VARIÁVEIS INSTRUMENTAIS E INFERÊNCIA CAUSAL

Luiz Guilherme Scorzafave

FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DE RIBEIRÃO PRETO
UNIVERSIDADE DE SÃO PAULO

BIBLIOGRAFIA

- Angrist, J.; Imbens, G. W.; and Rubin, D. (1996) Identification of Causal Effects Using Instrumental Variables (with discussion). *Journal of the American Statistical Association*, vol 91, no 434, 444-472.
- Morgan, S.L. (2001) Counterfactuals, Causal Effect Heterogeneity, and the Catholic School Effect on Learning. Sociology of Education v74, n4, 341-374.
- Morgan, S.L.; Winship C. (2008) Counterfactuals and causal inference. Methods and Principles for Social Research. Cambridge: Cambridge University Press
- ANGRIST, Joshua D.; PISCHKE, Jörn-Steffen. **Mostly harmless econometrics: An empiricist's companion**. Princeton university press, 2008

OBJETIVOS

- Vamos interpretar os estimadores de variáveis instrumentais permitindo que o efeito da variável endógena seja heterogêneo.
- Por que estimar um efeito heterogêneo é importante?
  – Validade Interna: se dada estratégia conseguiu obter os efeitos causais de interesse para certa população.
  – Validade Externa: se o valor encontrado para o efeito é válido em contextos diferentes do estudo em questão
- Exemplo: experimentos aleatórios tem geralmente pequena validade externa e grande validade interna.

Objetivos

- Um instrumental econométrico com efeito heterôgeneo nos ajuda a ter uma ideia da validade interna e externa das estimativas baseadas em variáveis instrumentais.

- Para estimar o efeito populacional médio causal, precisamos de hipóteses fortes sob o efeito do instrumento na variável endógena: efeito do tratamento é constante (homogêneo).

- Sem esta hipótese, é possível identificar efeitos médios para subpopulações que são induzidas pelo instrumento a ter uma mudança no valor das variáveis endógenas (LATE)

VARIÁVEIS INSTRUMENTAIS

- O estimador de VI é dado por:

- $\hat{\beta}_{vi} = (z'x)^{-1}\ zy$, onde z é matriz de instrumentos para z (modelo exatamente identificado).

- $\hat{\beta}_{Gvi} = (x'\mathcal{P}_z x)^{-1}\ x'\mathcal{P}_z y$, onde $\mathcal{P} = z(z'z)^{-1}z'$

(modelo sobreidentificado)

VARIÁVEIS INSTRUMENTAIS

- Considere que:

(1) plim $(\frac{1}{n}$ z'u) = 0

(2) plim$(\frac{1}{n}$ z'x) = M_n → positiva semidefinida, com posto k

VARIÁVEIS INSTRUMENTAIS

- Consistência de $\hat{\beta}_{vi}$

$$plim\,(\hat{\beta}_{vi}) = plim\,[(z'x)^{-1}\, z'y] =$$
$$\beta + plim\,[(z'x)^{-1}\, z'u]$$

- Equação de Slutsky (permite fatorar *plim*)

$$= \beta + plim\,\left[\frac{(z'x)^{-1}}{n}\right] plim\,[\frac{(z'u)}{n}],$$

- Por (1) → $plim\ \hat{\beta}_{vi} = \beta$

VARIÁVEIS INSTRUMENTAIS

- A prova é análoga para $\hat{\beta}_{Gvi}$:

- $\text{plim}\ \hat{\beta}_{Gvi} = \beta + \text{plim}\left\{\frac{x'z}{n}\left(\frac{z'z}{n}\right)^{-1}\frac{z'u}{n}\right\} = \beta$

- $\hat{\beta}_{Gvi}$ é consistente!

ESTIMADOR DE VI DE EFEITOS CAUSAIS

- A ideia é de que se houver alguma fonte exógena de variação que afete Y somente via D, variável de tratamento, essa variável pode ajudar na identificação do efeito causal de D em Y.

- Instrumento deve ser altamente correlacionado com D, mas não com não observáveis que afetam Y.

ESTIMADOR DE VI DE EFEITOS CAUSAIS

- Motivação: estimação de efeitos causais com variável instrumental binária.

- Seja: $Y = \alpha + \delta . D + \varepsilon$ (7.1)

- Tratamento ⇒ D = 1 → Tratado

  D = 0 → Não tratado

Estimador de VI de efeitos causais

- Suponha que P(D=1) = f(z), z é binária. A figura abaixo representa 2 formas possíveis pelas quais z pode se relacionar com ambos D e Y.

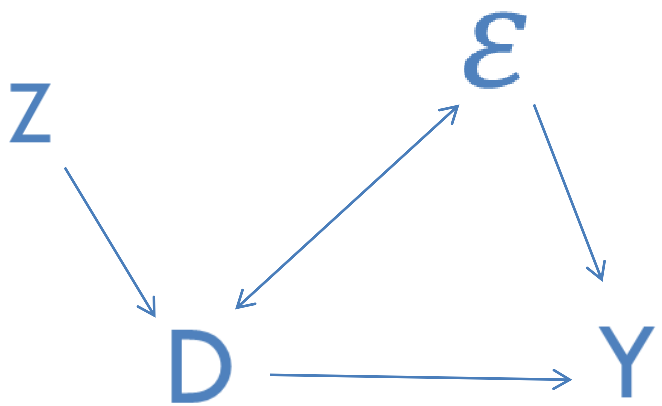

Figura A

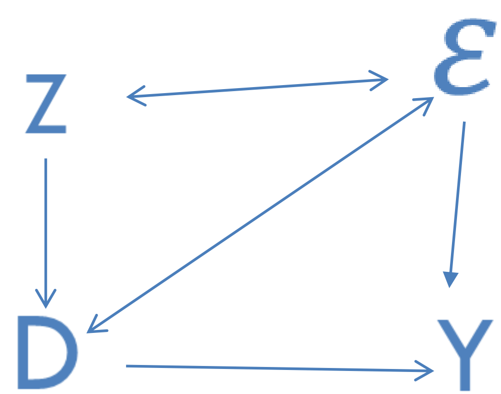

Figura B

ESTIMADOR DE VI DE EFEITOS CAUSAIS

- Como ε e D estão relacionados, em qualquer caso uma regressão de Y em D não captura de modo consistente o efeito de D em Y.
- Em A, associação de Z com Y é apenas via D. Em B também é via ε.
- No painel B, como Z e Y tem causas comuns (via ε), estratégia de efeito causal falha. Assim, a variação que Z parece gerar em D e Y pode, ao invés disso, ser gerada por não observáveis que causam tanto Z quanto Y.

Estimador de VI de efeitos causais

- Mais formalmente, de (7.1):
- E(y) = $\alpha + \delta E(D) + E(\varepsilon)$

$$\begin{cases} E(Y/Z=1) = \alpha + \delta\, E(D/Z=1) + E(\varepsilon/Z=1) \\ E(Y/Z=0) = \alpha + \delta\, E(D/Z=0) + E(\varepsilon/Z=0) \end{cases}$$

- $E(Y/Z=1) - E(Y/Z=0) = \delta\, [E(D/Z=1) - E(D/Z=0)] + [E(\varepsilon/Z=1) - E(\varepsilon/Z=0)]$ (7.2)

ESTIMADOR DE VI DE EFEITOS CAUSAIS

- Dividindo por E(D/Z=1) – E(D/Z=0), vem:

$$\frac{E(Y/Z=1) - E(Y/Z=0)}{E(D/Z=1) - E(D/Z=0)} =$$

$$= \frac{\delta[E(D/Z=1) - E(D/Z=0)\,] + [E(\varepsilon/Z=1) - E(\varepsilon/Z=0)]}{E(D/Z=1) - E(D/Z=0)}$$

- Se vale a Figura A, Z e ε são não correlacionados, de modo que $E(\varepsilon/Z=1) = E(\varepsilon/Z=0)$ e o lado direito fica $=\delta$.

ESTIMADOR DE VI DE EFEITOS CAUSAIS

- Assim, o efeito causal de D em Y é dado pela razão das associações entre (Y e Z) e (D e Z). Assim, o estimador será:

$$\hat{\delta}_{VI,WALD} = \frac{E_n(y_i/z_i=1) - E_n(y_i/z_i=0)}{E_n(d_i/z_i=1) - E_n(d_i/z_i=0)}$$

Chamamos assim quando o instrumento é binário!

- A abordagem trata δ como um parâmetro fixo comum a todos os indivíduos.

Estimador de VI de efeitos causais

- Do ponto de vista da abordagem de resultados potenciais, na qual geralmente se assume que os efeitos causais variam significativamente entre indivíduos, essa hipótese é limitante e não razoável.

- Assim, um dos objetivos aqui é explicar quando e como tal hipótese pode ser relaxada.

Estimador de VI de efeitos causais

- No caso do painel B:

$$\frac{E(Y/Z=1) - E(Y/Z=0)}{E(D/Z=1) - E(D/Z=0)} = \delta + \overbrace{\frac{E(\varepsilon\ /Z=1) - E(\varepsilon\ /Z=0)}{E(D/Z=1) - E(D/Z=0)}}^{\text{Inconsistência que decorre da relação entre } \varepsilon \text{ e Z}}$$

EXEMPLO

- Avaliar efeito da escola privada na proficiência via programa que dá voucher via sorteio a alunos (sistema público e privado) via loteria.

* D=1(privado) e D=0(público)

- Sendo que 10% dos estudantes elegíveis ganham voucher.
- Imagine que a configuração final de estudantes e escolas é a seguinte:

EXEMPLO

	Escola pública (d_i=0)	Escola privada (d_i=1)
Perde Voucher ($Z_i = 0$)	8000	1000
Ganha Voucher ($Z_i = 1$)	800	200

EXEMPLO

- Como a V.I., Z= ganhar voucher é
  - Aleatorizada, independe de ε em Y =$\alpha + \delta.D + \varepsilon$
  - cov(Z,D)$\neq$ 0, a princípio é bom instrumento.

Quem ganha voucher tem mais chance ir para a privada!

$$\hat{\delta}_{VI,WALD} = \frac{51.6-51.111}{0.2-0.111} = \frac{E_n((y_i/z_i=1) - E_n(y_i/z_i=0)}{E_n(d_i/z_i=1) - E_n(d_i/z_i=0)} = 5.5$$

Exemplo

- Sendo que rodando um MQO:

  $\hat{y}$ = 50,0 + 9,67 D → Viesado e inconsistente

- De acordo com a teoria de VI tradicional, pode ser dada uma interpretação causal a este coeficiente. No entanto, o efeito causal que essa VI identifica é ligeiramente diferente: quando há <u>heterogeneidade do efeito causal entre os indivíduos</u>, não identifica efeito para todos os indivíduos.

EXEMPLO

- Iremos mostrar que identifica somente o efeito causal médio para aqueles estudantes que vão para a escola privada com voucher, mas que sem ele, ficariam na escola pública.
- Assim, por ex., não diz nada sobre o efeito causal entre os que poderiam se matricular na escola privada sem o voucher (que são a maioria da amostra da privada, 83% $= \frac{1000}{1200}$)
- Assim, a literatura de resultados potenciais fornece o *insight* para entender tal resultado.

Limitações da análise tradicional de VI

- Na abordagem tradicional, desde que exista um instrumento que seja de “boa qualidade”, então um estimador de VI pode ser usado para efetivamente estimar um efeito causal.
- Mas mesmo na VI tradicional, há algumas limitações nessa estratégia de estimação.

LIMITAÇÕES DA ANÁLISE TRADICIONAL DE VI

1) Muitas vezes é difícil defender que Z não tem efeito direto sobre Y.

2) VI é viesado em amostras finitas

3) Viés pode ser grande se a correlação entre Z e D for fraca (instrumento fraco)

CASO (1)

- Mesmo “experimentos naturais” podem não estar imunes a esse problema. Suponha o caso em que a data de nascimento seja instrumento (Z) para status de veterano (D) para se estimar efeito do serviço militar no Vietnã sobre rendimento(Y).
- Mesmo Z sendo gerado aleatoriamente, depois do sorteio ter sido feito, empregadores podem ter comportamento diferente de acordo com o número sorteado (data de nascimento), investindo mais naqueles com menos chance de recrutamento.

CASOS (2) E (3)

- Mesmo que, na população, cov(Z,D) = 0, quando calculo isso na amostra pode dar ligeiramente $\neq 0$, o que permitirá que o computador cuspa um número.
- Ao mesmo tempo, a fórmula para calcular erro padrão das estimativas de VI falha, dando valores artificialmente pequenos (o erro padrão deveria ser infinito!)
- Se o instrumento também afetar Y diretamente, mesmo que "pouco" e o instrumento for fraco, o viés explode!

- Esse último ponto pode ser visto no Wald:

$$\hat{\delta}_{VI,WALD} = \frac{\text{E(Y/Z=1) – E(Y/Z=0)}}{\text{E(D/Z=1) – E(D/Z=0)}} = \delta + \frac{\text{E}(\varepsilon\ \text{/Z=1) – E}(\varepsilon\ \text{/Z=0)}}{\text{E(D/Z=1) – E(D/Z=0)}}$$

$$\frac{E(\varepsilon|Z=1) - E(\varepsilon|Z=0)}{E(D|Z=1) - E(D|Z=0)} = inconsistência$$

- Se viola $\text{cov}(\varepsilon, z) = 0$, numerador é diferente de zero, pois Z é associado com Y via ε. Quanto mais fraco o instrumento [menor $\text{cov}(D, z)$], menor o denominador e maior o viés.

Angrist, Imbens e Rubin (1996)

- **Inovação:** novo parâmetro para efeito do tratamento – *local average treatment effect* (LATE).
- Reconciliar VI com abordagem de resultados potenciais.

Rearranjando: $Y = D \cdot Y^1 + (1 - D) \cdot Y^0$

$$Y = Y^0 + \underbrace{(Y^1 - Y^0)}D = Y^0 + \delta \cdot D = \mu^0 + \delta \cdot D + v^0$$

δ não é mais um parâmetro estrutural como em (7.1)

Onde $\mu^0 \equiv E(Y^0)$ e $v^0 \equiv Y^0 - E(Y^0)$

LATE

Como interpretar um estimador de VI como estimador de efeito causal?

Imagine que D e Z sejam variáveis binárias. Pode-se definir 4 grupos:

Compliers $(\hat{C} = c)$: $D^{Z=0} = 0$ *e* $D^{Z=1} = 1$

Defiers $(\hat{C} = d)$: $D^{Z=0} = 1$ *e* $D^{Z=1} = 0$

Always takers $(\hat{C} = a)$: $D^{Z=0} = 1$ *e* $D^{Z=1} = 1$

Never takers $(\hat{C} = n)$: $D^{Z=0} = 0$ *e* $D^{Z=1} = 0$

LATE

No exemplo do efeito da escola privada (D) sobre nota (Y), loteria para receber o voucher (Z) é a V.I.

- **Compliers:** Estudantes que só ficam na escola privada se tiverem o voucher, senão ficam na pública;
- **Defiers:** Estudantes que só ficam na privada se NÃO receberem voucher.
- **Always:** Estudantes que sempre ficam na privada, independente do voucher
- **Never:** Estudantes que nunca ficam na privada, independente do voucher.

LATE

Analogamente à definição de Y, a variável D pode ser definida como:

$$D = D^{Z=0} + \underbrace{(D^{Z=1} - D^{Z=0})}_{} \cdot Z$$

$$\boldsymbol{D = D^{Z=0} + K \cdot Z}$$

- ***K*** representa o efeito causal individual do instrumento Z em D. Varia entre os indivíduos se $(D^{Z=1} - D^{Z=0})$ varia entre indivíduos.
- Para *compliers*, note que ***K***=1; para *defiers*, ***K***=-1; e para os demais, ***K***=0.

LATE

- Dadas essas definições, um instrumento Z válido para o efeito causal de D em Y deve satisfazer as seguintes hipóteses para se identificar um LATE:

LATE

1. **Independência:** $(Y^1, Y^0, D^{Z=1}, D^{Z=0}) \perp\!\!\!\perp Z$;

- Hipótese de independência é análoga a hipótese que E(u|Z) = 0 na literatura de V.I. tradicional.
- Saber o valor do instrumento para o Δi não dá nenhuma informação, nem sobre o resultado potencial sob ambos os tratamentos, nem sobre a probabilidade de ser tratado sob valores alternativos hipotéticos do instrumento.

LATE

2. **Efeito do Instrumento ≠ 0 (restrição de exclusão): K ≠ 0, para algum indivíduo i**

- Hipótese diz que instrumento Z deve predizer atribuição ao tratamento para pelo menos alguns indivíduos (deve haver pelo menos alguns *compliers* ou *defiers*);
- Z só opera via D

LATE

3. **Monotonicidade: ou K ≥ 0 para todo mundo, ou K ≤ 0 para todos indivíduos.**

- *Compliers* e *defiers* não podem coexistir na população.

EXEMPLO ESCOLA PRIVADA

No caso da escola privada, como o instrumento Z é aleatório, vale a **hipótese 1** e como Z prediz D, vale a **hipótese 2** (ou seja, quem recebe voucher está mais na privada).

Para **hipótese 3**, deve excluir possibilidade de *defier*: razoável excluir caso em que evita escola privada se recebe voucher.

LATE

Assim, se K ≥ 0 para todo i, então, sob as três hipóteses:

$$\widehat{\boldsymbol{\delta}}_{\boldsymbol{VI,WALD}} \xrightarrow{\mathrm{P}} \boldsymbol{E}[\boldsymbol{\delta}|\widetilde{\boldsymbol{C}} = \boldsymbol{c}],$$

que é o efeito causal médio para os *compliers*.

- Analogamente, se K ≤ 0, $\hat{\delta}_{VI,WALD} \xrightarrow{\mathrm{P}} E\left[\delta\middle|\tilde{C} = d\right]$.

Sob monotonicidade, exclui defier:

$$\Pr[D=1|Z=0] = fração\ de\ "always"$$

$$\Pr[D=0|Z=1] = fração\ de\ "\text{never}"$$

$$\Pr[D=1|Z=1] - \Pr[D=1|Z=0]$$
$$= \Pr[D=0|Z=0] - \Pr[D=0|Z=1]$$
$$= fração\ de\ "compliers"$$

	Z=0	Z=1
D=0	Never Complier	Never ~~Defier~~
D=1	Always ~~Defier~~	Always Complier

LATE - DEMONSTRAÇÃO

$$E(Y|Z=1)$$
$$= E(Y|Z=1, D=1) \cdot \Pr(D=1|Z=1)$$
$$+ E(Y|Z=1, D=0) \cdot \Pr(D=0|Z=1)$$

$$E(Y|Z=0)$$
$$= E(Y|Z=0, D=1) \cdot \Pr(D=1|Z=0)$$
$$+ E(Y|Z=0, D=0) \cdot \Pr(D=0|Z=0)$$

Numerador do Wald

Assim, $\boxed{E(Y|Z=1) - E(Y|Z=0)} =$

compliers

= E(Y|Z=1,D=1)·{Pr(D=1|Z=1)-Pr(D=1|Z=0)} -

compliers

E(Y|Z=0,D=0)·{Pr(D=0|Z=0)-Pr(D=0|Z=1)} +

always

{E(Y|Z=1,D=1)-E(Y|Z=0,D=1)}·Pr(D=1|Z=0) +

never

{E(Y|Z=1,D=0)-E(Y|Z=0,D=0)}·Pr(D=0|Z=1)

$$D = D^{Z=0} + (D^{Z=1} - D^{Z=0}) \cdot Z$$
$$\boldsymbol{D = D^{Z=0} + K \cdot Z}$$

Note que, pela restrição de exclusão, se K=0, então $D^{Z=1} = D^{Z=0}$. Isso implica que:

$$* \begin{cases} E(Y|Z=1, D=0) = E(Y|Z=0, D=0) \\ E(Y|Z=1, D=1) = E(Y|Z=0, D=1) \end{cases}$$

Ou seja, nesses casos Z não afeta D. Assim, E(Y|•) não depende de Z, mas apenas de D.

LATE - Demonstração

Colocando * na equação anterior, somem os termos *always* e *never*.

Assim, $E(Y|Z=1) - E(Y|Z=0) =$

= E(Y|Z=1,D=1)·{Pr(D=1|Z=1)-Pr(D=1|Z=0)} -

E(Y|Z=0,D=0)·{Pr(D=0|Z=0)-Pr(D=0|Z=1)} +

~~{E(Y|Z=1,D=1)-E(Y|Z=0,D=1)}·Pr(D=1|Z=0)~~ +

~~{E(Y|Z=1,D=0)-E(Y|Z=0,D=0)}·Pr(D=0|Z=1)~~

LATE - Demonstração

$$E(Y|Z=1) - E(Y|Z=0)$$
$$= [E(Y|Z=1, D=1)$$

LATE - Demonstração

O efeito médio do tratamento para os *compliers* é, rearranjando:

$$E(Y|Z=1,D=1)-E(Y|Z=0,D=0)=$$

$$=\frac{E(Y|Z=1)-E(Y|Z=0)}{\Pr(D=1|Z=1)-\Pr(D=1|Z=0)}$$

Estimador $\hat{\delta}_{VI,WALD}$!!!

LATE

- Note que *always* e *never* não respondem ao instrumento (sua atribuição ao tratamento não é determinada por Z).
- Se só há *compliers*, estes contribuem com toda a variação que gera a estimativa de V.I., pois só seu comportamento responde ao instrumento.
- Se tanto *compliers* quanto *defiers* estão presentes, a estimativa de Wald não tem interpretação causal bem definida. Em geral, assume-se que K≥0 e não há *defiers*.

Exemplo (p.204 – Morgan e Winship)

	Di=0	Di=1
Zi=0	$8000^{0,8}_{E(Y\|\bullet)=50}$	$1000^{0,1}_{E(Y\|\bullet)=60}$
Zi=1	$800^{0,08}_{E(Y\|\bullet)=50}$	$200^{0,02}_{E(Y\|\bullet)=58}$

$$E(Y|Z=1)=51{,}6$$
$$E(Y|Z=0)=51{,}111$$
$$E(D|Z=1)=0{,}2$$
$$E(D|Z=0)=0{,}111$$

EXEMPLO – MORGAN E WINSHIP

$$\left.\begin{cases} E(Y|D=1) = 60 \cdot \dfrac{1000}{1200} + 58 \cdot \dfrac{200}{1200} = 59{,}667 \\ E(Y|D=0) = 50 \cdot \dfrac{8000}{8800} + 50 \cdot \dfrac{800}{8800} = 50{,}00 \end{cases}\right\} \text{- OLS} = 9{,}667$$

$$\begin{cases} E(Y|Z=1) = 50 \cdot \dfrac{800}{1000} + 58 \cdot \dfrac{200}{1000} = 51{,}6 \\ E(Y|Z=0) = 50 \cdot \dfrac{8000}{9000} + 60 \cdot \dfrac{1000}{9000} = 51{,}111 \end{cases}$$

Exemplo – Morgan e Winship

Por independência, mesma distribuição de a, n e c para Z=0 e Z=1!!!

$$\Pr(a) = \frac{1000}{9000} = 0{,}111$$

Entre Z=0, 11,1% são *always*

Entre Z=1, 11,1% são *always*

$$\Pr(n) = \frac{800}{1000} = 0{,}8$$

Entre Z=0, 80% são *never*

Entre Z=1, 80% são *never*

EXEMPLO – MORGAN E WINSHIP

	D=0	D=1
Z=0	7200 *never* 800 *complier*	1000 *always*
Z=1	800 *never*	111 *always* 89 *compliers*

Razão A/N

$$\frac{1000}{7200} \approx \frac{111}{800}$$

Razão C/Total

$$\frac{800}{9000} \approx \frac{89}{1000}$$

EXEMPLO – MORGAN E WINSHIP

Como, nesse exemplo, mostrar que WALD é efeito sobre os *compliers*, vimos que $\hat{\delta}_{WALD} = 5{,}5$. Mas note que:

$$E\left(\delta \middle| \hat{C} = c\right) = E\left(Y^1 \middle| \hat{C} = c\right) - E(Y^0|\hat{C} = c)$$

Como achar?

$$E(Y|D = 1, Z = 1) = \frac{\Pr(c)}{\Pr(c) + \Pr(a)} \cdot E\left(Y \middle| \hat{C} = c\right) +$$

$$\frac{\Pr(a)}{\Pr(c) + \Pr(a)} \cdot E(Y|\hat{C} = c)$$

EXEMPLO – MORGAN E WINSHIP

$$58 = \frac{0{,}089}{0{,}089 + 0{,}111} \cdot E\left(Y \middle| \hat{C} = c\right) + \frac{0{,}111}{0{,}089 + 0{,}111} \cdot 60$$

$$\Longrightarrow E\left(Y^1 \middle| \hat{C} = c\right)_{D=1,Z=1} = 55{,}5$$

$$E(Y|D = 0, Z = 0) = \frac{\Pr(n)}{\Pr(n)+\Pr(c)} \cdot E(Y|D = 0, Z = 0, n) +$$

$$\frac{\Pr(c)}{\Pr(n)+\Pr(c)} \cdot E(Y^0|D = 0, Z = 0, c)$$

Exemplo – Morgan e Winship

$$50 = \frac{0{,}8}{0{,}8+0{,}089} \cdot 50 + \frac{0{,}089}{0{,}8+0{,}089} \cdot E(Y^0|D = 0, Z = 0, c) \Longrightarrow$$
$$E(Y^0|D = 0, Z = 0, c) = 50$$

$$\therefore\ \delta_{LATE} = E\left(Y^1\middle|\hat{C} = c\right) - E\left(Y^0\middle|\hat{C} = c\right) = 55{,}5 - 50$$
$$= 5{,}5$$

CONCLUSÕES

- Instrumento fraco é muito ruim;
- Se efeito heterogêneo causal ocorre, só deve usar V.I. se hipóteses 1 e 3 valerem, especialmente monotonicidade;
- Se efeito heterogêneo está presente, não deve combinar instrumentos via MQ2E e sim dar medida de LATE separada para os instrumentos que satisfazem monotonicidade.

Críticas e Extensões

Críticas

- LATE depende do instrumento, pois dependendo do instrumento, posso estar em *complier* ou *never* em cada caso.

Extensões

- Heckman e Vytlacil (2005): LATE é média ponderada de efeitos tratamento marginais.

EFEITO TRATAMENTO MARGINAL (MTE)

- Heckman e Vytlacil (2005) mostram que LATE pode ser visto como uma média ponderada de efeito tratamento marginal mais fundamentais.

Exemplo: Suponha que ao invés de 10% receberem um voucher de $3000, 10% recebem voucher que varia entre $1 e a mensalidade mais cara, sendo retirado de distribuição uniforme aleatoriamente.

Efeito Tratamento Marginal (MTE)

- Nesse caso, o tamanho do voucher é um instrumento válido Z, mantendo as mesmas hipóteses:
  - Z é atribuído aleatoriamente
  - Z possui efeito diferente de zero em D
  - Efeito Z em D é monotônico (nesse caso, essa hipótese estipula que a probabilidade real de receber o tratamento é maior para Δ’s com Z=Z” do que do que para os com Z=Z’ se Z”>Z’).

Efeito Tratamento Marginal (MTE)

- Heckman e Vytlacil (2005) define dois conceitos relacionados:

  a) Local instrumental variable (LIV): “VI local”

  É o caso limite de uma retirada dupla de VI de Z na qual Z” se aproxima de Z’, para quaisquer dois valores de Z tais que Z”>Z’.

  b) **Efeito Tratamento Marginal (MTE)**

  Cada LIV define um efeito tratamento marginal, o qual é a forma limite do LATE, no qual a VI é uma LIV.

Efeito Tratamento Marginal (MTE)

Exemplo: Pode-se formar LIV’s a partir de Z estratificando os dados pelos valores de Z e então considerando estratos vizinhos.

Dado uma amostra grande e programa voucher grande, LIV’s podem ser construídas para cada aumento de 1 dólar no voucher.

Cada LIV pode ser usada para estimar um LATE e estes podem ser considerados MTE.

O problema dessa abordagem é ter as LIV’s.