

Chapter 3

-Instrumental Variables-

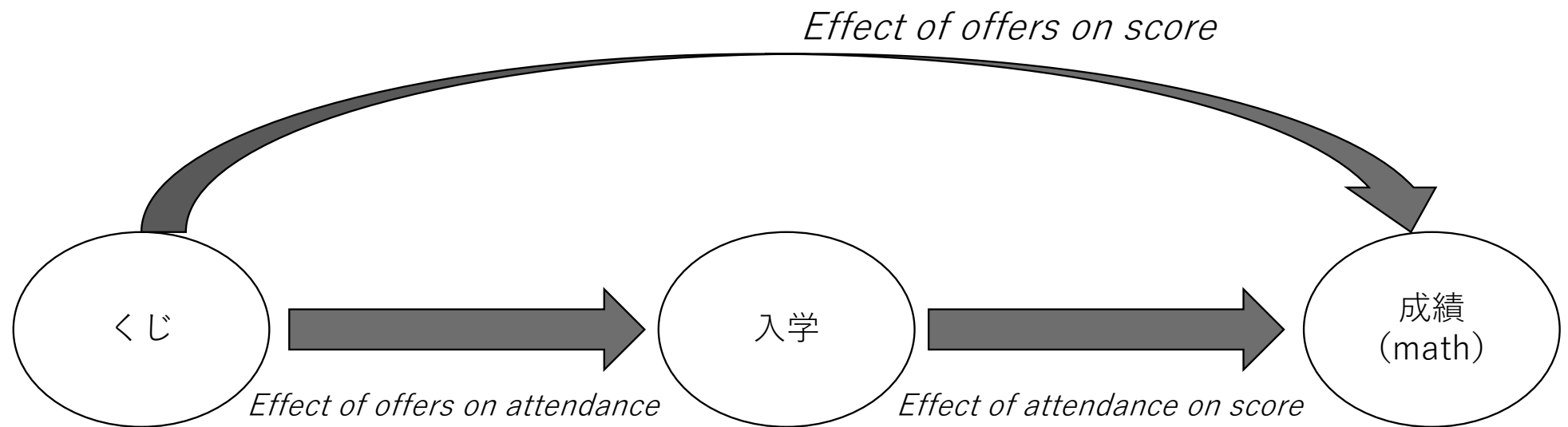
pp.109-123

(今回の内容)

- 操作変数方の事例①：K I P P への入学がもたらす
教育効果の測定（前回の続き）
- 操作変数方の事例②：D V に対する警察の最適措置に関する実験
(M D V E)

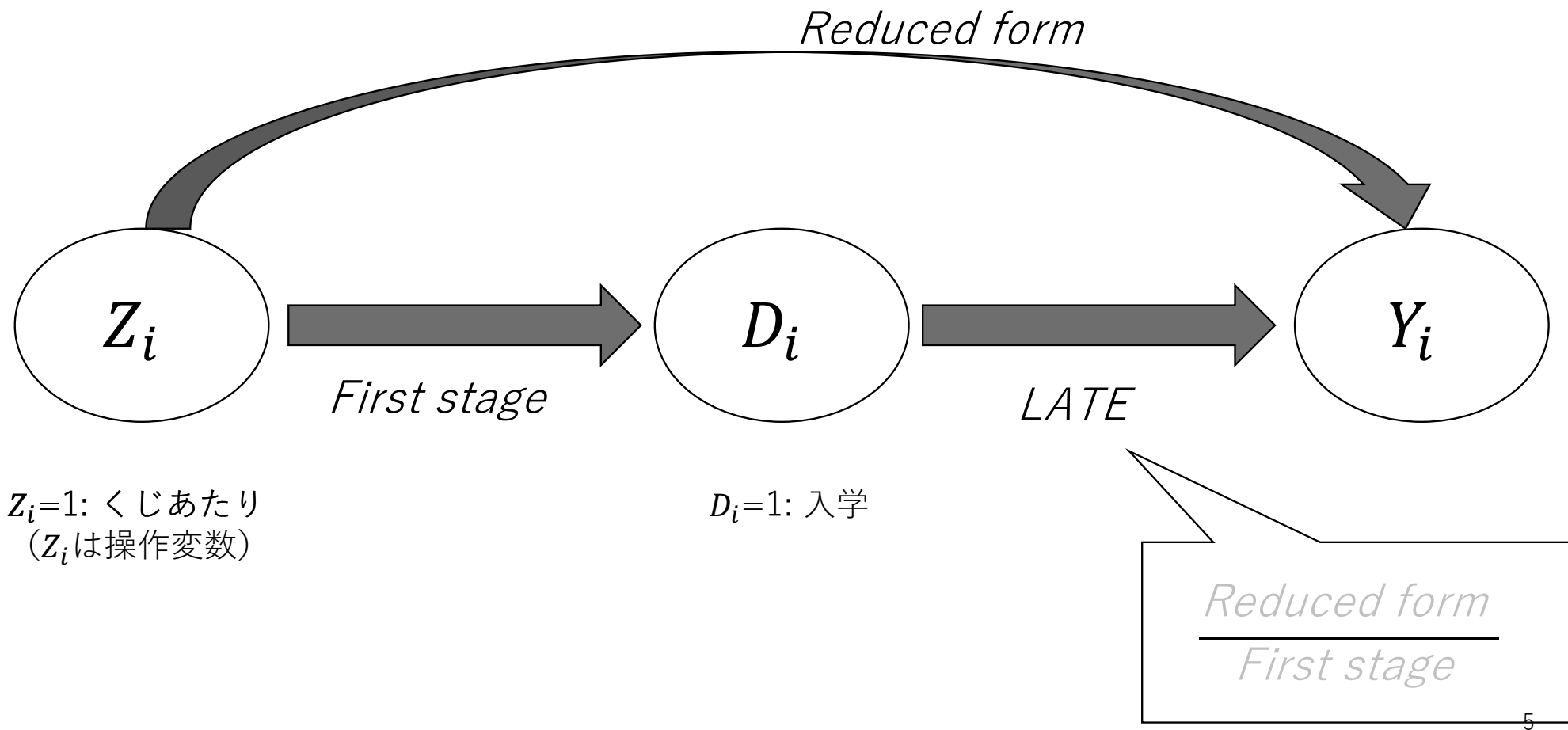
操作変数方の事例①：K I P P（続き）

操作変数方の概念的的理解（前回の内容）



前回はここまで→今回はこれの理論的なモデル化からスタートする。

操作変数方のモデル化



操作変数方のモデル化

以下のようにパラメータを設定する

- First stage(ϕ)

$$\phi = E\{D_i|Z_i = 1\} - E\{D_i|Z_i = 0\}$$

- Reduced form(ρ)

$$\rho = E\{Y_i|Z_i = 1\} - E\{Y_i|Z_i = 0\}$$

- LATE(λ)

$$\lambda = \frac{\rho}{\phi} = \frac{E\{Y_i|Z_i = 1\} - E\{Y_i|Z_i = 0\}}{E\{D_i|Z_i = 1\} - E\{D_i|Z_i = 0\}}$$

実例との対応関係を見てみる

前回のデータを用いた実際の算出例と先ほどのモデルを対応させる

- First stage(ϕ)

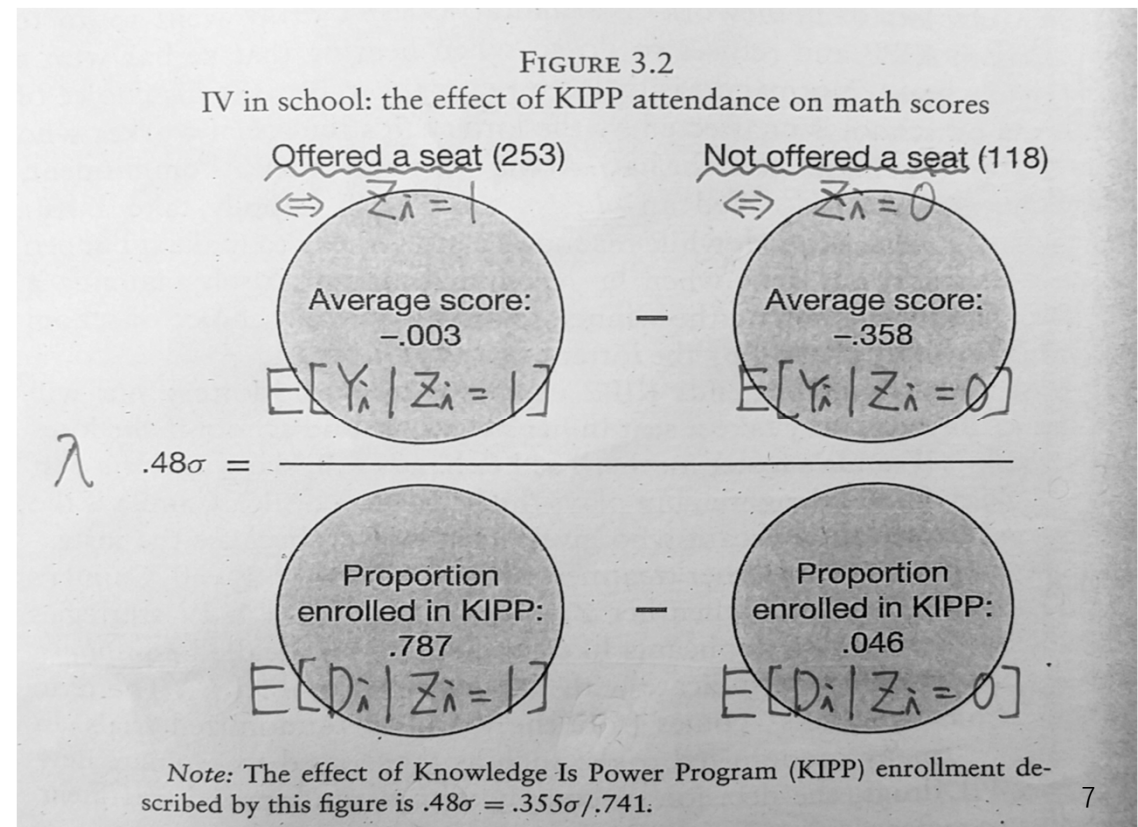
$$\phi = E\{D_i|Z_i = 1\} - E\{D_i|Z_i = 0\}$$

- Reduced form(ρ)

$$\rho = E\{Y_i|Z_i = 1\} - E\{Y_i|Z_i = 0\}$$

- LATE(λ)

$$\lambda = \frac{\rho}{\phi} = \frac{E\{Y_i|Z_i = 1\} - E\{Y_i|Z_i = 0\}}{E\{D_i|Z_i = 1\} - E\{D_i|Z_i = 0\}}$$



操作変数方の解釈

先ほどの図からも分かるように、
いま調査対象である局所的な平均介入効果（L A T E）を推定するには、
4つの期待値が分かればよい。（=方法1）

ただ、実際には、L A T Eの導出方法として、
「2 S L S（=二段階最小二乗法）」がよく使われている。（=方法2）
（詳細は、次週3.3節にて）

なお、どちらの方法にせよ、考え方やメソッドは同じ。

L A T E の解釈

L A T E は一体誰の効果を表しているの？

↑

(背景)

そもそも子どもによってK I P P入学による効果の大きさは異なってくると考えられる。

Ex)親が教育熱心な家庭で育った子は、元々勉学に励んでいるはずだから、

K I P Pに入学することが成績に及ぼす因果効果は0 (ないしは微少)。など

しかしL A T Eの算出時には、こうした子ども別の便益の違いが考慮されていない

⇒ L A T E を解釈するには、まず子どもをカテゴライズする必要あり。

子どもの分類

以下の4タイプに分類できることが知られている。

- 何が何でもK I P Pに入学しようとする（させられる）子〔**always-takers**〕
→万が一くじで外れても、教育熱心な母などによって何とか入学にこぎ着けるケース
- 何が何でもK I P Pへの入学を拒む子〔**never-takers**〕
→勉学にあまり意欲的でない(?)ので、長時間授業かつ膨大な宿題を課すK I P Pを拒むケース
- くじで当たったら入るし、外れたときは素直に入学を諦める子〔**compliers**〕
→外れたとしても与えられた環境で頑張るなどといったケース
- くじの結果を受け入れようとしない、ひねくれ者〔**defiers**〕
※分類上設けられたカテゴリーであり、実際そういう子が存在するかどうかは疑問。

子どもの分類

先ほどの4分類は、ダミー変数 Z , D を使うことで以下の表のようにまとめることが出来る

TABLE 3.2
The four types of children

		Lottery losers $Z_i = 0$	
		Doesn't attend KIPP $D_i = 0$	Attends KIPP $D_i = 1$
Lottery winners $Z_i = 1$	Doesn't attend KIPP $D_i = 0$	Never-takers (<i>Normando</i>)	Defiers
	Attends KIPP $D_i = 1$	Compliers (<i>Camila</i>)	Always-takers (<i>Alvaro</i>)

Note: KIPP = Knowledge Is Power Program.

<結果>

操作変数 Z_i が行動の変化をもたらすのは
実際にはCompliersのみ！

(Defiersの存在は非現実的なので不問)

↓

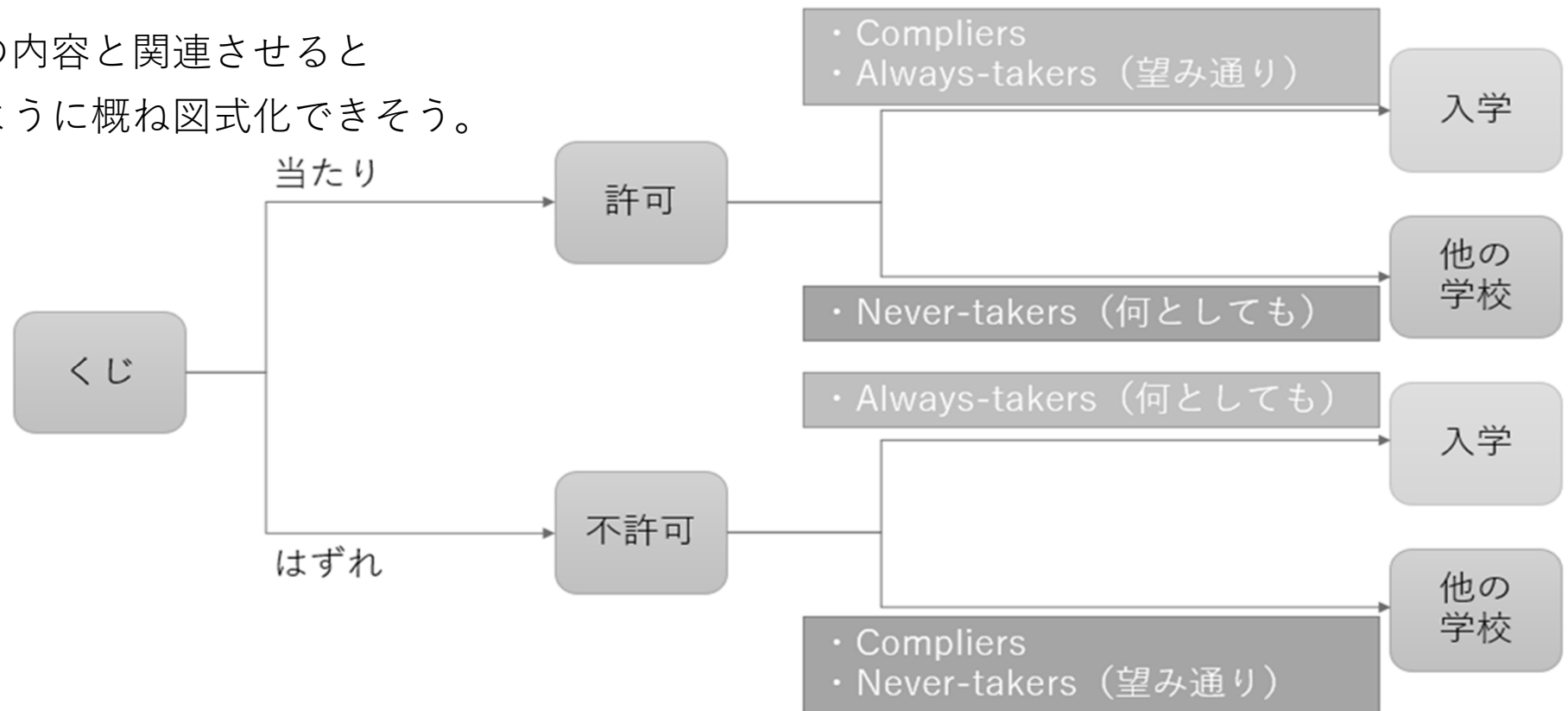
操作変数方におけるLATEは
Compliersの平均的な因果効果を測定して
いる

⇔逆に言うと

操作変数方では、Always-takersやNever-takersに対しては有効な分析を行えない。

子どもの分類（参考）

前回の内容と関連させると
このように概ね図式化できそう。



※子どものタイプとくじ以外の要因（引っ越し、近くの学校がいい、など）は単純化のため無視している。

備考：Defiersについて

理論上、操作変数はCompliersだけでなくDefiersの行動も変えうるので、L A T E にはDefiersの効果も含む可能性がある

<もし仮にDefiersが存在したら・・・>

DefiersはCompliersと対称的な存在なので、仮に全員にK I P Pの入学による効果があったとしても、2群の効果が相殺されてL A T Eの値が0になる可能性だってある

⇔Defiersが存在すると、操作変数方の正しい解釈が困難に

しかし、それでもやはりDefiersの存在は滅多にないので、分析に際してDefiersは存在しないと仮定する。（この仮定を“Monotonicity”と呼ぶ）

続) L A T E の解釈

一旦ここまでの議論を整理する...

- ・操作変数方では、操作変数 (Z_i) が原因変数 (D_i) を変化させ、それが転じて結果変数 (Y_i) に影響を及ぼすという一連のつながりについて考える
- ・以上の観点より、操作変数方の主な分析対象となるのは、操作変数の変化が原因変数の変化をもたらすケース (すなわちCompliers)

するとL A T Eは次のようにも表せる→

L A T E の解釈～結果～

<仮定>

- ・ I V はランダムに割り当てられ、なおかつ割り当ての結果参加者のうち最低限1人以上は介入を受ける。
- ・ Defiersは存在しないものとする (Monotonicity)
- ・ I V が結果変数に影響を及ぼすルートはただ1つのみ (Exclusive Restriction)
- ・ C_i : Compliersであるかどうかを示すダミー変数
($C_i = 1$: その者はCompliers)

$$\Rightarrow \lambda = \frac{\rho}{\varphi} = E\{Y_{1i} - Y_{0i} | C_i = 1\}$$

(局所的な平均介入効果) = (Compliersの介入効果の期待値)

(補足)第1章参考

Y_{1i} : 介入時のiさんの結果
 Y_{0i} : 非介入時のiさんの結果

操作変数方が示唆する結論と新たな課題

<ランダム化比較試験の場合>

介入効果を測定するために「定数効果仮定」という比較的強い仮定を置いて話を進めていた。（第1章参考）

<操作変数方>

「定数効果仮定」を置かずとも、分析が可能

⇒ 今回の結果： $\lambda = 0.48$ ⇒ Compliersへの介入効果が正

<推論>

少なくとも、Compliersのより多い地域でKIPPの受け入れ枠を増やすと教育効果を高めることができるということが推測できる。

↓

実際に、どの地域に、どのくらい定員数を増やすかは政策立案上の課題。

TOT ~ もう一つの指標 ~

介入を受けたすべての生徒 (Complier + Always-takers) を対象とする指標もよく用いられる

$$\text{TOT (Treatment Effect on the Treated)} = E\{Y_{1i} - Y_{0i} | D_i = 1\}$$

・ここで、くじに当選してKIPPに入学したCompliersは全Compliers(落選した人なども含む)を代表する

・入学者のうち、Compliersの効果とAlways-takersの効果は一致する必要はない。

⇒LATEとTOTは必ずしも一致しない。

(Compliersの効果 > Always-takersの効果)

先述のとおり、教育熱心な家庭のケース。など

(Compliersの効果 < Always-takersの効果)

KIPPへの入学がその後の人生に大きな影響を与えた。など

外的妥当性の問題

L A T E を考えていく上で、表題の問題が現れる。

外的妥当性：ある因果効果の推定実験で得られた推定結果が、サンプルが変わったとしても（時代、土地、国、人の性質などの要素が変わっても）同様に推定されること

↑

実際にはある特定の場合、L A T E の推定値が大きい（小さい）ことがあり得る。

⇒操作変数をうまく利用することで問題の分析ないし解消が可能かもしれない。

最もよいL A T E の外的妥当性の証明方法は

“全く同様の実験”を“異なる分布を持つ様々なサンプル”で行い、

そこから得られた複数のL A T E 推定値を比較検討することである。

操作変数方の事例②：M D V E

アメリカの有名な D V の事例の紹介

「O.J.シン普森事件」

(概要)

O.J.シン普森…アメリカで非常に有名なアメフト選手だが、結婚後に D V の疑いで度々逮捕はされなかったものの、よく警察の世話になっていた

1989年： O.J.シン普森初めて D V の罪で逮捕される。



罰則 = 少額の罰金、社会奉仕活動への参加、カウンセリング (≠ 拘留、懲役)

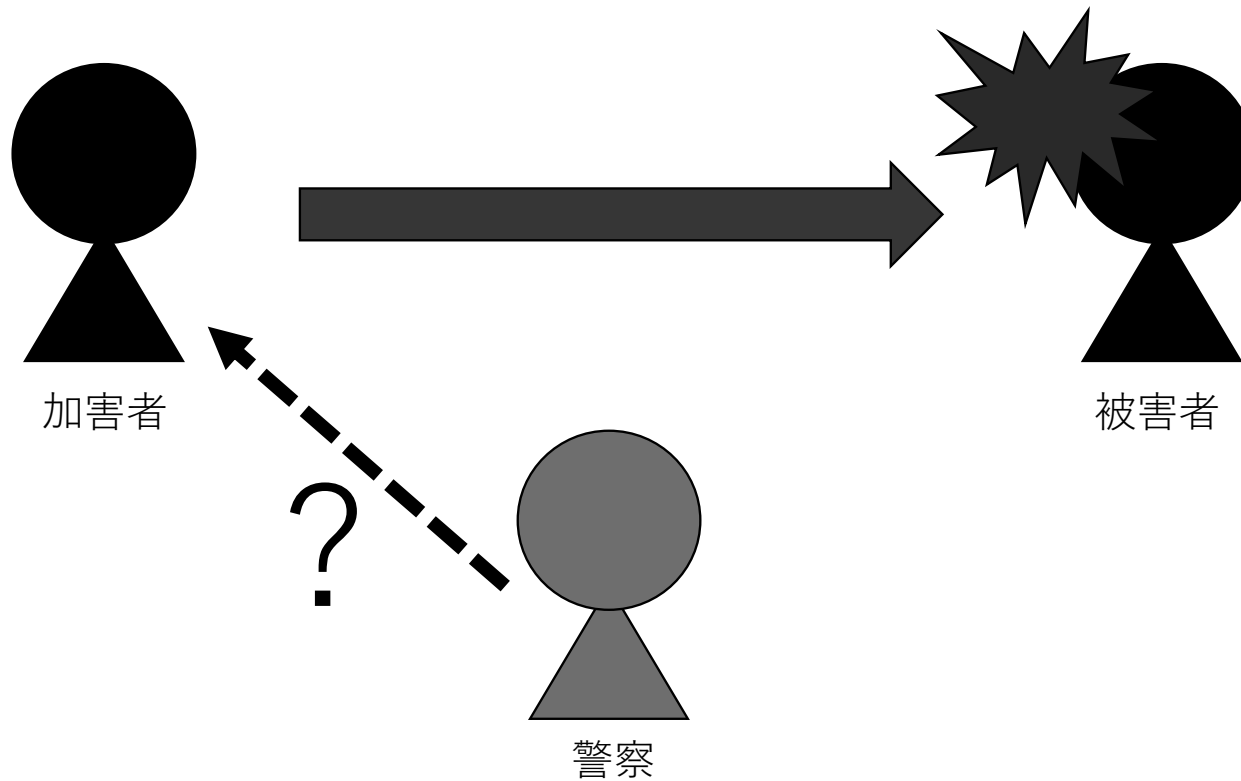
1994年：元妻とその友人が何者かによって殺害される事件が発生

O.J.シン普森が容疑者として逮捕された。(⇔ O.J.の再犯…?)

(考察)

D V の初犯時に下した“軽い罰則”はO.J.の再犯を未然に防ぐに至らなかったのでは？

Q：警察はDVに対して
どのような措置をとるのが最適なのか??



考えられる措置とそれに対する批判

○多くの警察や第三者の意見

「福祉機関」の設置による再犯防止策が最適措置だ。

↑

被害者側の批判：軽い罰則では不満

○被害者側の意見

「重い罰則」を行使することが最適だ。

↑

警察・第三者側の意見：ただ逮捕するだけでは無意味、再犯の可能性大

⇒両者の意見は対立・・・一体どちらの意見がより正しい??

D V への対応措置に関する実験

先ほど述べた議論にエビデンスを与え得る実験が行われた。

「M D V E (The Minneapolis Domestic Violence Experiment)」

1980年代初頭、

アメリカのミネアポリスにて市が警察と連携して行われた
D V に関する実験。

→ここからはM D V E の実験内容を見ていく。

M D V E の実験方法

<対象>

被害者と加害者の両者が生存している（⇔生命が脅かされたケースを除いた）D V 事件

<加害者への措置>

①Arrest：逮捕

②Advice：カウンセリングや警官による調停

③Separation：8時間、容疑者を被害者のもとから引き離す

※先ほどの議論と照合すると、①は“重い罰則”、②③は社会機関を通じた“弱い罰則”

⇒②③をまとめて“Coddled”とする。

<結果の測定基準>

措置後6ヶ月以内に再犯を起こすか否か。常習性を測定。

M D V E の実験方法

- 1, 警官がD Vの現場に駆けつける
- 2, 警官がもつ端末にある3色のうちからランダムに選ばれた1色が表示される。
※各色はそれぞれ各措置に対応
(Ex:赤色が表示される⇒①Arrest ; 容疑者を逮捕する)
- 3, 警官は端末が表示した色に対応する措置をとる

⇒ランダム化比較試験に極めて近いのでは・・・！？

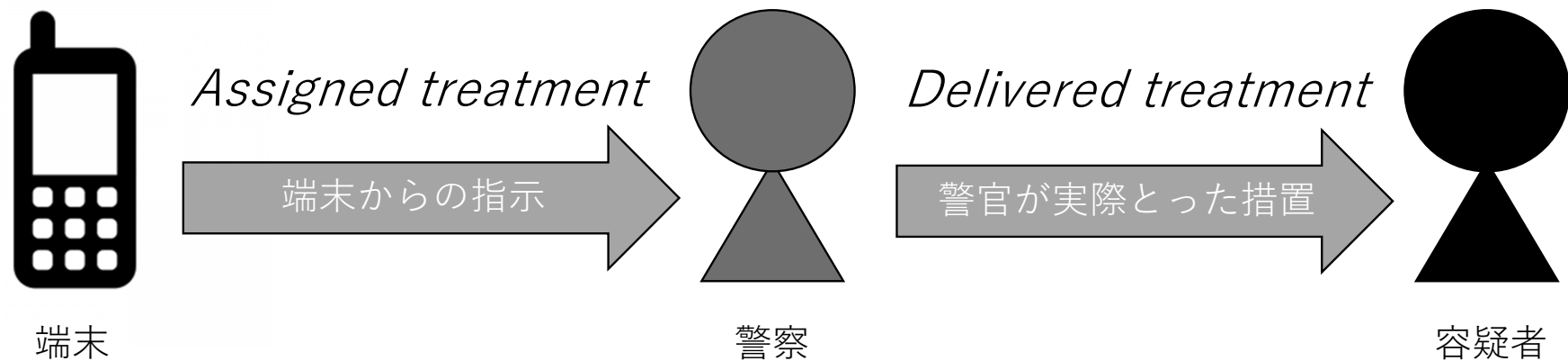
しかし、今回もK I P Pと同様に実際にはランダム化の不完全性が生じる

ランダム化の不完全性

以下の要因のせいで、完全なランダム化は達成されなかった

- (1) 警官が端末の指示に逆らって容疑者を逮捕した
 - －公務執行妨害
 - －被疑者が容疑者の逮捕を請求した場合
 - －D V の被疑者と容疑者が両方負傷している場合
- (2) 警官が実験のことを忘れて普通に業務をこなしてしまった
 - ⇒単なる人為的ミス

ランダム化の不完全性



<警察官が端末の指示に従っていたら>

Assigned treatment = *Delivered treatment*

<逆に指示から逸脱していたら>

Assigned treatment ≠ *Delivered treatment*

ランダム化の不完全性～実際のデータ

TABLE 3.3
Assigned and delivered treatments in the MDVE

Assigned treatment	Delivered treatment			Total
	Arrest	Coddled		
		Advise	Separate	
Arrest	98.9 (91)	0.0 (0)	1.1 (1)	29.3 (92)
Advise	17.6 (19)	77.8 (84)	4.6 (5)	34.4 (108)
Separate	22.8 (26)	4.4 (5)	72.8 (83)	36.3 (114)
Total	43.4 (136)	28.3 (89)	28.3 (89)	100.0 (314)

Notes: This table shows percentages and counts for the distribution of assigned and delivered treatments in the Minneapolis Domestic Violence Experiment (MDVE). The first three columns show row percentages. The last column reports column percentages. The number of cases appears in parentheses.

青丸

⇒警察が端末の指示通り動いたケース

赤丸

⇒警察が端末の指示から逸脱したケース

ランダム化してるので、
各項目が33.3%になる
のが理想的

データより・・・
概ね端末の指示通りの行動をとっているが、Coddledの指示を受けたのに実際にはArrestしたケースがかなりある。

不完全性を乗り越えるために

ランダム化が不完全なので、警察官がとった措置を基にして結果を分析しよう
とすると本来の結果とは異なるものが導き出されてしまう。
⇒ランダム化比較試験を想定したMDVEは失敗に終わる…？

Nonrandom!!
警察官が容疑者に対してとる措置



容疑者が6ヶ月以内に再犯するか否か



ここで操作変数を導入して操作変数法で分析すれば、この実験から調べたい介入効果が測定できそう。

⇒ここからいよいよ、操作変数法を用いてMDVEの効果进行分析していく

M D V E への操作変数方の導入

目標：セレクション・バイアスの排除

(ここまでの議論)

今回、ランダム化の完全性が成立しなかった主要な原因は

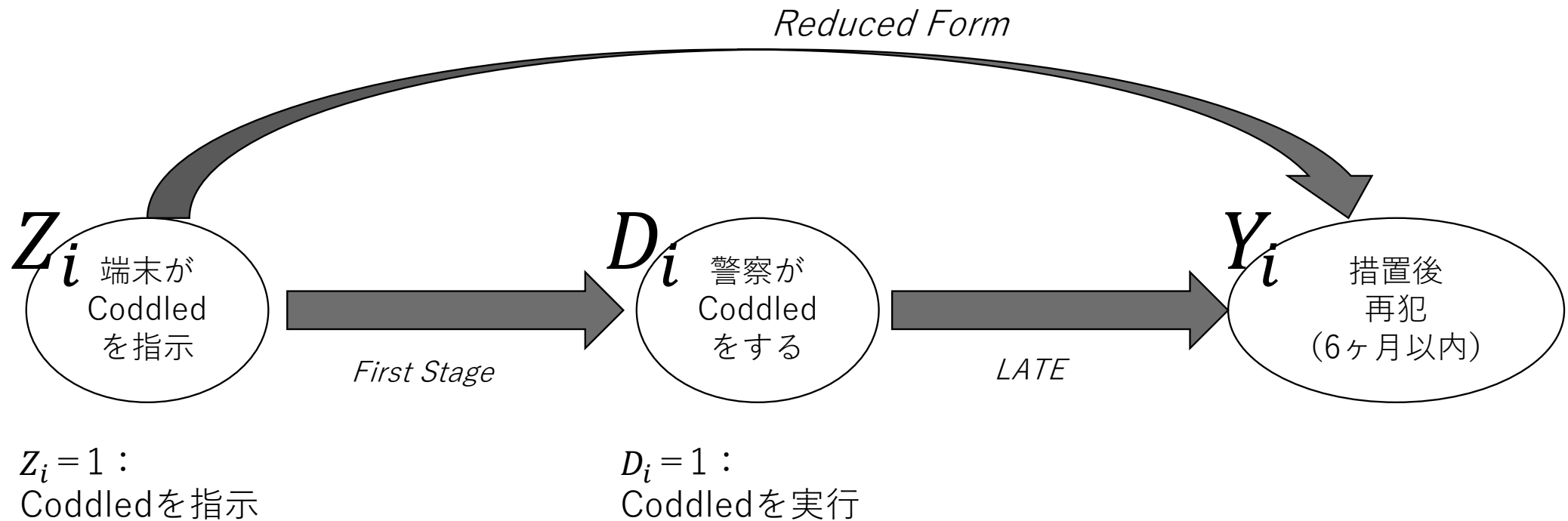
(1) 警官が端末の指示に逆らって容疑者を逮捕した。だった。

↑

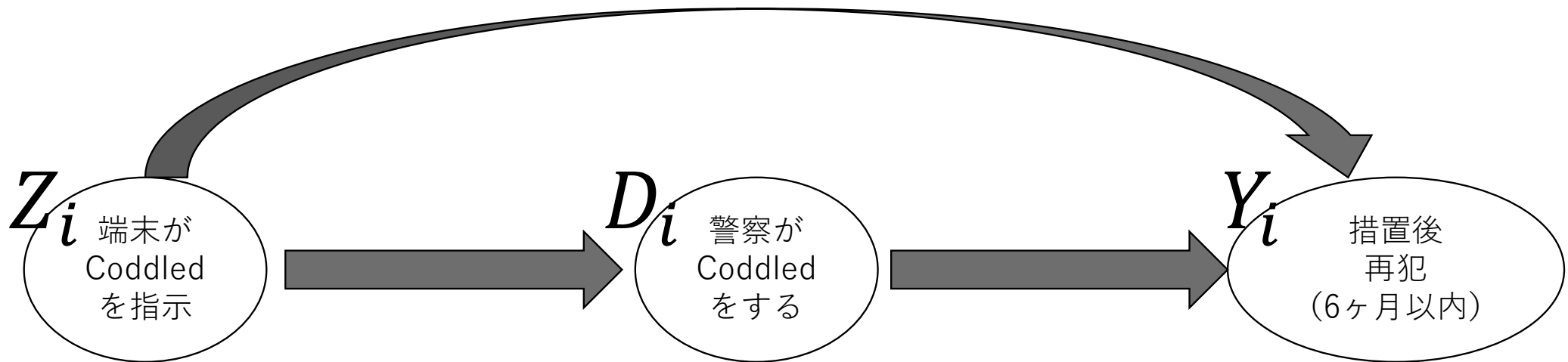
このケースに当てはまる容疑者は、比較的攻撃的・暴力的な性質が強かったからこそ、出動した警察官が自ら逮捕の判断を下したと考えられる。

⇒ここにセレクション・バイアスが存在している！

M D V E を操作変数法に当てはめる



備考：操作変数の3条件を満たすか？



- (1) 端末の指示は、警察の行動に影響を及ぼしている。
- (2) 端末の指示はランダムであり、なおかつ容疑者側は実験のことを何も知らない。
⇒ Z が Y に直接影響を及ぼすことはない。
- (3) 端末の指示は、現場の警察の行動にのみ影響を及ぼしている。
⇒ Z が Y に影響を及ぼすルートはただ1つのみ。

⇒以上より、今回のケースは「操作変数の3条件」を（おそらく）満たしていると思われる。

各パラメータを機械的に算出してみる

- First stage(ϕ) ←

$$\begin{aligned}\phi &= E\{D_i|Z_i = 1\} - E\{D_i|Z_i = 0\} \\ &= 0.797 - 0.011 = 0.786\end{aligned}$$

- Reduced form(ρ) ←

$$\begin{aligned}\rho &= E\{Y_i|Z_i = 1\} - E\{Y_i|Z_i = 0\} \\ &= 0.211 - 0.097 = 0.114 (\text{約}11\%) \end{aligned}$$

- LATE(λ)

$$\begin{aligned}\lambda &= \frac{\rho}{\phi} = \frac{E\{Y_i|Z_i = 1\} - E\{Y_i|Z_i = 0\}}{E\{D_i|Z_i = 1\} - E\{D_i|Z_i = 0\}} \\ &= 0.114 / 0.786 = 0.145 (\text{約}15\%) \end{aligned}$$

TABLE 3.3
Assigned and delivered treatments in the MDVE

Assigned treatment	Delivered treatment			Total
	Arrest	Coddled		
		Advise	Separate	
Arrest	98.9 (91)	0.0 (0)	1.1 (1)	29.3 (92)
Advise	17.6 (19)	77.8 (84)	4.6 (5)	34.4 (108)
Separate	22.8 (26)	4.4 (5)	72.8 (83)	36.3 (114)
Total	43.4 (136)	28.3 (89)	28.3 (89)	100.0 (314)

Notes: This table shows percentages and counts for the distribution of assigned and delivered treatments in the Minneapolis Domestic Violence Experiment (MDVE). The first three columns show row percentages. The last column reports column percentages. The number of cases appears in parentheses.

<再犯率データ>

- 全サンプルにおける再犯率 = 0.18
- Coddled指示対象の再犯率 = 0.211
- Arrest指示対象の再犯率 = 0.097

導出結果への評価・考察

- ・ランダム化の不完全性を考慮せずに導出した再犯率はおよそ11%だった。

↑

全サンプルでの再犯率は18%なので、11%という数字は妥当な値と言えそう。

このように導出した効果を ITT効果(intention-to-treat effect) という。

すなわち $ITT = \text{Reduced Form}$ と言える。

- ・操作変数方によってこの不完全性を考慮して導出した再犯率はおよそ15%だった。
この値が、求めたかった因果効果をうまく推定できているといえる。

(結論)

Coddledの場合の再犯率(15%)は、Arrestの場合の再犯率(10%(?))より高いと推定される
⇒ 再犯を防ぐDVへの最適措置は、Arrestだと推定できそう!

操作変数法じゃなきゃダメなのか？

操作変数法によってLATEを導出するのではなく、ランダム化の不完全性などは特に考慮しないでDからYへの因果効果を分析してみたらどうなるか・・・？
すなわち、以下の式で因果効果は測定できないだろうか？

$$E\{Y_i|D_i = 1\} - E\{Y_i|D_i = 0\} = 0.216 - 0.129 = 0.087 (\text{約}9\%) \leftarrow \text{LATE (約}15\%) \text{ と乖離}$$

結果：この方法では因果効果を過小評価してしまうことに。

理由：セレクション・バイアス（容疑者の攻撃的性格・常習性の違い）が排除されていないから。

⇒もしランダム化が不完全ならば、第1章より以下のようになることが分かっている。

$$\{2\text{群の平均値の差} = \text{介入効果} + \text{セレクション・バイアス}\}$$

⇒バイアスを除く操作変数法の方が優れた推定方法である。

M D V E のある特徴

K I P P の際・・・

子どもを4タイプに分類した上で、
操作変数方ではCompliersの平均因果効果を推定する
という流れで話を進めていた。

M D V E では・・・

分類してみると、Always-takersはほぼ存在しないという特徴が現れる。

警察官の分類

KIPPの際に子どもを分類したのと同様の考え方で警察官も分類してみる。

- 〔always-takers〕 いかなる端末の指示があってもCoddledの措置をとる警官
- 〔never-takers〕 いかなる端末の指示があってもArrestの措置をとる警官
- 〔compliers〕 端末の指示通りに動く警官

※ 〔defiers〕 に関しては不問としておく。



右の表との対応関係より...

〔always-takers〕は今回ほぼ存在しない



「Coddledの指示に背いてArrestしたケース」
は多く観察されたが、逆に

「Arrestの指示に背いてCoddledしたケース」
はほとんど観察されなかったことが反映

TABLE 3.3
Assigned and delivered treatments in the MDVE

Assigned treatment	Delivered treatment			Total
	Arrest	Coddled		
		Advise	Separate	
Arrest	98.9 (91)	0.0 (0)	1.1 (1)	29.3 (92)
Advise	17.6 (19)	77.8 (84)	4.6 (5)	34.4 (108)
Separate	22.8 (26)	4.4 (5)	72.8 (83)	36.3 (114)
Total	43.4 (136)	28.3 (89)	28.3 (89)	100.0 (314)

Notes: This table shows percentages and counts for the distribution of assigned and delivered treatments in the Minneapolis Domestic Violence Experiment (MDVE). The first three columns show row percentages. The last column reports column percentages. The number of cases appears in parentheses.

続) M D V E のある特徴

Always-takersがほぼ存在しないという結果から以下のように考察できる。

M D V E では...

$$L A T E = T O T$$

$$\Leftrightarrow \lambda = E\{Y_{1i} - Y_{0i} | C_i = 1\} = E\{Y_{1i} - Y_{0i} | D_i = 1\}$$

理由：K I P P の際に言及したのと同様に考えると

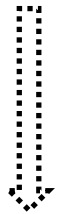
(Coddledした警官) = {Compliers + Always-takers(=0)} = (Compliers)
となるから。

まとめ～MDVE～

- セレクション・バイアスが存在したせいでランダム化が不完全になったMDVEだったが、操作変数方を使うことで推定したい因果効果を導き出すことが出来た。
- 今回のケースはAlways-takersがほぼ存在しないという点である意味特殊なケース。
- また、今回のケースはセレクション・バイアスの存在が大きな問題点となるので、バイアスを除去できる操作変数方が有効な推定方法だった。

おわりに

1980年代初頭：M D V E が実施される



⇒結果、D V の再犯を防ぐ最適措置は、
逮捕という重い罰則の方だった。

1989年：O.J.への処罰は、軽い罰則の方だった。

1994年：結果、O.J.による再犯が起こってしまった…。

もしM D V E の結果が操作変数方によって正確に導き出されたうえで、いち早く法制度に反映されれば、O.J.の再犯は防げたのでは…？



社会が変化するには、どうしても時間がかかってしまう。