

計量経済学応用

12. 操作変数法

矢内 勇生

2018年5月24日

高知工科大学 経済・マネジメント学群

今日の目標

- 操作変数法の使い方を身につける

説明変数の内生性

- 回帰分析では、因果効果の推定にバイアスが生じることがある
 - ▶ 交絡変数を統制し損ねる（欠落変数バイアス）
 - ▶ 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
- こういう状況下の説明変数を内生変数 (endogenous variable) と呼ぶ

例：教育が所得に及ぼす影響

- 結果変数：所得 Y_i
- 説明変数：教育（修学年数） D_i
- 交絡：能力 A_i
- 母回帰モデル（母集団における本当の関係）：

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

- β を推定したい！
- **問題**： A_i が観察できない
- 妥協案：交絡変数を除いた回帰分析 (short regression)

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

Short Regression のバイアス

- 最小二乗推定量 $\hat{\beta}_s$ の期待値：

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

- $\gamma \frac{\text{Cov}(A, D)}{\text{Var}(D)}$ がバイアス
- $\gamma \neq 0$ かつ $\text{Cov}(A, D) \neq 0$ であれば、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 (instrumental variable [IV] method)

- バイアスなしに β を推定するために、IV法を使う
- 操作変数 (instrument) : Z_i
- 操作変数は、以下の条件を満たす必要がある
 - ▶ 原因となる説明変数（内生変数）と相関がある
 - ▶ 結果変数には、**説明変数（内生変数）を通してのみ**影響を与える（除外制約）
- 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 $\text{Cov}(\epsilon, Z) = 0$
- 説明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

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

- ρ と ϕ 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段階で得られた D_i の予測値

$$\hat{D}_i = \hat{\alpha}_1 + \hat{\phi} Z_i$$

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（統制変数）を含める（共変量を X とする）
 - ▶ 第1段階： $D_i = \alpha_1 + \phi Z_i + \lambda X_i e_{1i}$
 - ▶ 第2段階： $Y_i = \alpha_2 + \beta_{2\text{SLS}} \hat{D}_i + \lambda X_i + e_{2i}$

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！
- 操作変数法用に用意されたコマンドを使う

AER::ivreg()

- AER パッケージの `ivreg()` 関数を使う
- 使い方

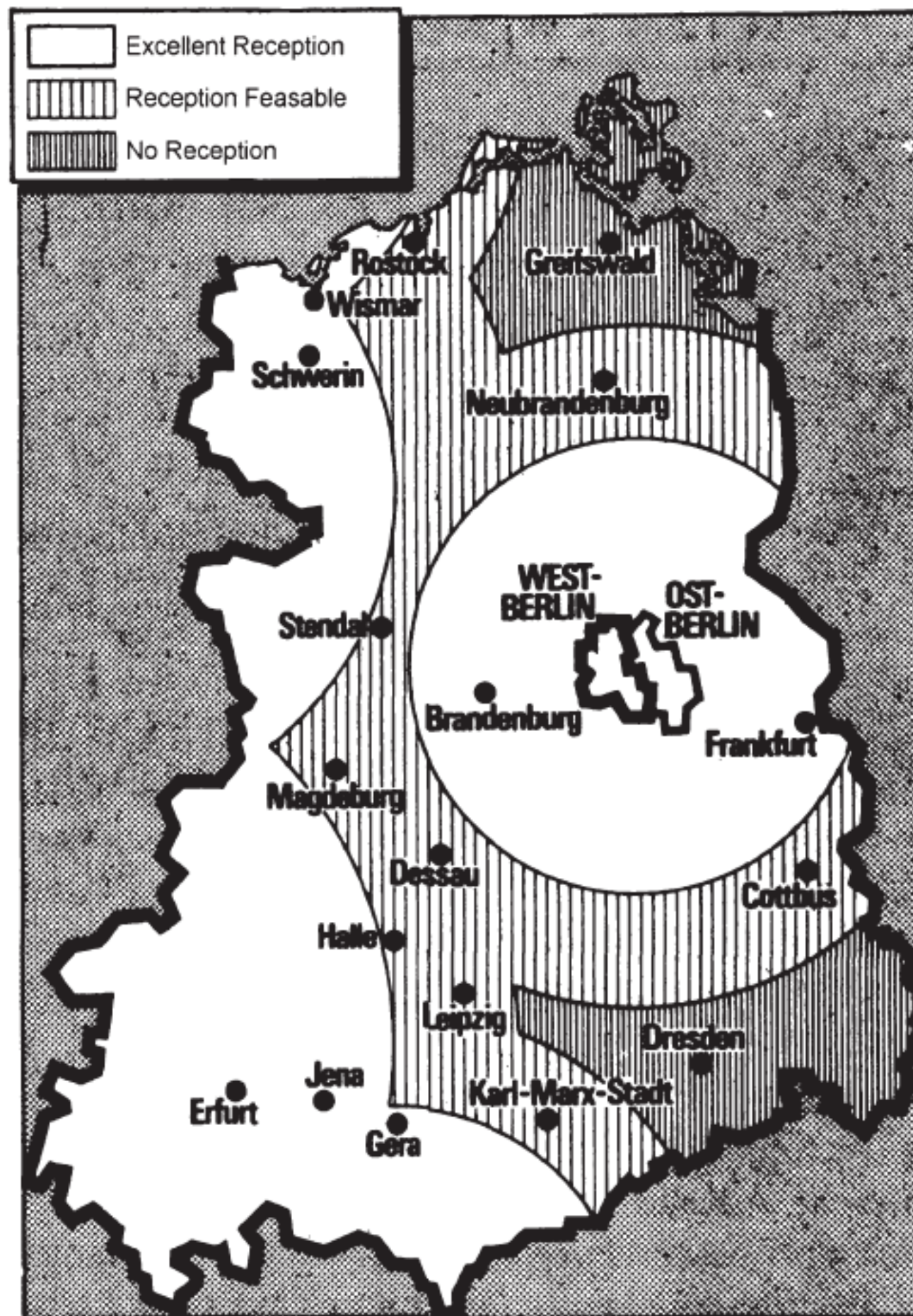
```
fit <- ivreg(Y ~ D + X | X + Z, data = ...)
```

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？
 - エビデンス（科学的方法で明らかにされた因果関係）がない

何が問題になり得るか

- 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.

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