

# Econometria com Dados em Painel: Fundamentos e Aplicações

**Adriano Vargas Saldanha**

Mestrando no Programa de Pós-Graduação em  
Organizações e Mercados - UFPel

**Pelotas/RS, novembro de 2025**

# Índice

- Regressão com Dados em Painel
- Modelos de Regressão com Dados em Painel
- Modelo Pooled
- Modelo de Efeitos Fixos
- Modelo de Efeitos Aleatórios
- Comparativo
- Teste de Hausman
- Exercícios no VSCode

# Regressão com Dados em Painel

## →Características

- O modelo de regressão com dados em painel combina:
  - Dimensão temporal: evolução ao longo do tempo;
  - Dimensão espacial: diferentes unidades de análise (famílias, empresas, estados, países).
- A mesma unidade de corte transversal é observada em vários períodos;
- Permite estudar simultaneamente efeitos no tempo entre unidades.
- Exemplo: Produção industrial mensal dos estados brasileiros em função da taxa de juros no período 2023-2024.
  - 26 estados (excluindo DF) x 24 meses = 624 observações;
  - Cada estado têm uma série temporal própria.

# Regressão com Dados em Painel

## → Benefícios

- Segundo Gujarati & Porter (2011):
  - Captura a heterogeneidade individual (diferenças entre indivíduos, empresas, estados, países);
  - Proporciona mais informação, maior variabilidade e menor colinearidade entre variáveis;
  - Adequado para estudar a dinâmica da mudança (emprego, renda, crescimento);
  - Melhora a detecção e medição de efeitos em relação a estudos puramente transversais ou de séries temporais;
  - Permite modelos comportamentais mais complexos;
  - Reduz o viés de agregação (evita generalizações excessivas por médias agregadas).

# Modelos de Regressão com Dados em Painel

- A partir da estrutura de dados em painel, diferentes especificações podem ser utilizadas para lidar com heterogeneidade e efeitos não observados.

## → Pooled (MQO Pooled)

- Intercepto e coeficientes angulares ( $\beta$ ) são constantes no tempo e entre as unidades;
- Diferenças entre indivíduos e no tempo ficam totalmente no termo de erro.

## → Efeitos Fixos (FE)

- Coeficientes angulares constantes;
- Intercepto varia entre indivíduos (captando heterogeneidade não observada que não muda no tempo).

# Modelos de Regressão com Dados em Painel

## → Efeitos Aleatórios (RE)

- Intercepto = valor médio comum + componente aleatório que varia entre indivíduos;
- Coeficientes angulares podem variar no tempo e entre indivíduos.

# Modelo Pooled

## → Definição

- “Empilha” todas as observações da base, ignorando a estrutura de dados em painel;
- Trata todas as observações como não correlacionadas entre indivíduos;
- Assume erros heterocedásticos entre indivíduos;
- Forma mais simples → desconsidera dimensões combinadas de tempo e espaço;
- Estimado por Mínimos Quadrados Ordinários (MQO).
  - Especificação Geral:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \varepsilon_{it}$$

$$i = 1, 2, \dots, n ; t = 1, 2, \dots, t$$

→ Onde:

- $i$  = unidade de corte transversal (empresa, estado, país, etc.);
- $t$  = período de tempo.

# Modelo Pooled

## → Exemplos Aplicados

- Economia Regional: Estimar o efeito da taxa de investimento (%) sobre o crescimento do PIB para diferentes estados brasileiros, tratando todos os estados como se fossem um único “estado agregado”;
- Finanças: Analisar o impacto da taxa Selic e da inflação sobre o retorno mensal de ações, ignorando que cada ação pertence a uma empresa distinta;
- Macroeconomia: Avaliar a relação entre gasto público e crescimento econômico para diversos países, ignorando diferenças estruturais entre eles.

# Modelo Pooled

## → Limitações

- Ignora heterogeneidade não observada → risco de viés de variável omitida;
- Pode superestimar significância estatística dos coeficientes.

# Modelo de Efeitos Fixos (FE)

## → Definição

- O intercepto ( $a_i$ ) varia entre as unidades, mas é constante ao longo do tempo para cada unidade;
- Os coeficientes de inclinação ( $\beta_k$ ) são os mesmos para todas as unidades;
- Captura características específicas e não observadas de cada unidade, que não mudam no tempo.
  - Especificação Geral:

$$y_{it} = \beta_1 X_{1it} + \beta_2 X_{2it} + a_i + \varepsilon_{it}$$
$$i = 1, 2, \dots, n ; t = 1, 2, \dots, t$$

→ Onde:

- $y_{it}$  = variável independente para a unidade  $i$  no tempo  $t$ ;
- $X_{kit}$  = variável explicativa  $k$  para a unidade  $i$  no tempo  $t$ ;
- $a_i$  = intercepto específico de cada unidade (efeito fixo);
- $\varepsilon_{it}$  = erro aleatório.

# Modelo de Efeitos Fixos (FE)

## → Exemplos Aplicados

- Economia Regional: Estimar o efeito da taxa de investimento (%) e da escolaridade média sobre o PIB per capita de estados brasileiros, controlando por características fixas (clima, localização geográfica, estrutura produtiva);
- Finanças: Avaliar o efeito do endividamento e do tamanho da firma sobre o retorno anual de ações, controlando por características fixas de cada empresa (setor, modelo de negócios);
- Macroeconomia: Medir o impacto da abertura comercial e da taxa de câmbio sobre o crescimento econômico, controlando por características históricas e culturais de cada país.

# Modelo de Efeitos Fixos (FE)

## →Vantagem

- Elimina o viés de variáveis omitidas que não variam no tempo.

# Modelo de Efeitos Aleatórios (RE)

## → Definição

- Os efeitos individuais ( $\alpha_i$ ) são tratados como variáveis aleatórias, e não como parâmetros fixos;
- O intercepto de cada unidade = valor médio comum ( $\beta_1$ ) + termo aleatório  $\varepsilon_i$ ;
- Útil quando as diferenças entre unidades são vistas como sorteadas de uma população maior.
  - Especificação Geral:

$$y_{it} = \beta_1 X_{1it} + \beta_2 X_{2it} + \alpha_i + W_{it}$$

$$W_{it} = \varepsilon_i + \mu_{it}$$

→ Onde:

- $\varepsilon_{it}$  = componente de corte transversal (efeito de cada unidade), média zero e variância  $\sigma^2$ ;
- $\mu_{it}$  = componente idiosincrático (tempo + corte transversal);
- Assume-se não correlações:  $\varepsilon_{it}$  não correlacionado com  $X_{kit}$  e  $\mu_{it}$ .

# Modelo de Efeitos Aleatórios (RE)

## → Exemplos Aplicados

- Economia Regional: Estimar o impacto da infraestrutura de capital humano sobre o PIB per capita de municípios, assumindo que diferenças não observadas entre eles sejam aleatórias;
- Finanças: Avaliar a relação entre endividamento e rentabilidade de empresas listadas, supondo que fatores específicos da empresa (ex: cultura organizacional);
- Macroeconomia: Medir o efeito da taxa de câmbio e da abertura comercial sobre o crescimento econômico de países, considerando que diferenças institucionais sejam aleatórias.

# Modelo de Efeitos Aleatórios (RE)

## →Vantagem

- Mais eficiente que FE se a hipótese de não correlação entre  $\varepsilon_i$  e  $X_{kit}$  for verdadeira.

# Comparativo

## → Resumo

Modelo	Hipótese Principal	Intercepto
Pooled	Ignora a heterogeneidade; dados “empilhados”.	Igual para todas as unidades e períodos.
Efeitos Fixos (FE)	Heterogeneidade correlacionada com regressores.	Varia entre unidades, constantes no tempo.
Efeitos Aleatórios (RE)	Heterogeneidade não correlacionada com regressores.	Intercepto médio + componente aleatório.

# Comparativo

## → Vantagens e Limitações

- Pooled:
  - Simples, referência inicial;
  - Ignora a heterogeneidade.
- Efeitos Fixos (FE)
  - Controla fatores não observados invariantes no tempo;
  - Não estima efeito de variáveis invariantes, menos eficiente que RE.
- Efeitos Aleatórios (RE)
  - Mais eficiente que FE se hipótese válida, estima variáveis invariantes;
  - Viés se hipótese de não correlação falhar, precisa do teste de Hausman.

# Teste de Hausman

## → Definição

- O teste de Hausman avalia se existe correlação entre os efeitos individuais não observados e as variáveis explicativas do modelo. É fundamental para determinar qual especificação (efeitos fixos ou efeitos aleatórios) é mais apropriada para a estimação;
- O teste compara os estimadores de ambos os modelos, explorando o fato de que o estimador FE é consistente sob ambas as hipóteses, enquanto o estimador RE só é consistente quando não há correlação.

# Teste de Hausman

## → Hipóteses do Teste

- Hipótese Nula ( $H_0$ ):
  - Os efeitos individuais não são correlacionados com as variáveis explicativas. Sob  $H_0$ , tanto FE quanto RE são consistentes, mas RE é mais eficiente.
- Hipótese Alternativa ( $H_a$ ):
  - Os efeitos individuais são correlacionados com as variáveis explicativas. Sob  $H_a$ , apenas FE é consistente.

# Teste de Hausman

## → Critérios de Decisão:

- Compare o valor da estatística H com o valor crítico  $\chi^2(k)$  ao nível de significância escolhido;
- Também podemos comparar o p-valor com o nível de significância (geralmente 0,05);
- Rejeite  $H_0$  se  $H > \chi^2$  crítico ou se p-valor < 0,05.

# Teste de Hausman

## → Aplicação e Interpretação

- Quando usar o teste?
  - Aplique o teste de Hausman sempre que estiver trabalhando com dados em painel e precisar escolher entre modelos de efeitos fixos e aleatórios. É importante principalmente em estudos onde a correlação entre efeitos individuais e regressores pode comprometer a validade dos resultados.
- Interpretação dos resultados:
  - Se  $H_0$  for rejeitada: Existe correlação significativa. Use o modelo de efeitos fixos (consistente);
  - Se  $H_0$  não for rejeitada: Não há evidência de correlação entre efeitos individuais e regressores. Use o modelo de efeitos aleatórios (mais eficiente).

# Teste de Hausman

## → Exemplo:

- Em um estudo sobre salários de trabalhadores ao longo do tempo, suponha que  $H = 15,3$  com 4 graus de liberdade (p-valor = 0,004).
  - Como o p-valor < 0,05, rejeitamos  $H_0$  e concluímos que características individuais não observadas (como habilidades) estão correlacionadas com educação e experiência, favorecendo o modelo de efeitos fixos.

# Regressão com Dados em Painel

## → Resumo de decisão:

- Passo a passo:
  - Estime um Pooled (MQO) como referência;
  - Estime FE e RE;
  - Faça o teste de Hausman:  
 $H_0 : RE \text{ é consistente } (\alpha_i \text{ não correlacionada com } X_{it}).$
- Decisão:
  - Hausman rejeita  $H_0 \rightarrow$  use FE;
  - Hausman não rejeita  $H_0 \rightarrow$  use RE.

# Regressão com Dados em Painel

## → Resumo de decisão:

- Quando preferir FE:
  - Heterogeneidade  $\alpha_i \uparrow$  e  $\text{corr}(\alpha_i, X_{it}) \neq 0$ ;
  - Quer remover viés de omitidas  $\uparrow$  invariantes no tempo.
- Quando preferir RE:
  - $\text{Corr}(\alpha_i, X_{it}) = 0$  é plausível;
  - Precisa estimar efeitos de variáveis invariantes no tempo;
  - Eficiência maior que FE sob  $H_0$ .
- Boas práticas:
  - Erros robustos cluster por unidade ( $i$ );
  - Checar efeitos fixos de tempo (ano) quando houver choques comuns.

# Exercícios VSCode