2022.5.23

Ch.7
Instrumental Variables
(操作変数)

大湾ゼミ
椚田菜月　富澤陽仁

AGENDA

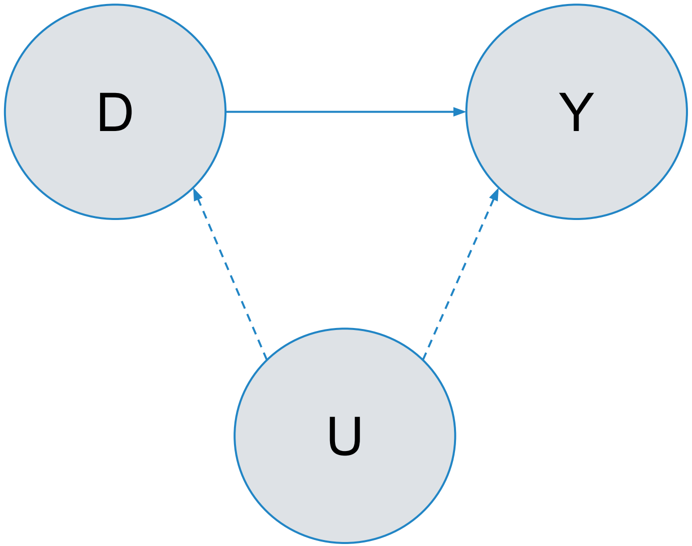
- ▷ 操作変数の直感的な理解
- ▷ 均一な処置効果のもとでの操作変数
- ▷ 親の覚せい剤乱用と里親制度
- ▷ 弱い操作変数が引き起こす問題

7.2

Intuition of Instrumental Variables

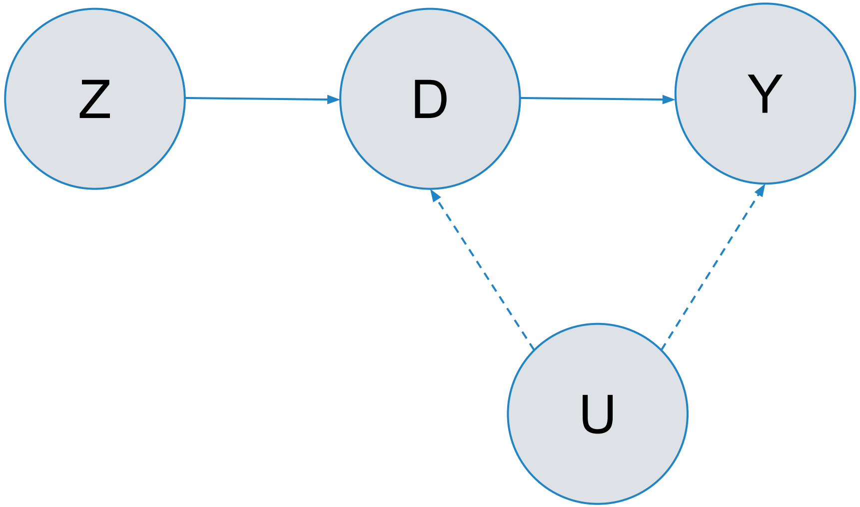
操作変数の直感的な理解

7.2.1 Canonical IV DAG



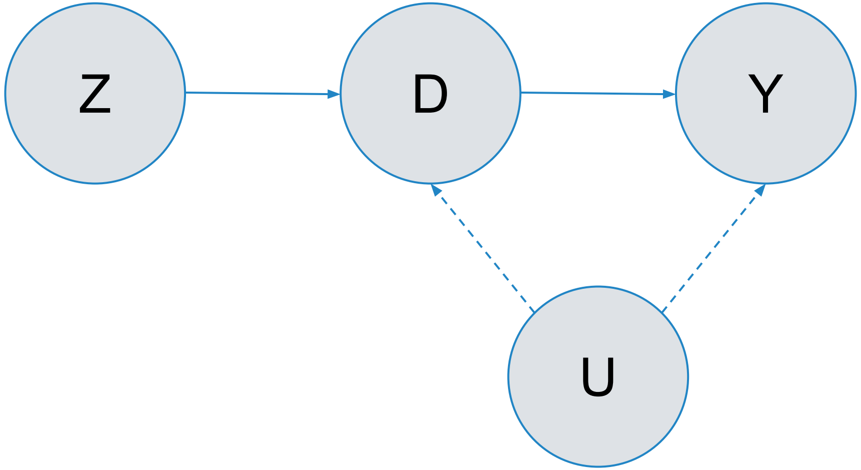
- Uは観測できない
- バックドアパスは開かれたまま
- DとYの直接的な因果関係はわからない

7.2.1 Canonical IV DAG



- 操作変数Zの導入
- ZはDと相関するが、Uとは相関していない
- Z affects Y “only through” D.

7.2.1 Canonical IV DAG

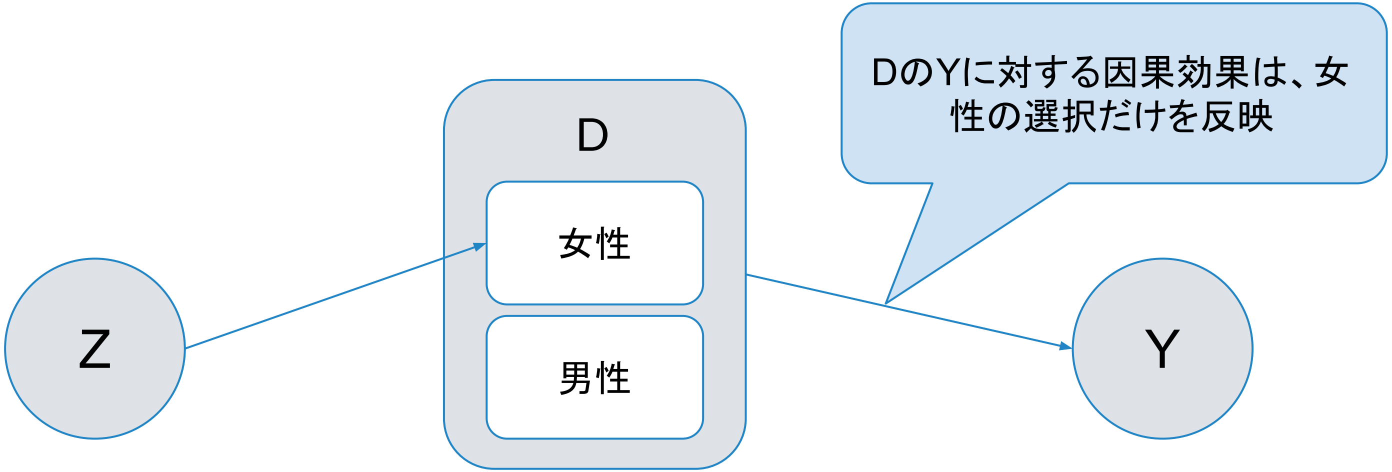


- Exclusion restriction

DがZとUのColliderであるためZとUは独立である
これは検証することができない

⇒IV推定量は、ZがD以外のYを決定する変数から独立しているということを仮定している。

7.2.1 Canonical IV DAG



問題①操作の影響に異質性がある場合（男性がYに与える影響が女性とは異なる）、ZはYに対するDの因果効果のうち女性のものだけしか特定することができない。

問題②女性のデータしかとることができないため、実際のDとYの因果関係を特定するには、データが少なくなってしまう。

7.2.1 Canonical IV DAG

▷ 操作変数の難しさ

①操作変数は、操作の結果として行動が変化したグループ(コンプライアー)についての因果効果のみを識別
＝コンプライアーの因果効果” the causal effect of the complier population”

②操作変数は、通常より大きな標準誤差を持つので、棄却に失敗することが多い

7.2.1 Canonical IV DAG

▷　操作変数の難しさ

②操作変数は、通常より大きな標準誤差を持つので、検出力不足以外の理由がない多くの場合、棄却に失敗する

問題②より
　使えるデータが少ない
→データのばらつき（分散）が大きい
→推定量の分散が大きい
→標準誤差が大きい
→検定の統計量が小さくなる
→棄却に失敗する

7.2.2 Good instruments should feel weird

▷ 良い操作変数とは

操作変数に対する結果（被説明変数）の話をしたときに、人々がすぐにはその関係を理解できない操作変数

e.g.
家族の大きさ（子どもの人数の大小）が、女性の労働供給量に影響する
→子どもが多い女性の方が家庭外で働く機会が少ない

最初の2人の子どもが同性であることが、女性の労働供給量に影響する
→？？？

7.2.2 Good instruments should feel weird

- ▷ 良い操作変数とは

最初の2人の子どもが同性であった母親は、2人の子どもの男女比が均衡していた母親よりも家庭外で働いていない
→性別の多様性を重視する家族にとって同性の子供は3人目を出産する（家族のサイズ）インセンティブとなる

良い操作変数Z	同性の子供
説明変数(内生変数)D	家族のサイズ
被説明変数Y	女性の労働供給

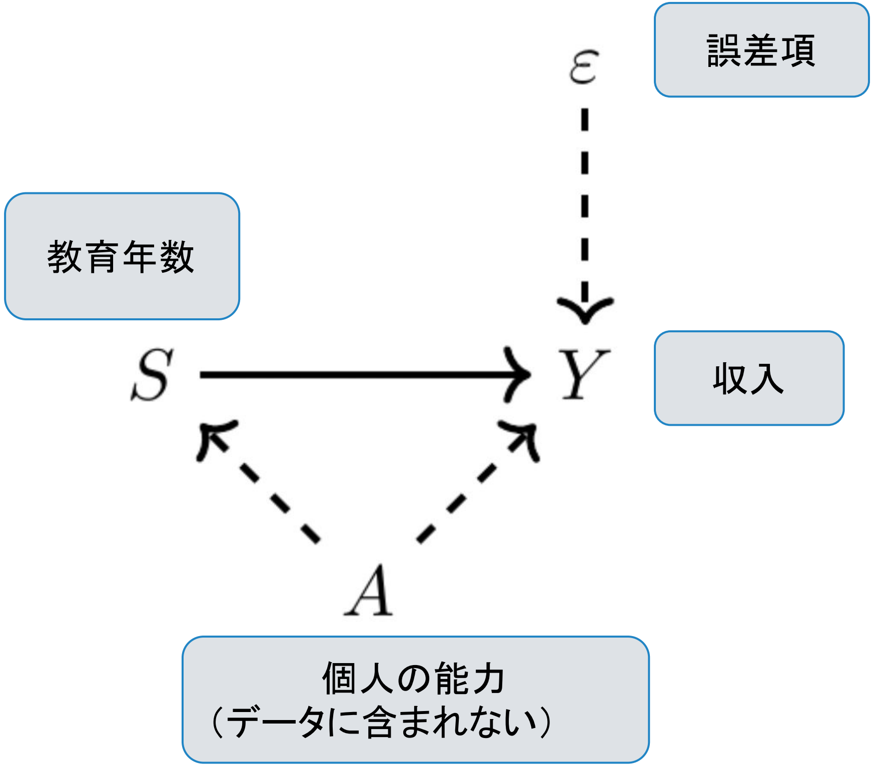
7.3 Homogeneous Treatment Effects

均一な処置効果のもとでの操作変数

この節のポイント

- 内生性の問題がある時にはOLS推定量はバイアスを持つ
- 操作変数法を使うときは2つの仮定が満たされている必要がある
- 二段階最小二乗法
  - 操作変数推定量は、説明変数の外生的な変動を通じた被説明変数への影響を測っている

学校教育が収入に与える影響



推定するモデル

➢ 真のモデル（実際にはAのデータがないので成り立たない）

$$Y_i = \alpha + \delta S_i + \cancel{\gamma A_i} + \varepsilon_i$$

➢ 実際のモデル

$$Y_i = \alpha + \boxed{\delta S_i + \eta_i} \qquad \eta_i = \gamma A_i + \varepsilon_i$$

内生性の問題！

➢ 処置効果δは一定と仮定

OLS推定量はバイアスを持つ

➢ モデルのOLS推定量

$$\hat{\delta} = \frac{C(Y, S)}{V(S)} = \frac{E[YS] - E[Y][S]}{V(S)}$$

欠落変数バイアス

$$\hat{\delta} = \frac{E[\alpha S + S^2\delta + \gamma SA + \varepsilon S] - E[\alpha + \delta S + \gamma A + \varepsilon][S]}{V(S)}$$

$$= \frac{\boxed{\delta E(S^2) - \delta E(S)^2} + \boxed{\gamma E(AS) - \gamma E(S)E(A)} + \boxed{E(\varepsilon S) - E(S)E(\varepsilon)}}{V(S)}$$

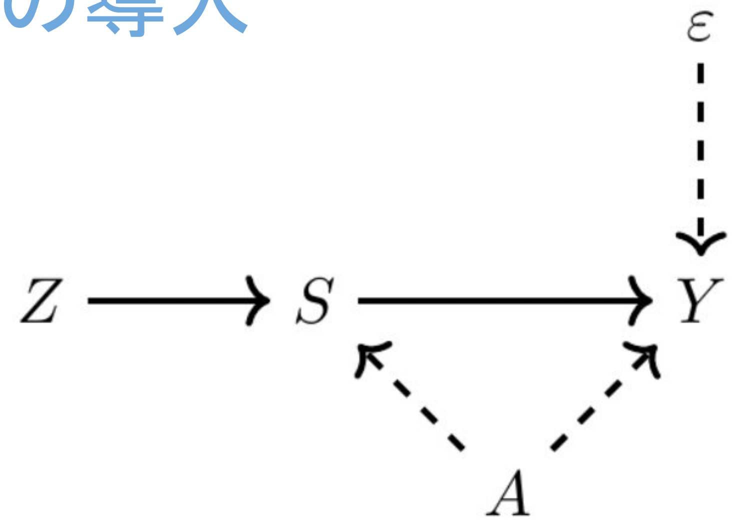
= COV(ε, S) = 0

$$= \delta + \gamma \frac{C(AS)}{V(S)}$$

能力と教育年数の
相関は正

能力が収入に与える
影響は正

操作変数の導入



- Zは教育年数Sを通してのみ収入Yに影響を与える
- Zは能力Aや誤差項とは無相関（exclusion restrictions）

操作変数推定量

$$\hat{\delta} = \frac{C(Y, Z)}{C(S, Z)}$$

Reduced form
（誘導形）

First stage
（一段階）

導出

Sを通してYに影響を与えているので≠0

$$
\begin{aligned}
\boxed{C(Y,Z)} &= C(\alpha + \delta S + \gamma A + \varepsilon, Z) \\
&= E[(\alpha + \delta S + \gamma A + \varepsilon), Z] - E(S)E(Z) \\
&= \{\alpha E(Z) - \alpha E(Z)\} + \delta\{E(SZ) - E(S)E(Z)\} \\
&\quad + \gamma\{E(AZ) - E(A)E(Z)\} + \{E(\varepsilon Z) - E(\varepsilon)E(Z)\} \\
&= \delta \boxed{C(S,Z)} + \boxed{\gamma C(A,Z) + C(\varepsilon, Z)}
\end{aligned}
$$

≠0である必要！

Exclusion Restrictions より0

操作変数が満たすべき2つの条件

$$Y_i = \alpha + \delta S_i + \eta_i$$ において、

$$C(S, Z) \neq 0$$

$$C(\eta, Z) = 0$$

$$\eta_i = \gamma A_i + \varepsilon_i$$

を満たすZを**操作変数**という

7.3.1 Two-stage least squares
（2SLS: 二段階最小二乗法）

- 2SLS推定量はIV推定量と一致
- IV推定量は説明変数の外生的な変動が被説明変数に与える影響を捉えたものである
- IV推定量は2つの条件のもとで一致推定量

7.3.1 Two-stage least squares（2段階最小二乗法）

➢ **二段階最小二乗法**：操作変数推定量のもう一つの推定方法

1. 説明変数Sを操作変数Zに回帰（一段階）
   これをOLS推定して、Sの予測値を取得

$$S_i = \gamma + \beta Z_i + \epsilon_i \xrightarrow{\text{OLS}} \hat{S}_i = \hat{\gamma} + \hat{\beta} Z_i$$

2. 被説明変数を説明変数の予測値に回帰

$$Y_i = \alpha + \delta \hat{S}_i + \varepsilon_i$$

7.3.1 Two-stage least squares
（2SLS: 二段階最小二乗法）

$$Y_i = \alpha + \delta S_i + \varepsilon_i$$

$$S_i = \gamma + \beta Z_i + \epsilon_i$$

$$\text{where } \boxed{C(Z, \varepsilon) = 0} \text{ and } \boxed{\beta \neq 0}$$

exclusion restriction

Non-zero first stage

2SLS推定量はIV推定量と一致

$$\hat{\delta}_{2SLS} = \frac{C(\hat{\beta}Z, Y)}{V(\hat{\beta}Z)} = \frac{C(\hat{S}, Y)}{V(\hat{S})} = \hat{\delta}_{IV}$$

Ⅳ推定量はSの外生的な変動を取り出す

- $\hat{\beta}Z$ は $\hat{S}_i = \hat{\gamma} + \hat{\beta}Z_i$ より、操作変数によって引き起こされた**外生的な変動**

- つまり操作変数Zを使うことで内生変数Sの外生的な変動による被説明変数Yへの因果効果を測ることができる

IV推定量は仮定のもとで一致推定量

- Exclusion RestrictionsとNon-zero first stageが成り立つとき、

$$p\lim \hat{\delta} = \delta$$

となり**一致性**を持つ

- ただし操作変数と説明変数（内生変数）の相関が弱いとき、バイアスを持つ（7.5参照）

IV推定量は仮定のもとで一致推定量(参考)

$$\hat{\delta} = \frac{C(Y,Z)}{C(S,Z)}$$

$$= \frac{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z})(Y_i - \bar{Y})}{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z})(S_i - \bar{S})}$$

$$= \frac{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z})Y_i}{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z})S_i}$$

IV推定量は仮定のもとで一致推定量（参考）

$$\hat{\delta} = \frac{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z})\{\alpha + \delta S + \varepsilon\}}{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z})S_i}$$

$$= \delta + \boxed{\frac{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z})\varepsilon_i}{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z})S_i}}$$

サンプルサイズが大きくなるとこ
の項は小さくなる

操作変数推定量の正しい導出(参考)

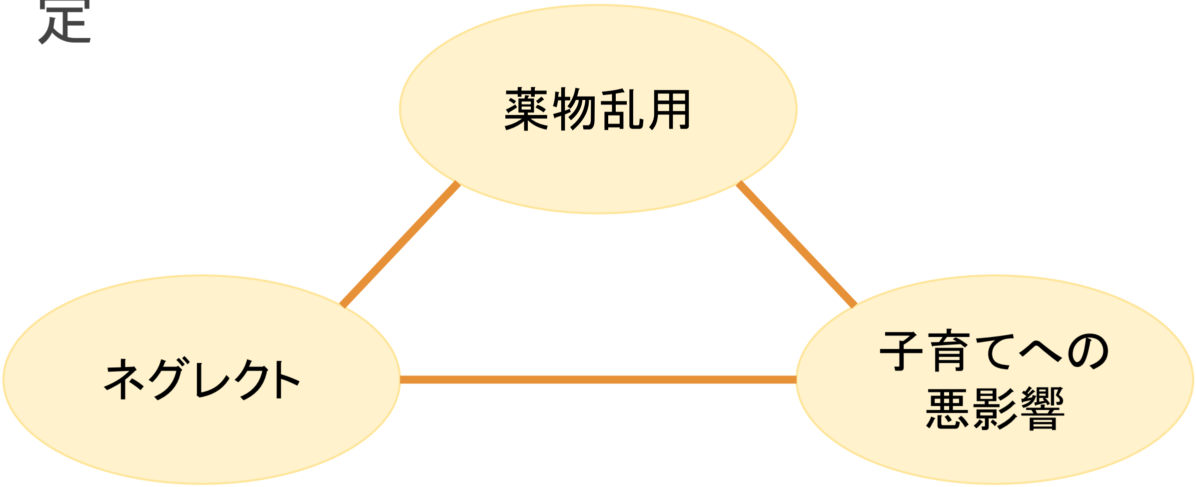
$$
\begin{aligned}
\mathbf{Cov}(Y,Z) &= \mathbf{E}[YZ] - \mathbf{E}[Y]\mathbf{E}[Z] \\
&= \mathbf{E}[(\alpha + \delta S + \gamma A + \varepsilon)Z] - \mathbf{E}[\alpha + \delta S + \gamma A + \varepsilon]\mathbf{E}[Z] \\
&= \Big(\mathbf{E}[\alpha Z + \delta SZ + \gamma AZ + \varepsilon Z]\Big) - \Big(\mathbf{E}[\alpha + \delta S + \gamma A + \varepsilon]\mathbf{E}[Z]\Big) \\
&= \Big(\mathbf{E}[\alpha Z] + \mathbf{E}[\delta SZ] + \mathbf{E}[\gamma AZ] + \mathbf{E}[\varepsilon Z]\Big) - \Big(\Big[\mathbf{E}[\alpha] + \mathbf{E}[\delta S] + \mathbf{E}[\gamma A] + \mathbf{E}[\varepsilon]\Big]\mathbf{E}[Z]\Big) \\
&= \Big(\alpha\mathbf{E}[Z] + \delta\mathbf{E}[SZ] + \gamma\mathbf{E}[AZ] + \mathbf{E}[\varepsilon Z]\Big) - \Big(\Big[\alpha + \delta\mathbf{E}[S] + \gamma\mathbf{E}[A] + \mathbf{E}[\varepsilon]\Big]\mathbf{E}[Z]\Big) \\
&= \Big(\alpha\mathbf{E}[Z] + \delta\mathbf{E}[SZ] + \gamma\mathbf{E}[AZ] + \mathbf{E}[\varepsilon Z]\Big) - \Big(\alpha\mathbf{E}[Z] + \delta\mathbf{E}[S]\mathbf{E}[Z] + \gamma\mathbf{E}[A]\mathbf{E}[Z] + \mathbf{E}[\varepsilon]\mathbf{E}[Z]\Big) \\
&= \Big(\delta\mathbf{E}[SZ] + \gamma\mathbf{E}[AZ] + \mathbf{E}[\varepsilon Z]\Big) - \Big(\delta\mathbf{E}[S]\mathbf{E}[Z] + \gamma\mathbf{E}[A]\mathbf{E}[Z] + \mathbf{E}[\varepsilon]\mathbf{E}[Z]\Big) \\
&= \delta\mathbf{Cov}(S,Z) + \gamma\mathbf{Cov}(A,Z) + \mathbf{Cov}(\varepsilon,Z).
\end{aligned}
$$

7.4
親の覚醒剤乱用と里親制度

（Cunninghum and Finlay, 2012）

7.4 親の覚醒剤乱用と里親制度

- 親の覚醒剤(methamphetamine)乱用による子どもへの虐待から里親への委託に与える影響を推定



7.4 親の覚醒剤乱用と里親制度

- methamphetamine（＝meth）
  心身に有害な毒、高い依存性
  警察などの発表では、メスの蔓延が里親募集の増加を引き起こす

⇒しかし、どうやって相関関係と因果関係を区別するか
⇒メスの製造方法によって可能

7.4 親の覚醒剤乱用と里親制度

▷　methamphetamine（＝meth）

市販の風邪薬の有効成分でもあるエフェドリンまたはプソイドエフェドリンの還元物から合成される

①米国議会がエフェドリンを主成分とする医薬品に規制
②しかしすぐにプソイドエフェドリンから合成されたメスが出回るように
③その後プソイドエフェドリンも規制対象に

7.4 親の覚醒剤乱用と里親制度

操作変数Z	小売価格の変動により計量した規制ショック
説明変数D	メス依存の治療のための入院者数
被説明変数Y	実の親元を離れ里親に預けられる子どもの数

メス、ヘロイン、コカインの小売価格の推移

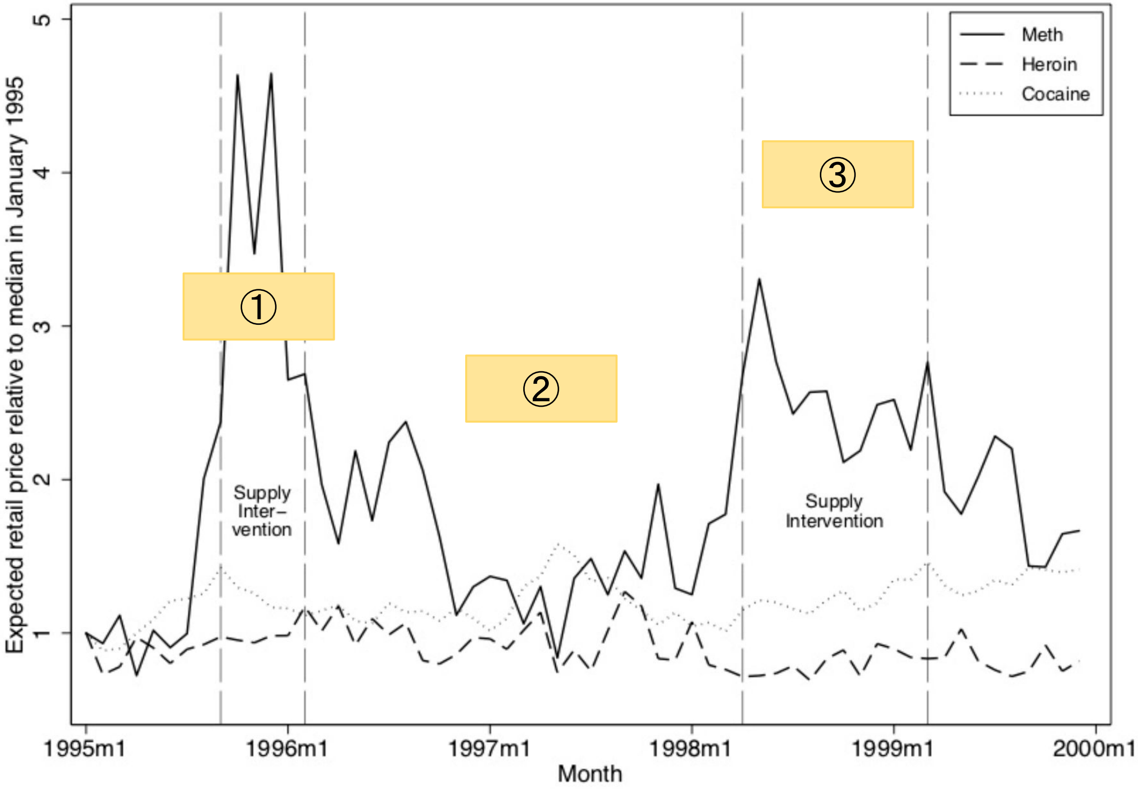


Figure 7.1: Ratio of Median Monthly Expected Retail Prices of Meth, Heroin, and Cocaine Relative to Their Respective Values in 1995, STRIDE 1995–1999. Reprinted from Cunningham and Finlay (2012).

- ▷ 米国議会による規制がメス市場にしか影響を与えていないことがわかる。
- ▷ 規制はメス市場に独自の影響を与え、その他の覚醒剤市場とは切り離してメスの影響を分析できる

連邦政府が資金援助しているクリニックで薬物依存の治療を受ける人の数の推移

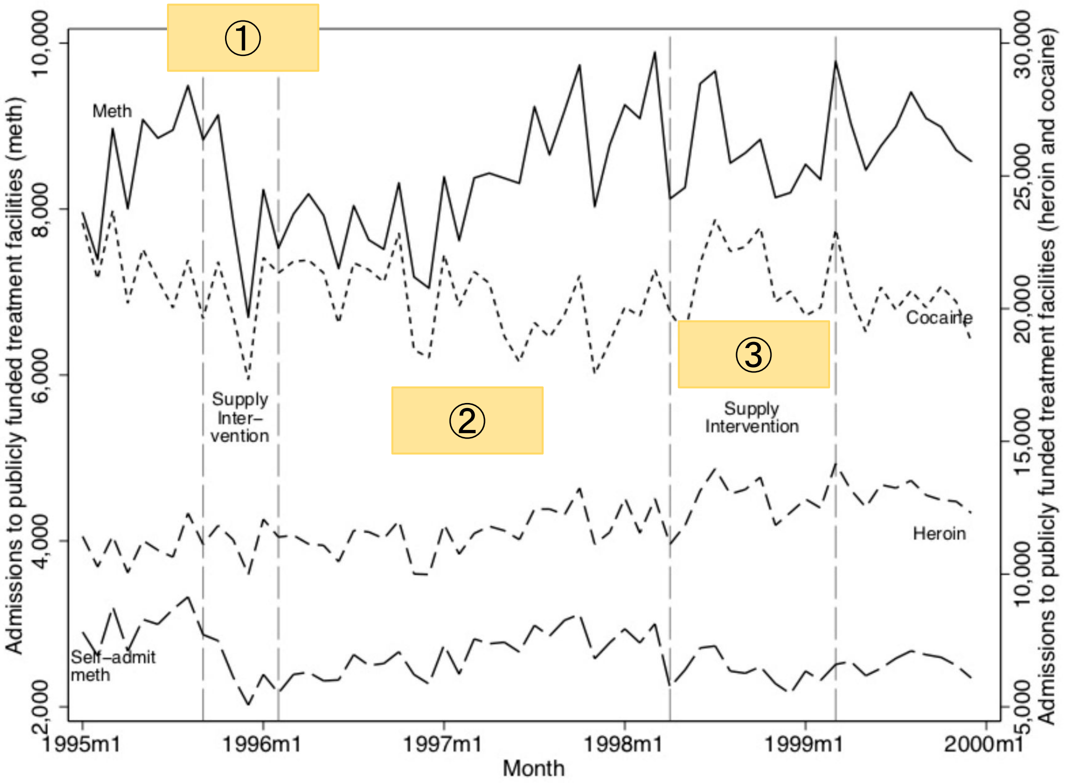


Figure 7.2: Visual representation of the equivalent of the first stage. Reprinted from Cunningham and Finlay (2012).

- ▷ エフェドリンの規制の際の効果は劇的である。
- ▷ プソイドエフェドリンのときは、わかりにくいもののメス依存による治療者数の増加は鈍化。
- ▷ ヘロインとコカインには影響がないこともわかる。

家族から引き離され里親に預けられた子どもの数の推移

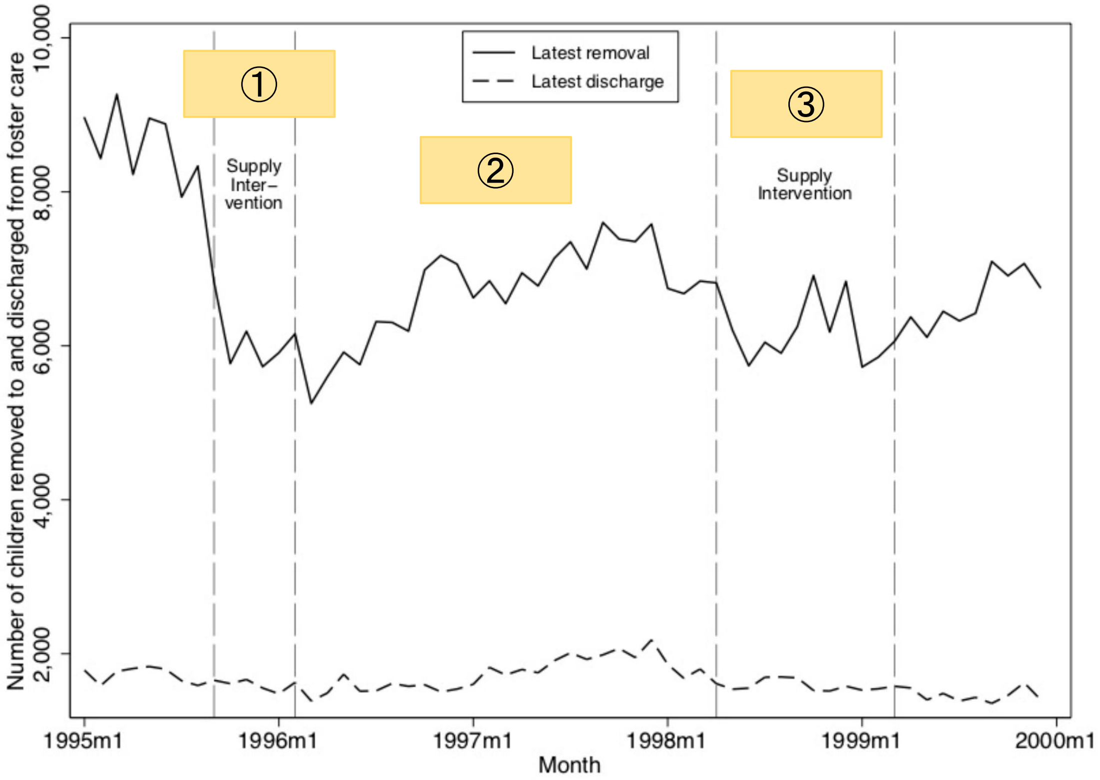


Figure 7.3: Figure 4 from Cunningham and Finlay (2012) showing reduced form effect of interventions on children removed from families and placed into foster care. Cunningham and Finlay (2012)

- 一回目の規制後は約8000人から約6000人に減少
- 二回目の介入の効果が薄い理由は
①価格への効果が一回目の半分程度だったこと
②メキシコからメス輸入が増えたこと

	Log Latest Entry into Foster Care		Log Latest Entry into Child Neglect		Log Latest Entry into Physical Abuse	
Covariates	**OLS**	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**
Log self-referred	0.001	1.54***	0.03	1.03***	0.04	1.49**
Meth treatment rate	(0.02)	(0.59)	(0.02)	(0.41)	(0.03)	(0.62)
Month-of-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State controls	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State linear time trends	Yes	Yes	Yes	Yes	Yes	Yes
First Stage Instrument						
Price deviation instrument		-0.0005^{***}		-0.0005^{***}		-0.0005^{**}
		(0.0001)		(0.0001)		(0.0001)
F-statistic for IV in first stage		17.60		17.60		17.60
N	1,343	1,343		1,343		1,343

Table 7.1: Log Latest Entry into Foster Care

▷ OLSで回帰した結果を見ると…
メス依存による治療者数の変動は、里親に預けられる子どもの数や、ネグレクト、虐待に対してほとんど影響していない。

▷ *(統計的に有意であるかを示す)がついていないので、このデータによってわかることは何もない。

	Log Latest Entry into Foster Care		Log Latest Entry into Child Neglect		Log Latest Entry into Physical Abuse	
Covariates	**OLS**	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**
Log self-referred	0.001	1.54***	0.03	1.03***	0.04	1.49**
Meth treatment rate	(0.02)	(0.59)	(0.02)	(0.41)	(0.03)	(0.62)
Month-of-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State controls	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State linear time trends	Yes	Yes	Yes	Yes	Yes	Yes
First Stage Instrument						
Price deviation instrument		-0.0005^{***}		-0.0005^{***}		-0.0005^{**}
		(0.0001)		(0.0001)		(0.0001)
F-statistic for IV in first stage		17.60		17.60		17.60
N	1,343	1,343		1,343		1,343

Table 7.1: Log Latest Entry into Foster Care

- 2SLSで回帰した結果を見ると…
- まずは first stage
  価格が一単位上がると、メス依存による治療者数は0.05%減少する。
- *統計的に有意
- F統計量が一般的に10より上であるため、強い操作変数であると言える。

	Log Latest Entry into Foster Care		Log Latest Entry into Child Neglect		Log Latest Entry into Physical Abuse	
Covariates	**OLS**	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**
Log self-referred	0.001	1.54**	0.03	1.03***	0.04	1.49**
Meth treatment rate	(0.02)	(0.59)	(0.02)	(0.41)	(0.03)	(0.62)
Month-of-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State controls	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State linear time trends	Yes	Yes	Yes	Yes	Yes	Yes
First Stage Instrument						
Price deviation instrument		-0.0005^{***}		-0.0005^{***}		-0.0005^{**}
		(0.0001)		(0.0001)		(0.0001)
F-statistic for IV in first stage		17.60		17.60		17.60
N	1,343	1,343		1,343		1,343

Table 7.1: Log Latest Entry into Foster Care

- 2SLSで回帰した結果を見ると…
  メス依存による治療者の数の変動は、里親に預けられる子どもの数や、ネグレクト、虐待に影響している。

（注）変数のどちらも対数を取っている。

7.5 The Problem of Weak Instruments

弱い操作変数が引き起こす問題

Angrist and Krueger(1991)

- 学校教育が将来の収入に与える影響を測った研究

- 米国の教育制度と義務教育法を利用
  - 教育制度:12/31以前に生まれた子供は1年生、1/1生まれは幼稚園
  - 義務教育法:16歳まで学校に通わなければならない

- 1月生まれの人は16歳の時点で12月生まれより少ない学校教育（1日しか誕生日が違わないのに！）

Angrist and Krueger(1991)

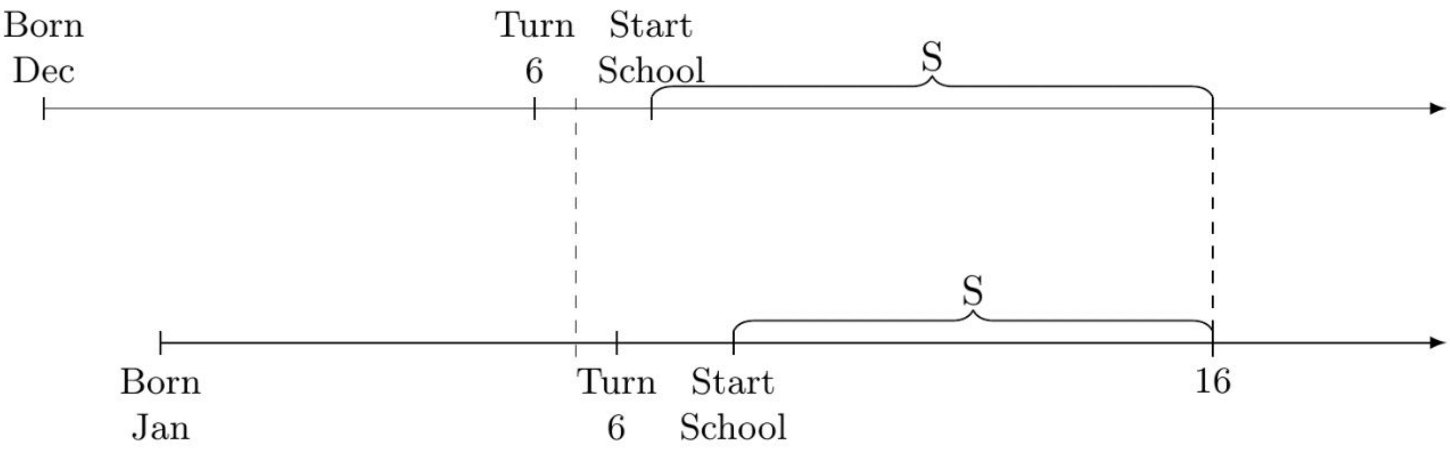


Figure 7.4: Compulsory schooling start dates by birthdates.

Angrist and Krueger(1991)

- 被説明変数Y : log(週給)
  説明変数　X: 教育年数
  操作変数　Z: 出生四半期(1-3 / 4-6 / 7-9 / 10 -12月)

- 出生四半期が収入に影響を与える直感的な理解ができないと言う点で、この操作変数は良い操作変数
  (7.2.2 Good Instruments should feel weird)

出生四半期と教育年数の関係 C(S,Z)

-

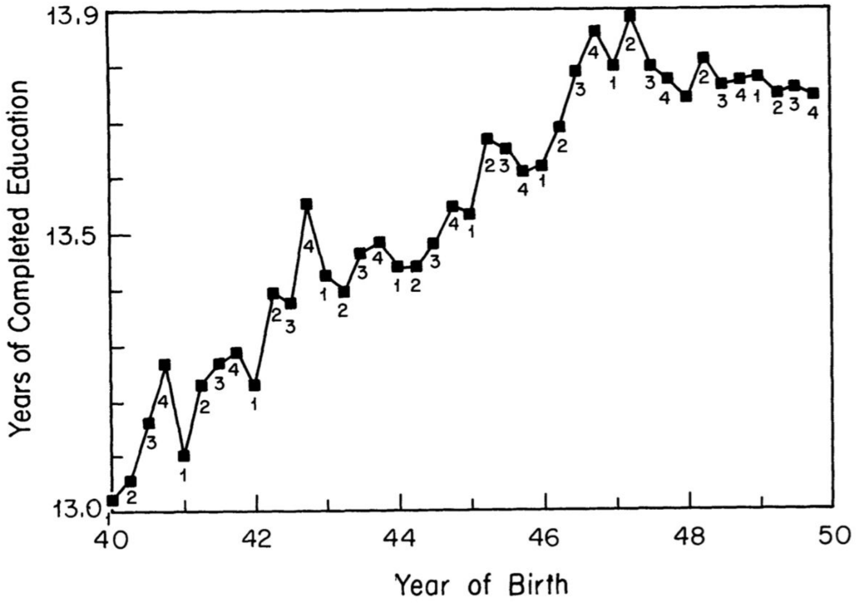


Figure 7.5: First stage relationship between quarter of birth and schooling. Reprinted from Angrist and Krueger (1991).

log(週給)と出生四半期の関係 C(Y,Z)

-

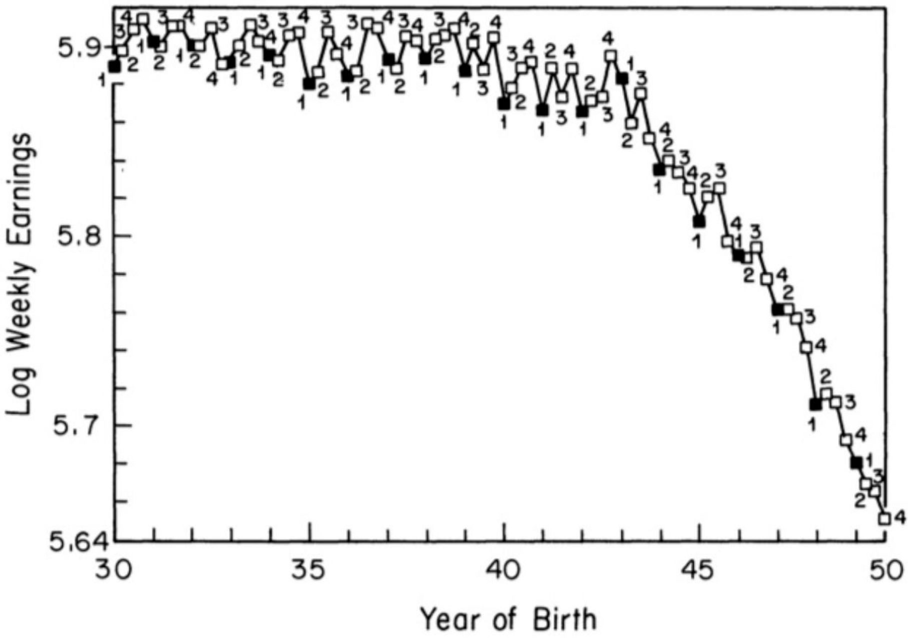


Figure 7.6: Reduced form visualization of the relationship between quarter of birth and log weekly earnings. Reprinted from Angrist and Krueger (1991).

第一段階

$$S_i = X\pi_{10} + Z_1\pi_{11} + Z_2\pi_{12} + Z_3\pi_{13} + \eta_1$$

- 教育年数を出生四半期ダミーに回帰

- 基準は第4四半期（最も長い教育年数）
  →ダミー変数の係数は負であることが予想される

第一段階の結果

Outcome variable	Birth cohort	Mean	Quarter-of-birth effect[a] I	II	III	F-test[b] [P-value]
Total years of education	1930–1939	12.79	−0.124	−0.086	−0.015	24.9
			(0.017)	(0.017)	(0.016)	[0.0001]
	1940–1949	13.56	−0.085	−0.035	−0.017	18.6
			(0.012)	(0.012)	(0.011)	[0.0001]
High school graduate	1930–1939	0.77	−0.019	−0.020	−0.004	46.4
			(0.002)	(0.002)	(0.002)	[0.0001]
	1940–1949	0.86	−0.015	−0.012	−0.002	54.4
			(0.001)	(0.001)	(0.001)	[0.0001]
Years of educ. for high school graduates	1930–1939	13.99	−0.004	0.051	0.012	5.9
			(0.014)	(0.014)	(0.014)	[0.0006]
	1940–1949	14.28	0.005	0.043	−0.003	7.8
			(0.011)	(0.011)	(0.010)	[0.0017]
College graduate	1930–1939	0.24	−0.005	0.003	0.002	5.0
			(0.002)	(0.002)	(0.002)	[0.0021]
	1940–1949	0.30	−0.003	0.004	0.000	5.0
			(0.002)	(0.002)	(0.002)	[0.0018]

第二段階の結果

Independent variable	**OLS**	**2SLS**
Years of schooling	0.0711	0.0891
	(0.0003)	(0.0161)
9 Year-of-birth rummies	Yes	Yes
8 Region-of-residence dummies	No	No

Weak Instrument（弱い操作変数）

- **弱い操作変数**とは説明変数Sと操作変数Zの相関が小さい操作変数のこと

- つまり、 $C(S, Z)$

  が十分に大きくない時の操作変数が引き起こすバイアスが弱い操作変数のバイアス

弱い操作変数のバイアス

- 考えているモデル

$$y = \beta s + \varepsilon$$

$$s = Z'\pi + \eta$$

- εとηが相関している場合、OLSのバイアスは

$$E[\hat{\beta}_{OLS} - \beta] = \frac{C(\varepsilon, S)}{V(S)} \xrightarrow{\text{置き換える}} \frac{\sigma_{\varepsilon\eta}}{\sigma_\eta^2}$$

弱い操作変数のバイアス

- 説明変数と操作変数の相関が弱い時、バイアスは

$$E[\hat{\beta}_{2SLS} - \beta] \approx \frac{\sigma_{\varepsilon\eta}}{\sigma_\eta^2}\frac{1}{F+1}$$

- 相関が弱い時、**Weak Instrument**と呼ぶ
- F統計量が小さくなるので、バイアスが大きくなる

弱い操作変数のバイアス

$F \to 0$ のとき、2SLSのバイアスは $\dfrac{\sigma_{\varepsilon\eta}}{\sigma_\eta^2}$

$F \to \infty$ のとき、2SLSのバイアスは0

弱い操作変数

Independent variable	OLS	2SLS	OLS	2SLS	OLS	2SLS
Years of schooling	0.063	0.142	0.063	0.081	0.063	0.060
	(0.000)	(0.033)	(0.000)	(0.016)	(0.000)	(0.029)
First stage F		13.5		4.8		1.6
Excluded instruments						
Quarter of birth		Yes		Yes		Yes
Quarter of birth × year of birth		No		Yes		Yes
Number of excluded instruments		3		30		28

Table 7.4: Effect of completed schooling on men's log weekly wages

Standard errors in parenthesis. First stage is quarter of birth dummies.

Independent variable	OLS	2SLS	OLS	2SLS
Years of schooling	0.063	0.083	0.063	0.081
	(0.000)	(0.009)	(0.000)	(0.011)
First stage F		2.4		1.9
Excluded instruments				
Quarter of birth		Yes		Yes
Quarter of birth × year of birth		Yes		Yes
Quarter of birth × state of birth		Yes		Yes
Number of excluded instruments		180		178

Table 7.5: Effect of completed schooling on men's log weekly wages controlling for state of birth

Standard errors in parenthesis.

➢ 相関の弱いダミーをモデルに入れることでF統計量が小さくなる

弱い操作変数の対処法

- Use just-identified model with your strongest IV
- Limited-Information Maximum Likelihood (LIML、制限情報最尤推定量) 推定量を使う
- 結局は**強い操作変数を探す**のが一番

2022.5.30

Ch.7
Instrumental Variables
(操作変数)

大湾ゼミ
椚田菜月　富澤陽仁

前回の復習

- 操作変数は内生性の問題がある時に使われる
- 処置効果が均一であると仮定すると操作変数推定量は

$$\hat{\delta} = \frac{C(Y, Z)}{C(S, Z)}$$

- 操作変数が満たすべき2つの条件は

Exclusion RestrictionsとNon-zero first stage

AGENDA

- 不均一な処置効果の場合の操作変数
- 実証例：Card(1995)
- 人気の操作変数デザイン
  - Lotteries
  - Judge and effects
  - Bartik instruments

7.6
Heterogeneous Treatment Effects
不均一な処置効果

不均一な処置効果とは

$$Y_i^1 - Y_i^0 = \delta_i$$

不均一

➢ 処置効果δiが各個人iについて異なる

$$Y_i^1 - Y_i^0 = \delta$$

均一

➢ 均一な場合は処置効果が定数δとなる

内的妥当性と外的妥当性

- **内的妥当性**:その実験集団内で因果効果が認められる
- **外的妥当性**:その因果効果が実験集団以外の一般的な集団に対しても認められる
- 均一な処置効果の場合これらの違いについて検討する必要はなかったが、不均一な場合は重要!

7.6節の焦点

(1) 操作変数推定量はどのように表されるか

(2) 推定量はどのような処置効果を捉えているか

(3) 推定量が識別できるために必要な条件

識別 : A population parameter, or set of parameters, can be consistently estimated (wooldridge, 2015).

表記の準備

- Y_i : 個人iの潜在アウトカム
- D_i : 個人iが処置を受けたかどうかの2値変数
  - 処置を受けた場合 D = 1
  - 処置を受けなかった場合 D = 0
- Z_i : 個人iに処置を受けるように促すかの2値変数
  - 処置を受けるように促す場合 Z = 1
  - 処置を受けるように促さない場合 Z = 0

表記の準備

- 潜在アウトカム YはDとZの関数

$Y_i(D_i = 0, Z_i = 1)$ を $Y_i(0, 1)$ と書く

- Potential Treatment Status

$$D_i^1 = i\text{'s treatment status when } Z_i = 1$$
$$D_i^0 = i\text{'s treatment status when } Z_i = 0$$

処置を受けるように促す

Potential Treatment Statusがとる値

Z = 1 のとき	Z = 0 のとき
D1 = 1	D0 = 1
D1 = 0	D0 = 0

- D1 = 1 は 処置を受けるよう促されて、かつ処置を受けたことを意味する
- 実際にある個人について観察できるのはこのうちの**1つ**だけ！

Observed Treatment Status

➢ 個人iについて実際に観測されたDの値

$$D_i = D_i^0 + (D_i^1 - D_i^0)Z_i$$
$$= \pi_{0i} + \pi_{1i}Z_i + \phi_i$$

Z = 1 の時 D1
Z = 0 の時 D0

$$\text{where} \quad \pi_{0i} = E[D_i^0], \pi_{1i} = (D_i^1 - D_i^0)$$

個人iについて不均一なZのDへの処置効果

Potential treatment status に基づく4つのグループ

Complier : $D_i^1 = 1 \text{ and } D_i^0 = 0$

Defier : $D_i^1 = 0 \text{ and } D_i^0 = 1$

Always-taker : $D_i^1 = 0 \text{ and } D_i^0 = 0$

Never-taker : $D_i^1 = 1 \text{ and } D_i^0 = 1$

Potential treatment status に基づく4つのグループ

- **Complier** : 処置を受けるように促されたら処置を受け、促されなかったら処置を受けない
- **Defier** : 処置を受けるように促されたら処置を受けず、促されなかったら処置を受ける
- **Never-taker** : 処置を受けるように促されてもされなくても処置を受けない
- **Always-taker** : 処置を受けるように促されてもされなくても処置を受ける

Potential treatment status に基づく4つのグループ

- Z : 抽選で当選したか（当選 = 1、落選 = 0）（当選したら基本的に入隊）
  D : 軍隊への所属（所属 = 1, 所属しない = 0）
  Y : 将来の収入

- **Complier** : 当選したら入隊し、落選したら入隊
- **Defier** : 当選したら理由をつけて入隊しないが、落選したら自ら進んで入隊
- **Never-taker** : （病気などの理由で）当選に関わらず入隊しない
- **Always-taker** : （強い愛国心などの理由で）当選に関わらず入隊

LATE（局所的平均処置効果）

$$\delta_{IV,LATE} = \frac{\text{Effect of } Z \text{ on } Y}{\text{Effect of } Z \text{ on } D}$$

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

$$= E[(Y_i^1 - Y_i^0) \mid \boxed{D_i^1 - D_i^0 = 1}]$$

D1 = 1, D0 = 0
なのでcomplier

識別に必要な5つの条件

1. SUTVA（Stable Unit Treatment Value Assumption ）
2. Independent Assumption（独立性）
3. Exclusion Restrictions（除外制約）
4. First stage
5. Monotonicity Assumption（単調性）

1. SUTVA（Stable Unit Treatment Value Assumption ）

1. 個人iのPotential outcome Yは他の個人が処置を受けた結果によって変化しない

   （ex. ある個人の頭痛の発生は他の個人が頭痛薬を飲んだという処置から影響を受けない）

2. 処置の内容は1通りである

   （満たす例：同じ内容のダイレクトメールを処置群へ送付）

   （ダメな例：デザインの異なるダイレクトメールを送付）

2. Independent Assumption（独立性）

- 操作変数ZがPotential outcome とPotential treatment statusと独立

$$\{Y_i(D_i^1, 1), Y_i(D_i^0, 0), D_i^1, D_i^0\} \perp\!\!\!\perp Z_i$$

- Zがランダムに割り当てられていることを仮定

(ex. 宝くじの当選, 裁判官の裁判への割り当て）

2. Independent Assumption（独立性）

- ➢ Reduced form：Y, Zの関係

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

$$
\delta_{IV,LATE} = \frac{\boxed{E[Y_i(D_i^1, 1) - Y_i(D_i^0, 0)]}}{E[D_i^1 - D_i^0]}
$$

2. Independent Assumption（独立性）

➢ First stage : D, Zの関係

$$E[D_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0]$$

$$= E[D_i^1 \mid Z_i = 1] - E[D_i^0 \mid Z_i = 0]$$

$$= E[D_i^1 - D_i^0]$$

$$\delta_{IV,LATE} = \frac{E[Y_i(D_i^1, 1) - Y_i(D_i^0, 0)]}{\boxed{E[D_i^1 - D_i^0]}}$$

独立性の仮定による式変形

$$
\begin{aligned}
\delta_{IV,LATE} &= \frac{\text{Effect of } Z \text{ on } Y}{\text{Effect of } Z \text{ on } D} \\
&= \frac{E[Y_i(D_i^1, 1) \mid Z_i = 1] - E[Y_i(D_i^0, 0) \mid Z_i = 0]}{E[D_i^1 \mid Z_i = 1] - E[D_i^0 \mid Z_i = 0]} \\
&= \frac{E[Y_i(D_i^1, 1) - Y_i(D_i^0, 0)]}{E[D_i^1 - D_i^0]}
\end{aligned}
$$

3. Exclusion Restrictions（除外制約）

- ZはDを通してのみYに影響を与えることを仮定
（Potential outcome YはDの関数）

$$Y_i(D_i, 0) = Y_i(D_i, 1) \quad \text{for } D = 0, 1$$

- Z = 1 だろうが Z = 0 だろうが Potential Outcomeは同じ

4. First stage

- ZはDと相関している(Dに影響を与える)ことを仮定

$$E[D_i^1 - D_i^0] \neq 0$$

- Always-taker(D1=1 & D0 = 1)と
Never-taker(D1=0&D0=0)が存在しない

$$\delta_{IV,LATE} = \frac{E[Y_i(D_i^1, 1) - Y_i(D_i^0, 0)]}{\boxed{E[D_i^1 - D_i^0]}}$$

5. Monotonicity Assumtion（単調性）

- ZはDに対して同じ方向に影響を与えることを仮定

$$D_i^1 \geq D_i^0 \text{ for all } i = 1, \ldots, N$$

- Defierが存在しないことを意味

- 理論をもとに満たされることを示すしかない

LATE（局所的平均処置効果）

$$\delta_{IV,LATE} = \frac{\text{Effect of } Z \text{ on } Y}{\text{Effect of } Z \text{ on } D}$$

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

$$= E[(Y_i^1 - Y_i^0) \mid \boxed{D_i^1 - D_i^0 = 1}]$$

D1 = 1, D0 = 0
なのでcomplier

まとめ

(1) 操作変数推定量はどのように表されるか

→ $$\delta_{IV,LATE} = E[(Y_i^1 - Y_i^0) \mid D_i^1 - D_i^0 = 1]$$

(2) 推定量はどのような処置効果を捉えているか

→ Complier についての局所平均処置効果(= LATE)

(3) 推定量が識別できるために必要な条件

→ SUTVA / 独立性 / Exclusion Restrictions / first stage / 単調性

これまでの議論は意味があるのか？

- 全ての人を対象とした平均処置効果(ATE)を見たいけれども、操作変数法では困難
- 操作変数が全ての人に影響を与えるとは考えづらいため
（7.2.1 Canonical IV DAG スライド7の男性と女性の例）
- そのため、ある集団に対する局所的平均処置効果(LATE)を見るのが現実的

7.7.1

College in the country

Card (1995)

7.7.1 College in the country

▷ データ

- 1966年～1981年
- 分析では1966年のデータを使用
- 14歳～24歳
- 男性5525人
- 自分の住む県内に大学があるか

7.7.1 College in the country

- ▷ 回帰式

$$Y_i = \alpha + \delta S_i + \gamma X_i + \varepsilon_i$$

Y : log収入
S : years of schooling (就学年数)
X : 外生共変量の行列
ε : 誤差項（未観測値・能力を含む）

7.7.1 College in the country

- ε：誤差項（未観測値・能力を含む）

  能力がschooling（学校教育）と相関していると
  すると

  $C(S, \varepsilon) \neq 0$

  より、schoolingにはバイアスがかかる

7.7.1 College in the country

▷ 操作変数
▷ 県内に大学があるかというダミー変数を導入する

・大学の遠さにかかわらず大学に行く...D1=D0=1

・大学の遠さにかかわらず大学に行かない...D1=D0=0

・県内に大学があるので大学に行く

（コンプライアー）...D1＝1、D0＝0

⇒コンプライアーに関するATE、つまりLATEを観測する

⇒コストの問題ではなく、単なる距離の問題の可能性もある

7.7.1 College in the country

Dependent variable	OLS	2SLS
educ	0.071^{***}	0.124^{**}
	(0.003)	(0.050)

First Stage Instrument		
College in the county		0.327^{***}
Robust standard error		(0.082)
F statistic for IV in first stage		15.767
N	3,003	3,003
Mean Dependent Variable	6.262	6.262
Std. Dev. Dependent Variable	0.444	0.444

Table 7.6: OLS and 2SLS regressions of Log Earnings on Schooling

Standard errors in parenthesis. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

- OLSの結果を見ると…
- 就学年数が1年増えると、収入は7.1％増加する

7.7.1 College in the country

Dependent variable	OLS	2SLS
educ	0.071***	0.124**
	(0.003)	(0.050)
First Stage Instrument		
College in the county		0.327***
Robust standard error		(0.082)
F statistic for IV in first stage		15.767
N	3,003	3,003
Mean Dependent Variable	6.262	6.262
Std. Dev. Dependent Variable	0.444	0.444

Table 7.6: OLS and 2SLS regressions of Log Earnings on Schooling

Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- 一段階を見ると...
- F統計量が15を超えているため、弱い操作変数である心配はない

- 2SLSの結果を見ると...
- 就学年数が1年増えると、収入は12.4%増加する

7.7.1 College in the country

- 2SLSの方がOLSよりも大きいのはなぜ？
  単なる能力バイアスであれば、過大評価されるためOLSの統計量の方が大きくなるはず…

理由

OLS…general population（区別なし）の教育効果を測定

2SLS…complier(D1=1, D0=0)の教育効果を測定

⇒complierの教育効果はgeneral populationの教育効果よりも高いので、係数も大きい

7.8.1
Popular IV Designs
Lotteries

7.8.1 Lotteries

- ▷ 無作為化実験の問題

  処置群に無作為に選ばれた人の実験への参加は任意であること。

⇒処置から利益を得る可能性が高い人だけが処置を受けることになる。

⇒正の選択バイアスにつながる。

7.8.1 Lotteries

- 解決策

  処置が適用されたか否か(くじに当たるか否か)を操作変数として、LATEを推定する。

7.8.1 Lotteries

▷　オレゴン州のMedicade Studies

・2000年代に入り、オレゴン州は貧困層の成人に対　する Medicade Programを手厚くした

・プログラムの受給資格の障壁はかなりひくいもの

・プログラムは大々的に宣伝された

・約5週間人々はこのプログラムにサインアップする　ことが可能

7.8.1 Lotteries

▷　オレゴン州のMedicade Studies

- サインアップした人数…約85000人
- 無作為に抽選でプログラムを適用される権利に当選…30000人
- 申請し、最終的に保険が適用された人数…10000人

7.8.1 Lotteries

▷ オレゴン州のMedicade Studies

操作変数	抽選結果	D1…当選した人	D0当選しなかった人
説明変数	保険加入	D1＝1…当選かつ加入	D1＝0…当選したが加入していない D0＝0…当選しない
被説明変数	研究による		

7.8.2
Popular IV Designs
Judge fixed effects

7.8.2 Judge fixed effects

- ▷ 裁判官の判決には系統的な違いがある
  Gaudet、Harris、John（1933）

・被告人は裁判官に対して無作為に割り当てられる
・割り当てられた被告人の特性は平均すると同じくら　いになると想定

7.8.2 Judge fixed effects

▷ 裁判官の判決には系統的な違いがある
Gaudet、Harris、John（1933）

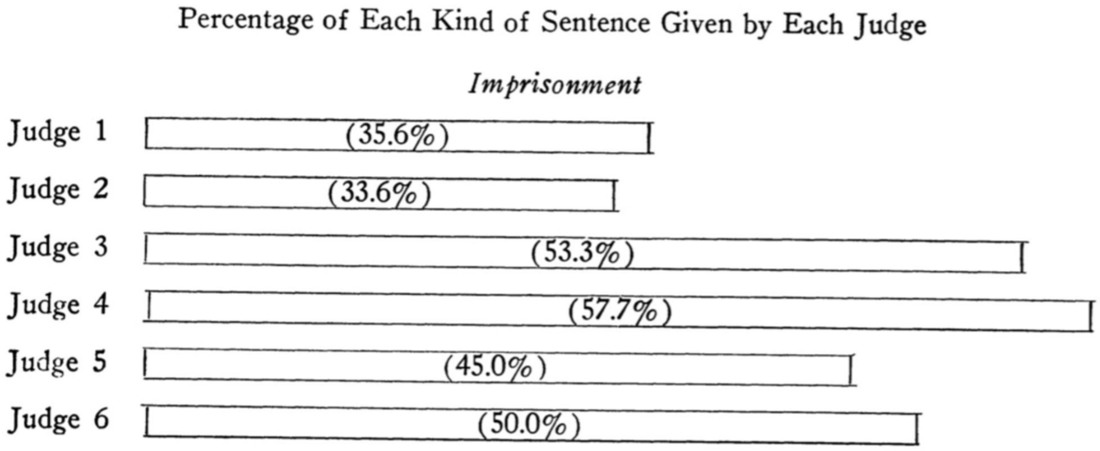


Figure 7.8: Variation in judge sentencing outcomes. This figure originally appeared in Gaudet, Harris, and John (1933). Reprinted by special permission of Northwestern University Pritzker School of Law, *Journal of Criminal Law and Criminology*

▷ 裁判官によって禁固刑に処す割合が異なる

⇒量刑の違いは、被告人の罪の重さだけではなく、裁判官の厳しさに依存する

7.8.2 Judge fixed effects

▷ モデル具体例

$$Y_i = \alpha + \beta_1 S_i + \beta_2 W_i + \varepsilon_i$$

Y : 収入
S : 刑期
W : その他のコントロール変数
ε : 誤差項（観測できないやる気等を含む）

刑期と誤差項は相関する
⇒内生性の問題

7.8.2 Judge fixed effects

▷ 操作変数：裁判官ダミーを導入

⇒刑期には影響を与えるが、誤差項とは相関しない
⇒裁判官は被告人に対して無作為に割り当てられるから

7.8.2 Judge fixed effects

▷ 重要ポイント

・裁判官は無作為に割り当てられる
⇒内生性の問題を含む誤差項と無相関になる

・個々の裁判官は厳しさの傾向を持つ
⇒説明変数（刑期）との相関を持つ

7.8.2 Judge fixed effects

- モデルが満たすべき5つの条件

1. SUTVA
2. independence assumption（独立性）
3. exclusion restrictions（除外制約）
4. first stage
5. monotonicity assumption（単調性）

7.8.1 Judge fixed effects

- モデルが満たすべき5つの条件

1. SUTVA
2. **independence assumption（独立性）**
3. **exclusion restrictions（除外制約）**
4. first stage
5. **monotonicity assumption（単調性）**

7.8.2 Judge fixed effects

- ▷ independence assumption（独立性）

無作為に裁判官は、割り当てられているので満たしている？

⇒裁判官の特性に応じて、被告人が戦略的な行動をとる可能性

⇒独立性を損なう

7.8.2 Judge fixed effects

- exclusion restrictions（除外制約）
  Dobbie, Goldin, and Yang (2018)

操作変数	裁判官の厳しさ
説明変数	公判前勾留期間
被説明変数	有罪率

・公判前勾留期間を決定する裁判官は、裁判を行う裁判官とやり取りを　することがない。
・被告人自身とやり取りすることもな　い。
⇒条件を満たしている…？

7.8.2 Judge fixed effects

- exclusion restrictions（除外制約）

公判前の勾留期間を決める際に厳しい裁判官が割り当てられると、実際の裁判で罰を決める裁判官はその裁判官とは異なるが、被告人自身の期待刑が重罰になる。期待刑よりも罰が軽ければ、その罰を受け入れる。

⇒公判前の勾留期間を決める裁判官の厳しさが有罪率に影響している

⇒条件を満たさない

7.8.2 Judge fixed effects

- monotonicity assumption（単調性）

人間の思考と経験によってバイアスがかかる

e.g.

ある寛大な裁判官は被告人が黒人である場合や犯罪が薬物犯罪である場合、厳しい判決を下す傾向にある。

この場合、厳しい裁判官よりも寛大な裁判官の方が重罰を科す可能性がある。

7.8.2 Judge fixed effects

- Stevenson (2018)
  現金保釈が事件の結果にどのような影響を与えるかを推定

・大都市の行政データを使用

⇒IVの有限サンプルバイアスを改善売るのに有効な大規模サンプルである

⇒データの多くは多くの場合オンラインで公開されているた　め、データを集めやすい

7.8.2 Judge fixed effects

▷ Stevenson (2018)

- 保釈金の額を決める保釈判事が無作為に割り当てられる
- 判事によって額設定の傾向には大きな差がある

操作変数	保釈判事ダミー
説明変数	公判前勾留期間
被説明変数	有罪確率

7.8.2 Judge fixed effects

▷ Stevenson (2018)

▷ 結果

・勾留期間が伸びると、有罪判決を受ける確率は13%増加する

・勾留期間が長いと、収監期間は42%増加し、保釈金以外にか　　かる費用は41%増加する

⇒保釈金を払えず勾留されることは、有罪率の上昇・裁判費用　　の上昇・再犯率の上昇につながり、負の連鎖を招く結果とな　　る

7.8.3
Popular IV Designs

Bartik instruments

Bartik instrumentsとは

- Bartik(1991)によって一般的になった手法
- 移民や貿易などのショックがある地域の雇用率・賃金などに与える影響を測る際に使われることが多い
- 移民・貿易・労働などの分野で幅広く使われる
- Shift-Share instrumentとも呼ばれる

例. 移民の流入が賃金に与える影響

➢ ある地域への移民の流入がその地域の賃金に与える影響を測りたい

$$Y_{l,t} = \alpha + \delta I_{l,t} + \rho X_{l,t} + \varepsilon_{l,t}$$

地域lのt期におけるlog(賃金)

地域lのt期における移民の流入数

コントロール変数と時間固定効果

➢ 移民の流入数は内生変数(Sharpe (2019))

例. 移民の流入が賃金に与える影響

➢ 移民の流入数は内生変数
-> 2SLSを使う-> 操作変数としてBartik instrument

$$B_{l,t} = \sum_{k=1}^{K} z_{l,k,t^0} m_{k,t}$$

Share 変数

Shift 変数

地域lへのt0期からt期の移民流入数の予測値

t0期における地域 lの移民流入数が国全体のk国からの移民流入数に占める割合

k国から当該国へのt期までの移民流入数

Bartik instrumentsとは

- ある地域のt0期からt期における変動
  ＝その地域のt0期におけるシェア x
  　国全体のt0期からt期における変動（シフト）
  として操作変数化

- t0期からt期の間の国全体の変動の影響はt0期におけるその地域のシェアに依存

7.9
Conclusion

まとめ

- ➢ 操作変数法は内生性の問題がある時に用いられる手法
- ➢ 良い操作変数を見つけることができれば非常に強力な一方で、操作変数が満たすべき条件は多い
  （処置効果が不均一な場合は5つ！）
- ➢ 良い操作変数を見つけられるかはどれだけその分野に親しんでいるか