



高知工科大学 経済・マネジメント学群

計量経済学応用

9. 操作変数法

やない ゆう き
矢内 勇生



<https://yukiyanai.github.io>



yanai.yuki@kochi-tech.ac.jp



このトピックの目標

- 操作変数法とは何かを理解する
 - ▶ 操作変数法は、どんな場面で使うのか？
- 操作変数法の仮定（のうち3つ）を理解する
- 操作変数法がどんな因果効果を推定するのか理解する
- 操作変数法の使い方を理解する

操作変数法

例：教育が所得に及ぼす影響

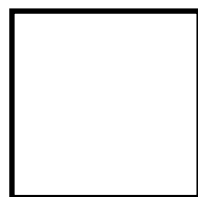
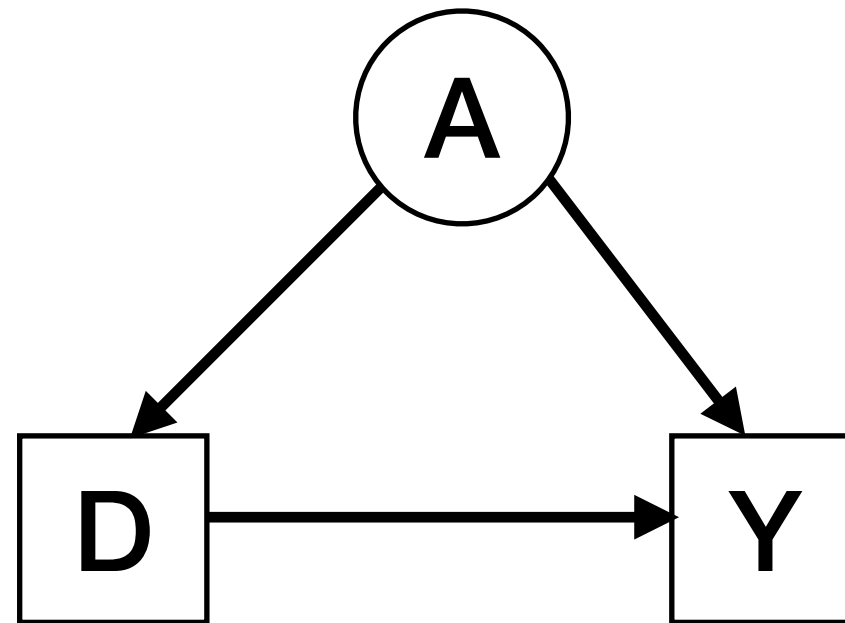
- ・ 結果変数：所得 Y_i
- ・ 処置変数：教育（修学年数） D_i
- ・ 交絡因子：能力 A_i
- ・ 正しいモデル（母回帰）：

$$Y_i = \alpha + \beta D_i + \gamma A_i + \epsilon_i$$

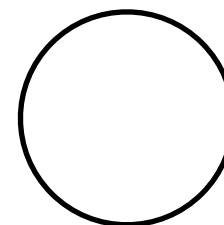
$$\epsilon_i \sim \text{Normal}(0, \sigma_\epsilon^2)$$

$$\text{Cor}(D, \epsilon) = 0$$

重回帰分析？



観測された変数



観測不能な
(観測していない) 変数

観察不能な交絡因子

- 問題：交絡因子 A が観測できない（観測されていない）
- 妥協案：交絡因子を無視した回帰分析 (**s**hort regression)

$$Y_i = \alpha_s + \beta_s D_i + \eta_i$$

Short Regression の誤差項

$$Y_i = \alpha_s + \beta_s D_i + \eta_i$$

- ここで、 $\eta_i = \gamma A_i + \epsilon_i$ とすると

$$Y_i = \alpha_s + \beta D_i + \eta_i$$

▶ $\text{Cor}(D, A) \neq 0$ だから、 $\text{Cor}(D, \eta_i) \neq 0$

– 説明変数と誤差項に相関がある

★ D は**内生変数**

説明変数の内生性

- 処置変数と誤差項に相関があると、因果効果の推定にバイアスが生じる
 - ▶ 交絡変数をコントロールし損ねる（脱落変数バイアス）
 - ▶ 結果変数が説明変数に影響を与える（逆の因果関係）
 - ▶ 測定誤差
 - ▶ 交絡変数が測定・観察不能
- 誤差項と相関がある説明変数：**内生変数** (endogenous variable)

Short Regression のバイアス

- 最小二乗推定量 $\hat{\beta}_s$ の期待値：

$$\begin{aligned}\mathbb{E}[\hat{\beta}_s] &= \frac{\text{Cov}(Y, D)}{\text{Var}(D)} = \frac{\text{Cov}(\alpha + \beta D + \gamma A + \epsilon, D)}{\text{Var}(D)} \\ &= \frac{\text{Cov}(\beta D, D) + \text{Cov}(\alpha + \gamma A + \epsilon, D)}{\text{Var}(D)} \\ &= \beta + \gamma \frac{\text{Cov}(A, D)}{\text{Var}(D)}\end{aligned}$$

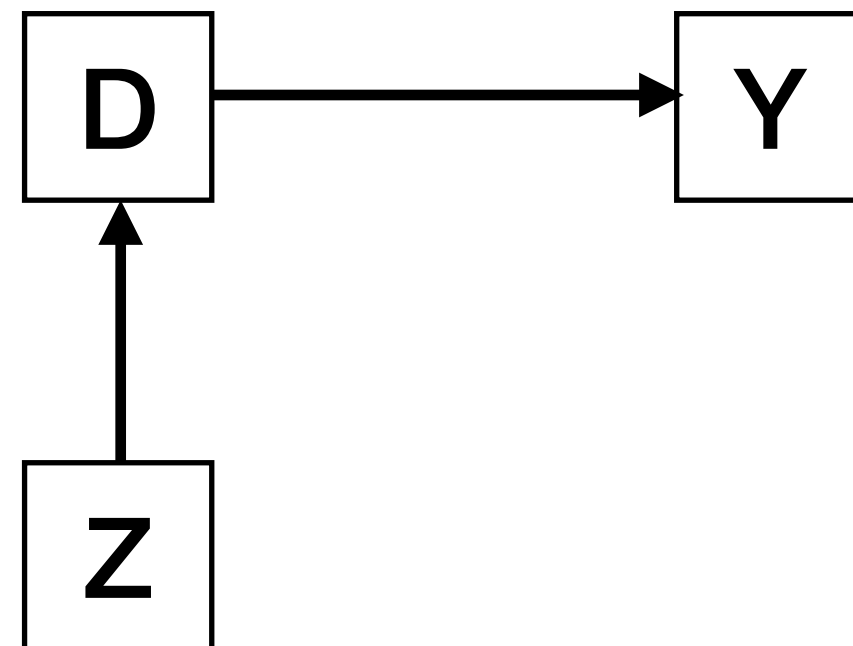
- $\gamma \frac{\text{Cov}(A, D)}{\text{Var}(D)}$ がバイアス
- $\gamma \neq 0$ かつ $\text{Cov}(A, D) \neq 0$ のとき、バイアスが生じる
- $\frac{\text{Cov}(A, D)}{\text{Var}(D)}$ は、 A を D に回帰したときの傾きパラメタ

操作変数法

- 操作変数法, instrumental variable method, IV法
- バイアスなしの β を推定するために、IV法を使う
- 操作変数 (instrument) : Z_i

★ 操作変数法の公式（内生変数も操作変数も1つずつのとき）

$$\beta_{IV} = \frac{\text{Cov}(Y, Z)}{\text{Cov}(D, Z)}$$



操作変数を使った推定量

- ・ 真のモデルを操作変数の公式に代入してみると

$$\begin{aligned}\beta_{IV} &= \frac{\text{Cov}(Y, Z)}{\text{Cov}(D, Z)} = \frac{\text{Cov}(\alpha + \beta D + \gamma A + \epsilon, Z)}{\text{Cov}(D, Z)} \\ &= \frac{1}{\text{Cov}(D, Z)} [\beta \text{Cov}(D, Z) + \gamma \text{Cov}(A, Z) + \text{Cov}(\epsilon, Z)]\end{aligned}$$

- ・ 除外制約があるので、 $\text{Cov}(A, Z) = 0$ かつ $\text{Cov}(\epsilon, Z) = 0$
- ・ 内生変数と操作変数には相関があるので $\text{Cov}(D, Z) \neq 0$
- ・ したがって、 $\beta_{IV} = \beta$

操作変数法の推定量は2つの回帰係数の比

- 操作変数法で得られる推定量は、2つの回帰係数の比である

- $$\beta_{IV} = \frac{\text{Cov}(Y, Z)}{\text{Cov}(D, Z)} = \frac{\text{Cov}(Y, Z)/\text{Var}(Z)}{\text{Cov}(D, Z)/\text{Var}(Z)} = \frac{\rho}{\phi}$$

- ϕ (phi) と ρ (rho) は以下の回帰モデルで推定できる

- ▶ 第1段階の回帰： $D_i = \alpha_1 + \phi Z_i + e_{1i}$

- ▶ 誘導型回帰： $Y_i = \alpha_0 + \rho Z_i + e_{0i}$

2段階回帰分析法

- 2段階回帰 (two-stage least squares, 2SLS, TSLS)
- 第1段階の回帰

$$D_i = \alpha_1 + \phi Z_i + e_{1i}$$

- 第2段階の回帰

$$Y_i = \alpha_2 + \beta_{2\text{SLS}} \hat{D}_i + e_{2i}$$

▶ ただし、 \hat{D}_i は第1段階の回帰で得られた D_i の予測値

$$\hat{D}_i = \alpha_1 + \phi Z_i$$

2段階回帰の推定量と操作変数法の推定量

- 操作変数の数と内生変数の数が一致するとき、2段階回帰の推定量と操作変数法の推定量は一致する

$$\begin{aligned}\beta_{2SLS} &= \frac{\text{Cov}(Y, \hat{D})}{\text{Var}(\hat{D})} \\ &= \frac{\text{Cov}(Y, \alpha_1 + \phi Z)}{\text{Var}(\alpha_1 + \phi Z)} \\ &= \frac{\phi \text{Cov}(Y, Z)}{\phi^2 \text{Var}(Z)} \\ &= \frac{\rho}{\phi} \\ &= \beta_{IV}\end{aligned}$$

共変量を伴う操作変数法

- 2段階回帰を行うなら、両方の回帰に共変量（交絡因子、コントロール変数）を含める（共変量を X とする）
 - ▶ 第1段階： $D_i = \alpha_1 + \phi Z_i + \lambda_1 X_i + e_{1i}$
 - ▶ 第2段階： $Y_i = \alpha_2 + \beta_{2\text{SLS}} \hat{D}_i + \lambda_2 X_i + e_{2i}$
- 共変量があるとき、 D と Z の相関条件は、「 X をコントロールしたうえで」判断する：つまり、 X を加えた重回帰の結果を使って検定を行う

操作変数の数

- 操作変数は複数存在しうる
- 操作変数の数が、処置変数が結果変数に与える影響を識別できるかどうかに関及ぼす影響を考える
- 内生変数の数を k , 操作変数の数を m とする
 - ▶ $k = m$: 丁度識別
 - ▶ $k < m$: 過剰識別
 - ▶ $k > m$: 識別できない
- $m \geq k$ であることが必要

Rで推定する場合

- 2つの回帰を別々には実行しない
 - ▶ ただし、第1段階だけは別の目的で実行する
- 2段階回帰を別々に実行すると、
 - ▶ 正しい（バイアスのない）回帰係数の推定値は得られる
 - ▶ **標準誤差が正しくない！**
- 操作変数法のために用意された関数を使う（補助教材を参照）

IV で何が推定できるのか？

- IV または 2SLS で β が推定できる
- しかし、 $\hat{\beta}_{IV}$ や $\hat{\beta}_{2SLS}$ は何の推定値？
 - ▶ ATE, ATT, ATC, それとも別のもの？
- IV で推定するもの：局所的平均処置効果 (local average treatment effect; **LATE**)
 - ▶ 「局所的」というのは、サンプルのうちのある「一部」についてという意味

設定

- 結果変数 Y_i
- 二値の操作変数 $Z_i \in \{0,1\}$
- 二値の処置変数 $D_i \in \{0,1\}$
- 潜在的処置：
 - ▶ $D_i(1) : Z_i = 1$ の場合の潜在的処置
 - ▶ $D_i(0) : Z_i = 0$ の場合の潜在的処置

サンプル中の個体を4タイプに分類

- D と Z が正の相関関係をもつ場合を考える
 - 潜在的処置によって、個体を4つのタイプに分類する
1. Complier[s] (従う者) : $D_i(1) = 1$ かつ $D_i(0) = 0$
 2. Always-taker[s] (常に処置) : $D_i(1) = 1$ かつ $D_i(0) = 1$
 3. Never-taker[s] (常に統制) : $D_i(1) = 0$ かつ $D_i(0) = 0$
 4. Defier[s] (反抗者) : $D_i(1) = 0$ かつ $D_i(0) = 1$

Complierに対する効果を推定する

- C_i : complier であることを示すダミー変数
 - ▶ $C_i = 1 \Leftrightarrow D_i(1) > D_i(0)$
- IV のLATE は、compliers についての平均処置効果
 - ▶ よって、complier average treatment effect; CATE
とも呼ばれる

$$\text{LATE} = \mathbb{E}[Y(1) - Y(0) \mid C_i = 1]$$

操作変数法の詳細

操作変数法の条件

- 操作変数法には、以下の3つ（本当は5つだが2つは割愛）の条件が必要

1. 操作変数 Z と内生変数 D （処置変数）の間に**相関**がある

$$\text{Cor}(Z, D) \neq 0$$

2. 操作変数 Z は、**内生変数 D を通してのみ**結果変数 Y に影響を与える（**除外制約** [exclusion restriction]、**唯一経路**）

$$\text{Cor}(Z, \eta) = 0$$

3. Defier が存在しない（単調性の仮定）

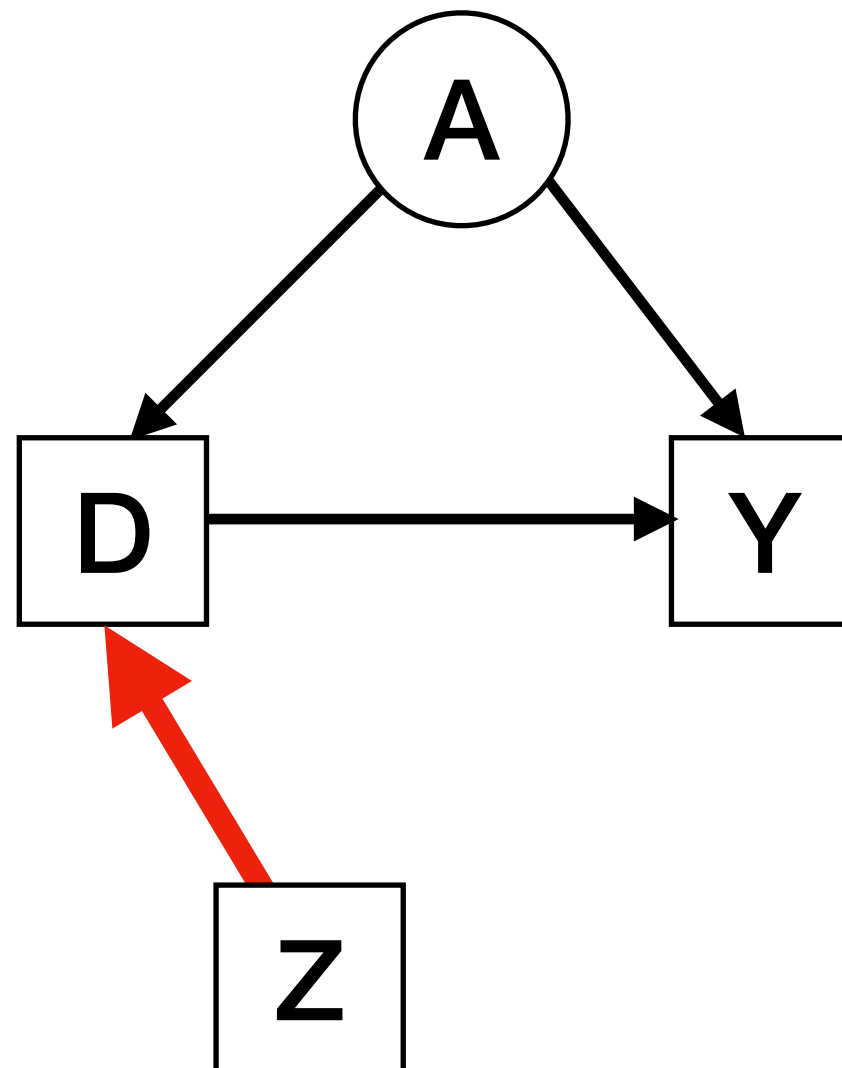
条件1：操作変数と内生変数の相関

- 操作変数 Z と内生変数 D には相関がなければならない

$$\text{Cor}(Z, D) \neq 0 \Leftrightarrow \text{Cov}(Z, D) \neq 0$$

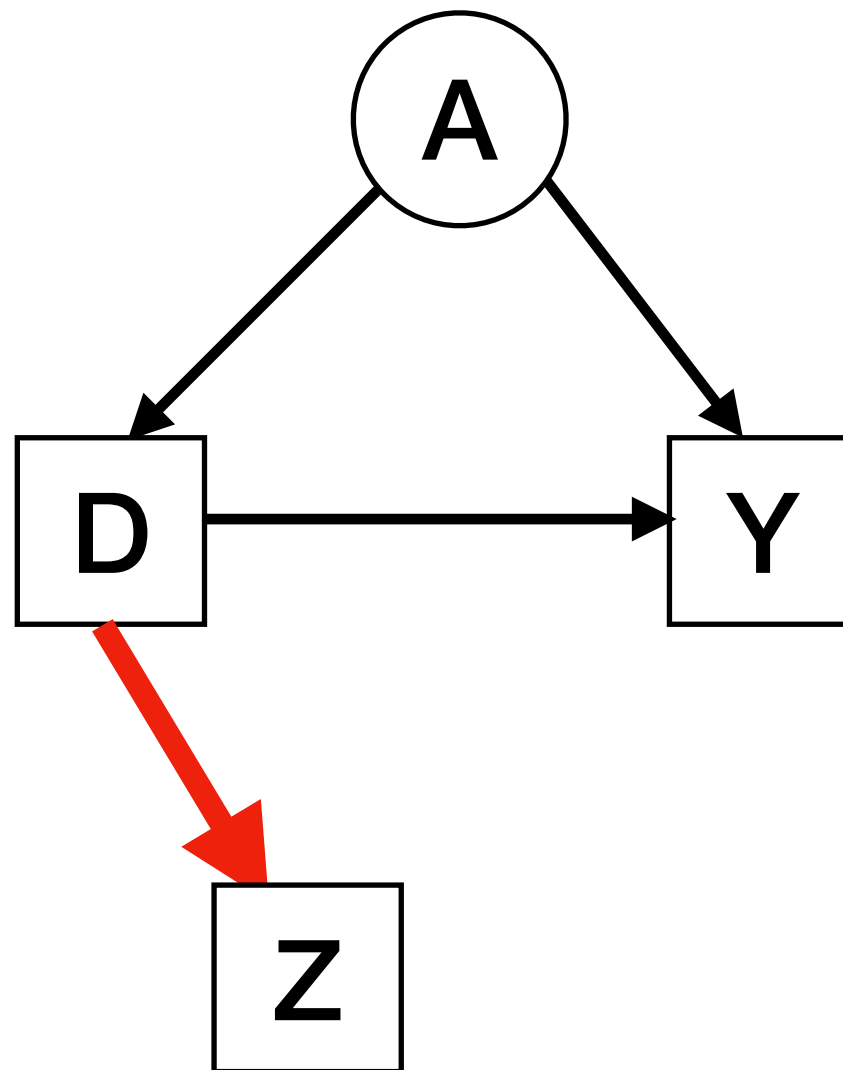
- ▶ 操作変数の値を変えれば、内生変数の値も変わる

操作変数と内生変数の相関



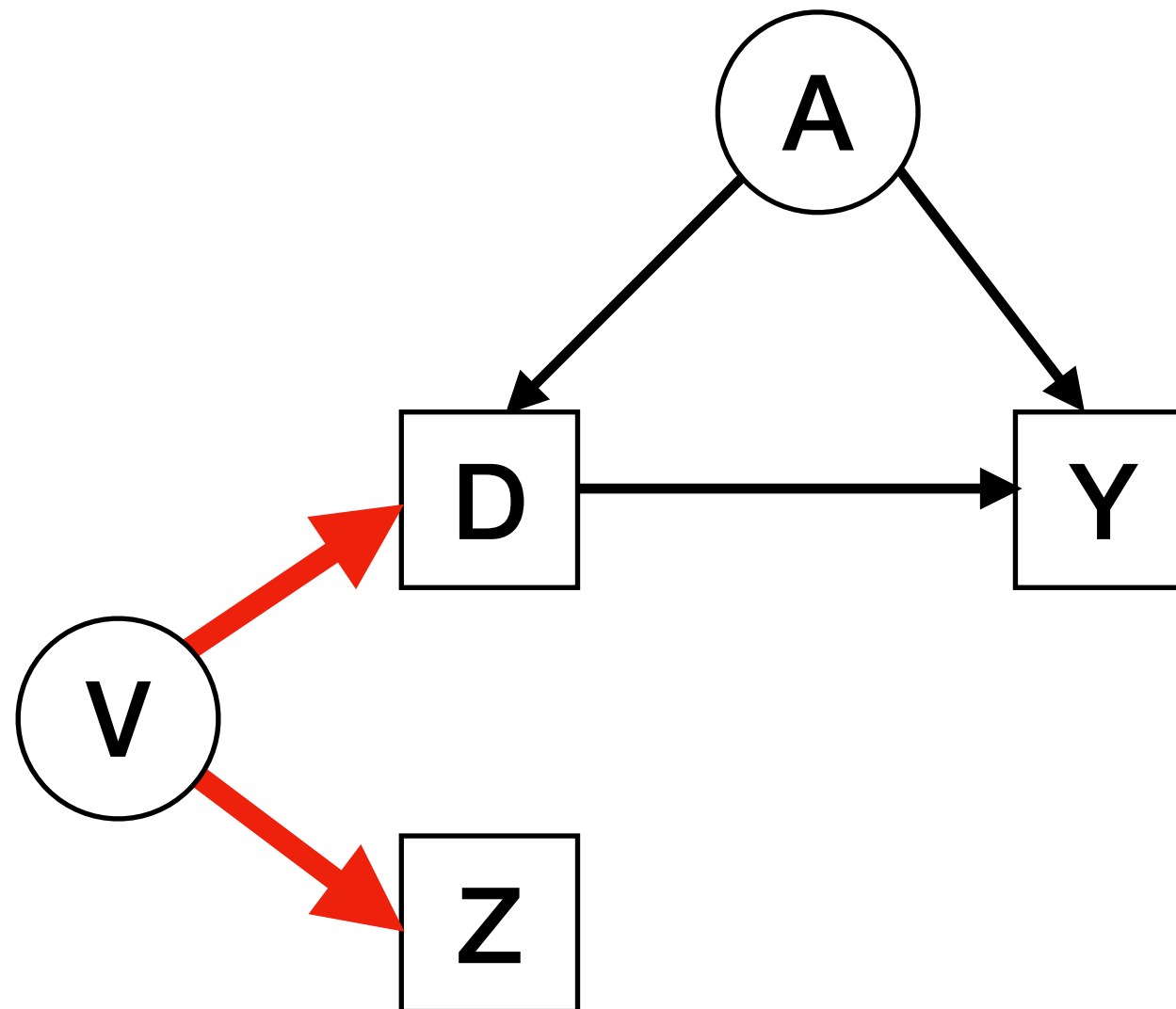
相関：操作変数 Z の値が変化すると、
内生変数 D の値も変化する

操作変数と内生変数の相関



相関：操作変数 Z の値が変化すると、
内生変数 D の値も変化する

操作変数と内生変数の相関



相関：操作変数 Z の値が変化すると、
内生変数 D の値も変化する

相関の確認方法

- 「 Z と D の相関はゼロ」という帰無仮説が棄却されるかどうか検定する
 - ▶ 操作変数が1つのとき： t 検定
 - ▶ 操作変数が複数あるとき： F 検定
- 実践的には「強い」相関が求められる（理由は後述）
- 目安
 - ▶ t 検定： t 値が3.2超（ p 値が 0.0016未満）
 - ▶ F 検定： F 値が10超

条件2：除外制約

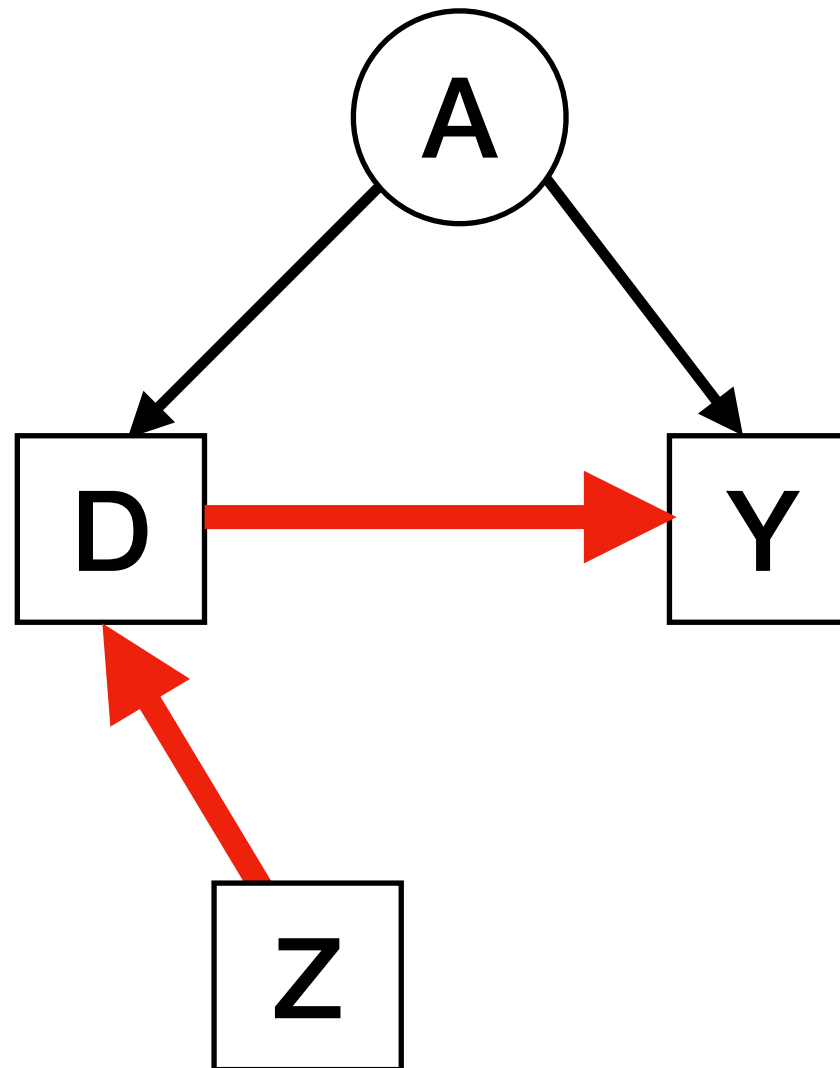
- 操作変数 Z は、内生変数 D を通じてのみ結果変数 Y に影響を与える
 - ▶ 唯一経路条件 (“only through” condition) とも呼ばれる
 - ▶ 除外制約が成り立つと：

$$Y_i(1) = Y_i(D_i(1) = 1, Z_i = 1) = Y_i(D_i(0) = 1, Z_i = 0)$$

かつ

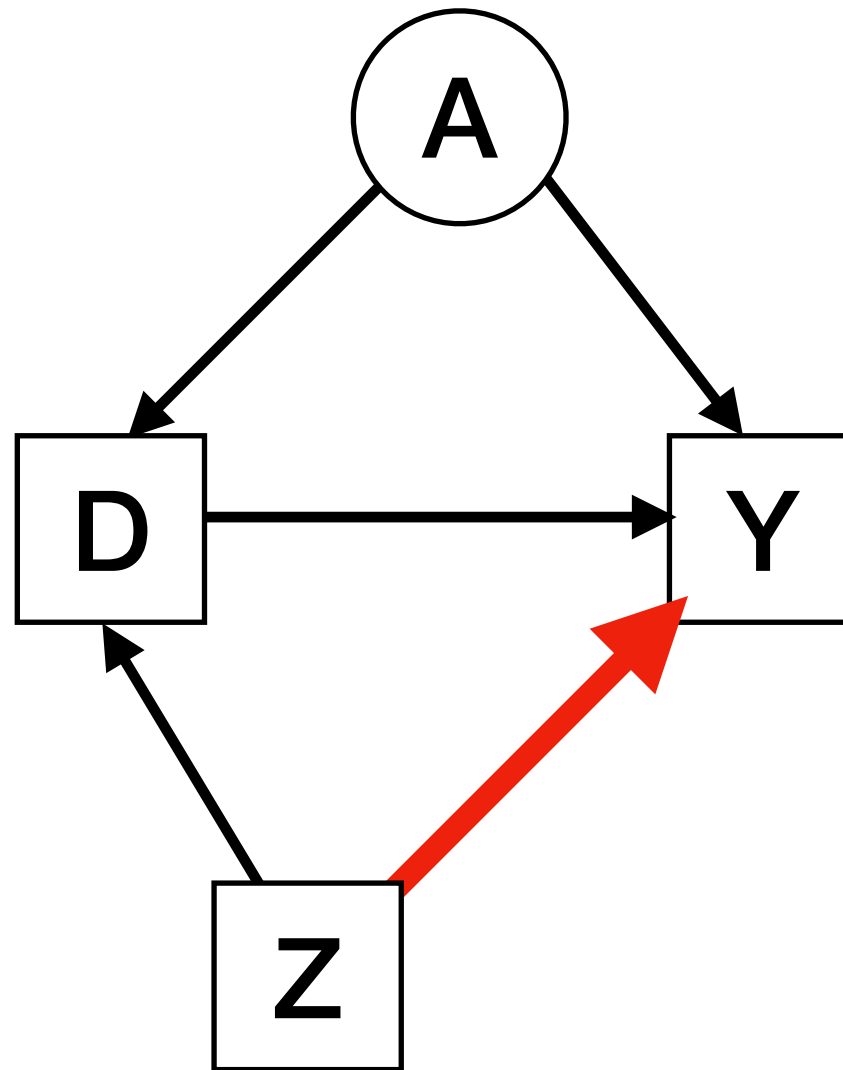
$$Y_i(0) = Y_i(D_i(1) = 0, Z_i = 1) = Y_i(D_i(0) = 0, Z_i = 0)$$

唯一経路（除外制約）



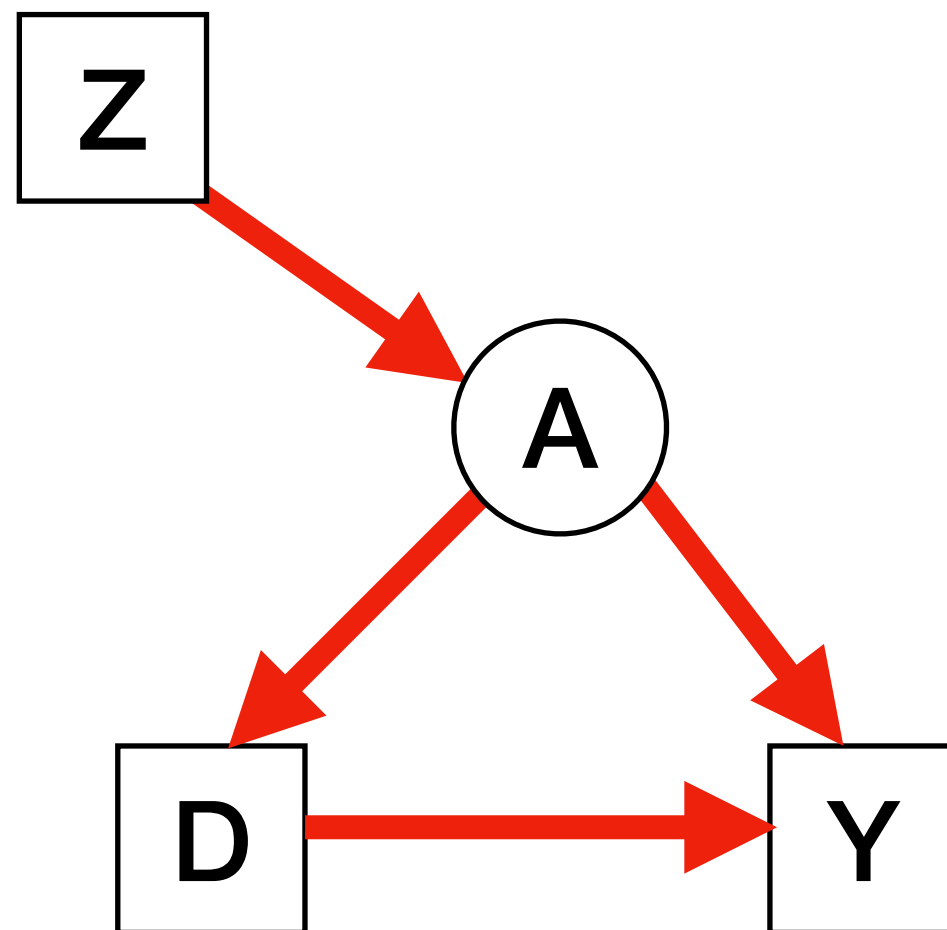
唯一経路：操作変数 Z は、内生変数 D を通してのみ
結果変数 Y に影響を与える (Z affects Y only through D)

除外制約が満たされないとき



Z は D と Y の交絡因子なので、
2段階目の回帰式から除外できない

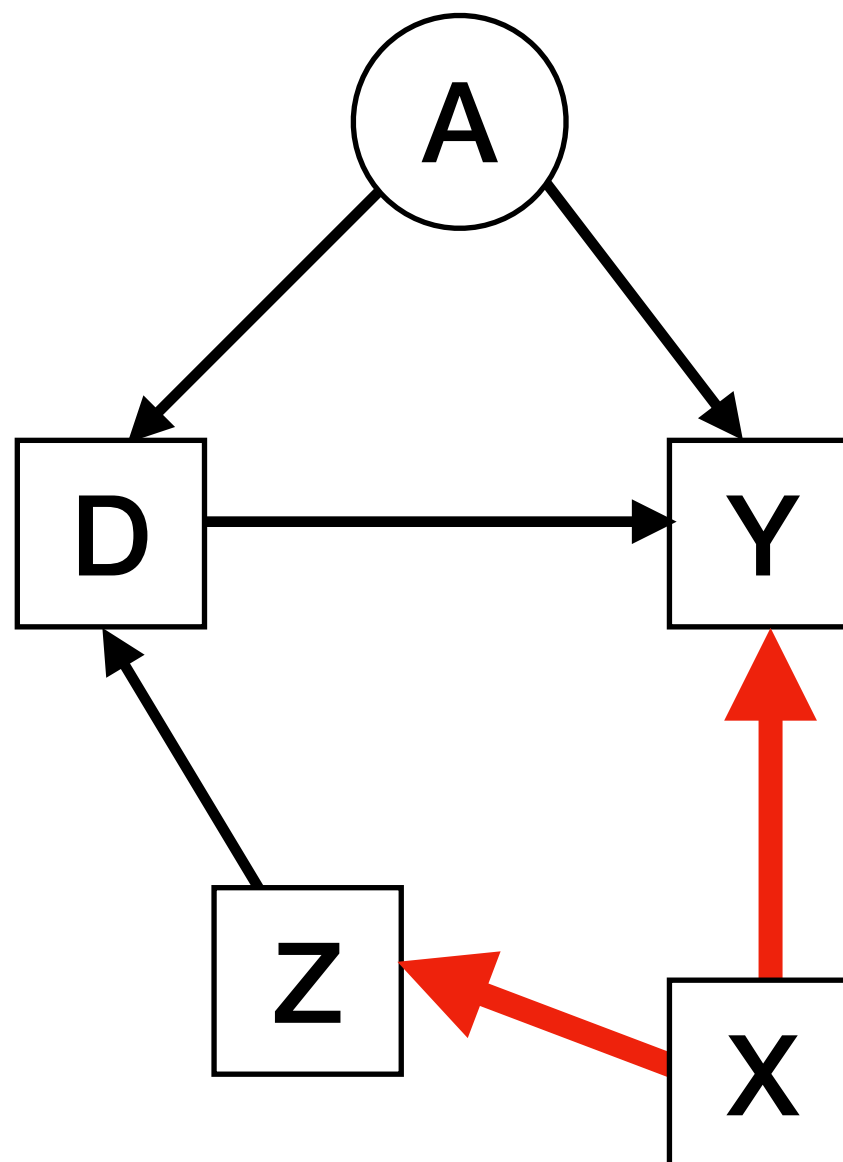
除外制約が満たされないとき



$$\text{Cor}(Z, \eta) \neq 0$$

A が観測されないので、Z が交絡に

除外制約が満たされないとき



X が観測されているので、 X をコントロールすれば、除外制約は満たされる

除外制約の確認法

- データでは確認できない！！！！
 - ▶ 「関係がないこと（値がゼロ）」は、統計的検定では示すことができないから
- では、どうするか？
 - ▶ 操作変数が結果変数に影響を与えるとすれば、内生変数を通じた効果しか考えられないことを、**データを使わずに説得する！**
 - ▶ 操作変数法を使った論文の成否は、この「説得力」に依存する

条件3：単調性

- LATE を推定するためには、
 - ▶ $D_i(1) - D_i(0) \geq 0$ ($\text{Cor}(D, Z) > 0$ のとき) または $D_i(1) - D_i(0) \leq 0$ ($\text{Cor}(D, Z) < 0$ のとき) であることが必要
 - ▶ つまり、defier がいてはいけない！
- 観測値による区別
 - ▶ $D_i(1) = 1 \mid Z_i = 1$: complier または always-taker
 - ▶ $D_i(1) = 0 \mid Z_i = 1$: never-taker または **defier**
 - ▶ $D_i(0) = 1 \mid Z_i = 0$: always-taker または **defier**
 - ▶ $D_i(0) = 0 \mid Z_i = 0$: complier または never-taker
- 単調性は、**理論**によって示すことが必要

操作変数の仮定：まとめ

- 操作変数と内生変数の相関：データを使って確認できる
 - ▶ 「相関がない」という帰無仮説が棄却されるかどうか
 - ▶ 相関が十分強いのか
 - 除外制約（唯一経路）：データからは確かめられない
 - ▶ **理論的に説得する**必要がある：操作変数法を使う論文で、最も重要なパートの一つ！
 - ▶ 自分が扱っている問題についての知識が求められる
 - 単調増加（defier の不在）の仮定：理論的に示す必要がある
- ★ これらの仮定が成り立ちそうであることを示すのが大切

IVを使ったLATEの推定 (1)

- 3つ (本当はさらに2つ必要だがここでは考えない) の仮定をおくと、IV によって LATE が推定できる
- IVによる係数の推定法 (再掲)

$$\beta_{IV} = \frac{\text{Cov}(Y, Z)}{\text{Cov}(D, Z)} = \frac{\text{Cov}(Y, Z)/\text{Var}(Z)}{\text{Cov}(D, Z)/\text{Var}(Z)} = \frac{\rho}{\phi}$$

▶ 第1段階の回帰 : $D_i = \alpha_1 + \phi Z_i + e_{1i}$

▶ 誘導型回帰 : $Y_i = \alpha_0 + \rho Z_i + e_{0i}$

IVを使ったLATEの推定 (2)

- $(D_i(0), D_i(1)) \perp\!\!\!\perp Z_i$ とすると

- 第1段階の回帰の係数 ϕ

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

- これは、サンプル内の compliers の割合である
 - ▶ サンプル内の全個体が complier なら、 $\phi = 1$ になる

IVを使ったLATEの推定 (3)

- ・誘導型回帰の係数 ρ

$$\begin{aligned}\rho &= \mathbb{E}[Y_i \mid Z_i = 1] - \mathbb{E}[Y_i \mid Z_i = 0] \\&= \mathbb{E}[Y_i(D_i(1), 1) \mid Z_i = 1] - \mathbb{E}[Y_i(D_i(0), 0) \mid Z_i = 0] \\&= \mathbb{E}[Y_i(D_i(1), 1)] - \mathbb{E}[Y_i(D_i(0), 0)] \\&= \mathbb{E}[Y_i(D_i(1))] - \mathbb{E}[Y_i(D_i(0))] \\&= \phi \mathbb{E}[Y_i(D_i(1)) - Y_i(D_i(0)) \mid C_i = 1] \\&\quad + (1 - \phi) \mathbb{E}[Y_i(D_i(1)) - Y_i(D_i(0)) \mid C_i = 0] \\&= \phi \mathbb{E}[Y_i(D_i(1)) - Y_i(D_i(0)) \mid C_i = 1] \\&= \phi \mathbb{E}[Y_i(1) - Y_i(0) \mid C_i = 1]\end{aligned}$$

IVを使ったLATEの推定 (4)

- IV が推定する係数 β_{IV}

$$\beta_{IV} = \frac{\rho}{\phi}$$

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

$$= \mathbb{E}[Y_i(1) - Y_i(0) \mid C_i = 1]$$

$$= \text{LATE}$$

IV で ATE を推定できる？

- 一般的にはできない
 - ▶ always-taker と never-taker の処置効果がわからない
 - 常に処置か常に統制かのいずれかだから
 - ▶ ATEを推定するためには、仮定を追加することが必要
 - 「全個体が complier」という仮定を追加
 - ◆ $LATE = ATE = ITT$ (intention-to-treat effect)
 - 「処置効果が個体によらず一定」という仮定を追加
 - ◆ $LATE = ATE = ITE$

IV で ATT を推定できる？

- 一般的にはできない

- $D_i = D_i(0) + [D_i(1) - D_i(0)]Z_i$

- ▶ 単調性の仮定により、 $D_i(0) = 1 \Rightarrow D_i(1) = 1$

- ▶ つまり、 $D_i(0) = 1$ なら i は always-taker

- 処置群にいるのは complier または always-taker

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

$$= p(C_i = 1 \mid D_i = 1)\mathbb{E}[Y(1) - Y(0) \mid C_i = 1]$$

$$+ p(D_i(0) = 1 \mid D_i = 1)\mathbb{E}[Y(1) - Y(0) \mid D_i(0) = 1]$$

- したがって、「always-taker は存在しない」と仮定できれば、 $LATE = ATT$

除外制約の不成立が疑われるとき

- 除外制約は検定できないので、少しくらいは「他の経路」がありそう： $\text{Cor}(Z, \eta) \neq 0$
- 操作変数法（2段階回帰）の推定にバイアスが生じる

$$\mathbb{E}[\hat{\beta}_{2\text{SLS}}] = \beta + \frac{\text{Cor}(Z, \eta)}{\text{Cor}(Z, D)} \frac{\text{SD}(\eta)}{\text{SD}(D)}$$

▶ バイアス：
$$\frac{\text{Cor}(Z, \eta)}{\text{Cor}(Z, D)} \frac{\text{SD}(\eta)}{\text{SD}(D)}$$

「弱い」 操作変数

- 内生変数との相関が弱い操作変数：弱い操作変数 (weak instrument[s])
- 除外制約が完全に満たされていないとき、弱い操作変数を使うと推定がうまくいかない
 - ▶ 弱すぎると、操作変数を使わない方がマシな場合も
 - バイアス：
$$\frac{\text{Cor}(Z, \eta)}{\text{Cor}(Z, D)} \frac{\text{SD}(\eta)}{\text{SD}(D)}$$

弱い操作変数とバイアス

- Short regression（内生性に対処しない単回帰）のバイアス：

$$\text{Cor}(D, \eta) \frac{\text{SD}(\eta)}{\text{SD}(D)}$$

- 除外制約が完全でないときの、操作変数法のバイアス：

$$\frac{\text{Cor}(Z, \eta)}{\text{Cor}(Z, D)} \frac{\text{SD}(\eta)}{\text{SD}(D)}$$

- ▶ 弱い操作変数を使うとバイアスが大きくなってしまう：弱い変数を使うくらいなら、short regressionのほうがマシかも

操作変数の利用例

- 研究課題：冷戦期、西側の文化は東側の体制崩壊に影響を与えたか？
- （かつての）一般的な見解：テレビやラジオで西側の文化に触れた東側の人間が、ソ連体制への不満を募らせ、それが体制崩壊を促進した
- 実験されたわけではない
 - ▶ 本当に因果関係があったか？

何が問題か

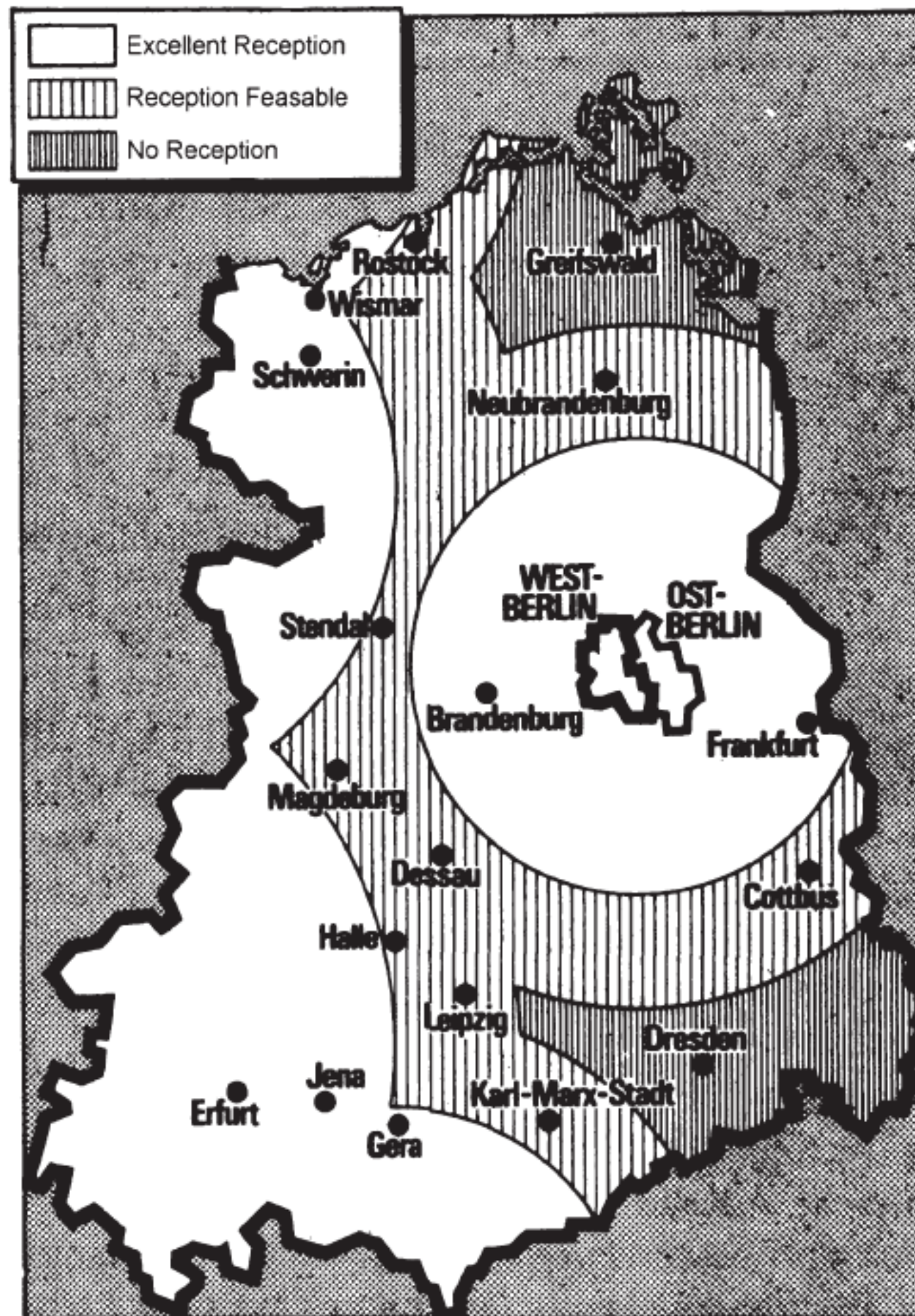
- 例：反体制だから西側のテレビを見るのでは？（逆の因果関係 [内生性] の可能性）
- 最良の解決策：実験する
 - ▶ 問題：実験できない
 - もう冷戦は終わった！
- 自然実験（操作変数法）を利用する

東ドイツにおける自然実験

地形によってテレビ電
波の強弱に違いがある
ことを利用

冷戦当時 (1988, 1989) の
世論調査を分析

Kern, H. L., and J. Hainmueller. 2009.
“[Opium for the Masses: How Foreign
Media Can Stabilize Authoritarian
Regimes](#).” *Political Analysis* 17:
377-399.



出典： Kern & Hainmueller (2009: 382), Fig. 1

分析結果（一部）

	単純な差 視聴者 - 非視聴者	操作変数を利用した分析 西側テレビの影響
マルクス・レーニン主義 を受け入れるか	-0.079	0.204
東ドイツに親近感を持っ ているか	-0.013	0.251

- ・西ドイツのテレビ番組を見ているほど、東ドイツ政府に好意的

▶ これまでの常識に反する結果

結果変数：東ドイツ政府に対する態度

内生変数：テレビの視聴

操作変数：西ドイツのテレビの電波

次回予告

Topic 10.

分析結果のプレゼンテーション