

Técnicas Econométricas para
Avaliação de Impacto

**Problemas de Contaminação
na Validação Interna**

Rafael Perez Ribas
Centro Internacional de Pobreza

Brasília, 18 de junho de 2008

Introdução

- Validação Interna é quando um estudo precisa ser consistente em termos dos parâmetros que estão sendo estimados.
- Problemas de Contaminação podem ocorrer principalmente por dois motivos:
 - Quanto o grupo de controle recebe um tratamento alternativo (Viés de Substituição)
 - Quanto o grupo de tratamento não quer receber o tratamento (Viés de Abandono / Desistência)
 - Casos onde ITT é diferente do ATE/ATT
- Em ambos os casos, o Efeito Médio do Tratamento é subestimado.

- Outro problema de Contaminação ocorre quando:
 - O grupo não-tratado também é afetado pelo programa.
 - *Peer Effect*
 - *Externality*
 - *Spill-Over Effect*
 - Duas razões:
 - Efeito de Equilíbrio Geral,
 - mudança nos “preços” da economia ou
 - mudança na probabilidade de eventos;
 - Efeito de Interações Sociais,
 - mudança nas preferências.

○ Exemplos:

- Distribuição de Ticket Alimentação aumenta a demanda por alimentos e modifica os preços dos alimentos;
- Treinamento aumenta a produtividade e o salário médio dos trabalhadores, mas gera demissões;
- Programa de transferência de renda aumenta a taxa de poupança das famílias e amplia as oportunidades de crédito;
- Tratamento contra vermes reduz o número de hospedeiros;
- Estudantes que recebem reforço facilitam o aprendizado de seus colegas;
- Mulheres conversam com uma vizinha participando de um curso sobre sexualidade.

Modelo de Escolha pelo Tratamento

- Para cada tratamento disponível (s), o retorno é representado por um valor presente (no período $t=0$):

$$V_s = \begin{cases} \sum_{t=0}^T \delta_t (Y_{s,t} - c_{s,t}) & s \in \{1, 2, \dots, S\} \\ \sum_{t=0}^T \delta_t Y_{0,t} & s = 0 \end{cases}$$

- Assim, cada indivíduo escolhe um tratamento D que maximiza o seu retorno esperado a valores presentes:

$$D = \arg \max_{s \in S} E[V_s]$$

$$s^* = D \Rightarrow E[V_D]$$

- Assumindo que estamos interessados em avaliar o tratamento $s=1$ e existe um grupo previamente aceito para participar do tratamento ($A=1$).

haverá um sorteio indicando quem terá a oportunidade de receber este tratamento e quem não terá:

$$R = 1 \Rightarrow D \in s$$

$$R = 0 \Rightarrow D_{-1} \in s_{-1}$$

- A diferença de médias entre os grupos de sorteados ($R=1$) e de não sorteados ($R=0$) representa um estimador do **Efeito Médio da Disponibilidade do Tratamento:**

$$\tau_0 = E[V_D \mid A = 1, R = 1] - E[V_{D_{-1}} \mid A = 1, R = 0]$$

- Este estimador não necessariamente representa o parâmetro do **Efeito Médio do Tratamento** em si:

$$\tau_1 = E[V_1 - V_0 \mid A = 1, D = 1]$$

- Os dois Efeitos serão iguais se:
 - **Não há abandono**, todos os sorteados continuarão participando do tratamento:
 $R = 1 \Rightarrow D = 1$
 - **Não há substituição**, todos os não-sorteados ficarão sem tratamento algum:
 $R = 0 \Rightarrow D_{-1} = 0$
- Caso contrário, $\hat{\tau}_0$ será um estimador enviesado de τ_1 .

Viés de Abandono e Viés de Substituição

- Mesmo em um experimento, a comparação entre os grupos de tratamento e controle estaria estimando o Efeito da Disponibilidade do Tratamento,
 - ao invés do Efeito do Tratamento em si.
- Quando não há opção de outro tratamento, mas há abandono no grupo de tratamento,
 - Assumimos a seguinte condição:
$$E[Y_0 | A = 1, R = 1, Q = 1] = E[Y_0 | A = 1, R = 0, Q = 1]$$
onde $Q=1$ representa permanecer recebendo o tratamento.

○ Assim o Efeito do Tratamento será:

$$\tau_1 = \frac{\tau_0}{p} = \frac{E[Y | A = 1, R = 1] - E[Y | A = 1, R = 0]}{p}$$

● Intuição:

○ Teorema Horvitz-Thompson (1952)

○ Para a proporção $(1 - p)$ do grupo de tratamento onde $Q = 0$:

$$E[Y | A = 1, R = 1, Q = 0] = E[Y | A = 1, R = 0, Q = 0]$$

$$\Rightarrow E[\tau_0 | Q = 0] = 0$$

○ Quanto menor for p (ou maior for o grupo $Q = 0$), maior será a subestimação do Efeito do Tratamento.

- Quando p não é aleatório, podemos condicioná-lo a X ,
 - Um *Propensity Score* para a probabilidade de permanecer com o tratamento.
 - Exemplo paralelo é o problema de atrito em painéis.

- Quando os dois problemas ocorrem (Substituição e Abandono),
 - Podemos assumir que:

$$E[Y_1 - Y_0 \mid D = 1, R = 1] = E[Y_{D_{-1}} - Y_0 \mid D_{-1} > 1, R = 0]$$
 - O efeito do tratamento entre os tratados é igual ao efeito do tratamento substituto entre os controles.

- Assim, a comparação entre os grupos fornece o seguinte estimador:

$$\tau_0 = \frac{p \cdot E[Y_1 - Y_0 \mid D = 1, R = 1]}{-q \cdot E[Y_{D_{-1}} - Y_0 \mid D_{-1} > 1, R = 0]}$$

onde q é a proporção do grupo de controle que recebe um tratamento alternativo.

- Assumindo que os tratamentos são iguais em termos de efeitos:

$$\tau_0 = (p - q) \cdot E[Y_1 - Y_0 \mid D = 1]$$

$$\Rightarrow \tau_1 = \frac{\tau_0}{(p - q)} = E[Y_1 - Y_0 \mid D = 1]$$

- Novamente, q pode ser estimado condicionado a X .

- **Importante:** Neste caso, sempre quando $p > 0$ ou $q > 0$, o impacto estará sendo subestimado.
- E quando o tratamento alternativo possui efeito distinto do tratamento de interesse?
 - Avaliação do Efeito de Tratamentos Múltiplos (aula passada)

Externalidade

- Suponha que um tratamento para vermes em uma escola:

$T_{ij} = 1$ se a criança i na escola j recebe o tratamento

$D_j = 1$ se a escola j recebe o tratamento

Importante: $T_{ij} = 1 \Rightarrow D_j = 1$

N_{di}^D é o número de crianças frequentando escolas tratadas a uma distância d da criança i .

N_{di} é o número de crianças frequentando qualquer escola a uma distância d da criança i .

- Modelo completo:

$$Y_{ij} = \alpha + \tau^p T_{ij} + \tau^{ew} D_j + \beta X_{ij} + \sum_d (\tau_d^{eb} N_{di}^D) + \sum_d (\lambda_d N_{di}) + u_j + e_{ij}$$

- Assim:

τ^p = Efeito Médio Direto do Tratamento sobre a Criança

τ^{ew} = Efeito Médio de Externalidade dentro da Escola

$\sum_d \tau_d^{eb} \bar{N}_{di}^D$ = Efeito Médio de Externalidade para fora da Escola

$\tau^{ew} + \sum_d \tau_d^{eb} \bar{N}_{di}^D$ = Efeito Médio sobre a Criança Não-Tratada na Escola Tratada

$\tau^p + \tau^{ew} + \sum_d \tau_d^{eb} \bar{N}_{di}^D$ = Efeito Médio sobre a Criança Tratada

- Se fosse estimada simplesmente a função:

$$Y_{ij} = \alpha + \tau \cdot T_{ij} + \beta X_{ij} + e_{ij}$$

τ representaria a diferença entre o grupo de tratamento e todas as outras crianças (afetadas ou não pelo tratamento) do grupo de controle.

- SUTVA violado.
- Se o grupo de controle é formado por crianças em áreas distantes das crianças tratadas, não há problema.
 - Mas, à medida que aumenta o número de crianças próximas no grupo de controle, o coeficiente perde a sua interpretação.
- Problema semelhante ao de Viés de Substituição,
 - pois o grupo de controle também está modificando seu resultado.