

# Regression Discontinuity Design

Pedro F., Pedro W., Rafael L., Stefano D.

LAMFO/UnB

30 de outubro de 2021



# Índice

- 1 Introdução e Background
- 2 Identificação e Interpretação
- 3 Inferência e Estimação
- 4 Casos Especiais e Aplicações
- 5 Exemplo Prático

# Table of Contents

## 1 Introdução e Background

## 2 Identificação e Interpretação

## 3 Inferência e Estimação

## 4 Casos Especiais e Aplicações

## 5 Exemplo Prático

# Introdução

- Fonte Bibliográfica: Regression Discontinuity Designs in Economics, David S. Lee and Thomas Lemieux
- Thistlethwaite and Campbell (1960): Analisar desempenho acadêmico de pessoas que receberam um prêmio por pouco e que quase o receberam.
- Hipótese de que, dado um limiar, os valores levemente inferiores e levemente superiores a ele assemelham-se
- Requer hipóteses mais leves comparado a outros modelos.

# Aplicações

- Wilbert van der Klaauw (2002): Impacto de auxílio estudantil.
- Cattaneo, Matias D., Rocío Titiunik, and Gonzalo Vazquez-Bare (2015): Auxílio sob a mortalidade infantil.
- David S. Lee (2007): Modelo causal das eleições.

# Sumário

- Seja  $X$  a variável atribuída, todos os valores de  $X$  maiores ou iguais a um limiar  $c$  receberam um tratamento.
- RD pode ser inválido se os indivíduos puderem manipular precisamente  $X$ .
- A variação do tratamento próximo ao limiar é aleatória se os indivíduos não forem capazes de manipular precisamente  $X$ .
- Portanto, RD pode ser testado como um experimento aleatório.

# Sumário

- Estimação não paramétrica não representa uma solução, mas um complemento à estimação paramétrica.
- Testes de Goodness of Fit são bons indicadores de especificações excessivamente restritas.

# Table of Contents

1 Introdução e Background

2 Identificação e Interpretação

3 Inferência e Estimação

4 Casos Especiais e Aplicações

5 Exemplo Prático



# Três Perguntas Relevantes

- 1) Como saber se RDD é a metodologia certa para o problema?
- 2) Como testar as premissas?
- 3) Até onde os resultados são generalizáveis?

# 1) Como saber se RDD é a metodologia certa para o problema?

- A RDD é apropriada em situações onde é plausível que as fatores não observáveis são contínuos em relação a variável de interesse
- Os agentes observados não tem controle na escolha em quem receberá o tratamento

## 2) Como testar as premissas?

- Deve-se avaliar se de fato a distribuição das variáveis observáveis ( $W$ ) é contínua em  $x$
- Ou seja,  $P(W = w|X = x)$  não apresenta descontinuidades em torno do ponto de quebra
- Isso implicaria que os indivíduos não teriam controle direto na variável de interesse

### 3) Até onde os resultados são generalizáveis?

- Os resultados da RD podem ser interpretados como uma média ponderada pelo efeito de "tratamento"
- Os pesos são determinados pela probabilidade de cada indivíduo antes do evento em questão estar em torno da região do ponto de quebra

# Table of Contents

1 Introdução e Background

2 Identificação e Interpretação

3 Inferência e Estimação

4 Casos Especiais e Aplicações

5 Exemplo Prático

# Estimação

- Duas regressões distintas em cada lado do ponto do corte.
- É adequado que subtraia o valor do corte da covariável.
- $X \rightarrow (X - c)$

# Interceptações

- As interceptações geram o valor das funções de regressão no ponto do corte.
- $X \leq c \rightarrow Y = \alpha_l + fl(X - c) + \epsilon$
- $X \geq c \rightarrow Y = \alpha_r + fr(X - c) + \epsilon$

# Interceptações

- O efeito pode ser calculado como a diferença entre as duas interceptações nos dois pontos do corte.
- $Y = \alpha I + \tau D + f(X - c) + \epsilon$



# Inferência

- MQO
- Erro padrão robusto
- 2SLS

# Table of Contents

1 Introdução e Background

2 Identificação e Interpretação

3 Inferência e Estimação

**4 Casos Especiais e Aplicações**

5 Exemplo Prático

# Variável de Atribuição Discreta

- Não compatibilidade da suposição de  $X$  contínua em certos casos (ex: idade como ponto de quebra)
- Uso de regressões para estimar a esperança condicional de  $Y$  no ponto de quebra por **extrapolação**
- Escolha da largura das bandas é simplificada
- Teste da especificação da regressão → **fitted model vs. dispersão bruta** em cada  $X$

# Dados em Painel e Efeitos Fixos

- A cada período, variável de tratamento é determinada de acordo com a realização da variável de atribuição  $X$
- Levar em consideração correlação serial dos erros no caso de cada indivíduo com *clustered* S.E.
- Impor uma dinâmica estrutural específica introduz restrições sem ganho para a identificação; esta é feita **diretamente** com um único cross-section
- Uso de variáveis defasadas como *baseline covariates* → diminui variância da amostragem

# O RDD aplicado em Economia

- Lee e Lamieux (2010) identificam que a maior parte dos estudos estão concentrados nas áreas de educação e de mercado de trabalho
- **Discussão:** alocação endógena → agentes econômicos reagem a incentivos ⇒ **manipulação da agenda?**
- **Discussão:** discontinuidade por tratamentos inevitáveis e por critérios geográficos ⇒ **incompatibilidade com RCT?**

# O RDD aplicado em Economia



1.

2.



# Principais contribuições do estudo em economia à literatura de RDD

- 1 Levar em consideração a possibilidade de **alocação endógena**
- 2 Reconhecimento do RDD como **experimento localmente randômico**
- 3 Importância da representação **gráfica**
- 4 Aumento do número de **áreas de aplicabilidade**

# Table of Contents

1 Introdução e Background

2 Identificação e Interpretação

3 Inferência e Estimação

4 Casos Especiais e Aplicações

5 Exemplo Prático



- Regression-Discontinuity Analysis: An Alternative to the Ex-Post Facto Experiment, Donald L. Thistlewaite and Donald T. Campbell (1960)
- Impacto do recebimento de um Certificado de Mérito sob a visibilidade pública: Alunos abaixo de um limiar de 10 não o recebem, ocasionando uma descontinuidade.
- Visibilidade pública afetará quais outras variáveis?

# Dados

- 5126 estudantes que receberam o certificado de mérito
- 2848 estudantes que receberam cartas de recomendação
- Certificados de Mérito são distribuídos para alunos que obtém o Qualifying Score de seu Estado

# Dados

Group	Scholarship qualifying test score interval <sup>b</sup>	Number of Merit Scholars	Number in designated sample <sup>a</sup>	Number of respondents	Percentage of designated sample responding	Percentage of C of M winners in each interval awarded Merit scholarships
Commended students	Below 1		419	322	76.8	
	1		318	256	80.5	
	2		368	281	76.4	
	3		320	258	80.6	
	4		407	338	83.1	
	5		324	259	79.9	
	6		333	267	80.2	
	7		280	213	76.1	
	8		301	248	82.4	
	9		256	201	78.5	
Totals	10		262	205	78.2	
			3,588	2,848	79.4	
Certificate of Merit winners	11	17	476	380	79.8	3.4
	12	22	466	370	79.4	4.5
	13	16	399	319	79.9	3.9
	14	17	371	298	80.3	4.4
	15	19	361	300	83.1	5.0
	16	34	358	289	80.7	8.7
	17	13	319	247	77.4	3.9
	18	18	345	256	74.2	5.0
	19	17	254	211	83.1	6.3
	20	23	301	237	78.7	7.1
Totals	Above 20	631	2,778	2,219	79.9	18.5
		827	6,428	5,126	79.7	11.4

Figure: Participantes do Programa

# Resultados

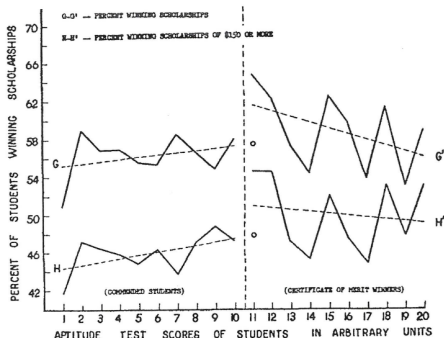


Figure: Regressão de Score sob o ganho de Bolsas

- Tanta a curva  $GG'$  quanto a  $HH'$  apresentam uma regressão visivelmente descontínua entre a pontuação 10 e 11.
- As curvas acima de 10 apresentam-se negativamente inclinadas se deve ao fato de bolsas de mérito não serem consideradas.
- Os testes indicam um p-valor menor que 0,025 para  $GG'$  e menor que 0,01 para  $HH'$ .

# Resultados

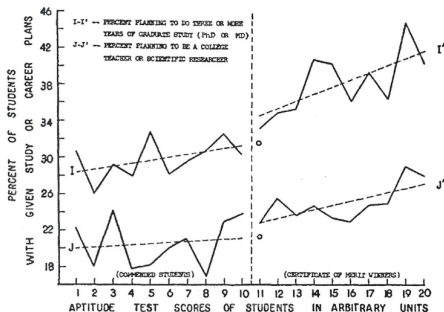


Figure: Regressão de Score sob a elaboração de Plano de Carreira

- Percebe-se um pequeno aumento entre as curvas  $II'$  e  $JJ'$ , podendo ser explicado por uma regressão curvilinear contínua.
- Os testes de significância demonstram não ser significativo o modelo descontínuo.
- Não há evidência para sustentar a descontinuidade sob a formação de planos de carreira devido a visibilidade do Certificado de Mérito.

# Resultados

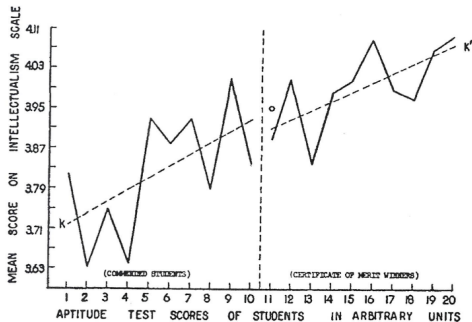


Figure: Regressão de Score sob o desempenho intelectual



- Não há qualquer indício de um deslocamento na curva  $KK'$ .
- Os testes de significância demonstram não ser significativo o modelo descontínuo.
- Não há evidência para dizer que há um aumento da vontade em seguir cursos superiores devido ao ganho do certificado de mérito.

# Referências

- Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2), 615-635.
- Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2), 281-355.
- Zhu, Z. (2021). An Introduction to Regression Discontinuity Design. *Towards Data Science*. Acessado em 13 de outubro de 2021. Disponível em <https://towardsdatascience.com/an-introduction-to-regression-discontinuity-design-f55075079def>