Gabarito Lista 1 (Exercícios adicionais) - Econometria II (2022)

Professora: Fabiana Rocha
Monitor: Thiago Pastorelli Rodrigues

Exercício 1. Sabendo que $ATE = \pi ATT + (1 - \pi)ATU$, com $ATT = E[Y^1|D = 1] - E[Y^0|D = 1]$ e $ATU = E[Y^1|D = 0] - E[Y^0|D = 0]$, com $D = \mathbb{1}\{Tratamento = 1\}$ e π é a parcela da amostra no grupo de tratamento, resolva os seguintes itens:

i. Demonstre a seguinte expressão:

$$\underbrace{E[Y^1|D = 1] - E[Y^0|D = 0]}_{\text{Diferença simples dos resultados observados}} = \underbrace{ATE}_{\text{Efeito médio do tratamento}} + \underbrace{(E[Y^0|D = 1] - E[Y^0|D = 0])}_{\text{Viés de seleção}} + \underbrace{(1 - \pi)(ATT - ATU)}_{\text{Viés de heterogeneidade}} \quad (1)$$

Primeiramente¹, usemos a seguinte relação dada no enunciado:

$$\begin{aligned} ATE &= \pi ATT + (1 - \pi)ATU \\ &= \pi(E[Y^1|D = 1] - E[Y^0|D = 1]) + (1 - \pi)(E[Y^1|D = 0] - E[Y^0|D = 0]) \\ &= [\pi E[Y^1|D = 1] + (1 - \pi)E[Y^1|D = 0]] - \\ &\quad - [\pi E[Y^0|D = 1] + (1 - \pi)E[Y^0|D = 0]] \\ &= \pi E[Y^1|D = 1] + E[Y^1|D = 0] - \pi E[Y^1|D = 0] + \\ &\quad + \pi E[Y^0|D = 1] - E[Y^0|D = 0] + \pi E[Y^0|D = 0] \\ &= \pi E[Y^1|D = 1] + E[Y^1|D = 0] - \pi E[Y^1|D = 0] + \\ &\quad + \pi E[Y^0|D = 1] - E[Y^0|D = 0] + \pi E[Y^0|D = 0] + \\ &\quad + (E[Y^1|D = 1] - E[Y^1|D = 1]) + (E[Y^0|D = 1] - E[Y^0|D = 1]) + \\ &\quad + (E[Y^0|D = 0] - E[Y^0|D = 0]) \end{aligned}$$

Passando o ATE para o lado direito da equação, temos:

$$\begin{aligned} 0 &= ATE - \pi E[Y^1|D = 1] - E[Y^1|D = 0] + \pi E[Y^1|D = 0] + \\ &\quad + \pi E[Y^0|D = 1] + E[Y^0|D = 0] - \pi E[Y^0|D = 0] + \\ &\quad + E[Y^1|D = 1] - E[Y^0|D = 1] + E[Y^0|D = 1] - E[Y^0|D = 0] \\ &= ATE - \pi E[Y^1|D = 1] - E[Y^1|D = 0] - \pi E[Y^1|D = 0] + \\ &\quad + \pi E[Y^0|D = 1] + E[Y^0|D = 0] - \pi E[Y^0|D = 0] - \\ &\quad - E[Y^1|D = 1] + E[Y^1|D = 1] - E[Y^0|D = 1] + \\ &\quad + E[Y^0|D = 1] - E[Y^0|D = 0] + E[Y^0|D = 0] \end{aligned}$$

Rearranjando os termos de modo que a diferença simples dos resultados obser-

¹Essa demonstração segue muito proximamente a demonstração disponível em Cunningham (2021) - Causal Inference - The Mixtape.

vados esteja no lado esquerdo, temos:

$$\begin{aligned}
E[Y^1|D=1] - E[Y^0|D=0] &= ATE + (E[Y^{0|D=1}] - E[Y^0|D=0]) + E[Y^1|D=1] - \\
&\quad - \pi E[Y^1|D=1] - E[Y^1|D=0] + \pi E[Y^1|D=0] - \\
&\quad - E[Y^0|D=1] + \pi E[Y^0|D=1] + E[Y^0|D=0] - \\
&\quad - \pi E[Y^0|D=0] \\
&= ATE + (E[Y^0|D=1] - E[Y^0|D=0]) + (1 - \pi)E[Y^1|D=1] - \\
&\quad - (1 - \pi)E[Y^1|D=0] + (1 - \pi)E[Y^0|D=0] - \\
&\quad - (1 - \pi)E[Y^0|D=1] \\
&= ATE + (E[Y^0|D=1] - E[Y^0|D=0]) + \\
&\quad + (1 - \pi)[E[Y^1|D=1] - E^1|D=0] - \\
&\quad - (1 - \pi)[E[Y^1|D=0] - E[Y^0|D=0]] \\
&= ATE + (E[Y^0|D=1] - E[Y^0|D=0]) + (1 - \pi)(ATT - ATU)
\end{aligned}$$

Conforme queríamos demonstrar.

- ii. Suponha que o efeito de tratamento seja constante entre os indivíduos. O que ocorre com o viés de heterogeneidade nesse caso?

O viés de heterogeneidade nesse caso desaparece, pois ele capta justamente os possíveis diferentes impactos do tratamento entre indivíduos no grupo de controle e tratamento. Mais precisamente, o viés de heterogeneidade é a diferença entre o impacto do tratamento entre os grupos de tratamento/controlado, multiplicado pela parcela da população (amostra geralmente) no grupo de tratamento.

- iii. Suponha agora que $(Y^1, Y^0) \perp\!\!\!\perp D$, isto é, o tratamento independe das consequências do tratamento sobre ambos os grupos de análise (controle e tratamento). Qual o impacto disso sobre a equação (1)? Adicionalmente, quais as hipóteses necessárias para que $E[Y^1|D=1] - E[Y^0|D=0] = ATE$.

Nesse caso, o viés da seleção desaparece, pois as diferenças existentes entre o grupo de controle e de tratamento são unicamente o fato de um receber o tratamento, enquanto ou outro não o recebe. Mais precisamente, o viés de seleção é a diferença existente entre o grupo de tratamento e controle caso eles nunca recebessem o tratamento. Como $(Y^1, Y^0) \perp\!\!\!\perp D$, então o grupo de tratamento e controle são idênticos sem o tratamento e, portanto, o viés de seleção se anula. Em relação à segunda questão, as hipóteses para que o ATE (efeito de tratamento médio) seja equivalente à diferença simples dos resultados observados são três segundo Cunningham (2021):

- Observa-se os resultados do experimento para grupo de controle e tratamento.
- Sabe-se quais indivíduos receberam ou não o tratamento.
- $(Y^1, Y^0) \perp\!\!\!\perp D$.

Exercício 2. Em um artigo recente, Evans e Schwab (1995) estudam os efeitos de frequentar uma escola de ensino médio católica sobre a probabilidade de cursar uma faculdade. Concretamente, defina *faculdade* como uma variável binária igual a um caso o estudante esteja na faculdade e zero caso contrário. Defina *EMCat* como uma variável binária igual a um caso o estudante frequente uma escola de ensino médio católica. Um modelo de probabilidade linear é:

$$faculdade = \beta_0 + \beta_1 EMCat + X'\theta + u$$

onde X é uma matriz de controles que inclui gênero, raça, renda familiar e educação dos pais.

a) Por que $EMCat$ pode ser correlacionado com u ?

De modo geral, alunos melhores e mais sérios tendem a ir para a faculdade, e estes mesmos alunos podem ter probabilidade maior ou menor de frequentarem escolas católicas. Então, a correlação entre $EmCat$ e u é um exemplo do problema de autoseleção: alunos se autoselecionam para escolas de ensino médio católicas, em lugar de serem atribuídos aleatoriamente a estes colégios.

b) Evans e Schwab tinham dados sobre a nota de um teste padronizado feito quando cada estudante cursava o 2º ano. O que pode ser feito com essa variável para melhorar a estimativa *ceteris paribus* de frequentar uma escola católica de ensino médio?

A nota num teste padronizado é uma medida de habilidade do aluno, logo podemos usá-la como *proxy* numa regressão por MQO. Utilizar esta medida deve trazer uma melhoria para o modelo, pois temos algum controle para habilidade do aluno.

c) Defina $RelCat$ como uma variável binária igual a um se o estudante for católico. Detalhe os requisitos necessários para que essa seja uma VI válida para $EMCat$ no modelo original. Qual destes requisitos pode ser testado?

Primeiramente, é necessário que $RelCat$ seja não correlacionada com características não-observáveis incluídas no termo de erro u (como motivação e habilidade). Em geral, isto é válido se somente ser católico não o transforma num aluno melhor, o que parece ser uma suposição razoável. Se o fato de ser católico é correlacionado com a motivação dos alunos, é mais controverso. Note que isto afeta a restrição de exclusão, que não é testável. O segundo requisito é que ser católico tenha algum efeito sobre frequentar uma escola de ensino médio católica, controlando pelos demais fatores exógenos que aparecem no modelo estrutural. Podemos testar isso estimando a seguinte forma reduzida:

$$EmCat = \pi_0 + \pi_1 RelCat + X'\theta + u$$

E testar a significância do coeficiente π_1 .

d) Não surpreendentemente, o fato de ser católico tem um efeito positivo significativo sobre frequentar o ensino médio numa escola católica. Você acha que $RelCat$ é um instrumento convincente para $EMCat$?

Como mencionado, os autores encontram uma relação positiva. Ainda, é razoável supor que $RelCat$ é exógena na equação estrutural, portanto é um instrumento válido.

Exercício 3. Considere um modelo simples para estimar o efeito da propriedade de um computador pessoal na nota média de graduação de formandos de uma grande universidade pública:

$$nmgrad = \beta_0 + \beta_1 PC + u$$

onde PC é uma variável binária indicando a propriedade de um computador pessoal.

a) Por que a propriedade de um PC pode estar correlacionada com u ?

É um fato bem estabelecido na literatura que o status socioeconômico afeta o desempenho dos alunos. O termo de erro u contém, entre outras coisas, a renda familiar, que espera-se ter um efeito positivo sobre a nota média na graduação, e também ser positivamente correlacionada com a propriedade de um PC.

b) Explique por que PC provavelmente está correlacionada com a renda anual dos pais. Isso significa que a renda anual dos pais seria uma boa VI para PC ? Por quê?

Famílias com renda mais alta terão mais recursos para comprar computadores para seus filhos. Logo, um dos requisitos para ser variável instrumental é satisfeito. No entanto, como discutimos acima a renda familiar tem efeito positivo sobre a nota média de graduação, então a renda média dos pais não é exógena na equação de interesse (isto é, não vale a restrição de exclusão). Caso tivéssemos uma variável de renda dos pais, usaríamos esta variável como controle na equação estrutural, e caso essa fosse a única variável omitida correlacionada com PC poderíamos estimar o modelo por MQO.

c) Suponha que, quatro anos atrás, a universidade tenha concedido subvenções para a compra de computadores a aproximadamente metade dos alunos novos, e que os alunos que receberam estas subvenções tenham sido escolhidos aleatoriamente. Explique cuidadosamente como você usaria essa informação para construir uma variável instrumental para PC .

Primeiramente note que este é um experimento natural que afeta somente alguns estudantes. Alguns estudantes que comprem computadores por conta da subvenção provavelmente não teriam comprado caso esta subvenção não existisse. Agora, para construir a variável instrumental defina uma *dummy* de concessão igual a um caso o estudante tenha recebido a concessão para compra de computadores e zero caso contrário. Caso esta concessão tenha sido distribuída aleatoriamente, ela será não correlacionada com o erro (ou seja, vale a restrição de exclusão). Além disso, esta *dummy* deve ser correlacionada com a probabilidade de ter um PC (particularmente, espera-se que a concessão aumente a probabilidade de possuir um PC).

Exercício 4. Considere o seguinte modelo para medir o efeito de um programa de escolha de escola sobre o desempenho em um teste padronizado:

$$nota = \beta_0 + \beta_1 escolha + \beta_2 renda_{fam} + u_1$$

onde $nota$ é a nota de um teste de âmbito estadual, $escolha$ é uma variável binária indicando se o aluno frequentou uma escola de sua própria escolha no último ano e $renda_{fam}$ é a renda familiar. A variável instrumental de $escolha$ é $conc$, o montante em dólares concedido aos alunos para ser usado como pagamento da anuidade da escola particular de sua escolha. O montante da concessão difere conforme o nível de renda familiar, razão pela qual controlamos $renda_{fam}$ na equação.

a) Mesmo com $renda_{fam}$ na equação, por que $escolha$ pode ser correlacionada com u_1 ?

Mesmo para um dado nível de renda, alguns alunos são mais motivados ou mais habilidosos do que outros, e podem receber maior apoio de suas famílias do que outros alunos. Ou seja, neste caso é possível que ocorra um problema de auto-seleção, que implica que estudantes que iriam melhor de qualquer forma são mais propensos a frequentarem uma escola de sua própria escolha.

b) Se no interior de cada classe de rendimento os montantes de concessão fossem atribuídos aleatoriamente, $conc$ seria não correlacionada com u_1 ?

Supondo que a forma como incluímos $renda_{fam}$ na equação estrutural está correta, então $conc$ seria não correlacionada com u_1 . Como estamos controlando pela renda das famílias, a atribuição aleatória de $conc$ dentro de cada classe de renda implica que a atribuição deste montante não está correlacionada com fatores não observáveis, como motivação do aluno, habilidade ou apoio familiar.

c) Escreva a forma reduzida da equação de *escolha*. O que é necessário para que *conc* seja parcialmente correlacionado com *escolha*?

Neste caso, a forma reduzida é dada por:

$$escolha = \pi_0 + \pi_1 conc + \pi_2 renda\ fam + v_1$$

Para que *conc* seja parcialmente correlacionada com *escolha* devemos ter $\pi_1 \neq 0$ na forma reduzida acima. Ou seja, após controlar pela renda familiar é necessário que o montante de concessão tenha algum efeito sobre a escolha da escola. É razoável supor que isto vale, uma vez que o montante de concessão varia dentro das classes de renda.

d) Escreva a equação da forma reduzida de *nota*. Explique por que esta equação é útil.

A forma reduzida para *nota* é dada por:

$$nota = \pi_3 + \pi_4 conc + \pi_5 renda\ fam + v_2$$

Esta equação é útil pois nos permite observar diretamente o efeito do montante de concessão sobre a nota do teste, mantendo a renda familiar constante, o que pode ser de interesse em alguns casos.

Exercício 5. Assinale se as afirmativas são verdadeiras (V) ou falsas (F) e justifique sua resposta.

a) O problema de endogeneidade surge quando as variáveis explicativas são correlacionadas entre si.

Falso. O problema de endogeneidade ocorre quando as variáveis explicativas incluídas no modelo são correlacionadas com o termo de erro não observável u .

b) Uma boa variável instrumental deve ser uma variável exógena excluída da equação estrutural e que tenha alguma correlação com a explicativa endógena.

Verdadeiro. Sendo u o termo de erro não-observável e x uma variável endógena num dado modelo, para que z seja uma boa variável instrumental para deve valer: i) $Cov(x, z) \neq 0$ e ii) $Cov(z, u) = 0$.

c) O estimador de variável instrumental será não-viesado sob as hipóteses usuais associadas a esse estimador.

Falso. Uma importante característica do método de variáveis instrumentais é que caso haja uma variável endógena (isto é, o instrumento seja de fato necessário), o estimador de variáveis instrumentais é em geral viesado, porém consistente. Este viés será maior em amostras menores, por isso é preferível utilizar amostras grandes ao utilizarmos o estimador de variáveis instrumentais.

d) O uso de uma variável *proxy* e o método de variáveis instrumentais são duas possíveis formas de lidar com o viés de variável omitida.

Verdadeiro. Mas atente para a diferença entre os dois métodos. Ao utilizar uma variável *proxy* buscamos resolver o problema de variável omitida substituindo a variável não observada por uma variável proxy, ou seja esperamos que a variável proxy seja fortemente correlacionada com a variável não observada. No caso do método de variáveis instrumentais, utilizamos um instrumento que não deve ser correlacionado com a variável não observada. Neste caso reconhecemos a existência do problema de variável

omitida, e buscamos identificar os parâmetros de interesse utilizando fatores exógenos.

e) Um instrumento fraco terá boas propriedades assintóticas.

Falso. O problema de instrumento fraco ocorre quando o instrumento utilizado tem uma baixa correlação com a variável endógena. Este problema pode levar a um grande viés assintótico (inconsistência) do estimador de variáveis instrumentais, mesmo se a correlação entre o erro e o instrumento for baixa.