



Exercícios do Wooldridge 01  
Variáveis Instrumentais e o  
Método de Mínimos Quadrados em 2 Estágios (TSLS)

Vítor Wilher  
Mestre em Economia  
[analisemacro.com.br](http://analisemacro.com.br)

---

## Pacotes

```
### Pacotes  
library(foreign)  
library(AER)  
library(stargazer)
```

# 1 Introdução

Variáveis instrumentais são uma poderosa ferramenta para identificar e estimar relações causais. Para ilustrar, considere o modelo simples de regressão linear abaixo especificado

$$y = \beta_0 + \beta_1 x + u \quad (1)$$

O estimador de MQO para o parâmetro de inclinação é dado por

$$\hat{\beta}_1^{MQO} = \frac{Cov(x, y)}{Var(x)} \quad (2)$$

Se supormos, por suposto, que o regressor  $x$  é correlacionado com o termo de erro  $u$ , esse estimador será viesado e inconsistente. Se tivermos um instrumento válido  $z$ , nós podemos estimar  $\beta_1$  de forma consistente usando o estimador de variáveis instrumentais

$$\hat{\beta}_1^V = \frac{Cov(z, y)}{Cov(z, x)} \quad (3)$$

Isto é, um instrumento válido é correlacionado com o regressor  $x$ , o que implica que o denominador de 3 é diferente de zero. Ele também deve ser não correlacionado como o termo de erro  $u$ .

## 2 Retorno da educação para mulheres casadas

Para ilustrar a implementação de variáveis instrumentais no **R**, vamos considerar o Exemplo 15.1 de Wooldridge (2013), sobre o retorno da educação para mulheres casadas.

### Exemplo 15.1 do Wooldridge (2013)

Utilizamos os dados sobre mulheres casadas que trabalham contidos no arquivo `mroz.raw` para estimar o retorno da educação no modelo de regressão simples

$$\ln(\text{salário}_i) = \beta_0 + \beta_1 \text{educ} + u \quad (4)$$

O código de **R** abaixo lê o arquivo Stata `mroz.dta`, utilizando para isso o pacote `foreign`.

```
mroz <- read.dta("http://fmwww.bc.edu/ec-p/data/wooldridge/mroz.dta")
```

Se o leitor observar o *data frame*, vai ver que há *missing values* nas observações de salários. Tratamos o problema com o comando abaixo.

```
amostra <- subset(mroz, !is.na(wage))
```

Feito isso, podemos passar para a estimação de  $\beta_1$ . Faremos a mesma de duas formas. Via MQO normal e via IV, tomando a *educação do pai* como instrumento para a educação das mulheres da amostra. Para a estimação via instrumentos, utilizaremos a função `ivreg` do pacote AER. Isso é feito abaixo.

```
reg.ols <- lm(log(wage) ~ educ, data=amostra)
reg.iv <- ivreg(log(wage) ~ educ | fatheduc, data=amostra)
```

Os resultados das estimações estão postos na tabela 1. A equação `ivreg` implementa variáveis instrumentais por mínimos quadrados em dois estágios. Observe, por suposto, a diferença de  $\beta_1$  nos dois métodos. Enquanto no estimador de MQO, o retorno da educação no salário é de 11% para um ano adicional de estudo, no estimador IV, passa a ser de 5.9%, um pouco mais da metade.

Tabela 1

	<i>Dependent variable:</i>	
	log(wage)	
	<i>OLS</i>	<i>instrumental variable</i>
	(1)	(2)
educ	0.109*** (0.014)	0.059* (0.035)
Constant	-0.185 (0.185)	0.441 (0.446)
Observations	428	428
R <sup>2</sup>	0.118	0.093
Adjusted R <sup>2</sup>	0.116	0.091
Residual Std. Error (df = 426)	0.680	0.689
F Statistic	56.929*** (df = 1; 426)	

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## Referências

Wooldridge, J. M. *Introductory Econometrics: A Modern Approach*. Editora Cengage, 2013.