



Universidade de Brasília
Departamento de Administração - LAMFO

Diferenças em Diferenças

Oficina LAMFO

Daniel Tavares de Castro
João Pedro Rodovalho Figueiredo

30 de outubro de 2021



- ▶ Técnica econométrica com objetivo de identificar o efeito de uma política, um programa ou uma intervenção em um determinado grupo (análise causal)
- ▶ A intervenção dá origem a um experimento natural ou quasi-experimento
- ▶ A própria intervenção divide a amostra em um grupo de tratamento, sujeito ao evento, e um grupo de controle, que não é afetado
- ▶ Os integrantes dos grupos não são escolhidos aleatoriamente como em um experimento real
- ▶ É necessário observar os grupos antes e depois do evento

► Exemplos

- Construção de uma unidade de tratamento de lixo afeta preços de casas próximas (grupo de tratamento), mas não o preço de casas distantes (grupo de controle)
- Introdução de um programa de alimentação escolar em uma região (grupo de tratamento) e não em outra (grupo de controle)
- Introdução de limite de preço para um determinado produto (grupo de tratamento) e não para outro produto substituto (grupo de controle)

MDR (*merchant discount rate*) cobrada pelos credenciadores dos estabelecimentos comerciais por função (débito ou crédito) antes e depois de intervenção regulatória.

Credenciador	MDR	Depois	Função
A	1.40%	0	Débito
B	1.45%	0	Débito
C	1.35%	0	Débito
D	1.38%	0	Débito
A	1.10%	1	Débito
B	1.12%	1	Débito
C	1.09%	1	Débito
D	1.11%	1	Débito
A	2.70%	0	Crédito
B	2.68%	0	Crédito
C	2.64%	0	Crédito
D	2.79%	0	Crédito
A	2.20%	1	Crédito
B	2.18%	1	Crédito

Médias

Função	Pré-intervenção	Pós-intervenção
Crédito	2.70%	2.19%
Débito	1.40%	1.11%

Qual a estimativa do efeito da intervenção?

Diferença entre os grupos depois da intervenção não poderia ser atribuída apenas a ela:

- ▶ Em um experimento aleatório, características agregadas dos grupos de controle e de tratamento não apresentam diferenças significativas; no experimento natural, grupos podem ser diferentes desde o começo
- ▶ Características dos grupos podem mudar com o tempo

DD, empregando diferenças de corte transversal de conjuntos de dados empilhados (*pooled cross-sectional data sets*), reduz a questão relacionada a variáveis omitidas (endogeneidade)

Aplicando logaritmos para facilitar a comparação

Função	Antes	Depois	Diferença
Crédito	0.994	0.784	-0.210
Débito	0.333	0.100	-0.233
Diferença	-0.661	-0.684	-0.023

$$DD = (0.100 - (0.333)) - (0.784 - 0.994) = -0.233 - (-0.210) = -0.023$$

ou

$$DD = (0.100 - (0.784)) - (0.333 - (0.994)) = -0.684 - (-0.661) = \mathbf{-0.023}$$

Diferença é significativa?

Generalizando

Grupo	Antes	Depois	Diferença
Controle	A	B	$B - A$
Tratamento	C	D	$D - C$
Diferença	$C - A$	$D - B$	$(D - C) - (B - A)$

Estimador de diferenças em diferenças: $(D - C) - (B - A)$

Formula básica

$$y_{it} = \beta_0 + \beta_1 D_{it}^{pos} + \beta_2 D_{it}^{tratamento} + \beta_3 D_{it}^{pos} * D_{it}^{tratamento} + \epsilon_{it} \quad (1)$$

Variáveis independentes são variáveis dummy para indicar se determinada observação pertence ao grupo de tratamento ou controle ou ao período anterior ou posterior ao evento

Valores esperados condicionais:

$$\begin{aligned} E[y_{it} | D_{it}^{pos} = 0, D_{it}^{tratamento} = 0] &= \beta_0 \\ E[y_{it} | D_{it}^{pos} = 1, D_{it}^{tratamento} = 0] &= \beta_0 + \beta_1 \\ E[y_{it} | D_{it}^{pos} = 0, D_{it}^{tratamento} = 1] &= \beta_0 + \beta_2 \\ E[y_{it} | D_{it}^{pos} = 1, D_{it}^{tratamento} = 1] &= \beta_0 + \beta_1 + \beta_2 + \beta_3 \end{aligned}$$

Sistematizando

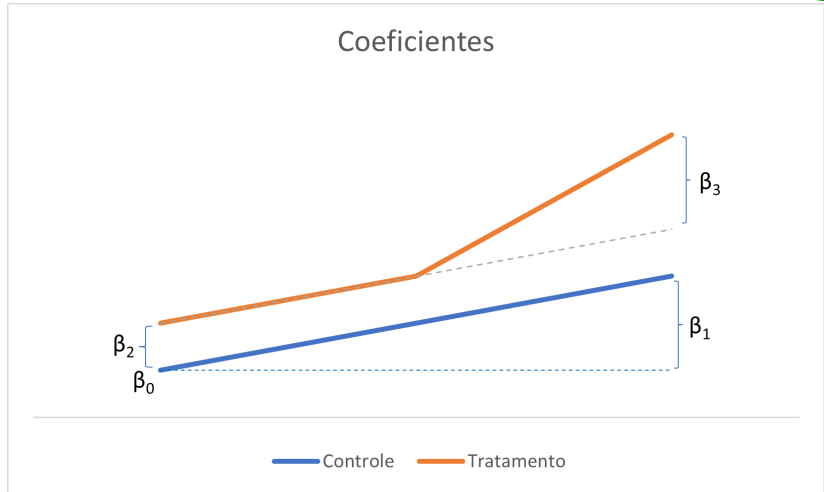
Grupo	Antes	Depois	Diferenças
Controle	β_0	$\beta_0 + \beta_1$	β_1
Tratamento	$\beta_0 + \beta_2$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_1 + \beta_3$
Diferenças	β_2	$\beta_2 + \beta_3$	β_3

β_0 é o nível original do grupo de controle

β_1 é a tendência original dos grupos

β_2 é a diferença inicial entre os grupos

β_3 é o efeito da intervenção, ou seja, é o parâmetro de interesse, o estimador de diferenças em diferenças ou efeito médio do tratamento (*average treatment effect - ATE*)



```
mdr <- read.csv('exemplo_dif_dif.csv', sep = ';')  
mdr['Tratamento'] <- mdr$Funcao == 'Debito'
```

```
mdr
```

##	i..Credenciador	MDR	Depois	Funcao	Tratamento
## 1	A	1.40	0	Debito	TRUE
## 2	B	1.45	0	Debito	TRUE
## 3	C	1.35	0	Debito	TRUE
## 4	D	1.38	0	Debito	TRUE
## 5	A	1.10	1	Debito	TRUE
## 6	B	1.12	1	Debito	TRUE
## 7	C	1.09	1	Debito	TRUE
## 8	D	1.11	1	Debito	TRUE
## 9	A	2.70	0	Credito	FALSE
## 10	B	2.68	0	Credito	FALSE
## 11	C	2.64	0	Credito	FALSE
## 12	D	2.79	0	Credito	FALSE
## 13	A	2.20	1	Credito	FALSE
## 14	B	2.18	1	Credito	FALSE
## 15	C	2.25	1	Credito	FALSE
## 16	D	2.13	1	Credito	FALSE

```
fit <- lm(log(MDR) ~ Depois + Tratamento + Depois * Tratamento, mdr)
summary(fit)
```

```
##
## Call:
## lm(formula = log(MDR) ~ Depois + Tratamento + Depois * Tratamento,
##     data = mdr)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.032451 -0.011258 -0.002552  0.006945  0.039008
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.99397    0.01144  86.854 < 2e-16 ***
## Depois        -0.21026    0.01618 -12.992 1.99e-08 ***
## TratamentoTRUE -0.66142    0.01618 -40.867 2.98e-14 ***
## Depois:TratamentoTRUE -0.02250    0.02289  -0.983  0.345
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02289 on 12 degrees of freedom
## Multiple R-squared:  0.9969, Adjusted R-squared:  0.9961
## F-statistic: 1277 on 3 and 12 DF, p-value: 2.717e-15
```

► Regressão MQO

- Relações lineares
- Observações independentes
- Ausência de colinearidade
- Erro distribuição normal com média zero e variância finita e não correlacionado com as variáveis independentes

$$E[\epsilon_{it} | D_{it}^{pos} * D_{it}^{tratamento}] = 0$$

- Tendências paralelas dos grupos (principal)
- Tendência do grupo de tratamento se manteria sem a intervenção (grupo de controle como contrafactual do grupo de tratamento)

Problemas

- ▶ Correlação dos grupos com variável que os afeta de forma diferente
- ▶ Composição dos grupos muda antes e depois do evento
- ▶ Observações não são independentes
- ▶ Tendência dos grupos diferente antes do evento

$$y_{it} = \beta_0 + (\beta_1 + \beta_4 D_{it}^{tratamento}) D_{it}^{pos} + \beta_2 D_{it}^{tratamento} + \beta_3 D_{it}^{pos} * D_{it}^{tratamento} + \epsilon_{it} \quad (2)$$

Soluções

- ▶ Usar erros robustos clusterizados para lidar com as observações dos mesmos indivíduos
- ▶ Justificar as tendências paralelas, mesmo que visualmente
- ▶ Testar econometricamente se tendências são paralelas via a seguinte regressão:

$$y = \beta_0 + \beta_1 D^{tratamento} + \sum_t (\beta_{2t} d_t + \beta_3 D^{tratamento} * d_t) \quad (3)$$

teste F para ver se todos os $\beta_{3t} = 0$

Considerando variáveis de controle para diferenças observáveis entre e dentro dos grupos

$$y_{it} = \beta_0 + \beta_1 D_{it}^{pos} + \beta_2 D_{it}^{tratamento} + \beta_3 D_{it}^{pos} * D_{it}^{tratamento} [+ \beta_4 X_{it}] + \epsilon_{it} \quad (4)$$

β_1 "controla" por mudanças não observáveis no tempo que afetam ambos os grupos

β_2 "controla" por diferenças não observáveis entre os grupos

β_4 "controla" por diferenças observáveis entre e dentro dos grupos

β_3 é o que "sobra"

Efeitos fixos de grupo (α) e de tempo (γ)

$$y_{igt} = \alpha_g + \gamma_t + \beta I_{gt} + \delta X_{igt} + \epsilon_{igt} \quad (5)$$

em que g denota um grupo e t um período.

β é o parâmetro de interesse e X são variáveis de controle.

Permite mais flexibilidade. Ex: adicionando os cartões pré-pagos como mais um grupo de controle.

Estudos especiais do Banco Central



Avaliação do resultado regulatório da introdução de limites para a tarifa de intercâmbio de cartão de débito

Estudo Especial nº 106/2021

Introdução

Este estudo avalia os resultados decorrentes da introdução, pela Circular 3.887, de 26 de março de 2018, de limites máximos (“cap”) para a tarifa de intercâmbio (TIC) incidente sobre as operações efetuadas com cartão de débito no âmbito dos arranjos de pagamento domésticos. A Circular, que entrou em vigor em 1º de outubro de 2018, estabeleceu os limites de 0,5% para a média ponderada pelo valor das transações e de 0,8% para o valor máximo a ser aplicado em qualquer transação, ambos por trimestre.

Exemplo

Mercado de dois lados

Z. Wang / Int. J. Ind. Organ. 28 (2010) 86–98

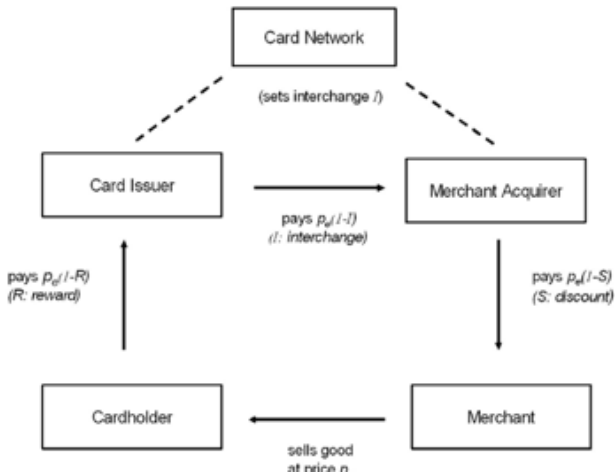
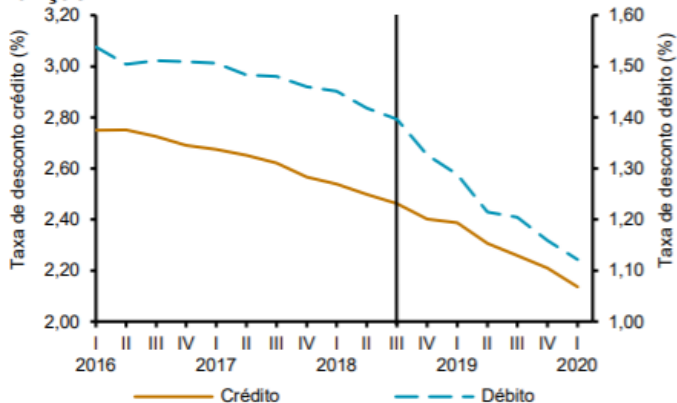


Gráfico 1 – Taxa de desconto média do mercado por função



Para avaliar se o *cap* provocou impactos distintos entre o MDR do débito (grupo de tratamento) e o MDR do crédito (grupo de controle), foi estimado um modelo econométrico de estimativa ponto a ponto dessas diferenças, conforme Equação (3).

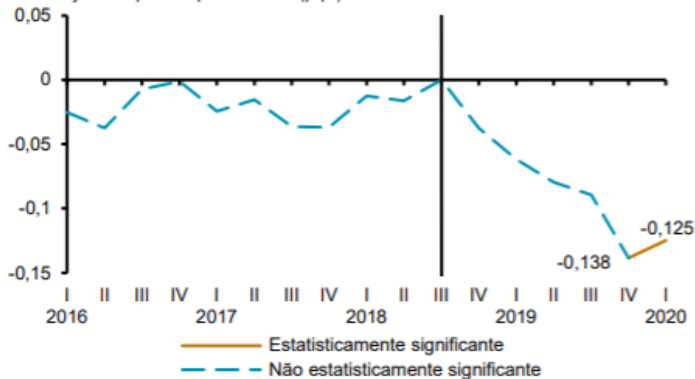
$$\ln(Y_{i,t}) = \beta_0 + \beta_1 \text{função}_i + \sum_{j=2, t=1}^{j=17, t=16} \beta_j \text{tri}_t + \sum_{j=18, t=1}^{j=33, t=16} \beta_j \text{tri}_t * \text{função}_i + \beta_{34} \ln(L_{i,t}) + \beta_{35} \ln(Mks_{i,t}) + \lambda_i + \varepsilon_{i,t}, \quad (3)$$

$\text{tri}_t * \text{função}_i$: interação entre cada um dos 16 trimestres avaliados, ou seja, de 2016T1 à 2020T1, exceto 2018T3, e a função.

Gráfico 3 – Diferença entre o MDR do crédito e do débito

Base = 2018T3

Variação em pontos percentuais (p.p.)



```
areg LN_DESC ib20183.ANO_TRIM_S##FUNCAO LN_PERC LN_LERNER if bandeira == 0, a( ispb )
```

Linear regression, absorbing indicators
Absorbed variable: ispb

```
Number of obs      =      204
No. of categories  =        8
F(   35,   161)    =      60.69
Prob > F           =      0.0000
R-squared          =      0.9390
Adj R-squared      =      0.9231
Root MSE          =      0.0852
```

	LN_DESC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
	ANO_TRIM_S						
	20161	.0575232	.0537585	1.07	0.286	-.0486395	.1636859
(...)	20193	-.0351478	.0457074	-0.77	0.443	-.1254111	.0551155
	20194	-.0402373	.0460621	-0.87	0.384	-.1312011	.0507266
	20201	-.0989816	.0492993	-2.01	0.046	-.1963382	-.001625
	FUNCAO						
	D	-.4974637	.045725	-10.88	0.000	-.5877618	-.4071656
	ANO_TRIM_S#FUNCAO						
	20161#D	-.0256206	.0755362	-0.34	0.735	-.1747901	.1235488
(...)	20184#D	-.0383409	.0637727	-0.60	0.549	-.1642798	.087598
	20193#D	-.0910659	.0645932	-1.41	0.161	-.218625	.0364932
	20194#D	-.137941	.0661891	-2.08	0.039	-.2686518	-.0072301
	20201#D	-.1245331	.0690152	-1.80	0.073	-.2608249	.0117586
	LN_PERC	.0142959	.0124465	1.15	0.252	-.0102836	.0388754
	LN_LERNER	.0224442	.0093771	2.39	0.018	.0039261	.0409622
	_cons	1.02946	.0516102	19.95	0.000	.9275401	1.131381

F test of absorbed indicators: F(7, 161) = 11.866

Prob > F = 0.000

Técnica empregada para identificar efeitos de intervenções em grupo(s) específico(s)

Requer que a amostra possa ser separada em quatro grupos (tratamento e controle, antes e depois da intervenção) e que as tendências dos grupos antes da intervenção seja paralela

Operacionalizada via MQO com dados transversais empilhados

Pode ser estendida para incorporar mais de um grupo e período