

Métodos Quantitativos em Ciência Política IV

Replicação do artigo “O impacto das leis de financiamento de campanha na competição partidária” (Potter & Tavits, 2013)

Luiz Henrique da Silva Batista (NUSP: 12687228)

Introdução

Este trabalho é uma replicação das análises empíricas que Potter e Tavits (2013) realizaram para embasar a hipótese de que as leis de financiamento de campanha impactam a competição partidária. Os autores reconhecem que o tamanho do sistema partidário é afetado pelas regras eleitorais e pelas clivagens sociais, conforme boa parte da literatura tem mostrado. No entanto, defendem que, mesmo controlando pelas variáveis institucionais e sociológicas, o conjunto de regras que estruturam o financiamento de campanhas apresenta forte efeito preditivo na competição entre os partidos.

Para mensurar a relação entre essas duas variáveis, os autores sintetizam a noção de competição partidária no *Número Efetivo de Partidos* (NEP)¹ e os regimes de financiamento de campanhas no conceito de *fund parity*. Este conceito busca caracterizar sistemas partidários que os autores chamam de *permissivos*, nos quais muitos partidos têm boas chances de ganhar espaço na legislatura, e *restritivos*, nos quais poucos partidos têm tais chances. A métrica que desenvolveram para este conceito varia de -2 (sistema restritivo) até 4 (sistema permissivo), conforme os critérios definidos abaixo:

- Se há limitação na quantidade de doação permitida (1), se não há limitação (0);
- Se há limitação dos gastos de campanhas dos partidos (1), se não há limitação (0);
- Se o acesso ao espaço na mídia é baseado no desempenho anterior do partido (-1), se o acesso é distribuído equitativamente (1), se o acesso não é permitido (0);
- Se o financiamento público das campanhas é baseado no desempenho anterior do partido (-1), se o financiamento público é distribuído equitativamente (1), se não tem financiamento público (0).

Assim, em síntese, o argumento dos autores é captar qual o efeito exclusivo do *fund parity*. Representando tal argumento em notação escalar, temos:

$$NEP = \alpha + \beta_1 \cdot X_{Fund\ Parity} + \beta_2 \cdot X_2 + \dots + \beta_k \cdot X_k + \epsilon$$

¹Desenvolvido por Laakso e Taagepera (1979) e Taagepera e Shugart (1989).

Descrição da base de dados

A base de dados utilizada pelos autores contém informações sobre 124 países. Além do número efetivo de cada país, dos critérios para a construção do índice de paridade de financiamento e do próprio índice, a base armazena algumas variáveis institucionais, como a forma do estado e o sistema de governo, e sociológicas, como as diferenças etnolinguísticas existentes no país.

Os dados também correspondem a dois momentos no tempo: 2003 e 2012. Mas isso não vale para todos os países. Isso porque os dados foram coletados por meio de fontes variadas. O número efetivo de partidos, por exemplo, foi obtido através de Gallagher e Mitchell (2008), complementado por dados disponibilizados por Adam Carr e pela European Elections Database. Já as observações a respeito do regime de financiamento das campanhas foram obtidas por meio do Internacional IDEA, pesquisa conduzida por Austin e Tjersnstrom (2003).

Replicação dos resultados

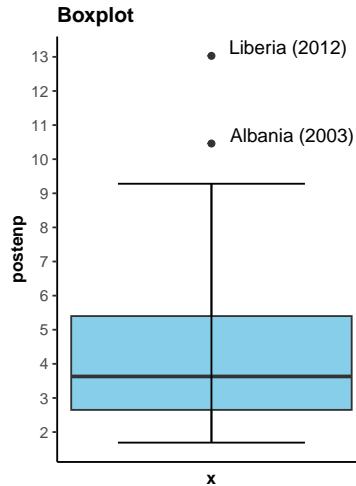
Os autores apresentam dois modelos no artigo. O primeiro com todas as democracias, exceto três casos considerados “outliers”, e o segundo deles somente com países que se tornaram democracias depois de 1974, recorte temporal justificado por ser uma das maneiras de se lidar com uma suposta endogeneidade dos dados (voltaremos a isso mais adiante). Vamos fazer a replicação desses dois modelos e de outros que foram apresentados nas notas de rodapé. Antes disso, no entanto, vamos fazer uma análise para detectar potenciais outliers na base de dados e verificar se chegamos às mesmas conclusões dos autores.

Análise dos outliers

Os três outliers retirados pelos autores foram Libéria (fund parity medido em 2012), Albânia (fund parity medido em 2003) e Brasil (fund parity medido em 2003). No entanto, os autores não evidenciam qual o método que utilizaram para a identificação desses casos. Embora existam diversas formas para identificar valores considerados outliers, aqui optamos por fazer essa detecção através de duas técnicas: plotamos um gráfico box-plot da variável *postenp* (variável dependente) e um gráfico que identifica valores que afetam as estimativas de um modelo de regressão.

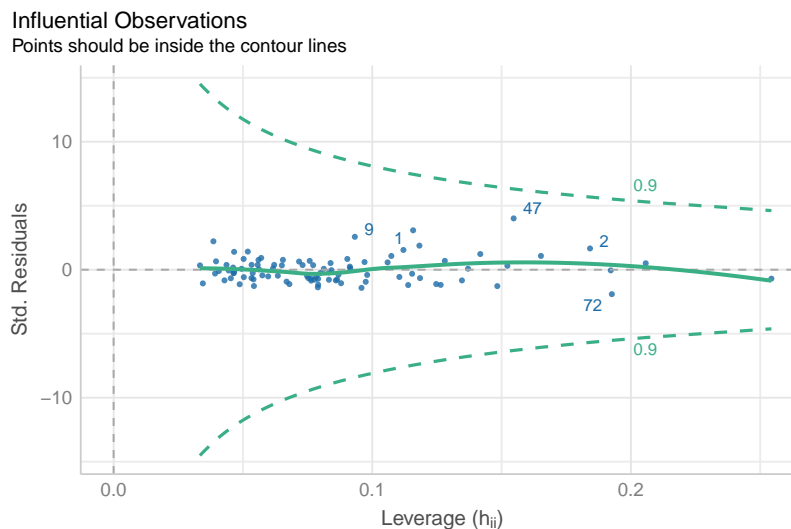
1. Box-plot

Os dois valores outliers da variável *postenp* são 10.5 e 13, os quais correspondem, respectivamente, à Albânia (2003) e à Libéria (2012). O Brasil (2003), considerado como outlier pelos autores, não o foi aqui. No entanto, devemos lembrar que, no R, o valor default para multiplicação do intervalo interquartil (IQR) é 1,5. Em outros softwares, o padrão pode ser outro.



2. Cook's distance

O método estatístico de “Cook's distance” identifica os valores influentes plotando-os depois da linha tracejada. Para fazer essa análise, é necessário, primeiro, rodar o modelo de regressão para, depois, gerar o gráfico. Assim, rodamos um modelo utilizando todos os dados. Pela análise do gráfico, nota-se a ausência de valores influentes.²



Modelos apresentados no artigo

Abaixo, temos a replicação dos dois modelos apresentados pelos autores no paper, excluindo a Libéria (2012), a Albânia (2003) e o Brasil (2003), países tidos como outliers pelos autores.

²Os valores considerados influentes não são, necessariamente, outliers.

Table 1: OLS Model Estimating the Effect of Fund Parity on ENP

	All democracies	1974 and later
	(1)	(2)
fundparity4	0.438*** (0.153)	0.454** (0.210)
demyears	0.013 (0.011)	-0.013 (0.033)
fed	-0.213 (0.476)	-0.228 (0.754)
pres	-0.172 (0.211)	-0.025 (0.284)
log(avemag)	0.601** (0.301)	0.343 (0.446)
fract	0.956 (1.290)	-0.427 (1.911)
log(avemag):fract	-0.750 (0.641)	-0.576 (0.889)
Constant	3.070*** (0.764)	4.383*** (1.217)
N	90	54
R ²	0.198	0.168

*p < .1; **p < .05; ***p < .01

Os resultados da replicação estão em conformidade com os resultados apresentados pelos autores.³ A única diferença reside no fato de que, aqui, os coeficientes estão arredondados em três casas decimais. Contudo, os resultados são essencialmente iguais.

Modelos mencionados nas notas de rodapé do artigo

No artigo, alguns modelos de regressão são mencionados em notas de rodapé. A seguir, replicaremos esses modelos indicando a que nota se referem e expondo brevemente o propósito de cada um deles.

1. Contribuição dos componentes do índice de *fund parity* (nota de rodapé 42)

Nesta nota de rodapé, os autores buscam verificar se alguma das quatros variáveis que compõem o fund parity trabalha contra as outras três, ou seja, se alguma variável apresenta um efeito oposto em relação às demais.

1.1 Algum componente trabalha contra os outros?

Os autores argumentam que todos os componentes estão positivamente associados à paridade de fundos (fund parity), embora somente a limitação da quantidade de doação permitida seja estatisticamente significativa. De fato, é isso o que acontece.

³Esses modelos foram apresentados na tabela 2 do artigo.

Table 2: Modelo da nota de rodapé 42 (1)

	Model
directelig	0.482 (0.308)
partyspend	0.167 (0.430)
donorlimit	1.145*** (0.412)
eligmedia	0.221 (0.225)
demyears	0.011 (0.011)
fed	-0.363 (0.492)
pres	-0.182 (0.215)
log(avemag)	0.642** (0.302)
fract	1.202 (1.296)
log(avemag):fract	-0.853 (0.639)
Constant	2.921*** (0.800)
N	90
R ²	0.236

*p < .1; **p < .05; ***p < .01

1.2 Mesmo quando cada componente é analisado individualmente?

Os autores também afirmam que, quando cada indicador da paridade de fundos é estimado individualmente, novamente somente coeficientes positivos são observados. Uma vez mais, os resultados da replicação correspondem ao que foi apresentado no artigo.

Table 3: Modelo da nota de rodapé 42 (2)

	directelig (1)	partyspend (2)	donorlimit (3)	eligmedia (4)
directelig	0.226 (0.318)			
partyspend		0.588 (0.392)		
donorlimit			1.108*** (0.365)	
eligmedia				0.187 (0.236)
demyears	0.009 (0.011)	0.010 (0.011)	0.007 (0.011)	0.008 (0.012)
fed	-0.193 (0.498)	-0.219 (0.483)	-0.209 (0.477)	-0.014 (0.516)
pres	-0.085 (0.222)	-0.018 (0.217)	-0.110 (0.209)	-0.131 (0.229)
log(avemag)	0.625** (0.305)	0.682** (0.305)	0.713** (0.295)	0.644** (0.321)
fract	0.789 (1.322)	0.959 (1.271)	1.345 (1.261)	1.197 (1.391)
log(avemag):fract	-0.560 (0.646)	-0.747 (0.641)	-0.946 (0.628)	-0.662 (0.678)
Constant	3.210*** (0.826)	2.687*** (0.795)	2.579*** (0.763)	2.986*** (0.822)
N	95	92	92	93
R ²	0.104	0.148	0.204	0.105

*p < .1; **p < .05; ***p < .01

1.3 E quando são analisados três componentes de cada vez?

Também afirmam que quando os indicadores são apresentados de 3 (três) em 3 (três), também temos somente coeficientes positivos. De novo, a replicação segue alinhada com os resultados do artigo.

Table 4: Modelo da nota de rodapé 42 (3)

	-eligmedia	-donorlimit	-partyspend	-directelig
	(1)	(2)	(3)	(4)
directelig	0.512* (0.306)	0.323 (0.315)	0.463 (0.305)	
partyspend	0.080 (0.420)	0.675* (0.405)		0.192 (0.433)
donorlimit	1.153*** (0.412)		1.241*** (0.370)	1.024** (0.408)
eligmedia		0.233 (0.235)	0.212 (0.219)	0.256 (0.226)
demyears	0.010 (0.011)	0.013 (0.011)	0.010 (0.011)	0.009 (0.011)
fed	-0.424 (0.488)	-0.233 (0.510)	-0.308 (0.485)	-0.226 (0.489)
pres	-0.149 (0.213)	-0.094 (0.222)	-0.197 (0.212)	-0.162 (0.217)
log(avemag)	0.702** (0.296)	0.652** (0.314)	0.676** (0.297)	0.639** (0.305)
fract	1.298 (1.292)	1.052 (1.348)	1.260 (1.288)	1.474 (1.296)
log(avemag):fract	-0.926 (0.635)	-0.796 (0.665)	-0.952 (0.629)	-0.866 (0.645)
Constant	2.881*** (0.798)	2.929*** (0.832)	2.926*** (0.789)	2.580*** (0.776)
N	90	90	91	90
R ²	0.227	0.162	0.235	0.212

*p < .1; **p < .05; ***p < .01

2. Diferenças entre as regras e a sua aplicação (nota de rodapé 43)

Na 43^a nota de rodapé, os autores chamam a atenção para as possíveis diferenças que podem existir entre as regras institucionais e os efeitos delas nas estratégias partidárias.

O debate surge em razão de uma possível diminuição do efeito preditor das leis de financiamento de campanha em países com reduzido nível de aplicação dessas leis.

Assim, os autores interagiram a variável *fundparity4* com a variável *rulelaw* (que mede o nível de aplicação das leis).

Table 5: Modelo da nota de rodapé 43

	Model
fundparity4	0.571*** (0.160)
rulelaw	0.417 (0.273)
demyears	0.004 (0.012)
fed	-0.566 (0.491)
pres	-0.078 (0.225)
log(avemag)	0.461 (0.300)
fract	0.997 (1.263)
fundparity4:rulelaw	-0.280* (0.152)
log(avemag):fract	-0.528 (0.634)
Constant	3.110*** (0.747)
N	90
R ²	0.251

*p < .1; **p < .05; ***p < .01

O argumento é que se a interação entre a paridade de fundos e a qualidade do Estado de direito fosse estatisticamente significativa o impacto da paridade de fundos sobre a competição partidária tenderia a diminuir com menores níveis de Estado de direito. De fato, como apresentam na nota de rodapé, essa relação não é estatisticamente significativa.

3. Controlando pelo limiar eleitoral legal (nota de rodapé 54)

Neste modelo, os autores incluem a variável *thresh*, que mede o limiar eleitoral legal. Em conformidade com o que afirmam, de fato a inclusão dessa variável não afeta a *fund parity*, que continua estatisticamente significante.

Table 6: Modelo da nota de rodapé 54

	Model
fundparity4	0.476*** (0.166)
thresh	0.084 (0.087)
demyears	0.013 (0.014)
fed	-0.537 (0.557)
pres	-0.331 (0.264)
log(avemag)	0.764** (0.345)
fract	2.966* (1.717)
log(avemag):fract	-1.505** (0.744)
Constant	2.651** (1.037)
N	69
R ²	0.256

*p < .1; **p < .05; ***p < .01

Lidando com a endogeneidade

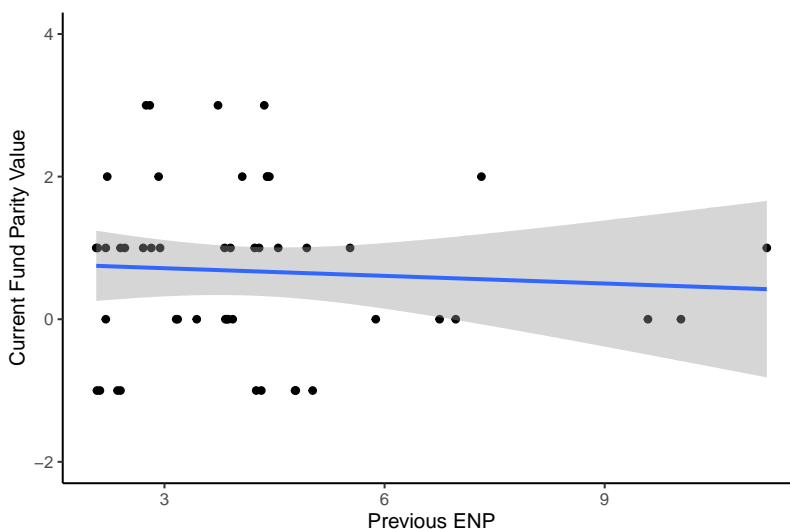
Os autores sugerem como direção causal que as regras de financiamento de campanha impactam a competição partidária. Ou seja, argumentam que:

$$Fund\ parity \rightarrow ENP$$

No entanto, para demonstrar que a direção é realmente essa e não o contrário, isto é, $ENP \rightarrow Fund\ parity$, precisam rodar análises adicionais.⁴

Em parte, os autores já lidaram com este problema ao rodarem a regressão somente com países que se tornaram democracias recentemente, porque partem do pressuposto de que as regras de financiamento de campanha foram implementadas antes do sistema partidário ou de forma exógena a ele (caso das pressões exercidas por organizações promotoras de democracia sobre as regras que deveriam ser adotadas).

Outra maneira que utilizaram para lidar com a endogeneidade foi estimar uma reta de regressão entre o NEP e o fund parity (figura 1).



O gráfico acima reproduz a figura 1 apresentada no artigo, demonstrando que, de fato, os resultados, aqui e lá, são idênticos.

Checagem dos pressupostos

Vamos agora fazer a checagem dos dois principais modelos apresentados no artigo, isto é, o modelo com todas as democracias (*all democracies*) e o somente com as democracias recentes (*1974 and later*).

⁴Apesar de falarem em causalidade na relação entre essas duas variáveis, cabe perguntar se as análises desenvolvidas realmente são capazes de evidenciar uma relação de causa e efeito.

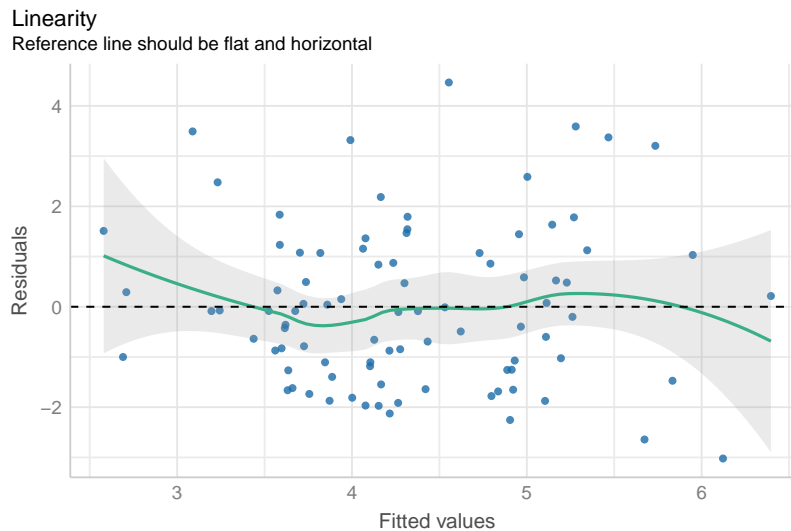
1. Linearidade

Supor linearidade significa que podemos descrever a relação entre a variável dependente e independente por meio de uma reta. Violar esse pressuposto implica na não consistência dos coeficientes.

Um gráfico dos resíduos pelos valores ajustados, embora útil para verificar a homocedasticidade, é também utilizado para verificar se temos uma relação linear entre a variável dependente e as variáveis independentes. Para este caso, a linha verde plotada no gráfico deve ser aproximadamente horizontal.

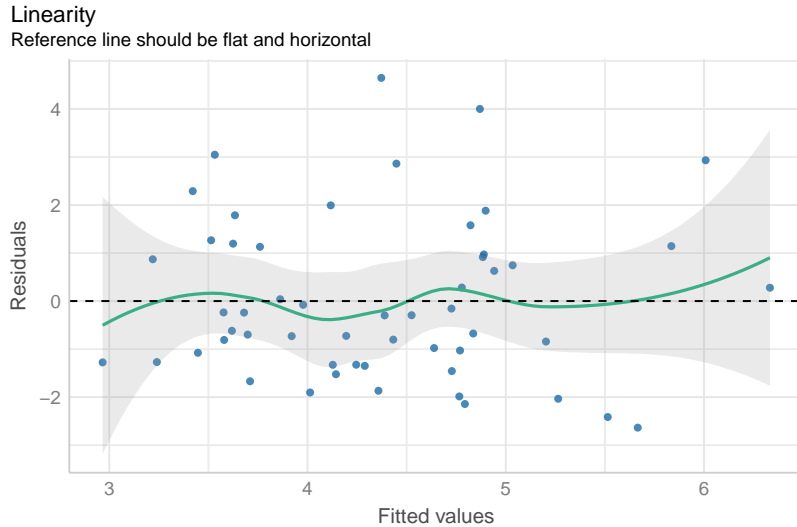
Modelo: *all democracies*

Ao analisar o gráfico abaixo, podemos perceber que a linha verde é aproximadamente horizontal, mas apresenta curvas suaves e se afasta um pouco nas extremidades. No entanto, se considerarmos o erro padrão, a linha verde está dentro do intervalo da reta horizontal em toda sua extensão.



Modelo: *1974 and later*

No caso do modelo somente com as democracias recentes, temos a mesma interpretação do modelo acima, com a exceção de que, aqui, a linha verde se afasta menos da reta horizontal nos extremos.



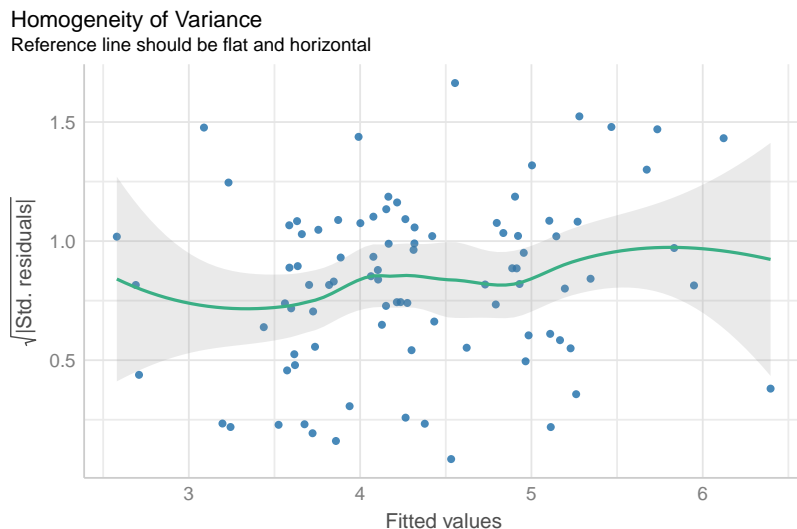
Podemos dizer que os dois modelos parecem apresentar uma relação entre a variável dependente e as variáveis independentes que não é perfeitamente linear.

2. Homocedasticidade

Supor homocedasticidade significa que os resíduos variam ao longo dos valores preditos sem apresentarem qualquer padrão sistemático de variação. Aqui, os pontos devem se espalhar aleatoriamente em torno da reta.

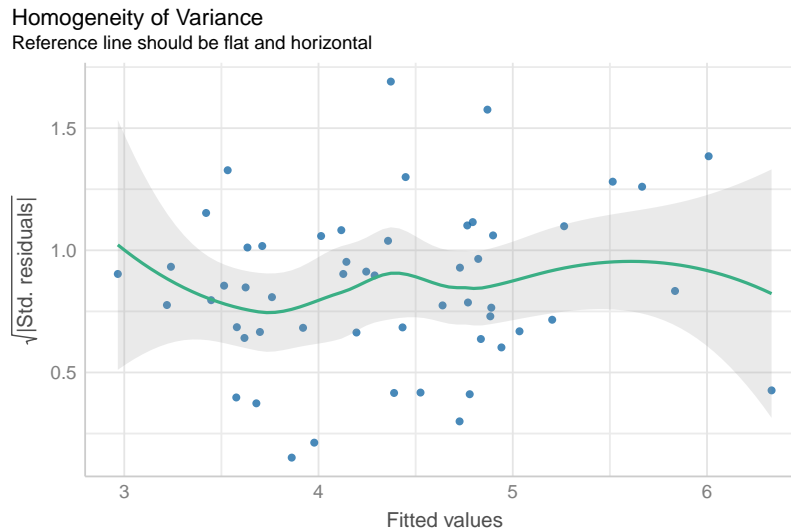
Modelo: *all democracies*

No gráfico acima, dos resíduos pelos valores ajustados, já foi possível notar que o modelo erra nos extremos. No gráfico que plotamos abaixo, temos a raiz quadrada dos resíduos pelos valores ajustados, que facilita a identificação de padrões nos resíduos. Novamente, notamos uma variação na distribuição dos resíduos nos extremos.



Modelo: *1974 and later*

Novamente, a distribuição dos pontos revela heterogeneidade na distribuição dos pontos. A linha verde plotada no gráfico demonstra que o modelo erra para mais ou para menos ao longo dos valores dos preditores.



Diante de heterocedasticidade, devemos rodar modelos com erro padrão robusto. Faremos isso mais adiante para verificar se os coeficientes estimados se alteram.

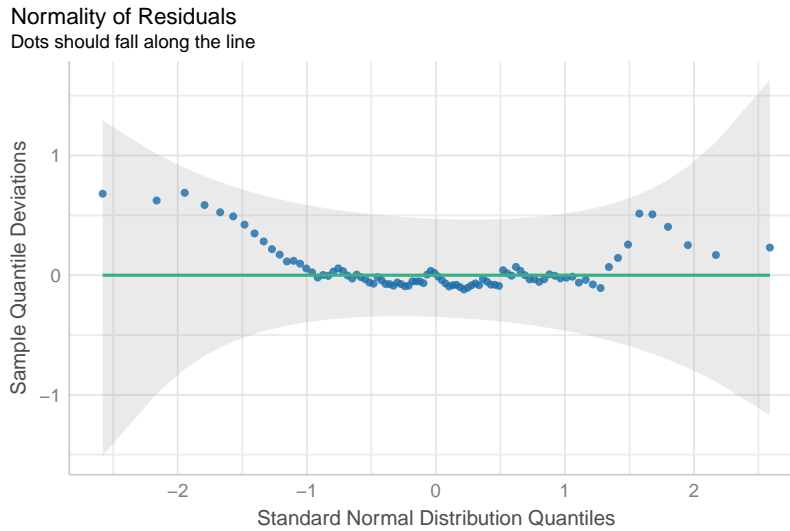
3. Normalidade dos resíduos

Satisfazer o pressuposto de que os resíduos do modelo seguem uma distribuição normal garante que as estimativas sejam consistentes. Aqui, esperamos encontrar os pontos sobrepostos na reta horizontal.

Modelo: *all democracies*

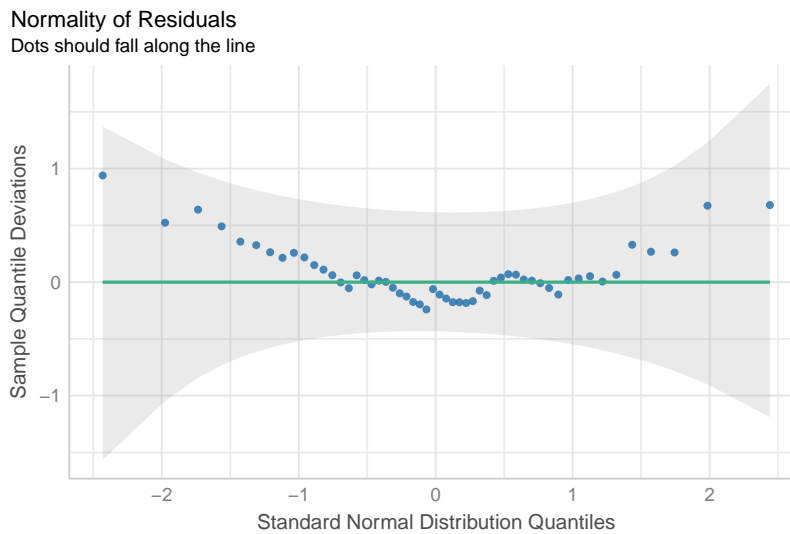
A distribuição dos pontos se sobrepõe à reta verde de referência, mas se afastam dela nos extremos. Isso acontece por serem “valores mais improváveis” e, portanto, “tendem a ficar um pouco mais distante da reta”.⁵

⁵Galdino. CEF (capítulo 6). In: _____. Introdução à Regressão para Ciências Sociais.



Modelo: 1974 and later

A mesma interpretação vale para este modelo. As estimativas são afetadas nos valores extremos.



4. Resíduos outliers ou pontos de alavancagem

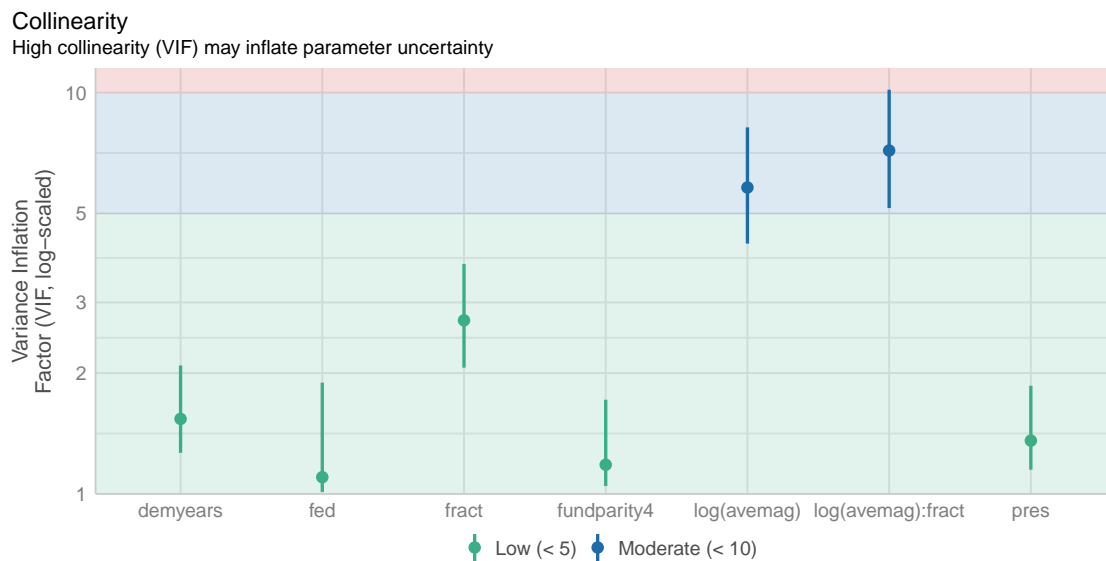
Já rodamos uma análise dos outliers mais acima, considerando a base de dados com todos os valores. Conforme foi demonstrado, nenhum valor influente foi detectado. Supõe-se, portanto, que o mesmo vale para o conjunto de dados sem aqueles valores considerados outliers pelos autores.

5. Ausência de multicolinearidade

Multicolinearidade se refere às variáveis independentes correlacionadas entre si. O pressuposto da regressão linear é de que isso não deve acontecer.

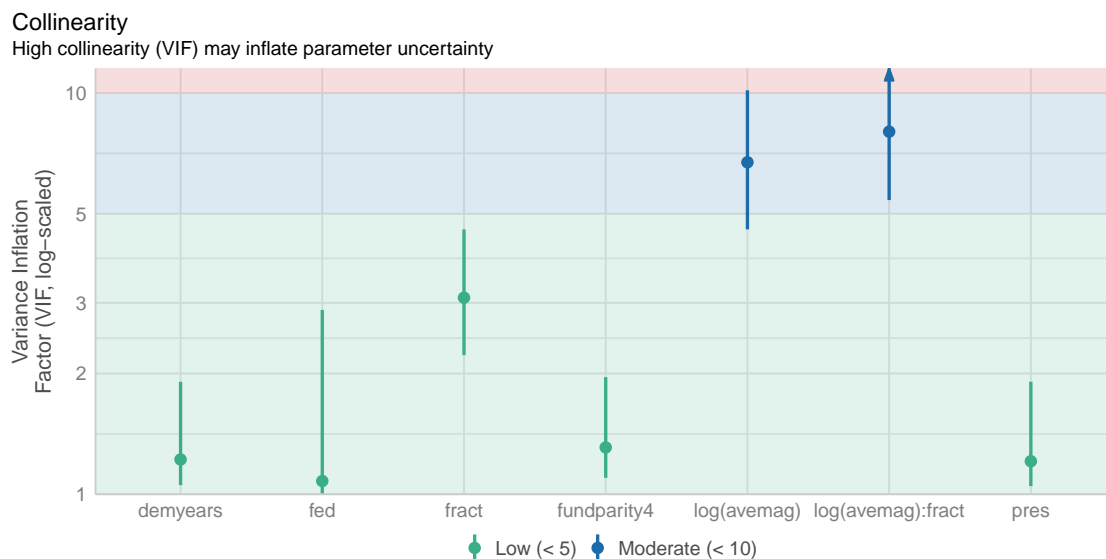
Modelo: *all democracies*

Todas as variáveis apresentam baixa colinearidade, com exceção das variáveis que foram interagidas. No entanto, essa situação é entre variáveis submetidas a interações ou transformações.



Modelo: *1974 and later*

A mesma interpretação vale aqui.



Modelos com erro padrão robusto

Os modelos de regressão linear com erro padrão robusto são uma forma de lidar com a heterocedasticidade. Esses modelos garantem estimativas consistentes mesmo em cenários de violação do pressuposto de variância constante dos erros. Abaixo, replicamos os modelos apresentados pelos autores, mas com erro padrão robusto.

Table 7: Modelos apresentados na tabela 2 (com erro padrão robusto)

	Todas as democracias	Pós-1974
	(1)	(2)
fundparity4	0.438** (0.182)	0.454** (0.182)
demyears	0.013 (0.011)	-0.013 (0.011)
fed	-0.213 (0.563)	-0.228 (0.563)
pres	-0.172 (0.217)	-0.025 (0.217)
log(avemag)	0.601* (0.351)	0.343 (0.351)
fract	0.956 (1.348)	-0.427 (1.348)
log(avemag):fract	-0.750 (0.719)	-0.576 (0.719)
Constant	3.070*** (0.773)	4.383*** (0.773)

*p < .1; **p < .05; ***p < .01

Mesmo com erro padrão robusto, a variável *fund parity* continua sendo estatisticamente significativa.

Análises adicionais

Finalmente, vamos rodar três modelos que não foram apresentados no artigo, mas que consideramos importantes para avaliar o papel dos “outliers”. São modelos com as mesmas variáveis dos modelos principais, mas com mudanças na base de dados. Os modelos são os seguintes:

- Modelo 1: dados completos (considerando os “outliers” excluídos pelos autores);
- Modelo 2: dados sem os outliers que foram revelados pelo gráfico box-plot da variável *postenp* (ou seja, sem os outliers que nossa análise revelou);
- Modelo 3: dados sem os outliers da variável *postenp* e somente com democracias recentes.

O modelo com todos os dados não só altera o valor estimado para a variável *fund parity*, como também a sua significância estatística, que passa de 0,01 para 0,1. Por outro lado, no modelo sem os outliers identificados por nós, a variável *fund parity* volta a ser estatisticamente significativa, mas ao nível de 0,05. Já no modelo sem os outliers e somente com as democracias recentes, a variável *fund parity* volta a perder significância estatística. Isso demonstra que os coeficientes parecem ser bastante afetados pela presença ou ausência de valores outliers no conjunto de dados.

Table 8: Verificando o peso dos outliers

	(1)	(2)	(3)
fundparity4	0.314* (0.185)	0.400** (0.160)	0.363 (0.223)
demyears	-0.001 (0.013)	0.013 (0.011)	0.003 (0.035)
fed	0.066 (0.578)	0.105 (0.489)	0.444 (0.768)
pres	-0.193 (0.259)	-0.089 (0.220)	0.098 (0.301)
log(avemag)	0.971*** (0.366)	0.567* (0.316)	0.291 (0.478)
fract	2.894* (1.525)	0.786 (1.355)	-0.448 (2.049)
log(avemag):fract	-1.731** (0.761)	-0.587 (0.671)	-0.333 (0.949)
Constant	2.951*** (0.940)	3.021*** (0.803)	3.916*** (1.293)
N	93	91	55
R ²	0.131	0.175	0.137

*p < .1; **p < .05; ***p < .01

Conclusões

Os resultados da replicação são consistentes com os resultados apresentados no artigo. No entanto, os principais modelos apresentados (*all democracies* e *1974 and later*) revelam violações em alguns pressupostos da regressão linear, como a linearidade e normalidade (nos valores extremos) e a homocedasticidade. Além disso, os outliers mostraram-se casos cruciais na estimativa dos coeficientes desses modelos. Rodar o modelo com todos os dados compromete a significância estatística da principal variável independente dos autores: a *fund parity* (paridade de fundos). Portanto, há uma lacuna a ser preenchida para compreender os critérios adotados pelos autores na identificação dos valores que consideram outliers.

Cabe também pontuar que, ao longo do artigo, não fica claro se o propósito é desenvolver um argumento causal ou preditivo. A interpretação que os autores fazem dos modelos sugere efeitos preditivos, mas ao lidarem com a questão da endogeneidade, a linguagem utilizada é a da causalidade: “[...] the direction of causality runs from fund parity to ENP” (Potter & Tavits, 2013).

Referências

Checking model assumption - linear models. Disponível em: https://easystats.github.io/performance/articles/check_model.html. Acesso em: 5 jan. 2024.

Figueiredo Filho, Dalson Britto. **Métodos quantitativos em ciência política.** Curitiba: InterSaberes, 2019.

Galdino, Manoel (2019). **Introdução à Regressão para Ciências Sociais.**

Potter, Joshua and Tavits, Margit. 2015. “The Impact of Campaign Finance Laws on Party Competition.” *British Journal of Political Science* 45(1): 73–95. doi: 10.1017/S0007123413000227.

Understanding Diagnostic Plots for Linear Regression Analysis. Disponível em: <https://library.virginia.edu/data/articles/diagnostic-plots>. Acesso em: 5 jan. 2024.