

Econometria II - IV

Prof: *Daniel Domingues dos Santos*

Monitor: Lucas Gerez Foratto

Sumário

1	Lembrar de perguntar aos alunos	1
2	Variáveis Instrumentais	1
2.1	Introdução	2
2.2	Intuição	2
2.3	Formalização	3
2.4	MQO2E: Estimando em 2 Estágios	4

1 Lembrar de perguntar aos alunos

Monitoria referente ao dia 26/09/2018.

- Todo mundo está usando o Wooldridge (2015)?
- Comentar sobre as notas de aula antigas do Daniel. Elas são extremamente rigorosas e bem completas. Aconselho dar uma olhada quem estiver muito confortável na matéria (**primeiro leiam o Wooldridge (2015)!**)
- Gente vcs viram que eu botei as resoluções do Wooldridge (2015) no *Moodle*? Fazer alguns exercícios depois de estudar a matéria ajuda muito a ter mais clareza sobre o assunto.
- Sobre aplicações dos modelos: eu deixei no *Moodle* um ótimo site com aplicações econométricas de praticamente todos os modelos estudados em aula, para quem se interessar, esse é uma excelente forma de aprender como estimar no computador e interpretar as saídas de cada software (R e STATA).
- Ver se todo mundo tem acesso ao drive com as provas dos anos anteriores.

2 Variáveis Instrumentais

A ideia aqui não é fornecer um material que compreenda todos os aspectos de Variáveis Instrumentais. O nosso intuito é esclarecer possíveis dúvidas que podem surgir na leitura do Wooldridge (2015), por isso, vamos tentar manter a notação do livro.

Mais uma vez, uma bom material *complementar* pode ser o site da *Ani Katchova* nele há uma breve discussão sobre o que são Variáveis Instrumentais, como seria um exemplo de aplicação desse caso e como poderíamos estimar tanto no R quanto no STATA.

2.1 Introdução

Estamos no mundo das Cross Section, com a presença de endogeneidade do termo de erro u , desta forma, **não** podemos mais garantir a hipótese de $E(u|X) = 0$. É fácil verificar sem a hipótese de exogeneidade, os estimadores serão viesados e inconsistentes.

Tendo em vista isso, vamos mostrar como a estimação por variáveis instrumentais pode nos ajudar a solucionar o problema de inconsistência dos estimadores. Segue um exemplo diferente do utilizado pelo Wooldridge (2015) para vocês terem mais opções de análise.

2.2 Intuição

Suponha que você é um pesquisador e queira estimar o efeito do consumo de café nas horas de sono na população. Assim, um primeiro passo seria montar um modelo linear sob o qual café é uma variável explicativa da horas dormidas:

$$\text{Horas de Sono} = \beta_0 + \beta_1 \text{café} + u \quad (1)$$

Suponha agora que a população decida dormir mais horas (saiu no JN que é bom pra saúde). Assim, as pessoas acreditam que o café diminui o sono e vão diminuir o seu consumo para dormir mais. Entretanto, além dessa diminuição os indivíduos adotam uma porção de outras coisas para dormir mais (comem coisas mais leves, fazem um relaxamento, etc..). Assim, dados que só observamos o consumo de café, podemos garantir que ele realmente afetou o número de horas dormidas?

Queremos então encontrar uma variável que explique variações da variável explicativa, no caso o consumo de café, mas **não** expliquem a produtividade. Isso no ajuda a inferir a direção da causalidade (eu digo direção porque os estimadores continuam viesados, porém com um bom instrumento, eles serão consistentes.)

Quem seria um bom candidato? **O preço do café!** Imagine que houve uma forte geada no sul de minas (só o Estado de minas produz 59% do café no Brasil, segundo o Censo Agro de 2018), de modo que o preço do café suba consideravelmente. Desta forma, temos um choque **exógeno** no consumo do café e, a partir disso, podemos observar se houve ou não alguma diferença nas horas de sono da população.

O preço do café (Variável Instrumental que chamaremos de z) só vai afetar as suas horas dormidas através da variável café. Vamos falar então em termos formais.

2.3 Formalização

Dado a equação 1, sabemos que:

$$Cov(cafe, u) \neq 0 \tag{2}$$

Mas, a partir da variável instrumental z , preço do café, temos que:

$$\begin{aligned} Cov(z, u) &= 0 \\ Cov(z, cafe) &\neq 0 \end{aligned} \tag{3}$$

Desta forma, se aplicarmos a covariância de z na equação 1, teremos:

$$\beta_1 = \frac{Cov(z, \text{Horas de Sono})}{Cov(z, cafe)} \tag{4}$$

Portanto, o estimador de variáveis instrumentais para este caso será:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (z_i - \bar{z})(\text{Horas de Sono}_i - \overline{\text{Horas de Sono}})}{\sum_{i=1}^n (z_i - \bar{z})(cafe_i - \overline{cafe})} \tag{5}$$

Note que, como a variável instrumental que garante ortogonalidade, nós faremos as hipóteses de exogeneidade, e de variância condicional condicionadas em z .

2.4 MQO2E: Estimando em 2 Estágios

Dado a solução de Variáveis Instrumentais, a pergunta agora é: como estimar?

Quando temos apenas uma endógena e múltiplos instrumentos

Seguindo a nomenclatura do Wooldridge (2015), chamamos de z as variáveis utilizadas como instrumentos e, para este caso, y_2 a variável endógena (que equivale ao consumo de café no nosso exemplo). Assim, teremos a chamada equação reduzida:

$$y_2 = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \pi_3 z_3 + v_2 \quad (6)$$

Onde $E(v_2) = 0$, $Cov(z_1, v_2) = 0$, $Cov(z_2, v_2) = 0$, $Cov(z_3, v_2) = 0$. Se isso for válido, todas as variáveis z são instrumentos e o melhor instrumento para y_2 é o \hat{y}_2 . Assim, temos os dois estágios:

Primeiro Estágio:

$$\hat{y}_2 = \hat{\pi}_0 + \hat{\pi}_1 z_1 + \hat{\pi}_2 z_2 + \hat{\pi}_3 z_3 \quad (7)$$

Segundo Estágio:

$$\sum_{i=1}^n \hat{y}_{i2} (y_{i1} - \hat{\beta}_0 - \hat{\beta}_1 y_{i2} - \hat{\beta}_2 z_{i1}) = 0 \quad (8)$$

Referências

WOOLDRIDGE, J. M. *Introductory econometrics: A modern approach*. [S.l.]: Nelson Education, 2015.