の評価...

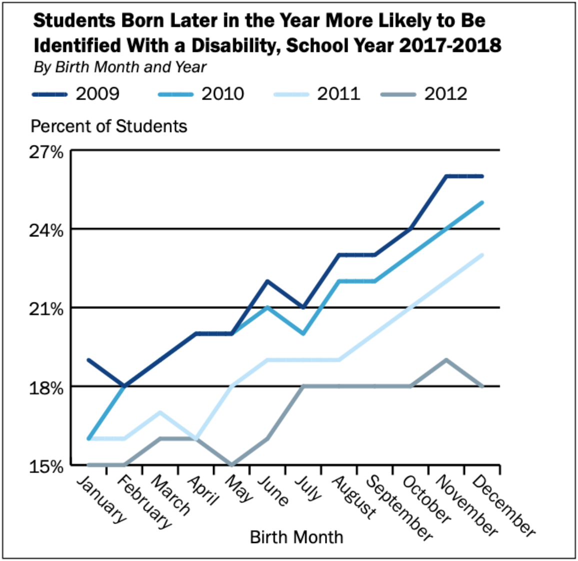
- 除外制約：出生四半期は教育年数を通じてのみ賃金に影響を与える。
- これは、出生四半期によって教育年数が変化しない人々にとっては、いつ生まれても賃金が同じであることを意味します。
- 一般的は妥当そうですが...

仮定の評価...

- 除外制約：出生四半期は教育年数を通じてのみ賃金に影響を与える。
- これは、出生四半期によって教育年数が変化しない人々にとっては、いつ生まれても賃金が同じであることを意味します。
- 一般的は妥当そうですが... 学年の中で相対的に年上か年下かということが、教育の質に直接影響を与えるかもしれません。



Birth Month of NHL Players Born Since 1980
12.00%
9.00%
6.00%
3.00%
0.00%
10.18%
10.93%
10.78%
9.86%
9.74%
8.15%
8.75%
6.84%
6.72%
6.60%
5.41%
6.04%
January
February
March
April
May
June
July
August
September
October
November
December


Students Born Later in the Year More Likely to Be Identified With a Disability, School Year 2017-2018
By Birth Month and Year
2009
2010
2011
2012
Percent of Students
27%
24%
21%
18%
15%
January
February
March
April
May
June
July
August
September
October
November
December
Birth Month

仮定の評価

- 単調性：第 1 四半期に生まれなかった方が、全員が少なくとも同等以上の教育年数を得る。

仮定の評価

- 単調性：第 1 四半期に生まれなかった方が、全員が少なくとも同等以上の教育年数を得る。
- すべての学校が 1 月 1 日を学年の区切りにしていれば妥当です。

仮定の評価

- 単調性：第 1 四半期に生まれなかった方が、全員が少なくとも同等以上の教育年数を得る。
- すべての学校が 1 月 1 日を学年の区切りにしていれば妥当です。
- もし 9 月 1 日を区切りにしている学校があれば、第 4 四半期生まれの方が教育年数が短くなる可能性もあります。

TABLE II

PERCENTAGE OF AGE GROUP ENROLLED IN SCHOOL BY BIRTHDAY AND LEGAL DROPOUT AGE[a]

Date of birth	Type of state law[b]: School-leaving age: 16 (1)	School-leaving age: 17 or 18 (2)	Column (1) − (2)
	Percent enrolled April 1, 1960		
1. Jan 1–Mar 31, 1944	87.6	91.0	−3.4
(age 16)	(0.6)	(0.9)	(1.1)
2. Apr 1–Dec 31, 1944	92.1	91.6	0.5
(age 15)	(0.3)	(0.5)	(0.6)
3. Within-state diff.	−4.5	−0.6	−4.0
(row 1 − row 2)	(0.7)	(1.0)	(1.2)

TABLE II

PERCENTAGE OF AGE GROUP ENROLLED IN SCHOOL BY BIRTHDAY AND LEGAL DROPOUT AGE[a]

Date of birth	Type of state law[b]: School-leaving age: 16 (1)	School-leaving age: 17 or 18 (2)	Column (1) − (2)
	Percent enrolled April 1, 1960		
1. Jan 1–Mar 31, 1944	87.6	91.0	−3.4
(age 16)	(0.6)	(0.9)	(1.1)
2. Apr 1–Dec 31, 1944	92.1	91.6	0.5
(age 15)	(0.3)	(0.5)	(0.6)
3. Within-state diff.	−4.5	−0.6	−4.0
(row 1 − row 2)	(0.7)	(1.0)	(1.2)

- 義務教育法（CSL）の対象年齢が異なる州を比較したところ、出生四半期の教育への影響は CSL を通じて生じていることが示唆されました。これは独立性に関する懸念をいくらか和らげるものです。

誰に対する処置効果か？

- IV の仮定が（近似的に）成り立つとしましょう。
- 推定された効果をどう解釈すべきでしょうか？

誰に対する処置効果か？

- IV の仮定が（近似的に）成り立つとしましょう。
- 推定された効果をどう解釈すべきでしょうか？
- それはコンプライヤーに対する局所平均処置効果 (LATE) です。この場合は、「1 年の後半に生まれていれば、もう 1 年長く学校に通ったであろう人々」に対する効果です。

誰に対する処置効果か？

- IV の仮定が（近似的に）成り立つとしましょう。
- 推定された効果をどう解釈すべきでしょうか？
- それはコンプライヤーに対する局所平均処置効果 (LATE) です。この場合は、「1 年の後半に生まれていれば、もう 1 年長く学校に通ったであろう人々」に対する効果です。
- なぜ LATE が全集団の ATE と一致しない可能性があるのでしょうか？

誰に対する処置効果か？

- IV の仮定が（近似的に）成り立つとしましょう。
- 推定された効果をどう解釈すべきでしょうか？
- それはコンプライヤーに対する局所平均処置効果 (LATE) です。この場合は、「1 年の後半に生まれていれば、もう 1 年長く学校に通ったであろう人々」に対する効果です。
- なぜ LATE が全集団の ATE と一致しない可能性があるのでしょうか？
  - 中退するかどうかの境界線上にいる人々にとっての教育の収益は、他の人々とは異なる可能性があるからです。

複数の操作変数

- これまでは、2 値の操作変数（当選 vs 落選）を考えてきました。
- しかし、操作変数が多値をとる場合や、複数の異なる操作変数がある場合もあります。
- 次に、多値の操作変数からの変動をどのように利用できるかを議論します。

例：Angrist (1990)

- Angrist (1990) は以下の問いに興味を持っています：軍隊に従事すること（特にベトナム戦争）は、その後の人生の労働賃金にどのような影響を与えるか？

例：Angrist (1990)

- Angrist (1990) は以下の問いに興味を持っています：軍隊に従事すること（特にベトナム戦争）は、その後の人生の労働賃金にどのような影響を与えるか？
- なぜ単に退役軍人と非退役軍人を比較するだけではいけないのでしょうか？

例：Angrist (1990)

- Angrist (1990) は以下の問いに興味を持っています：軍隊に従事すること（特にベトナム戦争）は、その後の人生の労働賃金にどのような影響を与えるか？
- なぜ単に退役軍人と非退役軍人を比較するだけではいけないのでしょうか？
  - 退役軍人ステータスはランダムではないからです。彼らはキャリアの目標、家庭環境、身体能力などの面で異なっている可能性があります。

例：Angrist (1990)

- Angrist (1990) は以下の問いに興味を持っています：軍隊に従事すること（特にベトナム戦争）は、その後の人生の労働賃金にどのような影響を与えるか？
- なぜ単に退役軍人と非退役軍人を比較するだけではいけないのでしょうか？
  - 退役軍人ステータスはランダムではないからです。彼らはキャリアの目標、家庭環境、身体能力などの面で異なっている可能性があります。
- IV を使うには何が必要でしょうか？

例：Angrist (1990)

- Angrist (1990) は以下の問いに興味を持っています：軍隊に従事すること（特にベトナム戦争）は、その後の人生の労働賃金にどのような影響を与えるか？
- なぜ単に退役軍人と非退役軍人を比較するだけではいけないのでしょうか？
  - 退役軍人ステータスはランダムではないからです。彼らはキャリアの目標、家庭環境、身体能力などの面で異なっている可能性があります。
- IV を使うには何が必要でしょうか？ 実質的にランダムで、軍への入隊を通じてのみ賃金に影響を与える操作変数です。

Angrist (1990) – 徴兵抽選

- Angrist (1990) は、ベトナム戦争中に（男性を対象とした）強制的な徴兵制があったことを利用しました。
- 1970 年代、軍は抽選を行い、各誕生日にランダムに優先番号を割り当てました。番号が小さい日に生まれた人から順に徴兵の対象となりました。

Angrist (1990) – 徴兵抽選

- Angrist (1990) は、ベトナム戦争中に（男性を対象とした）強制的な徴兵制があったことを利用しました。
- 1970 年代、軍は抽選を行い、各誕生日にランダムに優先番号を割り当てました。番号が小さい日に生まれた人から順に徴兵の対象となりました。
- 抽選番号は、入隊するかどうかを完全には決定しませんでした：
  - 番号が小さくても、医学的な免除を受けたり、国外へ逃れたりする人がいました。
  - 番号が大きくても、自ら志願して入隊する人がいました。

Angrist (1990) – 徴兵抽選

- Angrist (1990) は、ベトナム戦争中に（男性を対象とした）強制的な徴兵制があったことを利用しました。

- 1970 年代、軍は抽選を行い、各誕生日にランダムに優先番号を割り当てました。番号が小さい日に生まれた人から順に徴兵の対象となりました。

- 抽選番号は、入隊するかどうかを完全には決定しませんでした：
  - 番号が小さくても、医学的な免除を受けたり、国外へ逃れたりする人がいました。
  - 番号が大きくても、自ら志願して入隊する人がいました。

- Angrist はまず操作変数を 2 値化（徴兵対象の誕生日 vs 対象外）して考え、次に多値の操作変数（各誕生日ごとの変数）へと一般化しました。

IV 仮定の評価

- 独立性：

IV 仮定の評価

- 独立性：抽選番号が、潜在的な賃金や入隊傾向と系統的に関連していないこと。

IV 仮定の評価

- 独立性：抽選番号が、潜在的な賃金や入隊傾向と系統的に関連していないこと。
  - 抽選のランダム化により、この仮定は非常に妥当です。
  - 番号の大小で、観察可能な特徴に差はありません。
- 除外制約：

IV 仮定の評価

- 独立性：抽選番号が、潜在的な賃金や入隊傾向と系統的に関連していないこと。
  - 抽選のランダム化により、この仮定は非常に妥当です。
  - 番号の大小で、観察可能な特徴に差はありません。
- 除外制約：抽選番号が入隊を通じてのみ賃金に影響を与えること。

IV 仮定の評価

- 独立性：抽選番号が、潜在的な賃金や入隊傾向と系統的に関連していないこと。
    - 抽選のランダム化により、この仮定は非常に妥当です。
    - 番号の大小で、観察可能な特徴に差はありません。
- 除外制約：抽選番号が入隊を通じてのみ賃金に影響を与えること。
    - この文脈での除外制約は少し巧妙です。
    - ネバー・テイカーは、番号が小さいと（徴兵を避けるために）国外へ逃げなければならなかったかもしれません。
    - オールウェイズ・テイカーは、自ら志願した場合と徴兵された場合で、軍での経験が異なるかもしれません。
- 関連性：

IV 仮定の評価

- 独立性：抽選番号が、潜在的な賃金や入隊傾向と系統的に関連していないこと。
  - 抽選のランダム化により、この仮定は非常に妥当です。
  - 番号の大小で、観察可能な特徴に差はありません。

- 除外制約：抽選番号が入隊を通じてのみ賃金に影響を与えること。
  - この文脈での除外制約は少し巧妙です。
  - ネバー・テイカーは、番号が小さいと（徴兵を避けるために）国外へ逃げなければならなかったかもしれません。
  - オールウェイズ・テイカーは、自ら志願した場合と徴兵された場合で、軍での経験が異なるかもしれません。

- 関連性：抽選番号が入隊に影響を与えること。

IV 仮定の評価

- 独立性：抽選番号が、潜在的な賃金や入隊傾向と系統的に関連していないこと。
  - 抽選のランダム化により、この仮定は非常に妥当です。
  - 番号の大小で、観察可能な特徴に差はありません。
- 除外制約：抽選番号が入隊を通じてのみ賃金に影響を与えること。
  - この文脈での除外制約は少し巧妙です。
  - ネバー・テイカーは、番号が小さいと（徴兵を避けるために）国外へ逃げなければならなかったかもしれません。
  - オールウェイズ・テイカーは、自ら志願した場合と徴兵された場合で、軍での経験が異なるかもしれません。
- 関連性：抽選番号が入隊に影響を与えること。
  - 妥当であり、実際に影響していました。

IV 仮定の評価

- 独立性：抽選番号が、潜在的な賃金や入隊傾向と系統的に関連していないこと。
    - 抽選のランダム化により、この仮定は非常に妥当です。
    - 番号の大小で、観察可能な特徴に差はありません。
- 除外制約：抽選番号が入隊を通じてのみ賃金に影響を与えること。
    - この文脈での除外制約は少し巧妙です。
    - ネバー・テイカーは、番号が小さいと（徴兵を避けるために）国外へ逃げなければならなかったかもしれません。
    - オールウェイズ・テイカーは、自ら志願した場合と徴兵された場合で、軍での経験が異なるかもしれません。
- 関連性：抽選番号が入隊に影響を与えること。
    - 妥当であり、実際に影響していました。
- 単調性：番号が大きくても入隊する人は、番号が小さければ必ず入隊する。

IV 仮定の評価

- 独立性：抽選番号が、潜在的な賃金や入隊傾向と系統的に関連していないこと。
  - 抽選のランダム化により、この仮定は非常に妥当です。
  - 番号の大小で、観察可能な特徴に差はありません。
- 除外制約：抽選番号が入隊を通じてのみ賃金に影響を与えること。
  - この文脈での除外制約は少し巧妙です。
  - ネバー・テイカーは、番号が小さいと（徴兵を避けるために）国外へ逃げなければならなかったかもしれません。
  - オールウェイズ・テイカーは、自ら志願した場合と徴兵された場合で、軍での経験が異なるかもしれません。
- 関連性：抽選番号が入隊に影響を与えること。
  - 妥当であり、実際に影響していました。
- 単調性：番号が大きくても入隊する人は、番号が小さければ必ず入隊する。
  - 妥当と思われます。

- Angrist はまず、誕生日を徴兵対象（eligible）と対象外（non-eligible）にグループ分けした 2 値版の操作変数を考えました。

- 以下はその第一段階の結果です（ $\hat{p}^e - \hat{p}^n$ を参照）。

TABLE 2—VETERAN STATUS AND DRAFT ELIGIBILITY

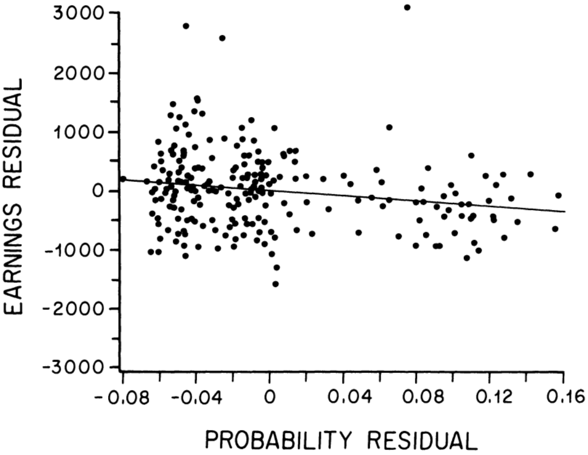
Whites						
Data Set	Cohort	Sample	P(Veteran)	$\hat{p}^e$	$\hat{p}^n$	$\hat{p}^e - \hat{p}^n$
SIPP (84)[a]	1950	351	0.2673	0.3527	0.1933	0.1594
			(0.0140)	(0.0325)	(0.0233)	(0.0400)
	1951	359	0.1973	0.2831	0.1468	0.1362
			(0.0127)	(0.0390)	(0.0180)	(0.0429)
	1952	336	0.1554	0.2310	0.1257	0.1053
			(0.0114)	(0.0473)	(0.0146)	(0.0495)
	1953	390	0.1298	0.1581	0.1153	0.0427
			(0.0106)	(0.0339)	(0.0152)	(0.0372)
DMDC/CWHS[b]	1950	16119	0.0633	0.0936	0.0279	0.0657
			(0.0019)	(0.0032)	(0.0019)	(0.0037)
	1951	16768	0.1176	0.2071	0.0708	0.1362
			(0.0025)	(0.0053)	(0.0024)	(0.0059)
	1952	17703	0.1515	0.2683	0.1102	0.1581
			(0.0027)	(0.0065)	(0.0027)	(0.0071)
	1953	17749	0.1343	0.1548	0.1268	0.0280
			(0.0026)	(0.0053)	(0.0029)	(0.0060)

TABLE 3—WALD ESTIMATES

Cohort	Year	Draft-Eligibility Effects in Current $: FICA Earnings (1)	Adjusted FICA Earnings (2)	Total W-2 Earnings (3)	$\hat{p}^e - \hat{p}^n$ (4)	Service Effect in 1978 $ (5)
1950	1981	−435.8	−487.8	−589.6	0.159	−2,195.8
		(210.5)	(237.6)	(299.4)	(0.040)	(1,069.5)
	1982	−320.2	−396.1	−305.5		−1,678.3
		(235.8)	(281.7)	(345.4)		(1,193.6)
	1983	−349.5	−450.1	−512.9		−1,795.6
		(261.6)	(302.0)	(441.2)		(1,204.8)
	1984	−484.3	−638.7	−1,143.3		−2,517.7
		(286.8)	(336.5)	(492.2)		(1,326.5)
1951	1981	−358.3	−428.7	−71.6	0.136	−2,261.3
		(203.6)	(224.5)	(423.4)	(0.043)	(1,184.2)
	1982	−117.3	−278.5	−72.7		−1,386.6
		(229.1)	(264.1)	(372.1)		(1,312.1)
	1983	−314.0	−452.2	−896.5		-2,181.8
		(253.2)	(289.2)	(426.3)		(1,395.3)
	1984	−398.4	−573.3	−809.1		−2,647.9
		(279.2)	(331.1)	(380.9)		(1,529.2)
1952	1981	−342.8	−392.6	−440.5	0.105	−2,502.3
		(206.8)	(228.6)	(265.0)	(0.050)	(1,556.7)
	1982	−235.1	−255.2	−514.7		−1,626.5
		(232.3)	(264.5)	(296.5)		(1,685.8)
	1983	−437.7	−500.0	−915.7		−3,103.5
		(257.5)	(294.7)	(395.2)		(1,829.2)
	1984	−436.0	−560.0	−767.2		−3,323.8
		(281.9)	(330.1)	(376.0)		(1,959.3)

この例における LATE の可視化

図解は こちら を参照してください。



Notes: The figure plots the history of FICA taxable earnings for the four cohorts born 1950–53. For each cohort, separate lines are drawn for draft-eligible and draft-ineligible men. Plotted points show average real (1978) earnings of working men born in 1953, real earnings + \$3000 for men born in 1950, real earnings + \$2000 for men born in 1951, and real earnings + \$1000 for men born in 1952.

FIGURE 1. SOCIAL SECURITY EARNINGS PROFILES BY DRAFT-ELIGIBILITY STATUS

複数の操作変数を用いた 2SLS の定式化

- この手続きは、複数の操作変数を用いた二段階最小二乗法として定式化できます。
- 操作変数のベクトル $\boldsymbol{Z}_i$ があるとします（例：Z_1 は誕生日 1、Z_2 は誕生日 2、...）。

複数の操作変数を用いた 2SLS

- ステップ 1（第一段階）：OLS 回帰により $E[D_i|\boldsymbol{Z}_i]$ を推定する：

$$D_i = \gamma_0 + \boldsymbol{Z}_i'\boldsymbol{\gamma}_1 + u_i$$

（例：各誕生日ごとの平均入隊率を推定する）

複数の操作変数を用いた 2SLS

- ステップ 1（第一段階）：OLS 回帰により $E[D_i|\boldsymbol{Z}_i]$ を推定する：

$$D_i = \gamma_0 + \boldsymbol{Z}_i'\boldsymbol{\gamma}_1 + u_i$$

（例：各誕生日ごとの平均入隊率を推定する）

- ステップ 2：各ユニット i について $E[D_i|\boldsymbol{Z}_i]$ の予測値を構築する：

$$\hat{D}_i = \hat{\gamma}_0 + \boldsymbol{Z}_i'\hat{\boldsymbol{\gamma}}_1$$

（例：$\hat{D}_i$ はユニット i と同じ誕生日の人の平均入隊率）

複数の操作変数を用いた 2SLS

- ステップ 1（第一段階）：OLS 回帰により $E[D_i|\boldsymbol{Z}_i]$ を推定する：

$$D_i = \gamma_0 + \boldsymbol{Z}_i'\boldsymbol{\gamma}_1 + u_i$$

（例：各誕生日ごとの平均入隊率を推定する）

- ステップ 2：各ユニット i について $E[D_i|\boldsymbol{Z}_i]$ の予測値を構築する：

$$\hat{D}_i = \hat{\gamma}_0 + \boldsymbol{Z}_i'\hat{\boldsymbol{\gamma}}_1$$

（例：$\hat{D}_i$ はユニット i と同じ誕生日の人の平均入隊率）

- ステップ 3（第二段階）：Y_i を $\hat{D}_i$ に OLS 回帰する：

$$Y_i = \beta_0 + \hat{D}_i\beta_1 + \varepsilon_i$$

（例：賃金を誕生日ごとの平均入隊率に回帰する）

複数の操作変数を用いた 2SLS

- ステップ 1（第一段階）：OLS 回帰により $E[D_i|\boldsymbol{Z}_i]$ を推定する：

$$D_i = \gamma_0 + \boldsymbol{Z}_i'\boldsymbol{\gamma}_1 + u_i$$

（例：各誕生日ごとの平均入隊率を推定する）

- ステップ 2：各ユニット i について $E[D_i|\boldsymbol{Z}_i]$ の予測値を構築する：

$$\hat{D}_i = \hat{\gamma}_0 + \boldsymbol{Z}_i'\hat{\boldsymbol{\gamma}}_1$$

（例：$\hat{D}_i$ はユニット i と同じ誕生日の人の平均入隊率）

- ステップ 3（第二段階）：Y_i を $\hat{D}_i$ に OLS 回帰する：

$$Y_i = \beta_0 + \hat{D}_i\beta_1 + \varepsilon_i$$

（例：賃金を誕生日ごとの平均入隊率に回帰する）

- $\hat{\beta}_1$ が（加重）LATE の推定値となります。

TABLE 4—TWO-STAGE INSTRUMENTAL VARIABLES ESTIMATES

Whites			
Cohort	FICA Taxable Earnings	Adjusted FICA Earnings	Total W-2 Compensation
Model 1			
1950	−1709.2	−2093.7	−1895.0
	(946.8)	(1108.8)	(1333.1)
1951	−1457.1	−1983.7	−2431.4
	(959.3)	(1036.1)	(1152.1)
1952	−1724.0	−1943.0	−2058.7
	(863.1)	(927.2)	(1001.9)
1953	1223.8	900.7	−488.6
	(3232.1)	(3505.3)	(3936.0)

Martin and Yurukoglu (2017)

- Martin and Yurukoglu は以下の問いに興味を持っています：Fox News（保守的なテレビ局）への露出は、米国の投票行動にどのように影響するか？
- Fox News をよく見る地域とそうでない地域の投票行動を比較するだけで、因果的効果が得られるでしょうか？

Martin and Yurukoglu (2017)

- Martin and Yurukoglu は以下の問いに興味を持っています：Fox News（保守的なテレビ局）への露出は、米国の投票行動にどのように影響するか？
- Fox News をよく見る地域とそうでない地域の投票行動を比較するだけで、因果的効果が得られるでしょうか？
  - おそらく得られません！ 政治的傾向という強力な交絡因子があるからです（保守的な人ほど Fox News を見、かつ共和党に投票する）。
- 実質的にランダムで、視聴数を通じてのみ投票に影響する操作変数が必要です。何かアイデアはありますか？

Martin and Yurukoglu (2017)

- Martin and Yurukoglu は以下の問いに興味を持っています：Fox News（保守的なテレビ局）への露出は、米国の投票行動にどのように影響するか？

- Fox News をよく見る地域とそうでない地域の投票行動を比較するだけで、因果的効果が得られるでしょうか？
  - おそらく得られません！ 政治的傾向という強力な交絡因子があるからです（保守的な人ほど Fox News を見、かつ共和党に投票する）。

- 実質的にランダムで、視聴数を通じてのみ投票に影響する操作変数が必要です。何かアイデアはありますか？

- 彼らは、ケーブルテレビの「チャンネル・ラインナップにおける位置」を IV として使うことを提案しました。
  - 番組ガイドの早い位置にあるほど、Fox News はよく見られる傾向があります。
  - ラインナップの位置は、1990 年代のケーブルニュース導入時の特異な要因によって決まっており、ランダムと同様であると議論しました。

Martin and Yurukoglu の 2SLS 仕様

第一段階：

$$D_i = \pi Z_i + \boldsymbol{\gamma}' \boldsymbol{X}_i + u_i$$

ここで D_i は郵便番号 i における 1 週間あたりの Fox News 平均視聴時間、$\boldsymbol{X}_i$ は州・郡の固定効果、MSNBC の位置、人口統計学的変数（所得、年齢、人種など）を含むコントロール変数のベクトルです。

第二段階：

$$Y_i = \beta \hat{D}_i + \boldsymbol{\delta}' \boldsymbol{X}_i + \varepsilon_i$$

ここで Y_i は 2008 年選挙における共和党得票率、$\hat{D}_i$ は第一段階からの予測値です。

仮定の評価

- 独立性：

仮定の評価

- 独立性：Fox News の位置が（観察可能な特徴を条件として）実質的にランダムであること。

仮定の評価

- 独立性：Fox News の位置が（観察可能な特徴を条件として）実質的にランダムであること。これが最も議論の余地があります。需要の高い場所ほど Fox News を早い位置に置くのではないか、という懸念があります。

- 除外制約：

仮定の評価

- 独立性：Fox News の位置が（観察可能な特徴を条件として）実質的にランダムであること。これが最も議論の余地があります。需要の高い場所ほど Fox News を早い位置に置くのではないか、という懸念があります。

- 除外制約：チャンネルの位置が視聴数を通じてのみ投票に影響すること。

仮定の評価

- 独立性：Fox News の位置が（観察可能な特徴を条件として）実質的にランダムであること。これが最も議論の余地があります。需要の高い場所ほど Fox News を早い位置に置くのではないか、という懸念があります。
- 除外制約：チャンネルの位置が視聴数を通じてのみ投票に影響すること。比較的妥当に見えます。
- 関連性：

仮定の評価

- 独立性：Fox News の位置が（観察可能な特徴を条件として）実質的にランダムであること。これが最も議論の余地があります。需要の高い場所ほど Fox News を早い位置に置くのではないか、という懸念があります。
- 除外制約：チャンネルの位置が視聴数を通じてのみ投票に影響すること。比較的妥当に見えます。
- 関連性：位置が視聴数に影響すること。

仮定の評価

- 独立性：Fox News の位置が（観察可能な特徴を条件として）実質的にランダムであること。これが最も議論の余地があります。需要の高い場所ほど Fox News を早い位置に置くのではないか、という懸念があります。
- 除外制約：チャンネルの位置が視聴数を通じてのみ投票に影響すること。比較的妥当に見えます。
- 関連性：位置が視聴数に影響すること。実際に影響していました。
- 単調性：

仮定の評価

- 独立性：Fox News の位置が（観察可能な特徴を条件として）実質的にランダムであること。これが最も議論の余地があります。需要の高い場所ほど Fox News を早い位置に置くのではないか、という懸念があります。
- 除外制約：チャンネルの位置が視聴数を通じてのみ投票に影響すること。比較的妥当に見えます。
- 関連性：位置が視聴数に影響すること。実際に影響していました。
- 単調性：位置が早くなるほど、視聴数が増える（または減らない）こと。

仮定の評価

- 独立性：Fox News の位置が（観察可能な特徴を条件として）実質的にランダムであること。これが最も議論の余地があります。需要の高い場所ほど Fox News を早い位置に置くのではないか、という懸念があります。
- 除外制約：チャンネルの位置が視聴数を通じてのみ投票に影響すること。比較的妥当に見えます。
- 関連性：位置が視聴数に影響すること。実際に影響していました。
- 単調性：位置が早くなるほど、視聴数が増える（または減らない）こと。概ね妥当ですが、遅い位置の方が他の人気チャンネルに近くて視聴が増える、といったことがあれば違反になります。

第一段階

TABLE 2—FIRST-STAGE REGRESSIONS: NIELSEN DATA

	FNC minutes per week					
	(1)	(2)	(3)	(4)	(5)	(6)
FNC position	−0.146	−0.075	−0.174	−0.167	−0.097	−0.111
	(0.043)	(0.039)	(0.028)	(0.025)	(0.033)	(0.030)
MSNBC position	0.078	0.073	0.064	0.070	0.019	0.020
	(0.036)	(0.032)	(0.025)	(0.022)	(0.034)	(0.035)
Has MSNBC only	1.904	1.137	−3.954	−2.804	−1.220	−1.562
	(3.697)	(3.713)	(4.255)	(3.416)	(6.180)	(5.397)
Has FNC only	31.423	26.526	23.460	22.011	15.141	15.069
	(2.677)	(2.546)	(2.278)	(1.864)	(2.697)	(2.314)
Has both	24.859	23.118	18.338	16.168	15.159	14.486
	(2.919)	(2.687)	(2.361)	(1.991)	(3.216)	(2.842)
Satellite FNC minutes				0.197		0.173
				(0.013)		(0.015)
Fixed effects	Year	State-year	State-year	State-year	County-year	County-year
Cable controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	None	None	Extended	Extended	Extended	Extended
Robust F-stat	11.39	3.72	39.02	44.7	8.86	13.43
Number of clusters	5,789	5,789	4,830	4,761	4,839	4,770
Observations	71,150	71,150	59,541	52,053	59,684	52,165
R^2	0.030	0.074	0.213	0.377	0.428	0.544

- ラインナップの後ろの方にあるほど、視聴数が少なくなっています。ラインナップ位置が 1 標準偏差改善すると、視聴数は週に約 2.5 分増加します。

独立性に関する証拠

TABLE 5—FNC CABLE POSITION PLACEBO TESTS

	Predicted viewing		Predicted voting		1996 contributions		1996 vote
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FNC position	0.100	0.033	0.027	−0.0004	0.0002	0.001	−0.006
	(0.034)	(0.026)	(0.022)	(0.017)	(0.0002)	(0.0003)	(0.012)
Fixed effects	State-year	County-year	State-year	County-year	State-year	County-year	State-year
Demographics	Extended	Extended	Extended	Extended	Extended	Extended	Extended
Number of clusters	4,830	4,839	4,814	4,827	4,830	4,839	4,830
Observations	59,551	59,694	17,400	17,451	59,551	59,694	59,551
R^2	0.380	0.827	0.339	0.729	0.176	0.436	0.571

- Fox News の位置は、1996 年（ケーブルニュース普及前）の得票率や、ほとんどの人口統計学的特徴と有意な関連がありません。

衛星放送視聴者を用いた検証

TABLE 6—SATELLITE PLACEBO FIRST STAGE: NIELSEN DATA

	FNC minutes per week			
	(1)	(2)	(3)	(4)
FNC position × cable	−0.155	−0.264	−0.151	−0.219
	(0.043)	(0.035)	(0.048)	(0.051)
FNC position × sat	0.031	−0.050	0.037	0.045
	(0.049)	(0.041)	(0.063)	(0.067)
MSNBC position × cable	0.102	0.092	0.035	0.046
	(0.036)	(0.032)	(0.049)	(0.048)
MSNBC position × sat	−0.004	−0.029	−0.029	−0.033
	(0.040)	(0.033)	(0.072)	(0.074)
Fixed effects	State-year	State-year	County-year	County-year
Cable controls	Yes	Yes	Yes	Yes
Demographics	None	Extensive	None	Extensive
Chow test *p*-value	0	0	0.011	0.001
Number of clusters	5,786	4,830	5,786	4,830
Observations	127,072	107,829	127,072	107,829
R^2	0.032	0.077	0.232	0.278

- チャンネル位置が視聴数を予測するのは（そのラインナップを使っている）ケーブルテレビ利用者のみであり、別のラインナップを使っている衛星放送利用者には影響していません。

2SLS の結果

Table 4—Second Stage Regressions: Zip Code Voting Data

	2008 McCain vote percentage			
	(1)	(2)	(3)	(4)
Predicted FNC minutes	0.152	0.120	0.157	0.098
	(0.056, 0.277)	(0.005, 0.248)	(−0.126, 0.938)	(−0.121, 0.429)
Satellite FNC minutes		−0.021		−0.015
		(−0.047, 0.001)		(−0.073, 0.022)
Fixed effects	State	State	County	County
Cable system controls	Yes	Yes	Yes	Yes
Demographics	Extended	Extended	Extended	Extended
Number of clusters	4,814	3,993	4,729	4,001
Observations	17,400	12,417	17,283	12,443
R^2	0.833	0.841	0.907	0.919

チャンネル位置の 1 標準偏差の改善は視聴数を週 2.5 分増加させ、それにより共和党得票率が 0.3 ポイント上昇したと推定されました。

2SLS の推論

- 2SLS 推定値の標準誤差をどのように求めればよいでしょうか？

2SLS の推論

- 2SLS 推定値の標準誤差をどのように求めればよいでしょうか？
- 2SLS は、第一段階の予測値 $\hat{D}$ を説明変数として OLS を実行するのと同等であることを見ました。

2SLS の推論

- 2SLS 推定値の標準誤差をどのように求めればよいでしょうか？
- 2SLS は、第一段階の予測値 $\hat{D}$ を説明変数として OLS を実行するのと同等であることを見ました。
- しかし、単にその第二段階の OLS を実行して、その OLS 標準誤差をそのまま使うのは間違いです。

2SLS の推論

- 2SLS 推定値の標準誤差をどのように求めればよいでしょうか？
- 2SLS は、第一段階の予測値 $\hat{D}$ を説明変数として OLS を実行するのと同等であることを見ました。
- しかし、単にその第二段階の OLS を実行して、その OLS 標準誤差をそのまま使うのは間違いです。$\hat{D}$ 自体が推定値である（誤差を含んでいる）ことを考慮に入れていないからです。

2SLS の推論

- 2SLS 推定値の標準誤差をどのように求めればよいでしょうか？
- 2SLS は、第一段階の予測値 $\hat{D}$ を説明変数として OLS を実行するのと同等であることを見ました。
- しかし、単にその第二段階の OLS を実行して、その OLS 標準誤差をそのまま使うのは間違いです。$\hat{D}$ 自体が推定値である（誤差を含んでいる）ことを考慮に入れていないからです。
- 幸い、Stata などのソフトウェアを使えば、正しい 2SLS 標準誤差を簡単に得ることができます。
    - Stata では `ivregress 2sls y x (d=z), r` と入力します。

2SLS の推論

- 2SLS 推定値の標準誤差をどのように求めればよいでしょうか？
- 2SLS は、第一段階の予測値 $\hat{D}$ を説明変数として OLS を実行するのと同等であることを見ました。
- しかし、単にその第二段階の OLS を実行して、その OLS 標準誤差をそのまま使うのは間違いです。$\hat{D}$ 自体が推定値である（誤差を含んでいる）ことを考慮に入れていないからです。
- 幸い、Stata などのソフトウェアを使えば、正しい 2SLS 標準誤差を簡単に得ることができます。
  - Stata では `ivregress 2sls y x (d=z), r` と入力します。
- これらの標準誤差はどこから来ているのでしょうか？

第一段階と誘導型の正規性

- 単一の操作変数の場合、$\hat{\beta}_{IV} = \hat{\gamma}_1 / \hat{\pi}_1$ でした。ここで $\hat{\gamma}_1, \hat{\pi}_1$ は以下の OLS 推定値です：

$$Y_i = \gamma_0 + Z_i \gamma_1 + \varepsilon_i$$
$$D_i = \pi_0 + Z_i \pi_1 + u_i$$

第一段階と誘導型の正規性

- 単一の操作変数の場合、$\hat{\beta}_{IV} = \hat{\gamma}_1 / \hat{\pi}_1$ でした。ここで $\hat{\gamma}_1, \hat{\pi}_1$ は以下の OLS 推定値です：

$$
\begin{aligned}
Y_i &= \gamma_0 + Z_i \gamma_1 + \varepsilon_i \\
D_i &= \pi_0 + Z_i \pi_1 + u_i
\end{aligned}
$$

- OLS 推定値が漸近的に正規分布に従うことは既に示しました（これらは同時にも成立します）：

$$
\sqrt{N} \left(\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\pi}_1 \end{pmatrix} - \begin{pmatrix} \gamma_1 \\ \pi_1 \end{pmatrix} \right) \to_d \mathrm{N}(0, \Sigma)
$$

第一段階と誘導型の正規性

- 単一の操作変数の場合、$\hat{\beta}_{IV} = \hat{\gamma}_1 / \hat{\pi}_1$ でした。ここで $\hat{\gamma}_1, \hat{\pi}_1$ は以下の OLS 推定値です：

$$Y_i = \gamma_0 + Z_i \gamma_1 + \varepsilon_i$$
$$D_i = \pi_0 + Z_i \pi_1 + u_i$$

- OLS 推定値が漸近的に正規分布に従うことは既に示しました（これらは同時にも成立します）：

$$\sqrt{N}\left(\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\pi}_1 \end{pmatrix} - \begin{pmatrix} \gamma_1 \\ \pi_1 \end{pmatrix}\right) \to_d \mathrm{N}(0, \Sigma)$$

- これを用いて $\hat{\gamma}_1 / \hat{\pi}_1$ の漸近分布を導けるでしょうか？

デルタ法

- $g(\hat{\gamma}_1, \hat{\pi}_1) = \hat{\gamma}_1/\hat{\pi}_1$ とします。

デルタ法

- $g(\hat{\gamma}_1, \hat{\pi}_1) = \hat{\gamma}_1/\hat{\pi}_1$ とします。$(\hat{\gamma}_1, \hat{\pi}_1) \approx (\gamma_1, \pi_1)$ のとき、一次のテイラー展開により：

$$g(\hat{\gamma}_1, \hat{\pi}_1) \approx g(\gamma_1, \pi_1) + \nabla g(\gamma_1, \pi_1)(\hat{\gamma}_1 - \gamma_1, \hat{\pi}_1 - \pi_1)'$$

  ここで $\nabla g(\gamma_1, \pi_1)$ は (γ_1, π_1) における g の勾配（ベクトル）です。

- これは次を意味します：

$$\sqrt{N}\left(\underbrace{\frac{\hat{\gamma}_1}{\hat{\pi}_1}}_{g(\hat{\gamma}_1, \hat{\pi}_1)} - \underbrace{\frac{\gamma_1}{\pi_1}}_{g(\gamma_1, \pi_1)}\right) \approx \nabla g(\gamma_1, \pi_1)\sqrt{N}\left(\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\pi}_1 \end{pmatrix} - \begin{pmatrix} \gamma_1 \\ \pi_1 \end{pmatrix}\right)$$

デルタ法

- $g(\hat{\gamma}_1, \hat{\pi}_1) = \hat{\gamma}_1 / \hat{\pi}_1$ とします。$(\hat{\gamma}_1, \hat{\pi}_1) \approx (\gamma_1, \pi_1)$ のとき、一次のテイラー展開により：

$$g(\hat{\gamma}_1, \hat{\pi}_1) \approx g(\gamma_1, \pi_1) + \nabla g(\gamma_1, \pi_1)(\hat{\gamma}_1 - \gamma_1, \hat{\pi}_1 - \pi_1)'$$

  ここで $\nabla g(\gamma_1, \pi_1)$ は (γ_1, π_1) における g の勾配（ベクトル）です。

- これは次を意味します：

$$\sqrt{N}\left(\underbrace{\frac{\hat{\gamma}_1}{\hat{\pi}_1}}_{g(\hat{\gamma}_1, \hat{\pi}_1)} - \underbrace{\frac{\gamma_1}{\pi_1}}_{g(\gamma_1, \pi_1)} \right) \approx \nabla g(\gamma_1, \pi_1) \sqrt{N} \left(\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\pi}_1 \end{pmatrix} - \begin{pmatrix} \gamma_1 \\ \pi_1 \end{pmatrix} \right)$$

- 連続写像定理より、これは $\mathrm{N}(0, \nabla g(\gamma_1, \pi_1) \Sigma \nabla g(\gamma_1, \pi_1)')$ に分布収束します。

デルタ法

- $g(\hat{\gamma}_1, \hat{\pi}_1) = \hat{\gamma}_1 / \hat{\pi}_1$ とします。$(\hat{\gamma}_1, \hat{\pi}_1) \approx (\gamma_1, \pi_1)$ のとき、一次のテイラー展開により：

$$g(\hat{\gamma}_1, \hat{\pi}_1) \approx g(\gamma_1, \pi_1) + \nabla g(\gamma_1, \pi_1)(\hat{\gamma}_1 - \gamma_1, \hat{\pi}_1 - \pi_1)'$$

ここで $\nabla g(\gamma_1, \pi_1)$ は (γ_1, π_1) における g の勾配（ベクトル）です。

- これは次を意味します：

$$\sqrt{N}\left(\underbrace{\frac{\hat{\gamma}_1}{\hat{\pi}_1}}_{g(\hat{\gamma}_1, \hat{\pi}_1)} - \underbrace{\frac{\gamma_1}{\pi_1}}_{g(\gamma_1, \pi_1)}\right) \approx \nabla g(\gamma_1, \pi_1)\sqrt{N}\left(\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\pi}_1 \end{pmatrix} - \begin{pmatrix} \gamma_1 \\ \pi_1 \end{pmatrix}\right)$$

- 連続写像定理より、これは $\mathrm{N}(0, \nabla g(\gamma_1, \pi_1)\Sigma\nabla g(\gamma_1, \pi_1)')$ に分布収束します。

- 分散は標本対応物（例：$\nabla g(\hat{\gamma}_1, \hat{\pi}_1)$）を用いて推定できます。

弱識別 (Weak Identification)

- 注意：$\hat{\beta}_{IV} = \hat{\gamma}_1/\hat{\pi}_1$ が定義されるのは $\hat{\pi}_1 \neq 0$ のときのみです。
- $\hat{\pi}_1 \to 0$ となると、$|\hat{\beta}_{IV}| \to \infty$ となり、$\hat{\pi}_1 \approx 0$ のとき IV 推定量の挙動は非常に不安定になります。

弱識別 (Weak Identification)

- 注意：$\hat{\beta}_{IV} = \hat{\gamma}_1/\hat{\pi}_1$ が定義されるのは $\hat{\pi}_1 \neq 0$ のときのみです。
- $\hat{\pi}_1 \to 0$ となると、$|\hat{\beta}_{IV}| \to \infty$ となり、$\hat{\pi}_1 \approx 0$ のとき IV 推定量の挙動は非常に不安定になります。
- これまでの漸近理論では、$\pi_1 \neq 0$（関連性）を仮定し、N が大きくなれば $\hat{\pi}_1 \to_p \pi_1$ となるため、$\hat{\pi}_1$ が 0 に近くなる確率は 0 に収束し、問題になりませんでした。

弱識別 (Weak Identification)

- 注意：$\hat{\beta}_{IV} = \hat{\gamma}_1 / \hat{\pi}_1$ が定義されるのは $\hat{\pi}_1 \neq 0$ のときのみです。
- $\hat{\pi}_1 \to 0$ となると、$|\hat{\beta}_{IV}| \to \infty$ となり、$\hat{\pi}_1 \approx 0$ のとき IV 推定量の挙動は非常に不安定になります。
- これまでの漸近理論では、$\pi_1 \neq 0$（関連性）を仮定し、N が大きくなれば $\hat{\pi}_1 \to_p \pi_1$ となるため、$\hat{\pi}_1$ が 0 に近くなる確率は 0 に収束し、問題になりませんでした。
- しかし実務上、$\hat{\pi}_1$ が（標準誤差に対して）非常に 0 に近い場合があります。
- この場合、上記の正規分布による近似は、$\hat{\beta}_{IV}$ の分布を適切に捉えることができません。これは「弱識別」問題として知られています。

弱操作変数 – モンテカルロ・シミュレーション

モンテカルロ・シミュレーション：真の処置効果は 0

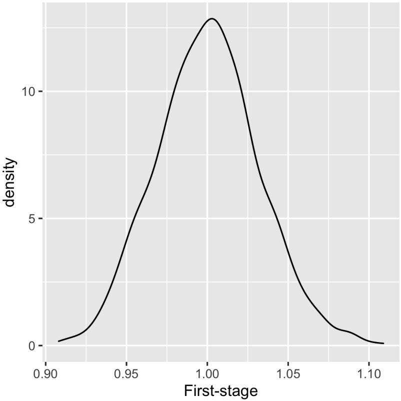
$Y_i(d) = \eta + \nu; \qquad D_i = \pi_1 Z_i + \eta; \qquad \eta, \nu, Z_i \sim N(0,1)$

弱操作変数 – モンテカルロ・シミュレーション

モンテカルロ・シミュレーション：真の処置効果は 0

$Y_i(d) = \eta + \nu; \qquad D_i = \pi_1 Z_i + \eta; \qquad \eta, \nu, Z_i \sim N(0,1)$

ケース 1**：強い第一段階** $(\pi_1 = 1) \rightarrow \pi_1 / SD(\pi_1) \approx 22$

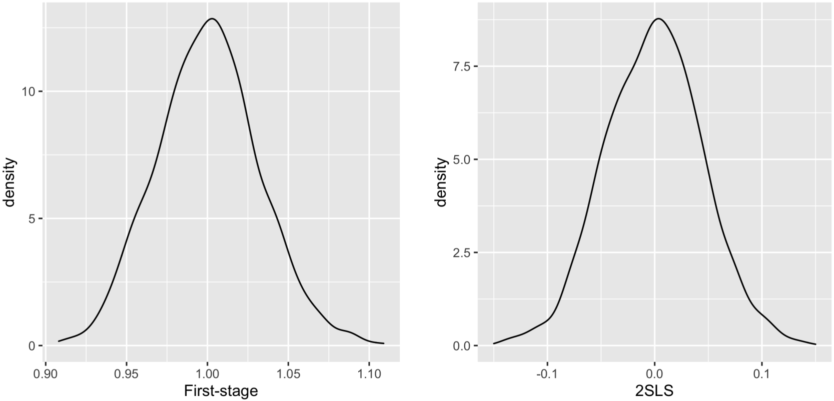


弱操作変数 – モンテカルロ・シミュレーション

モンテカルロ・シミュレーション：真の処置効果は 0

$Y_i(d) = \eta + \nu; \qquad D_i = \pi_1 Z_i + \eta; \qquad \eta, \nu, Z_i \sim N(0,1)$

ケース 1：強い第一段階 $(\pi_1 = 1) \rightarrow \pi_1 / SD(\pi_1) \approx 22$

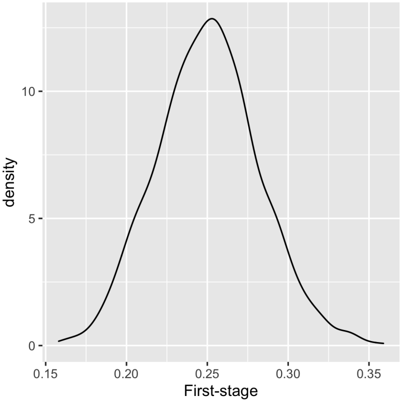


弱操作変数 – モンテカルロ・シミュレーション

モンテカルロ・シミュレーション：真の処置効果は 0

$Y_i(d) = \eta + \nu; \qquad D_i = \pi_1 Z_i + \eta; \qquad \eta, \nu, Z_i \sim N(0,1)$

ケース 2**：中程度の第一段階** $(\pi_1 = 0.25) \to \pi_1/SD(\pi_1) \approx 7$

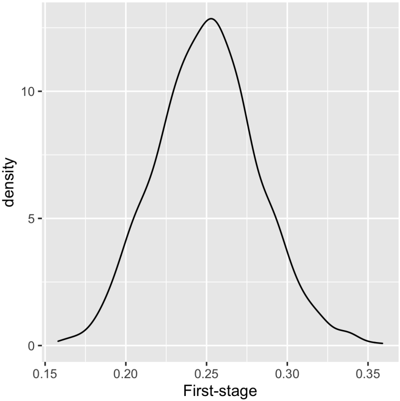


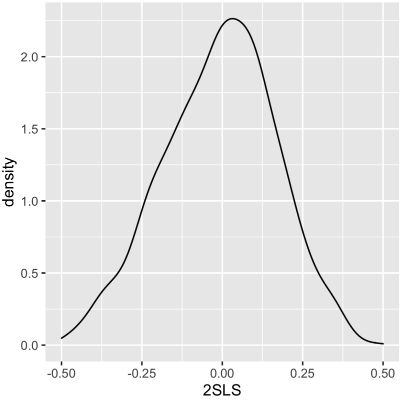
弱操作変数 – モンテカルロ・シミュレーション

モンテカルロ・シミュレーション：真の処置効果は 0

$Y_i(d) = \eta + \nu; \qquad D_i = \pi_1 Z_i + \eta; \qquad \eta, \nu, Z_i \sim N(0,1)$

ケース 2：中程度の第一段階 $(\pi_1 = 0.25) \rightarrow \pi_1 / SD(\pi_1) \approx 7$



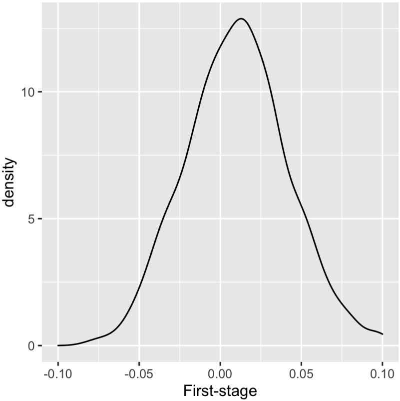


弱操作変数 – モンテカルロ・シミュレーション

モンテカルロ・シミュレーション：真の処置効果は 0

$Y_i(d) = \eta + \nu; \qquad D_i = \pi_1 Z_i + \eta; \qquad \eta, \nu, Z_i \sim N(0,1)$

ケース 3**：非常に弱い第一段階** $(\pi_1 = 0.01) \rightarrow \pi_1/SD(\pi_1) \approx 0.3$

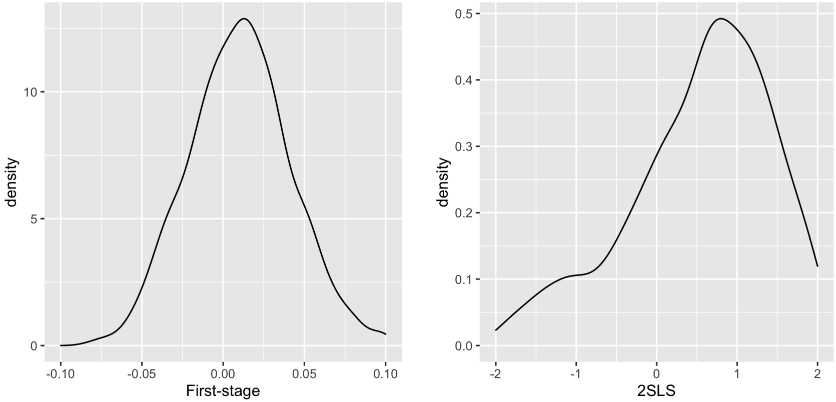


弱操作変数 – モンテカルロ・シミュレーション

モンテカルロ・シミュレーション：真の処置効果は 0

$Y_i(d) = \eta + \nu; \qquad D_i = \pi_1 Z_i + \eta; \qquad \eta, \nu, Z_i \sim N(0,1)$

ケース 3：非常に弱い第一段階 $(\pi_1 = 0.01) \rightarrow \pi_1 / SD(\pi_1) \approx 0.3$



複数の操作変数がある場合の弱 IV

多くの操作変数がある場合にも、概念的に関連した問題が生じます。

複数の操作変数がある場合の弱 IV

多くの操作変数がある場合にも、概念的に関連した問題が生じます。

- 典型的な例：Angrist and Krueger (1991) は、出生四半期と「出生年・州」の交差項を操作変数として用いました。

複数の操作変数がある場合の弱 IV

多くの操作変数がある場合にも、概念的に関連した問題が生じます。

- 典型的な例：Angrist and Krueger (1991) は、出生四半期と「出生年・州」の交差項を操作変数として用いました。
- これにより標本内の D_i の予測精度は上がりますが、標準誤差は減少します。

複数の操作変数がある場合の弱 IV

多くの操作変数がある場合にも、概念的に関連した問題が生じます。

- 典型的な例：Angrist and Krueger (1991) は、出生四半期と「出生年・州」の交差項を操作変数として用いました。
- これにより標本内の D_i の予測精度は上がりますが、標準誤差は減少します。
- しかし Bound et al. (1995) は、ランダムに生成した操作変数を使っても、彼らと同じ 2SLS 推定値が得られてしまうことを示しました。

複数の操作変数がある場合の弱 IV

多くの操作変数がある場合にも、概念的に関連した問題が生じます。

- 典型的な例：Angrist and Krueger (1991) は、出生四半期と「出生年・州」の交差項を操作変数として用いました。
- これにより標本内の D_i の予測精度は上がりますが、標準誤差は減少します。
- しかし Bound et al. (1995) は、ランダムに生成した操作変数を使っても、彼らと同じ 2SLS 推定値が得られてしまうことを示しました。

直感：第一段階で多くの操作変数を使うと、D_i を「過学習（overfit）」してしまいます。

複数の操作変数がある場合の弱 IV

多くの操作変数がある場合にも、概念的に関連した問題が生じます。

- 典型的な例：Angrist and Krueger (1991) は、出生四半期と「出生年・州」の交差項を操作変数として用いました。
- これにより標本内の D_i の予測精度は上がりますが、標準誤差は減少します。
- しかし Bound et al. (1995) は、ランダムに生成した操作変数を使っても、彼らと同じ 2SLS 推定値が得られてしまうことを示しました。

直感：第一段階で多くの操作変数を使うと、D_i を「過学習（overfit）」してしまいます。

- 極端なケースとして、各観測値に 1 つずつ操作変数を割り当てたとすると、第一段階のフィットは完璧になります：$\hat{D}_i = D_i$。すると数値的に 2SLS = OLS となります。
- つまり、弱操作変数の問題は第一段階での「過学習」から生じるのです。

弱 IV のテスト

- 一般的な経験則として、第一段階における操作変数の F 統計量が 10 未満であれば、弱操作変数を懸念すべきとされています。

- 第一段階の回帰

$$D_i = \pi_0 + \boldsymbol{Z}_i' \boldsymbol{\pi}_1 + \boldsymbol{X}_i' \boldsymbol{\pi}_2 + u_i$$

  を実行し、帰無仮説 $H_0 : \boldsymbol{\pi}_1 = 0$ に対する F 統計量を計算します。

- 操作変数が 1 つの場合、F 統計量は操作変数の t 統計量の 2 乗に等しいため、$F > 10$ は $|t| > 3.2$ に対応します。

```
. reg education D highnum_region_D post high_intensity highnum_region highnum_regi
> on_post highnum_region_high, r

Linear regression                               Number of obs     =     31,310
                                                F(7, 31302)       =     133.29
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0282
                                                Root MSE          =     3.8458

------------------------------------------------------------------------------
             |               Robust
   education |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           D |  -.1623344   .1236149    -1.31   0.189    -.4046245    .0799557
highnum_regio~D |   .5457673   .1761518     3.10   0.002     .2005028    .8910318
        post |   .3354358   .0763198     4.40   0.000     .1858459    .4850256
high_intensity |  -1.309835   .0951979   -13.76   0.000    -1.496426   -1.123243
highnum_region |  -.3988454   .0838969    -4.75   0.000    -.5632866   -.2344043
highnum_regio~t |  -.2082078   .1109533    -1.88   0.061    -.4256806     .009265
highnum_regio~h |   .0932633   .1348633     0.69   0.489    -.1710742    .3576008
       _cons |   10.05763   .0590164   170.42   0.000     9.941953     10.1733
------------------------------------------------------------------------------

. test D highnum_region_D

 ( 1)  D = 0
 ( 2)  highnum_region_D = 0

       F(  2, 31302) =    5.53
            Prob > F =    0.0040
```

この例では、第一段階の F 統計量は 5.53 です。

```
. ivreg2 log_wage (education=D highnum_region_D) post high_intensity highnum_regio
> n highnum_region_post highnum_region_high, r

IV (2SLS) estimation
--------------------

Estimates efficient for homoskedasticity only
Statistics robust to heteroskedasticity

                                                      Number of obs =    31310
                                                      F(  6, 31303) =   276.17
                                                      Prob > F      =   0.0000
Total (centered) SS     =  13459.19497                Centered R2   =  -0.2431
Total (uncentered) SS   =  4570266.616                Uncentered R2 =   0.9963
Residual SS             =  16731.05867                Root MSE      =     .731

------------------------------------------------------------------------------
             |               Robust
    log_wage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   education |   .1856081   .0584726     3.17   0.002     .0710038    .3002124
        post |  -.2773642   .0197758   -14.03   0.000     -.316124   -.2386044
high_intensity|  .0949207   .0819745     1.16   0.247    -.0657465    .2555878
highnum_region|  -.0947804  .0319689    -2.96   0.003    -.1574383   -.0321225
highnum_regio~t | -.0734455 .0165758    -4.43   0.000    -.1059335   -.0409576
highnum_regio~h |  .0379194 .0261998     1.45   0.148    -.0134312     .08927
       _cons |   10.45536   .5899054    17.72   0.000     9.299167    11.61155
------------------------------------------------------------------------------
Underidentification test (Kleibergen-Paap rk LM statistic):             11.057
                                                   Chi-sq(2) P-val =    0.0040
------------------------------------------------------------------------------
Weak identification test (Cragg-Donald Wald F statistic):                5.617
                         (Kleibergen-Paap rk Wald F statistic):          5.530
Stock-Yogo weak ID test critical values: 10% maximal IV size             19.93
                                         15% maximal IV size             11.59
                                         20% maximal IV size              8.75
                                         25% maximal IV size              7.25
Source: Stock-Yogo (2005).  Reproduced by permission.
NB: Critical values are for Cragg-Donald F statistic and i.i.d. errors.
------------------------------------------------------------------------------
```

Stata の ivreg2 コマンドなどから、この F 統計量を直接得ることができます。

弱操作変数に対して何ができるか？

- サンプルを分割して過学習を抑える方法があります（Split-sample IV や Jackknife IV）。
- 弱操作変数がある場合に、より適切な信頼区間を得る方法もあります。
  - 最も一般的なのは Anderson-Rubin 信頼集合ですが、この講義では扱いません。
- 操作変数の強度を高める努力をすることも考えられます：
  - サンプルサイズを大きくする。
  - D と相関する（が Z とは強く相関しない）コントロール変数を追加する。
  - 新しい操作変数を考える :)