Técnicas Econométricas para Avaliação de Impacto

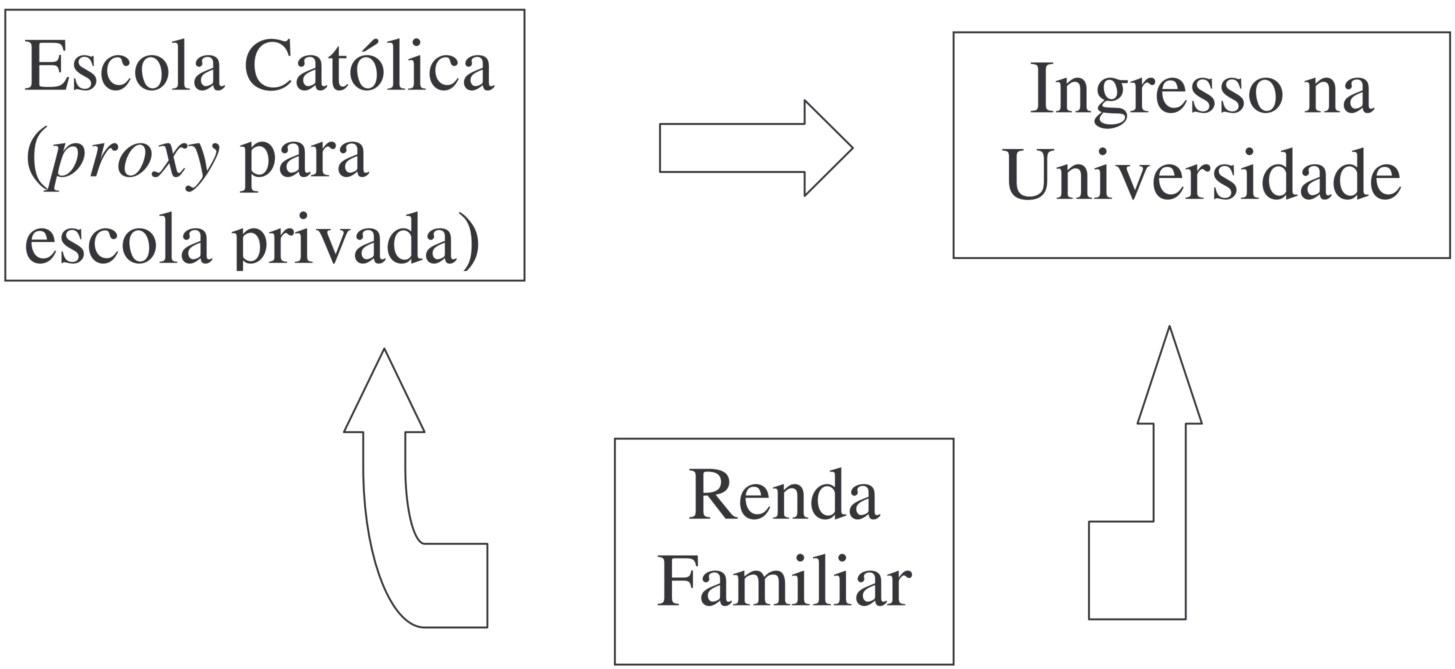
Variáveis Instrumentais

Guilherme Issamu Hirata
Centro Internacional de Pobreza (IPC/PNUD)

Brasília, 21 de maio de 2008.

Introdução

<u>Qualidade do Ensino</u>: Escola Pública X Escola Privada



<u>Salário e Educação:</u>

$$S = f(X, E) + u$$

O problema é que a habilidade determina salário e pessoas mais habilidosas buscam mais educação. Mesmo mecanismo anterior.

<u>Ocupação e Salário</u>

Incluir ou não ocupação na equação salarial? A escolha ocupacional pode não se dar por meio de fatores observáveis. Pode haver correlação espúria entre ocupação e salário.

Nos exemplos acima, temos:

$$Y = \beta_0 + \Sigma \beta_i X_i + \alpha D + v$$

$$E(v) = 0 \quad \mathrm{Cov}(\mathrm{X}, \mathrm{v}) = 0 \quad \mathrm{Cov}(D, \mathrm{v}) \neq 0$$

Isto é, há um fator determinando D que também determina Y ou é determinado por Y.

Basicamente, o método de Variável Instrumental busca eliminar do modelo essa correlação.

Nas três últimas aulas, a implementação dos métodos expostos tinham como pressupostos fundamentais:

- SUTVA
- Independência da média condicional

Variável observada:

$$Y = Y_1 D + Y_0 (1 - D)$$

onde

$$Y_1 = \mu_1 + u_1, \qquad \mathrm{E}(u_1) = 0$$

$$Y_0 = \mu_0 + u_0, \qquad \mathrm{E}(u_0) = 0$$

$$D = 1\left[D^* = Z\gamma + v > 0\right]$$

Assim,

$$Y = \mu_0 + (\mu_1 - \mu_0)D + u_0 + (u_1 - u_0)D$$

- Segundo termo representa o ganho médio relativo ao tratamento, para indivíduos com vetor X
- Quarto termo indica o ganho específico individual

Na estimação, viés pode aparecer em dois casos:

- Variável omitida que determina D e Y: seleção sobre observáveis
- Fatores não observáveis que determinam D e Y: seleção sobre não observáveis

<u>! No segundo caso, abandono do pressuposto de independência da média condicional</u>

Em outras palavras, $P(D=1 \mid Y, X) \neq P(D=1 \mid X) = p(x)$

Como contornar o problema: basicamente, a idéia é retirar de u o componente correlacionado com v

Métodos:

- Variável instrumental
- Heckman
- DD
- Modelos estruturais

Variáveis instrumentais

Exemplos de aplicação:

- Tostines: Vende mais porque é fresquinho ou é fresquinho porque vende mais?
- Oferta e demanda (Wright, 1928)
- Oferta de trabalho (Mroz, 1980)
- Educação (Altonji *et al*, 2002)
- Mortalidade (Jablon & Seltzer, 1975) dos veteranos de guerra

Obs: Método bastante empregado em casos de Omissão de Variáveis e Erros de Medida.

Identificação

Suposição 1: SUTVA

Suposição 2: Restrição de Exclusão

Seja Z uma variável aleatória. Z atende a esta suposição se $Y(z, D) = Y(z', D), \forall z, z' \in Z$.

Suposição não testável diretamente, uma vez que está baseada em um contrafactual.

Para facilitar o entendimento, suponha que Z é a variável de designação do tratamento, aleatoriamente distribuída na população, ou seja, a intenção do tratamento.

$Z = 1$ para o indivíduo elegível.
$Z = 0$ para o indivíduo não elegível.

Se todos os elegíveis recebem o tratamento, $Z = D$.
No entanto, esse é o caso ideal. Na realidade, temos quatro tipos de indivíduos:

	Z = 1, D = 0	Z = 1, D = 1
Z = 0, D = 0	*Never-taker*	*Complier*
Z = 0, D = 1	*Defier*	*Always-taker*

Suposição 3: Monotonicidade

As suposições 1 e 2 não garantem a identificação de um efeito de tratamento. Vejamos por que.

$$E[Y \mid Z = z] - E[Y \mid Z = z'] =$$

$$E[D(z)Y_1 + (1 - D(z))Y_0 \mid Z = z] - E[D(z')Y_1 + (1 - D(z'))Y_0 \mid Z = z'] =$$

$$E[Y_0 + (Y_1 - Y_0)D(z) \mid Z = z] - E[Y_0 + (Y_1 - Y_0)D(z') \mid Z = z']$$

Por causa da suposição 2, a equação acima torna-se:

$$E[D(z)-D(z')(Y_1-Y_0)]=$$

$$\Pr[D(z)-D(z')=1]\cdot E[Y_1-Y_0 \mid D(z)-D(z')=1]-$$

$$\Pr[D(z)-D(z')=-1]\cdot E[Y_1-Y_0 \mid D(z)-D(z')=-1]$$

A equação acima pode ser zero ou negativa, mesmo se, para todo indivíduo, $Y_1-Y_0>0$.

Para assegurar que isso não aconteça, assume-se uma restrição não verificável, mas bastante plausível: a monotonicidade do efeito de Z sobre D:

- Se, para $z>z'$, $D(z')=1$, então $D(z)=1$, para todos os indivíduos.

Observação: alternativamente, poderia ser assumido efeito de tratamento constante para todos os indivíduos, ou seja, restringir a possibilidade de heterogeneidade.

Suposição 4: $corr(Z, D) \neq 0$

Basicamente, essa suposição diz que Z apresenta um efeito sobre D. Dessa forma, Z também afeta Y, mas apenas indiretamente, via D. Isto é, pode-se definir um efeito médio causal de Z sobre D: $E[D(Z=1) - D(Z=0)]$.

Definição de **Variável Instrumental**: Z satisfaz as suposições 2 e 4.

Dadas as suposições 1 a 4:

Local Average Treatment Effect (LATE)

$$E\left[Y_1 - Y_0 \mid D(1) - D(0) = 1\right] = \frac{E\left[Y_1(1, D(1)) - Y(0, D(0))\right]}{E\left[D(1) - D(0)\right]}$$

LATE representa o impacto sobre os *compliers*, não sendo, em geral, representativo do efeito sobre todos os tratados.

LATE = ATE = ATT se o efeito de tratamento é homogêneo.

Importante: Não podemos identificar o grupo dos *compliers*, ou seja, os indivíduos que mudam de comportamento influenciados pelo instrumento.

Efeito de Tratamento Marginal

$$MTE = \left.\frac{\partial E[Y \mid X, Z]}{\partial \Pr[D = 1 \mid X, Z]}\right|_{Z=z}$$

- ATE = valor esperado de MTE, incluindo todos os indivíduos.
- ATT = valor esperado de MTE, excluindo os indivíduos que não participam do tratamento.
- LATE = valor esperado do MTE, para o intervalo de Z onde a taxa de participação difere.

Violação de pressupostos

SUTVA: se a variável de interesse de um indivíduo é afetada pela condição de tratamento das demais pessoas ou por algum outro tratamento que ocorre ao mesmo tempo, não podemos identificar o contrafactual (Rubin, 1986).

Restrição de exclusão:

$$Y = \alpha D + X'\beta + \varepsilon$$

$$\operatorname{cov}(X, e) = 0$$

$$\hat{Z} = X'\pi$$

$$\hat{D} = X'\beta + \lambda Z$$

$$v \equiv Z - X'\pi$$

$$u \equiv D - X'\beta - \lambda Z$$

$$Y = \alpha \hat{D} + X'\beta + \varepsilon$$

$$\hat{\alpha} = \alpha + \frac{\operatorname{cov}(v, \varepsilon)}{\lambda \operatorname{var}(v)}$$

- Quanto maior a correlação entre Z e D, isto é, quanto mais forte o instrumento, menos sensível é o estimador à violação da restrição de exclusão.
- Podemos ver também que, quanto maior o efeito de Z sobre Y, maior o viés do efeito do tratamento.

Monotonicidade:

$$LATE = \frac{E[Y_1(1, D(1)) - Y(0, D(0))]}{E[D(1) - D(0)]} =$$

$$E[Y_1 - Y_0 \mid D(1) > D(0)] +$$

$$\lambda\{E[Y_1 - Y_0 \mid D(1) < D(0)] - E[Y_1 - Y_0 \mid D(1) > D(0)]\}$$

$$\text{onde } \lambda = \frac{-P[D(1) - D(0) = -1]}{E[D(1) - D(0)]} < 0$$

- Mesmo se a proporção de *defiers* for baixa, o viés pode ser alto caso o instrumento seja fraco.

- Se o efeito médio do tratamento para *compliers* e *defiers* é o mesmo, então a violação da monotonicidade não produz viés.

$\underline{corr(Z,D)=0}$: Não há identificação do efeito do tratamento, uma vez que Z não representa um instrumento.

Estimação

ATE: Supondo ausência de heterogeneidade do impacto, MQ2E produz resultados consistentes e eficientes. No entanto, há controversa sobre a melhor maneira de estimar o primeiro estágio.

$$P(D=1 \mid X, Z) = G(X, Z; \gamma)$$

$$Y = \beta + \beta_i X_i + \alpha D + u$$

ATE incluindo interações

$$Y = \beta + \beta_i X_i + \alpha D + D(X - \overline{X})\delta + u$$

$$Instrumentos = \hat{G}, \hat{G}(X - \overline{X})$$

LATE: $Y = \delta + \alpha D + u$, usando Z como instrumento.

Múltiplos Instrumentos

Baseados no estimador de Wald, Angrist e Imbens (1995) identificam um ATE por meio da média ponderada do que eles chamam de Resposta Causal Média (*average causal response*, ACR).

Considerando k variáveis instrumentais discretas, pode-se construir uma única variável Z com $k+1$ valores.

O parâmetro de efeito de tratamento seria:

$$\beta = \sum_{Z} \mu_k \frac{E[Y \mid Z = k] - E[Y \mid Z = k-1]}{E[D \mid Z = k] - E[D \mid Z = k-1]}$$

onde $\sum_{k=1}^{K} \mu_k = 1$.

O peso é proporcional à $E(D \mid Z = k) - E(D \mid Z = k - 1)$, ou seja, à qualidade do instrumento. Além disso, β dá mais peso às estimativas de Wald que estão mais próximas ao centro da distribuição de Z, uma vez que μ_k é proporcional a $\Pr(Z > k)(1 - \Pr(Z > k))$.

Aleatorização e Variável Instrumental

Tanto a aleatorização da eligibilidade ao tratamento, quanto a aleatorização da participação dentre os elegíveis garante que as mesmas não sejam correlacionadas com fatores pessoais ou sociais e, portanto, tratados e controles não apresentam diferenças, na média, nas características não observáveis.

Por construção $X \perp D$.

$$E[Y_1 - Y_0 \mid X] = g(Y_1 \mid X) - g(Y_0 \mid X) - E(U_1 - U_0 \mid X)$$

ATE = ATT quando:

- Não há heterogeneidade
- Há heterogeneidade, mas há aleatorização da participação.

***Propensity Score* e Variável Instrumental**

Ichimura e Taber(2001) mostram como o método de *propensity score* pode ser considerado um caso especial do método de variável instrumental

Pressupostos

- Independência condicional: $Y_1, Y_0 \perp D \mid X$
- Suporte comum: $0 < p(\mathbf{X}) < 1$

Podemos assumir que D é uma restrição de exclusão em $E[Y_i \mid X]$.

Assim, suponha que Z seja uma variável instrumental, não necessariamente binária. Temos que $Y_1, Y_0 \perp Z \mid X$, ou seja:

$$E[Y \mid Z, X] = E[Y_0 \mid X] + E[Y_1 - Y_0 \mid X]P(Z, X)$$

onde $P(Z, X) \equiv \Pr(D = 1 \mid Z, X)$

Dois métodos de estimação

<u>Primeiro:</u> regressão

Defina $P(X) = E[P(Z,X) \mid X]$ e $Q(X) = E[P^2(Z,X) \mid X]$

Pode-se mostrar que:

$$E[Y_1 - Y_0 \mid P(X), Q(X)] = \frac{Cov[Y, P(Z,X) \mid P(X), Q(X)]}{Var[P(Z,X) \mid P(X), Q(X)]}$$

<u>Segundo:</u> Diferença

Defina $Q(X) = \Pr[Z = z \mid Z = z \text{ ou } \mathrm{Z} = \mathrm{z}', \mathrm{X}], \forall z, z'$

$$E[Y_1 - Y_0 \mid Q(X), Z = z, z'] =$$

$$E\left(\frac{Y}{P(z, X) - P(z', X)} \mid Z = z, Q(X) \right) - E\left(\frac{Y}{P(z, X) - P(z', X)} \mid Z = z', Q(X) \right)$$

Se Z = D, P(Z,X) = D, e Q(X) = P(X), ou seja, temos o propensity score.

Considerações finais

- Método promissor teoricamente, mas pode frustrar na prática
- Há um problema de eficiência em pequenas amostras
- Sensibilidade do LATE ao instrumento
- Controversa: Qual parâmetro é factível de ser estimado no contexto do método IV ?