Esta aula

- Plano
  - Heterocedasticidade
  - Variáveis binárias
  - Escalas
  - Formas Funcionais

- Bibliografia
  - Wooldridge, J. M. Introductory Econometrics: A modern Approach, 6th Ed.

Heterocedasticidade

Regressão Linear - Inferência

- Para realizar testes de hipóteses, assumimos que u tem distribuição normal com média 0 e variância constante.

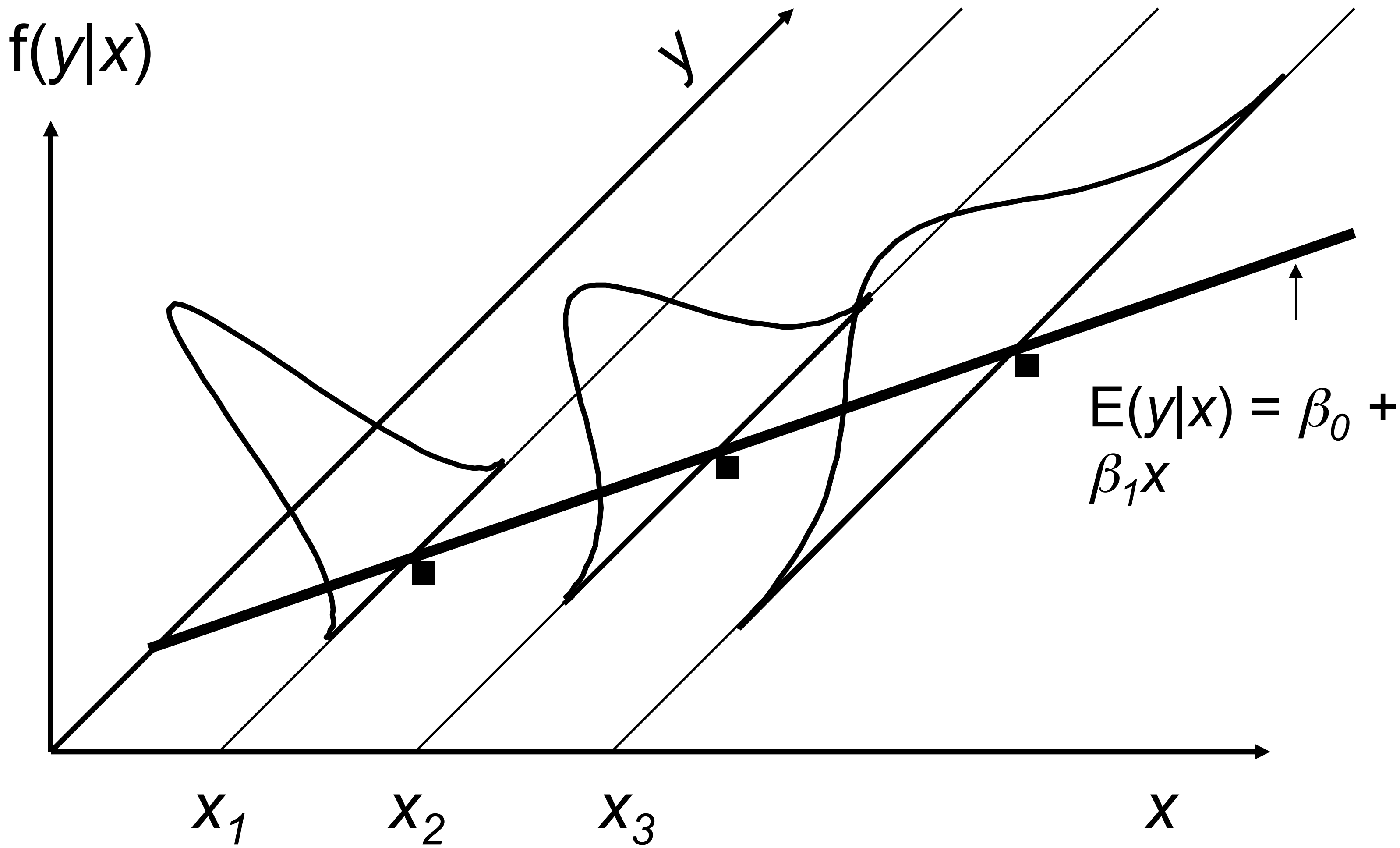
- Nesse caso, $$\left(\hat{\beta}_{\mathrm{j}} - \beta_j\right) \Big/ se\left(\hat{\beta}_j\right) \sim t_{n-k-1}$$

- IMPORTANTE: Mesmo que u não tenha distribuição normal, se a amostra for relativamente grande as estatística t e F usuais tem uma distribuição que converge para as distribuições t-student e F.

Revisão: Homocedasticidade

- O pressuposto de homocedasticidade significa que a variância do erro não-observável é constante e independente do valor das variáveis explicativas

Exemplo de Heterocedasticidade



Consequências da Heterocedasticidade

- MQO é não-enviesado mesmo na presença de heterocedasticidade.
- Porém, nesse caso, os erros padrões são enviesados.
- Portanto, as estatisticas t e F não são válidas.

Heterocedasticidade

Importante: se o modelo tiver heterocedasticidade e os demais 4 pressupostos de Gauss-Markov continuarem válidos, o estimador MQO continua não-enviesado porém não é mais BLUE!

Além disso, testes de hipóteses baseadas nas variâncias dos parâmetros estimadas por MQO não são mais válidos

Variância com Heterocedasticidade

- A variância e desvio-padrão robusto somente levarão a estatísticas t e F válidas se as amostras forem grandes.

Teste para Heterocedasticidade: Breusch-Pagan

- Queremos testar H_0: $Var(u|x_1, x_2,\ldots, x_k) = \sigma^2$
- Se assumirmos uma relação linear entre u^2 e x_j:

$$u^2 = \delta_0 + \delta_1 x_1 + \ldots + \delta_k x_k + v$$

, podemos testar:

- H_0: $\delta_1 = \delta_2 = \ldots = \delta_k = 0$
- Esse é o teste Breusch-Pagan.

Teste White

- O Teste White permite que a heterocedasticidade dependa das variáveis explicativas ao quadrado e produtos cruzados.
- Para simplificar o teste, procedemos da seguinte forma:

Teste White

- Fazemos a regressão dos resíduos ao quadrado nos valores de $\hat{y}$ e $\hat{y}^2$ e fazemos um teste F no R-quadrado

Mínimos Quadrados Ponderados

- Apesar de conseguirmos obter erros padrões robustos, se soubermos a forma funcional da heterocedasticidade, podemos obter estimativas mais eficientes.
- O métodos dos mínimos quadrados ponderados (MQP) ou weighted least squares (WLS) permite transformar o modelo tal que ele tenha erros homocedásticos, soubermos a forma exata da heterocedasticidade. Nesse caso WLS is BLUE!
- Assim como em MQO, os testes t e F são assintoticamente válidos ou exatamente válidos se os erros possuírem distribuição Normal.

FGLS

- FGLS não é não-enviesado, mas ainda consistente e assintoticamente eficiente.

Variável Binária ou Dummy

Variável Binária ou “Dummy”

- Uma variável binária é aquela que toma dois valores possíveis, geralmente 0 e 1.
- No nosso banco de dados trabalhado nas últimas aulas, female é uma variável binária.

Variável Binária ou "Dummy"

- Considere o seguinte modelo com uma variável binária:
- $y = \beta_0 + \delta_0 d + \beta_1 x + u$
- Nesse caso, a variável d representa uma mudança de intercepto quando se passa de um grupo para o outro do banco de dados.
- If $d = 0$, $y = \beta_0 + \beta_1 x + u$
- If $d = 1$, $y = (\beta_0 + \delta_0) + \beta_1 x + u$

Variável Binária ou "Dummy"

- Considere agora o seguinte modelo:
- $y = \beta_0 + \delta_1 d + \beta_1 x + \delta_2 d^* x + u$
- If $d = 0$, $y = \beta_0 + \beta_1 x + u$
- If $d = 1$, $y = (\beta_0 + \delta_1) + (\beta_1 + \delta_2)\, x + u$
- Nesse modelo, a variável dummy permite uma mudança de intercepto, bem como uma mudança de inclinação.

Modelo Linear de Probabilidade

- Se a variável dependente for binária, então $P(y = 1|x) = E(y|x)$. Nesse caso:
- $P(y = 1|x) = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k$
- Nesse caso, cada intercepto beta i mede o impacto de variações marginais em xi na probabilidade do evento 1 ocorrer.
- O valor previsto de y nesse é a probabilidade estimada do evento 1 ocorrer.

Modelo Linear de Probabilidade

- Problema: nada impede que o y previsto não esteja no intervalo [0,1].
- Geralmente, esse modelo viola também o pressuposto de homocedasticidade.

Avaliação de Políticas/Programas

- Podemos utilizar variáveis dummy para avaliar o impacto de políticas/programas
- Por exemplo, qual o impacto da participação no programa Bolsa Família no nível educacional da família?

Avaliação de Políticas/Programas

- Problema: variáveis que influenciam a participação no Bolsa Família, como a renda dos ascendentes, também podem explicar o nível educacional.
- Isso levaria a viés nas estimações

Estudo de Evento

Estudo de Evento (*Event Study)*

- Em um estudo de evento, estamos preocupados com o impacto de um evento particular em um resultado.
- Por exemplo, considere o seguinte modelo com uma variável binária "dt":
- $yt = \beta_0 + \delta_0 dt + \beta_1 xt + ut$

  em que a variavel "dt" toma o valor um 1 quando o evento ocorreu e zero nos demais períodos.

- Em um estudo de evento estamos interessados na magnitude e significancia do coeficiente delta0, associado à variável "dt".

Exemplo 2

- *Fair (1996) analisaram o efeito da performance economica nos resultados das eleições presidenciais.*
- *Para entender, utilize o arquivo stata denominado "FAIR".*

Escalas

Escala

Se o modelo populacional for dado por:

$$y = \beta_0 + \beta_1 x + u$$

Em que y é o peso do indivíduo em gramas e quisermos transformar a unidade de medida para kg: y/1000

Escala

Se dividirmos os dois lados da equação pela mesma constante, tem-se o mesmo modelo:

$$\frac{y}{1000} = \frac{\beta_0}{1000} + \frac{\beta_1}{1000}x + \frac{u}{1000}$$

Quando fazemos uma regressão de y/1000 contra x, obtemos uma estimativa de $\dfrac{\beta_1}{1000}$

Escala

Suponha agora que queremos alterar a escala da variável explicativa x, altura, que está em cm, para metros: x/100

Veja que podemos reescrever o modelo como:

$$y = \beta_0 + 100\beta_1 \frac{x}{100} + u$$

Quando fazemos uma regressão de y contra x/100, obtemos uma estimativa de $100\beta_1$

Formas Funcionais

Linearidade

- Até agora, consideramos o modelo linear:

$$y = \beta_0 + \beta_1 x + u$$

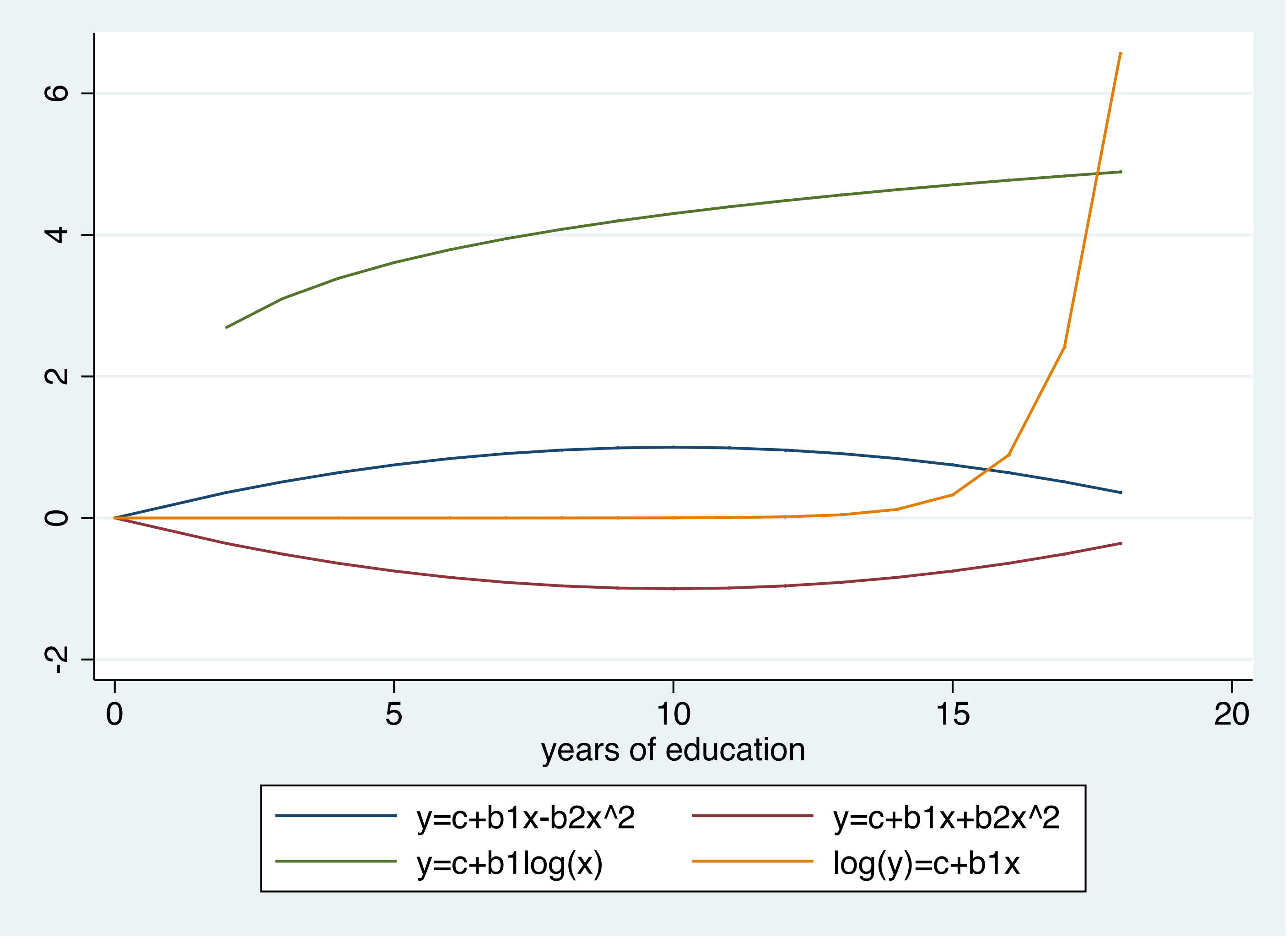
- Esse modelo preve um impacto constante β_1, também chamado de efeito marginal de x em y, que independe do valor inicial de x: $\Delta y = \beta_1 \Delta x$

- Em: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2.$

$$\Delta y = \beta_1 \Delta x_1 \text{ if } \Delta x_2 = 0$$

- β_1 é chamado de efeito parcial de x_1 em y

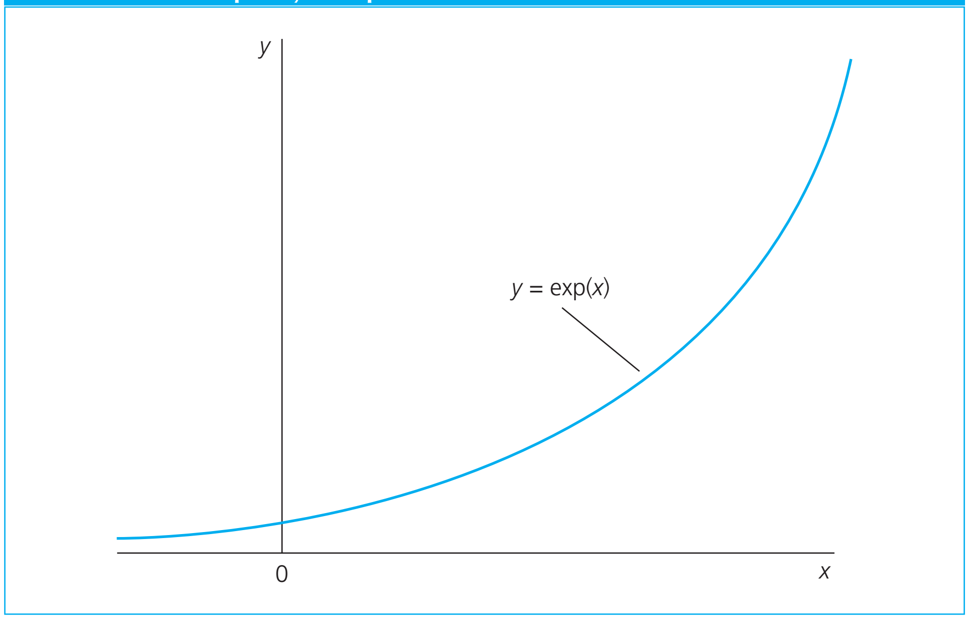
Linearidade



Linearidade

- É comum fazer-se a regressão de log(y) contra x. A interpretação do modelo faz mais sentido para muitos problemas em ciências sociais. Em nosso exemplo anterior, teríamos: $\log(wage) = \beta_0 + \beta_1 educ + u,$

FIGURE A.5 Graph of $y = \exp(x)$.



Linearidade

- Utilizando cálculo, é possível demonstrar que:

$$\log(x_1) - \log(x_0) \approx (x_1 - x_0)/x_0 = \Delta x/x_0$$

- Em $\log(wage) = \beta_0 + \beta_1 educ + u$,

, multiplicando-se por 100 tem-se:

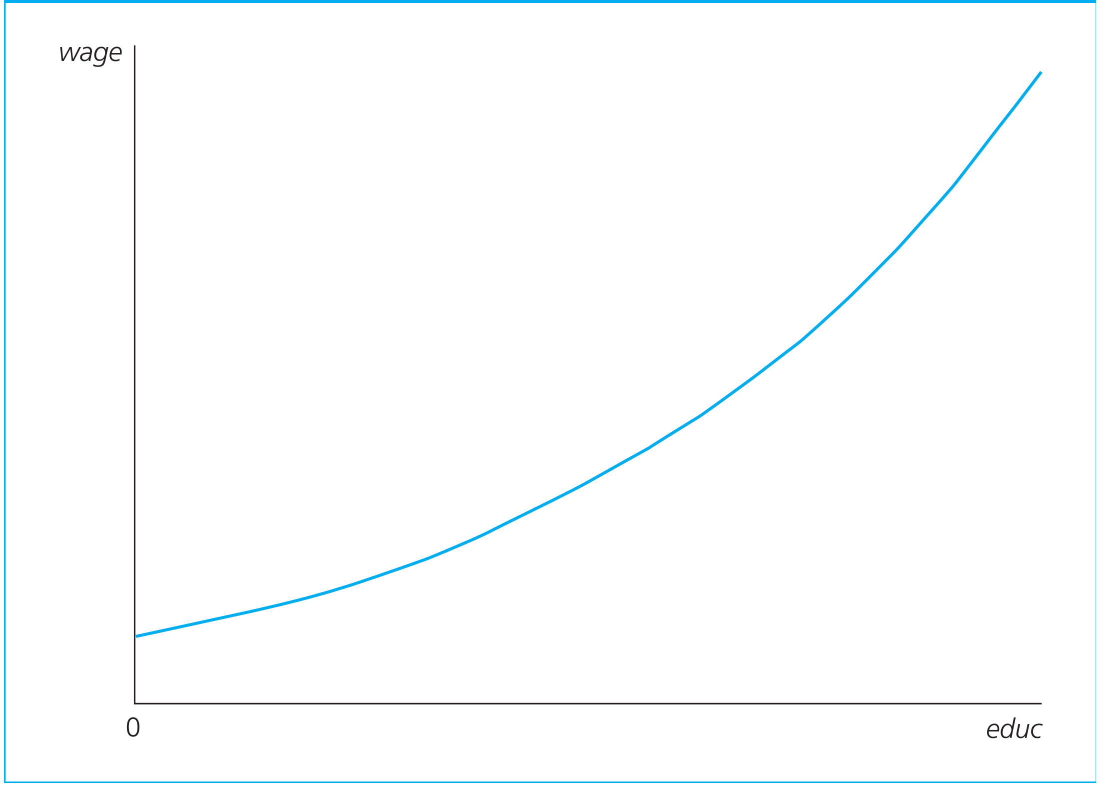
$$\%\Delta wage \approx (100 \cdot \beta_1)\Delta educ.$$

Linearidade

- Nesse caso, tem-se:

$$\%\Delta wage \approx (100 \cdot \beta_1)\Delta educ.$$

FIGURE 2.6 $wage = \exp(\beta_0 + \beta_1 educ)$, with $\beta_1 > 0$.

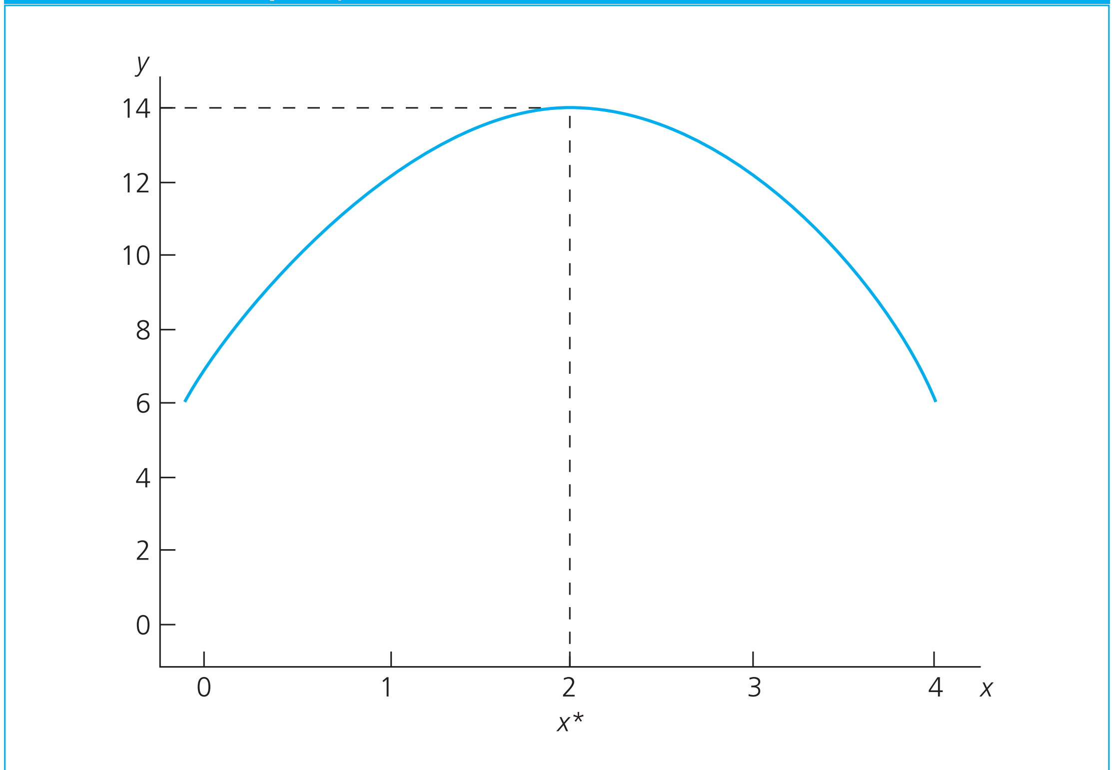


Linearidade

- Função quadrática:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2,$$

FIGURE A.3 Graph of $y = 6 + 8x - 2x^2$.

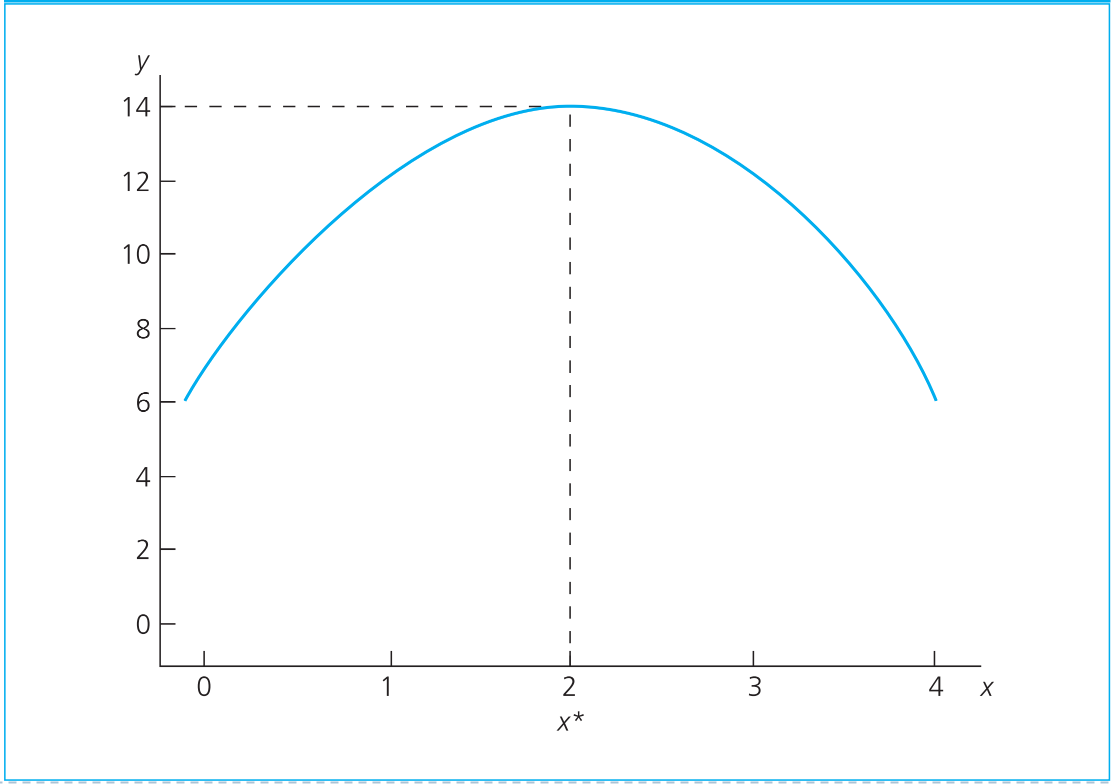


Linearidade

- Função quadrática:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2,$$

FIGURE A.3 Graph of $y = 6 + 8x - 2x^2$.



Linearidade

- Elasticidade é definida pela variação percentual de y dividida pela variação percentual de x:

$$\frac{\Delta y}{\Delta x} \cdot \frac{x}{y} = \frac{\%\Delta y}{\%\Delta x}$$

- (Cálculo) A elasticidade pode ser aproximada por:

$$\Delta \log(y) / \Delta \log(x)$$

Linearidade

- Portanto, em

$$\log(y) = \beta_0 + \beta_1 \log(x)$$

- $\beta_1 = \log(y) / \log(x)$, sendo aproximadamente a elasticidade de y com relação x

Linearidade

- Elasticidade de y com relação a x é definida pela variação percentual de y dividida pela variação percentual de x:

$$\frac{\Delta y}{\Delta x} \cdot \frac{x}{y} = \frac{\% \Delta y}{\% \Delta x}.$$

- (Cálculo) A elasticidade pode ser aproximada por:

$$\Delta \log(y) / \Delta \log(x)$$

Linearidade

- Nesse caso, tem-se:

TABLE 2.3 Summary of Functional Forms Involving Logarithms

Model	Dependent Variable	Independent Variable	Interpretation of β_1
Level-level	y	x	$\Delta y = \beta_1 \Delta x$
Level-log	y	$\log(x)$	$\Delta y = (\beta_1/100)\%\Delta x$
Log-level	$\log(y)$	x	$\%\Delta y = (100\beta_1)\Delta x$
Log-log	$\log(y)$	$\log(x)$	$\%\Delta y = \beta_1 \%\Delta x$

Obrigada!