

# **AULA 01**

# **Principais Conceitos em Econometria**

**Ernesto F. L. Amaral**

**02 de março de 2010**

**Métodos Quantitativos de Avaliação de Políticas Públicas (DCP 030D)**

**Fonte:**

**Wooldridge, Jeffrey M. “Introdução à econometria: uma abordagem moderna”. São Paulo:  
Cengage Learning, 2008. pp.1-17.**

# ECONOMETRIA

- A econometria evoluiu como uma disciplina separada da estatística matemática, porque enfoca problemas inerentes à coleta e à análise de dados econômicos não-experimentais.
- **Dados não-experimentais** não são acumulados por meio de experimentos controlados de indivíduos, firmas ou segmentos da economia.
- Dados não-experimentais são também chamados de **dados observacionais** para enfatizar o fato de que o pesquisador é um coletor passivo de dados.
- **Dados experimentais** são frequentemente coletados em ambientes de laboratório nas ciências naturais, mas são muito mais difíceis de serem obtidos nas ciências sociais.
- O método de análise da **regressão múltipla** é utilizado por econometristas e estatísticos matemáticos, mas o foco e interpretação pode diferir significativamente.

# ANÁLISE ECONÔMICA EMPÍRICA

- Os métodos econométricos são usados para testar uma teoria econômica ou para analisar relações que apresentam importância para análises de políticas públicas.
- Uma análise empírica usa dados para testar uma teoria ou estimar uma relação.
- O primeiro passo em qualquer análise empírica é a formulação cuidadosa da questão de interesse, a qual pode ser a de testar efeitos de uma política governamental ou, até mesmo, de testar hipóteses e teorias.
- O modelo econômico formal consiste em equações matemáticas que descrevem relações para testar teorias.

# MICROECONOMIA

- Os indivíduos fazem escolhas para maximizar seu bem-estar (**maximização da utilidade**), sujeitas às restrições de recursos.
- Isso oferece um arcabouço para criar modelos econômicos para fazer previsões entre variáveis.
- A maximização da utilidade leva a um conjunto de **equações de demanda**, no contexto das decisões de consumo.
- Em uma equação de demanda, a quantidade demandada de cada produto depende do seu próprio preço, do preço dos bens substitutos e complementares, da renda do consumidor e das características individuais que influem no gosto.

# MODELO ECONÔMICO

- O modelo econômico é a formulação teórica de uma relação entre variáveis econômicas.
- A quantidade de tempo gasto na atividade criminosa é uma função de vários fatores (Gary Becker 1968):

$$y=f(x_1, x_2, x_3, x_4, x_5, x_6, x_7),$$

$y$  = horas gastas em atividades criminosas.

$x_1$  = “salário” por hora ocupada em atividade criminosa.

$x_2$  = salário-hora em emprego legal.

$x_3$  = renda de outras atividades que não o crime ou um emprego legal.

$x_4$  = probabilidade de ser capturado.

$x_5$  = probabilidade de ser condenado se capturado.

$x_6$  = sentença esperada se condenado.

$x_7$  = idade.

# MODELO ECONOMETRICO

- Após elaborar o modelo econômico, é especificado um modelo econométrico, que será aplicado a dados existentes.
- A forma da função  $f(.)$  deveria ser especificada antes de realizar uma análise econométrica.
- Se uma variável não pode ser obtida, é possível utilizar uma variável que se aproxima desta que se quer medir (**proxy**).
- Outros fatores são considerados no termo de erro  $u$  (ou termo de perturbação):
  - **Erro amostral** é a diferença entre o resultado amostral e o verdadeiro resultado da população (devidos ao acaso).
  - **Erro não-amostral** ocorre quando os dados amostrais são coletados, registrados ou analisados incorretamente.
- Modelo econométrico de Becker (1968):

$$\textit{crime} = \beta_0 + \beta_1 \textit{salário} + \beta_2 \textit{outrenda} + \beta_3 \textit{freqpris} + \beta_4 \textit{freqcond} + \beta_5 \textit{sentmed} + \beta_6 \textit{idade} + u$$

# MODELO ECONOMÉTRICO NA PRÁTICA

- Na maioria dos casos, a análise econométrica começa pela especificação de um modelo econométrico, sem consideração de detalhes da criação do modelo econômico.
- É comum começar com um modelo econométrico e usar o raciocínio econômico e conhecimentos científicos como guias para escolher as variáveis.
- Após a especificação do modelo econométrico, várias hipóteses podem ser formuladas em termos das direções e influências dos parâmetros desconhecidos (independentes) sobre a variável de interesse (dependente).
- Após os dados terem sido coletados, os métodos econométricos são usados para estimar os parâmetros do modelo econométrico e para testar as hipóteses de interesse.

# ESTRUTURA DO DADOS ECONÔMICOS

- Dados de corte transversal.
- Dados de séries de tempo.
- Cortes transversais agrupados.
- Dados de painel ou longitudinais.



## DADOS DE CORTE TRANSVERSAL

- Um conjunto de dados de corte transversal consiste em uma amostra de uma unidade de análise, tomada em um determinado ponto no tempo.
- Esses dados são muito utilizados em economia e em outras ciências sociais.
- Dados em um determinado ponto do tempo são importantes para testar hipóteses e avaliar políticas.
- A ordenação das observações nos dados de corte transversal não importa para a análise econométrica.
- Os dados podem apresentar problemas, em casos que a amostragem aleatória não é a técnica apropriada:
  - Problema de seleção amostral: determinados indivíduos não revelam informações acuradas.
  - Observações não são extrações independentes: unidades próximas possuem características semelhantes.

## EXEMPLO DE DADOS DE CORTE TRANSVERSAL

– Conjunto de dados de corte transversal para o ano de 1976 de 526 trabalhadores (Wooldridge 2008):

Número da observação	Salário por hora	Anos de escolaridade	Anos de experiência no mercado de trabalho	Feminino	Estado civil (casado)
1	3,10	11	2	1	0
2	3,24	12	22	1	1
3	3,00	11	2	0	0
4	6,00	8	44	0	1
5	5,30	12	7	0	1
...	...	...	...	...	...
525	11,56	16	5	0	1
526	3,50	14	5	1	0

## DADOS DE SÉRIES DE TEMPO

- Um conjunto de dados de séries de tempo consiste em observações sobre uma ou mais variáveis ao longo do tempo.
- Como eventos passados podem influenciar eventos futuros, o tempo é uma dimensão importante em um conjunto de dados de séries de tempo.
- A ordenação cronológica das observações em uma série de tempo transmite informações importantes.
- A análise desses dados pode ser dificultada, porque observações econômicas não são independentes ao longo do tempo (variáveis possuem tendências temporais).
- As frequências mais comuns são: diária, semanal, mensal, trimestral e anual.
- Os pesquisadores devem estar atentos para séries que apresentam padrão sazonal.

## EXEMPLO DE DADOS DE SÉRIES DE TEMPO

– Conjunto de dados de séries de tempo sobre efeitos do salário mínimo em Porto Rico (apud Wooldridge 2008):

Número da observação	Ano	Salário mínimo médio no ano	Taxa de trabalhadores cobertos pela lei de salário mínimo	Taxa de desemprego	Produto Nacional Bruto (PNB)
1	1950	0,20	20,1	15,4	878,7
2	1951	0,21	20,7	16,0	925,0
3	1952	0,23	22,6	14,8	1.015,9
...	...	...	...	...	...
37	1986	3,35	58,1	18,9	4.281,6
38	1987	3,35	58,2	16,8	4.496,7

## CORTES TRANSVERSAIS AGRUPADOS

- Alguns conjuntos de dados têm características de corte transversal e de séries de tempo.
- Um mesmo conjunto de variáveis é coletado em diferentes períodos do tempo, em **distintas** amostras aleatórias de uma mesma população (Censo Demográfico, Pesquisa Nacional por Amostra de Domicílios – PNAD).
- Agrupar cortes transversais de diferentes anos é eficaz para analisar os efeitos de uma política pública.
- O ideal é coletar dados de anos anteriores e posteriores a uma importante mudança de política governamental.
- Além de aumentar o tamanho da amostra, a análise de corte transversal agrupada é importante para estimar como uma relação fundamental mudou ao longo do tempo.

# EXEMPLO DE CORTES TRANSVERSAIS AGRUPADOS

– Conjunto de dados sobre os preços da moradia em 1993 e 1995 nos Estados Unidos (Wooldridge 2008):

Número da observação	Ano	Preço comercializado	Impro	Arquad	Quantidade de dormitórios	Quantidade de banheiros
1	1993	85.500	42	1.600	3	2,0
2	1993	67.300	36	1.440	3	2,5
3	1993	134.000	38	2.000	4	2,5
...	...	...	...	...	...	...
250	1993	243.600	41	2.600	4	3,0
251	1995	65.000	16	1.250	2	1,0
252	1995	182.400	20	2.200	4	2,0
253	1995	97.500	15	1.540	3	2,0
...	...	...	...	...	...	...
520	1995	57.200	16	1.100	2	1,5

## DADOS DE PAINEL OU LONGITUDINAIS

- Um conjunto de dados de painel (ou dados longitudinais) consiste em uma série de tempo para **cada** membro do corte transversal do conjunto de dados.
- Dados de painel podem ser coletados para indivíduos, domicílios, instituições ou unidades geográficas.
- Os dados de painel são distintos dos dados de corte transversal agrupados, porque as **mesmas** unidades são acompanhadas ao longo de um determinado período.
- Assim como no corte transversal puro, a ordenação dos dados de painel não é importante.
- Ter múltiplas observações sobre as mesmas unidades permite controlar por características não observáveis.
- Esses dados permitem estudar defasagens de resultados (impacto de políticas pode ser sentido após algum tempo).
- Esses dados são mais difíceis e caros de se obter.

## EXEMPLO DE DADOS DE PAINEL OU LONGITUDINAIS

- Conjunto de dados de painel sobre crime e estatísticas relacionadas em 1986 e 1990 em 150 cidades nos Estados Unidos (Wooldridge 2008):

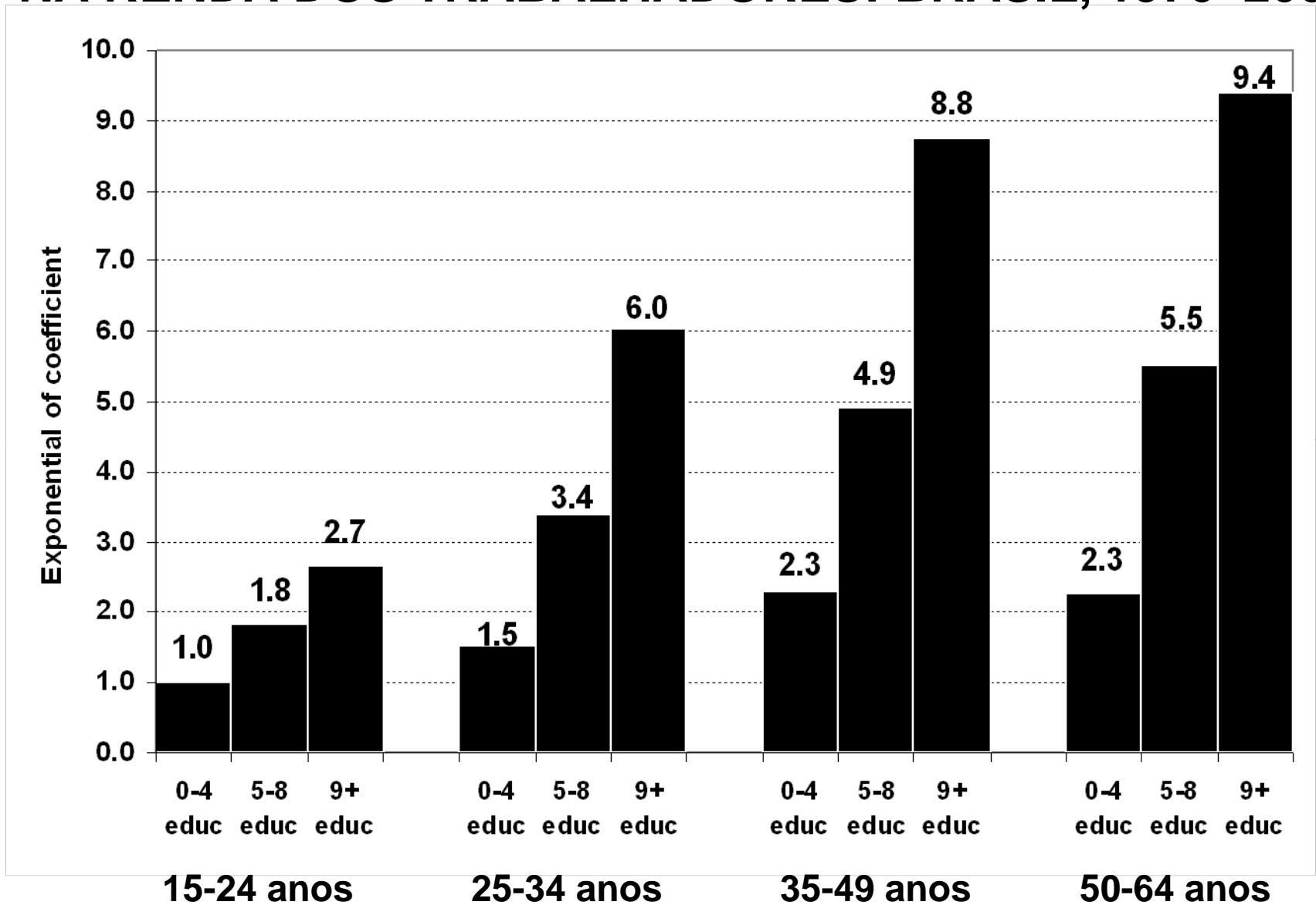
Número da observação	Cidade	Ano	Homicídios	População	Desemprego	Polícia
1	1	1986	5	350.000	8,7	440
2	1	1990	8	359.200	7,2	471
3	2	1986	2	64.300	5,4	75
4	2	1990	1	65.100	5,5	75
...	...	...	...	...	...	...
297	149	1986	10	260.700	9,6	286
298	149	1990	6	245.000	9,8	334
299	150	1986	25	543.000	4,3	520
300	150	1990	32	546.200	5,2	493



# CAUSALIDADE

- Na avaliação de políticas públicas, o objetivo do pesquisador é inferir que uma variável tem um **efeito causal** sobre outra variável.
- Encontrar uma associação entre duas ou mais variáveis pode ser sugestivo (correlação), mas somente será convincente se for possível estabelecer uma causalidade.
- A noção de ***ceteris paribus*** é importante, já que significa “outros fatores (relevantes) permanecendo iguais”.
- Se outros fatores não forem mantidos fixos, não poderemos conhecer o efeito causal de uma variável sobre outra.
- Como a maioria dos dados coletados nas ciências sociais são não-experimentais (não são experimentos controlados como nas ciências naturais), descobrir relações causais é uma tarefa complexa.

# EFEITOS DE GRUPOS DE IDADE-ESCOLARIDADE NA RENDA DOS TRABALHADORES: BRASIL, 1970–2000



Fonte: Censos Demográficos Brasileiros 1970 a 2000 (IBGE).