

Эконометрика. Лекция 9

Эндогенность и инструментальные переменные

Д. С. Терещенко

НИУ ВШЭ, Санкт-Петербург

18 ноября 2022 г.

Основные источники

- **SW, глава 12**
- Видео-материалы для самостоятельного освоения (скоро появятся?)

Содержание

- 1 Эндогенность (еще раз о внутренней валидности)
- 2 Инструментальные переменные
- 3 Двухшаговый метод наименьших квадратов («парный-парный» случай)
- 4 Примеры исследований

Содержание

- 1 Эндогенность (еще раз о внутренней валидности)
- 2 Инструментальные переменные
- 3 Двухшаговый метод наименьших квадратов («парный-парный» случай)
- 4 Примеры исследований

Повторение (предположения в регрессионной модели)

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

Мы бы хотели, чтобы выполнялось следующее:

- $\mathbb{E}(u_i | x_i) = 0$
- $\text{var}(u_i | x_i = x) = \sigma^2(x) < \infty, \forall x$
- $\text{var}(x_i) \neq 0$
- $\{(y_i, x_i, u_i)\} \sim i.i.d.$

Тогда можно было бы получить несмещенные, состоятельные и асимптотически нормальные МНК-оценки.

$$\hat{\beta}_1^{OLS} = \frac{\widehat{\text{cov}}(x, y)}{\widehat{\text{var}}(x)} \xrightarrow[n \rightarrow \infty]{p} \frac{\text{cov}(x, y)}{\text{var}(x)} = \beta_1$$

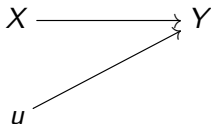
Эндогенность

Эндогенность

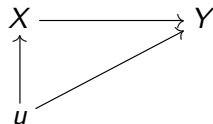
Ситуация, когда мы не уверены, что $\mathbb{E}(u_i|x_i) = 0$

$$\hat{\beta}_1^{OLS} = \frac{\widehat{cov}(x, y)}{\widehat{var}(x)} \xrightarrow[n \rightarrow \infty]{p} \dots$$

МНК можно использовать:



МНК нельзя использовать:



Причины эндогенности

- пропущенные переменные;
- ошибки в измерениях;
- одновременность и обратная причинность;
- самоотбор.

Ошибки измерения

Ошибки измерения зависимой переменной

- Мы хотим оценить: $y = \beta_0 + \beta_1 x + u$
- Но оцениваем: $y^* = \beta_0 + \beta_1 x + u$, где $y^* = y + \zeta$, $E(\zeta|x, y) = 0$
- Получаем состоятельную МНК-оценку:

$$\hat{\beta}_1 = \frac{\widehat{cov}(x, y^*)}{\widehat{var}(x)} \xrightarrow[n \rightarrow \infty]{p} \frac{cov(x, y^*)}{var(x)} = \frac{cov(x, y + \zeta)}{var(x)} = \frac{cov(x, y)}{var(x)} = \beta_1$$

Ошибки измерения

Ошибки измерения объясняющей переменной

- Мы хотим оценить: $y = \beta_0 + \beta_1 x + u$
- Но оцениваем: $y = \beta_0 + \beta_1 x^* + u$, где $x^* = x + \zeta$, $E(\zeta|x, y) = 0$
- Получаем несостоятельную МНК-оценку:

$$\hat{\beta}_1 = \frac{\widehat{cov}(x^*, y)}{\widehat{var}(x^*)} \xrightarrow[n \rightarrow \infty]{p} \frac{cov(x^*, y)}{var(x^*)} = \frac{cov(x + \zeta, y)}{var(x + \zeta)} = \frac{cov(x, y)}{var(x) + var(\zeta)}$$

- Оценка смещена в сторону нуля относительно истинного значения β_1

Пример для случая пропущенной переменной

Ability bias

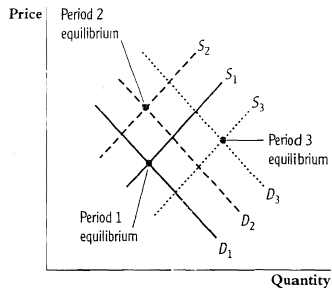
- Истинная модель: $w = \beta_0 + \beta_1 s + \beta_2 a + e$
- Оцениваем «модель»: $w = \beta_0 + \beta_1 s + u$, $u = \beta_2 a + e$
- Получаем потенциально несостоятельную МНК-оценку:

$$\hat{\beta}_1 = \frac{\widehat{cov}(w, s)}{\widehat{var}(s)} \xrightarrow[n \rightarrow \infty]{p} \frac{cov(s, w)}{var(s)} = \frac{cov(s, \beta_0 + \beta_1 s + \beta_2 a + e)}{var(s)} = \beta_1 + \beta_2 \frac{cov(s, a)}{var(s)}$$

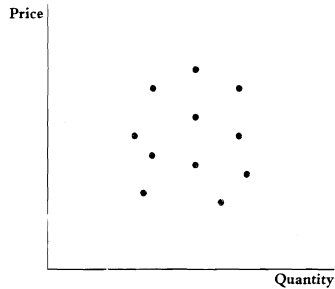
- В какую сторону смещение?

Пример одновременности

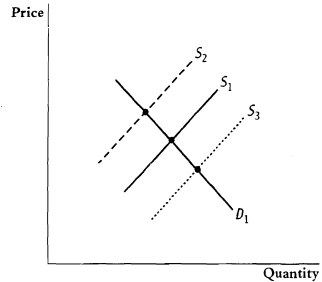
Оценка кривой спроса (Stock J. & Watson M. Introductory econometrics, ch. 12)



(a) Demand and supply in three time periods



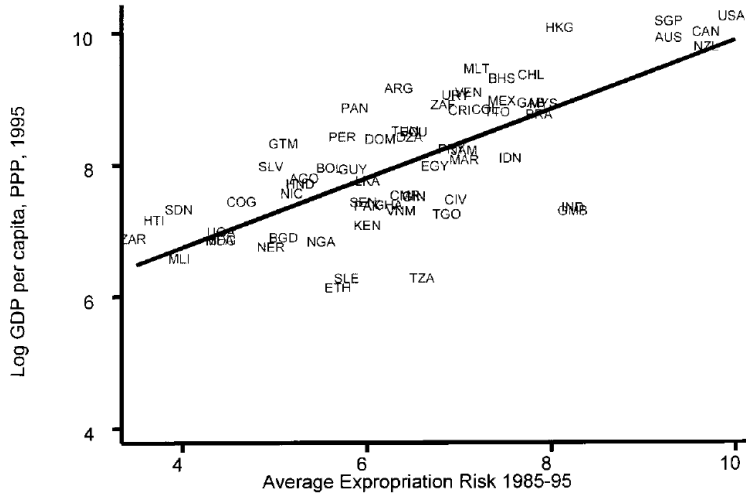
(b) Equilibrium price and quantity for 11 time periods



(c) Equilibrium price and quantity when only the supply curve shifts

Пример одновременности

Институты и экономическое развитие: Acemoglu, Johnson & Robinson (2001)



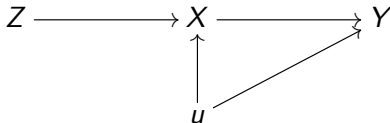
Содержание

- 1 Эндогенность (еще раз о внутренней валидности)
- 2 Инструментальные переменные
- 3 Двухшаговый метод наименьших квадратов («парный-парный» случай)
- 4 Примеры исследований

Борьба с эндогенностью — идея

- ① эндогенная переменная X раскладывается на две составляющие:
 - проблемная компонента, которая может быть коррелирована с ошибкой,
 - хорошая, очищенная компонента, которая не коррелирована с ошибкой;
- ② очищенная компонента используется для получения оценок.

Инструментальные переменные



Для хороших инструментов должно выполняться:

- релевантность сила: $\text{cov}(z, x) \neq 0$
- экзогенность годность: $\text{cov}(z, u) = 0$

При выборе инструментов необходимо полагаться на:

- экономическую теорию;
- экспертные мнения, институциональное знание, экономическую интуицию.

Содержание

- 1 Эндогенность (еще раз о внутренней валидности)
- 2 Инструментальные переменные
- 3 Двухшаговый метод наименьших квадратов («парный-парный» случай)
- 4 Примеры исследований

Двухшаговый МНК (2МНК, 2ШМНК, 2SLS, TSLS)

- ① эндогенная переменная X раскладывается на две составляющие:
 - проблемная компонента, которая может быть коррелирована с ошибкой,
 - хорошая, очищенная компонента, которая не коррелирована с ошибкой;
- ② очищенная компонента используется для получения оценок.

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

- ① $x_i = \pi_0 + \pi_1 z_i + v_i$
- ② $y_i = \beta_0 + \beta_1 \hat{x}_i + u_i$

Инструментальные переменные и двухшаговый МНК (2МНК, 2ШМНК, 2SLS, TSLS)

$$\hat{\beta}_1^{TSLS} = \frac{\widehat{cov}(z, y)}{\widehat{cov}(z, x)} \xrightarrow[n \rightarrow \infty]{p} \frac{cov(z, y)}{cov(z, x)} = \frac{cov(z, \beta_0 + \beta_1 x + u)}{cov(z, x)} = \beta_1$$

$$\hat{\beta}_1^{TSLS} = \frac{\widehat{cov}(z, y)}{\widehat{cov}(z, x)} \xrightarrow[n \rightarrow \infty]{p} \frac{cov(z, y)}{cov(z, x)} = \frac{cov(z, y)/var(z)}{cov(z, x)/var(z)} = \frac{\beta_1 \pi_1}{\pi_1} = \beta_1$$

Оценки 2МНК:

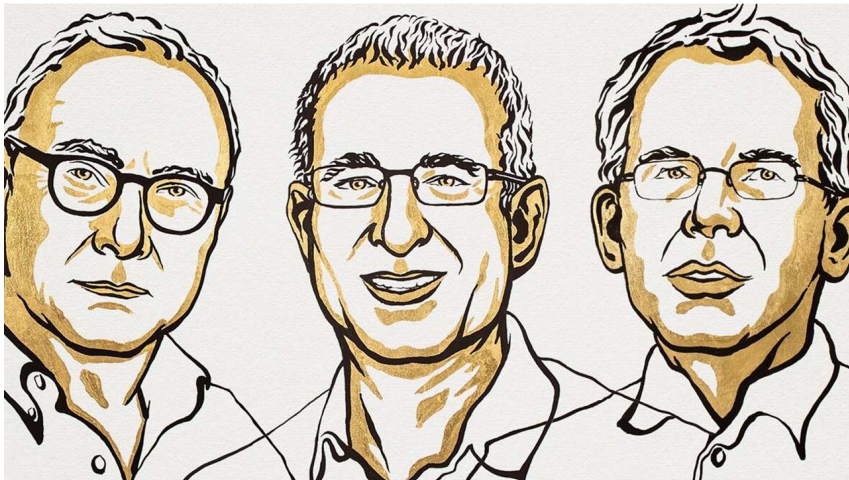
- состоятельные
- асимптотически нормальные

Содержание

- 1 Эндогенность (еще раз о внутренней валидности)
- 2 Инструментальные переменные
- 3 Двухшаговый метод наименьших квадратов («парный-парный» случай)
- 4 Примеры исследований

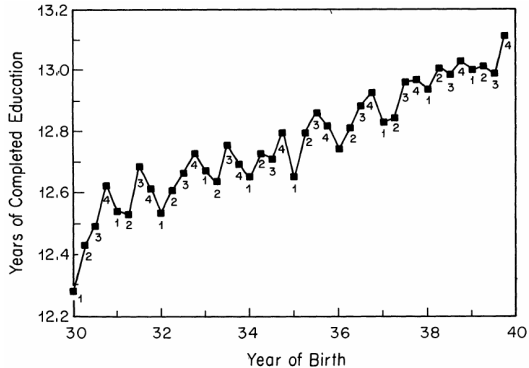
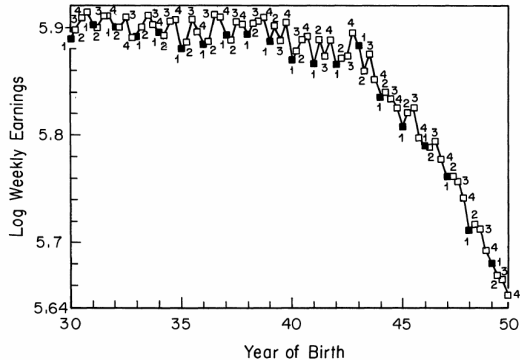
Нобелевская премия по экономике — 2021

Дэвид Кард («за эмпирический вклад в экономику труда»), Джошуа Ангрист и Гвидо Имбенс («за методологический вклад в анализ причинно-следственных связей»)



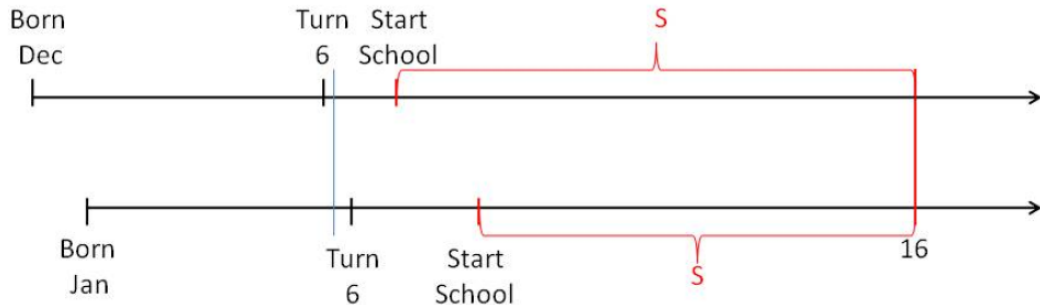
Примеры

Оценка влияния образования на зарплату: Angrist & Krueger (1991)



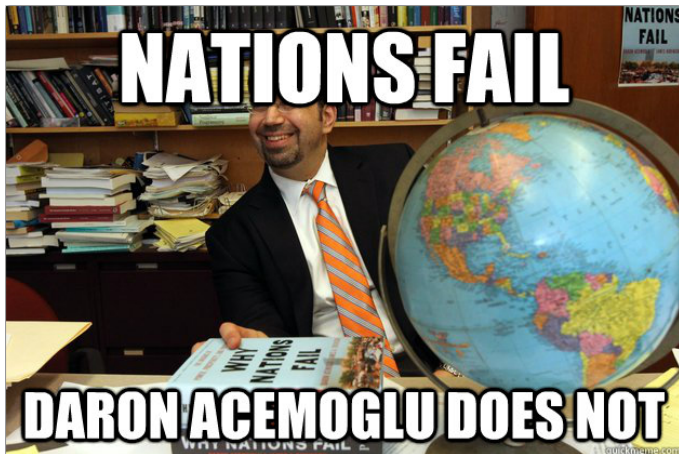
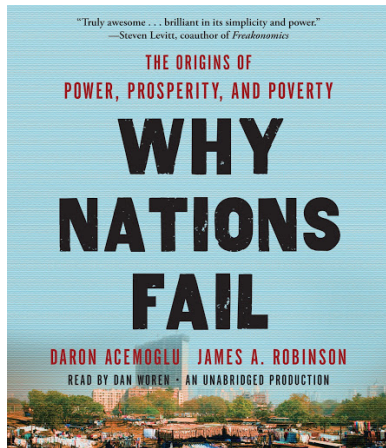
Примеры

Оценка влияния образования на зарплату: Angrist & Krueger (1991)



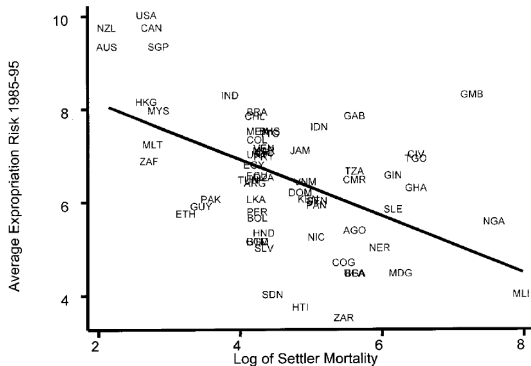
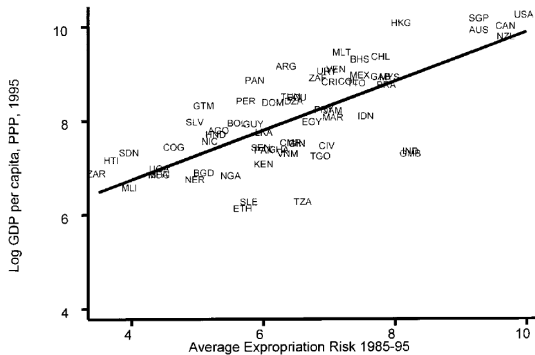
Why nations fail

Исследования Дарона Асемоглу и соавторов об экономическом росте и институтах



Примеры

Институты и экономическое развитие: Acemoglu, Johnson & Robinson (2001)

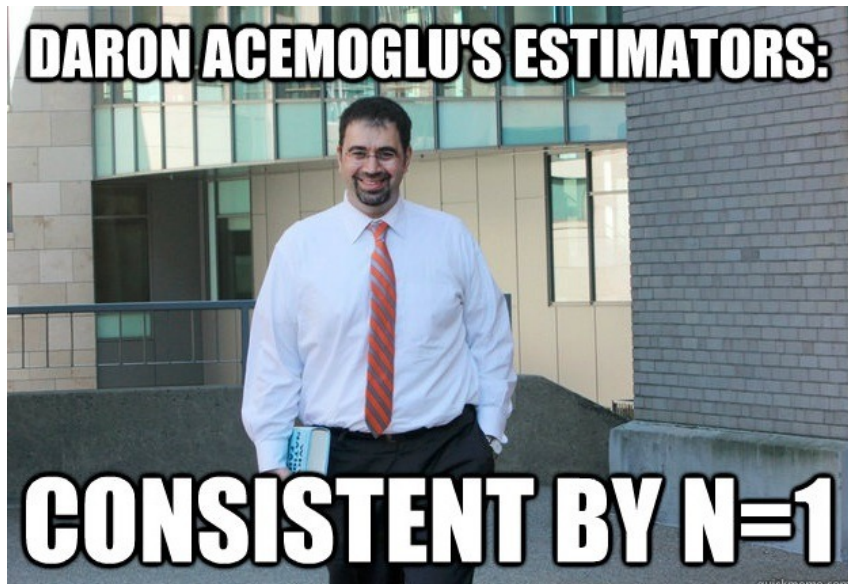


(potential) settler mortality \Rightarrow settlements \Rightarrow early institutions \Rightarrow current institutions \Rightarrow current performance.

Мемы про Дарона Асемоглу



Мемы про Дарона Асемоглу



Мемы про Дарона Асемоглу



Дополнительные источники

Примеры хороших статей с использованием инструментальных переменных

- ① Daron Acemoglu, Simon Johnson & James A. Robinson, 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation," American Economic Review, American Economic Association, vol. 91(5), pages 1369-1401.
- ② Joshua D. Angrist & Alan B. Krueger, 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?," The Quarterly Journal of Economics, Oxford University Press, vol. 106(4), pages 979-1014.