

# Ocena efektów zmian wprowadzonych w Kodeksie Wyborczym w zakresie organizacji wyborów organów wykonawczych gmin w 2018 roku - analiza kontrfaktyczna z zastosowaniem metody syntetycznej jednostki kontrolnej

Jabłonna 23-24 września 2021

---

Maciej Alimowski (Szkola Doktorska Nauk Społecznych Uniwersytet w Białymstoku)  
Sławomir Bartnicki (Uniwersytet w Białymstoku)  
Maciej Górecki (Uniwersytet Warszawski)  
Michał Pierzgalski (Uniwersytet Łódzki)

kontakt: [s.bartnicki@uwb.edu.pl](mailto:s.bartnicki@uwb.edu.pl)

# PROBLEM BADAWCZY

**Czy zmiany w organizacji wyborów gminnych organów wykonawczych w 2018 roku wpływają na anomalie wyborcze i wybrane zmienne charakteryzujące sytuację wyborczą?**

## **POJĘCIA:**

zmiany w organizacji wyborów organów wykonawczych gmin w 2018 roku:

- dwie komisje wyborcze, jedna odpowiedzialna za proces głosowania, druga za liczenie głosów
- formalne ograniczenie uprawnień organów wykonawczych w zakresie organizacji wyborów

anomalie wyborcze:

- brak walidacji efektu RDD w zakresie ciągłości rozkładu MSE i CER - w analizie uznawany przez zawężenie przewagi wyborczej wśród inkumbentów
- wtórnie, obecnie mniejsze znaczenie: brak walidacji w zakresie podobieństwa cech wśród ledwie wygrywających i przegrywających (randomizacja cech)

wybrane zmienne:

- charakterystyczne w analizach wyborczych i w zakresie poruszanego problemu, np.: głosy nieważne

## **HIPOTEZA**

Zmiany w organizacji wyborów gminnych organów wykonawczych nie wpływają na występowanie anomalii wyborczych.

# KONTEKST ANALIZY

Dotychczas w wyborach różnych typów ujawniono kilka przypadków anomalii danych wyborczych przy pomocy schematu RDD ostrej:

- Caughey i Sekhon 2011 wybory do Kongresu USA: szczegółowa walidacja danych Lee 2008, zbyt często ledwie wygrywają inkumbenci (randomizacja cech, przedziały arbitralne); przyczyna: nadzwyczajna mobilizacja wyborców inkumbentów w dniu wyborów
- Vogl 2014 wybory burmistrzów miast z południa i północy USA: zbyt często ledwie wygrywają czarni w południowych stanach (randomizacja cech, przedziały arbitralne); przyczyna: nadzwyczajna mobilizacja czarnych wyborców wskutek dziejowych krzywd
- Aksoy 2016 wybory burmistrzów w Turcji: zbyt często ledwie wygrywają kandydaci partii rządzącej (randomizacja cech, przedziały arbitralne); przyczyna: być: a) przypadek; b) manipulacja; c) inne np. kontekst wyborczy, w którym wyborcy na szczeblu lokalnym popierają kandydatów partii tworzącej rząd

Aksoy, O., 2016. Regression discontinuity diagnostics reveal statistical anomalies in Turkish elections. *Elect. Stud.* 44, 284–292.

Caughey, D., Sekhon, J.S., 2011. “Elections and the regression discontinuity design: lessons from close US House races, 1942–2008”. *Polit. Anal.* 19 (4), 385–408.

Lee, D.S., 2008. Randomized experiments from non-random selection in US House elections. *J. Econom.* 142 (2), 675–697.

Vogl, T.S., 2014. Race and the politics of close elections. *J. Publ. Econ.* 109, 101–113.

## ODNIESIENIA DO USA

- Eggers i in. 2015: weryfikacja walidacji na różnych typach wyborów: wnioski Caughey i Sekhon 2011 uznane za przypadek, wnioski Vogl 2014 uznane za nieadekwatne wskutek małych liczebności blisko c; dobór jednorodnych przedziałów wokół c dla różnych typów wyborów (lokalne, krajowe) - potencjalny błąd wnioskowania polegający na tym, że te same przedziały w różnych wyborach dają ostatecznie różną N głosów w przedziale
- Erikson i Rader 2017: weryfikacja Caughey i Sekhon 2011: brak walidacji uzyskany w skutek błędu obliczeniowego ( po prostu "fluke"), sugestia pominięcia bliskich wyników i uznania efektu za prawdziwy

Eggers, A.C., Fowler, A., Hainmueller, J., Hall, A.B., Snyder Jr., J.M., 2015. On the validity of the regression discontinuity design for estimating electoral effects: new evidence from over 40,000 close races. *Am. J. Polit. Sci.* 59 (1), 259–274.

Erikson, R.S., Rader, K., 2017. Much ado about nothing: rdd and the incumbency advantage. *Polit. Anal.* 25 (2), 269–275.

# WYBORY BEZPOŚREDNIE GMINNEJ EGZEKUTYWY W POLSCE

Bartnicki i in. 2021: wybory 2006-2018 do organów wykonawczych gmin: brak walidacji na ciągłości i randomizacji cech: koncentracja anomalii na gminach małych z inkumbentami i kandydatami sprawującymi funkcje wykonawcze w okresie wyborów pośrednich; przyczyna: strategie klientalne nie pozostawiające śladów w danych kwantytatywnych (walidacja zmiennych dotyczących: frekwencji, przemieszczeń wyborców między gminami i w ramach gminy, głosów nieważnych)

Systematyczne anomalie na licznych cechach osobniczych: inkumbent, wiek, sprawowanie funkcji przed 2002, poziom wykształcenia, płeć wykazywane na wiązaniach silnych CCT, CER

Bartnicki S., Alimowski M., Górecki A.M. (2021). *The anomalous electoral advantage: Evidence from over 17,000 mayoral candidacies in Poland*. *European Journal of Political Economy*, <https://doi.org/10.1016/j.ejpoleco.2021.102109>.

# PLAN BADANIA

Ocena, w jaki sposób interwencja w postaci modyfikacji procedury głosowania w wyborach w 2018 roku wpłynęła na występowanie zmniejszenia przewagi wyborczej i wartości wybranych zmiennych

## Dopasowanie grup:

- **naiwne: grupa kontrolna to przeciwieństwo cech grupy eksperymentalnej**
- bayesowskie: potencjalnie do zastosowania

Grupa eksperymentalna: rządzący od 2006 roku z doświadczeniem na stanowisku wykonawczym przed 2002 rokiem, gminy małe < 6199 uprawnionych do głosowania w roku wyborów

kontrolna 1: rządzący od 2006, bez doświadczenia na stanowisku sprzed 2002, gminy małe < 6199 uprawnionych

kontrolna 2: rządzący od 2006, bez doświadczenia na stanowisku sprzed 2002, gminy duże > 6199 uprawnionych

**model I: eksperymentalna x kontrolna 1 - zależna: vote margin**

**model II: eksperymentalna x kontrolna 2 - zależna: vote margin**

**model III: eksperymentalna x kontrolna 1 - zależna: głosy nieważne**

**model IV: eksperymentalna x kontrolna 2 - zależna: głosy nieważne**

schemat doboru predyktorów do modelu: algorytm Random Forest: Boruta (Kursa 2020)

Kursa M. B. (2020). Boruta for those in a hurry. <https://cran.r-project.org/web/packages/Boruta/vignettes/inahurry.pdf>.

# DANE

Macierz N=40384 startów wyborczych do gminnej egzekutywy od 2002 do 2018 roku z eliminacją III tur, wyborów przez radę gminy i losowań zwycięzców

Jednostka obserwacji: start wyborczy

Macierz opisana cechami kandydatów i gmin

Model I i III — 372 grupa kontrolna, 76 grupa eksperymentalna

Model II i IV — 942 grupa kontrolna, 78 grupa eksperymentalna

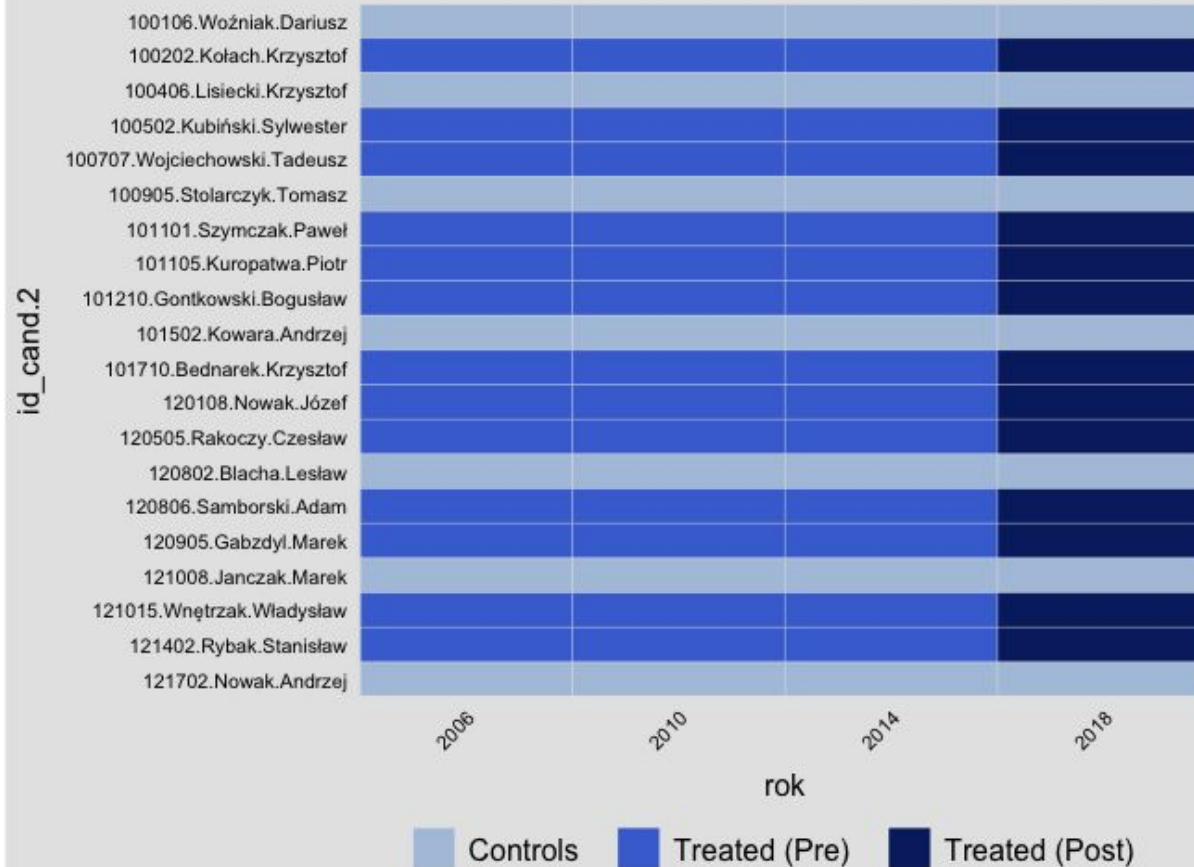
SCHEMAT *tjbal* (Hazlett, Xu 2018)

- minus 2 okres: 2006
- minus 1 okres: 2010
- okres 0: 2014
- bodziec: 2018

Hazlett, Chad and Yiqing Xu, 2018. "Trajectory Balancing: A General Reweighting Approach to Causal Inference with Time-Series Cross-Sectional Data." Working Paper, UCLA and Stanford. Available at SSRN:

[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3214231](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3214231).

## Treatment Status





# MODEL I (exp x kont 1)

## Model 1: MV ~ D

Call:

```
tjbal.default(data = data.f1.bal.3, Y = "mv", D = "D", X = c("bez",  
  "P2627_ogolem", "konk", "wmipc", "publwog", "kw"), X.avg.time = list(c(2006,  
  2010, 2014), c(2006, 2010, 2014), c(2006, 2010, 2014), c(2006,  
  2010, 2014), c(2006, 2010, 2014), c(2006, 2010, 2014)), index = c("id_cand.2",  
  "rok"), Y.match.time = c(2006, 2010, 2014), demean = T, estimator = "kernel",  
  vce = "jackknife", parallel = T, seed = 1234)
```

~ by Period (including Pre-treatment Periods):

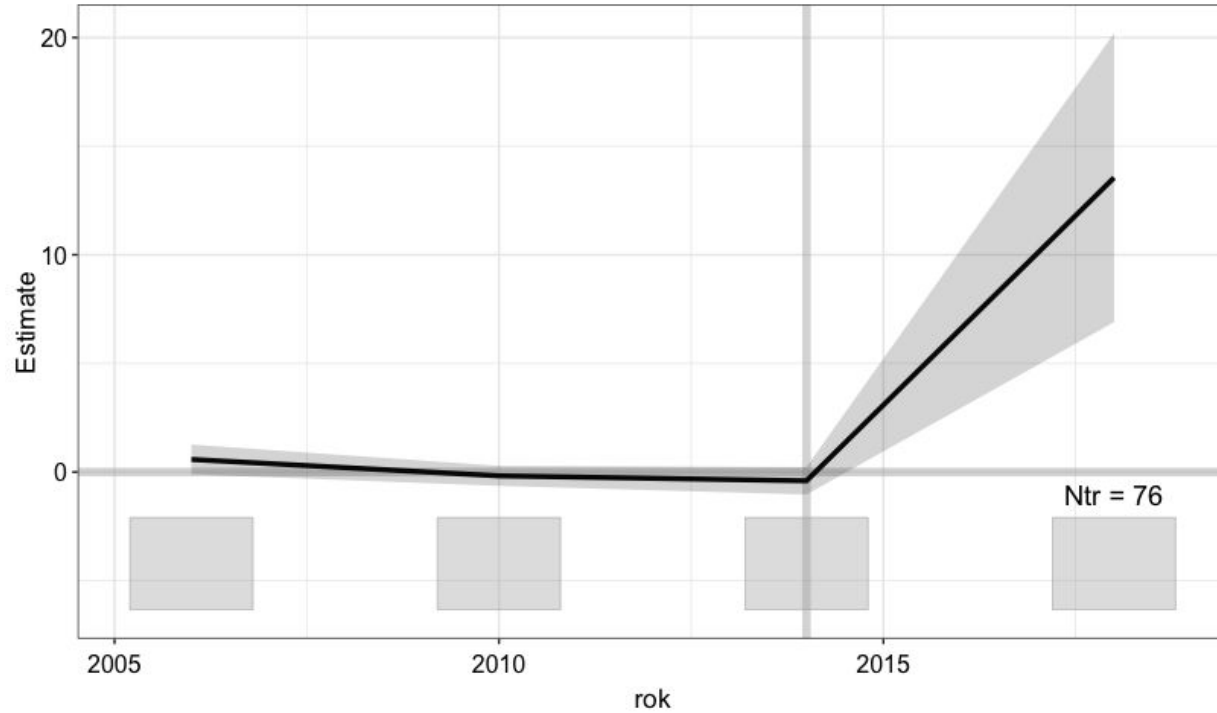
	ATT	S.E.	z-score	CI.lower	CI.upper	p.value	n.Treated
2006	0.5792	0.3490	1.6594	-0.1049	1.2633	0.0970	76
2010	-0.1751	0.2354	-0.7438	-0.6366	0.2863	0.4570	76
2014	-0.4041	0.3222	-1.2541	-1.0356	0.2274	0.2098	76
2018	13.5498	3.3919	3.9948	6.9019	20.1977	0.0001	76

Average Treatment Effect on the Treated:

	ATT	S.E.	z-score	CI.lower	CI.upper	p.value
[1,]	13.55	3.392	3.995	6.902	20.2	1e-04

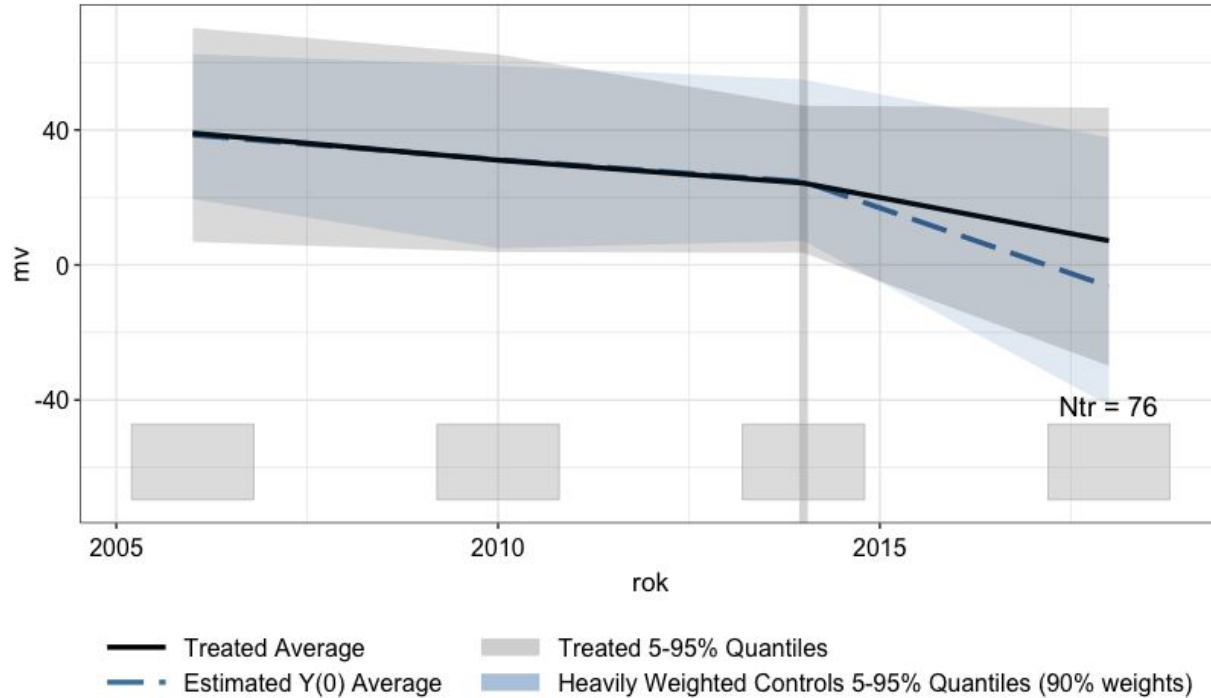
# MODEL I (exp x kont 1)

## Average Treatment Effect on the Treated



# MODEL I (exp x kont 1)

## Treated and Counterfactual Averages



# MODEL II (exp x kont 2)

## Model 2 MV ~ D

Call:

```
tjbal.default(data = data.f1.bal.3, Y = "mv", D = "D", X = c("bez",  
  "P2627_ogolem", "konk", "wmipc", "publwog", "kw"), X.avg.time = list(c(2006,  
  2010, 2014), c(2006, 2010, 2014), c(2006, 2010, 2014), c(2006,  
  2010, 2014), c(2006, 2010, 2014), c(2006, 2010, 2014)), index = c("id_cand.2",  
  "rok"), Y.match.time = c(2006, 2010, 2014), demean = T, estimator = "kernel",  
  vce = "jackknife", parallel = T, seed = 1234)
```

~ by Period (including Pre-treatment Periods):

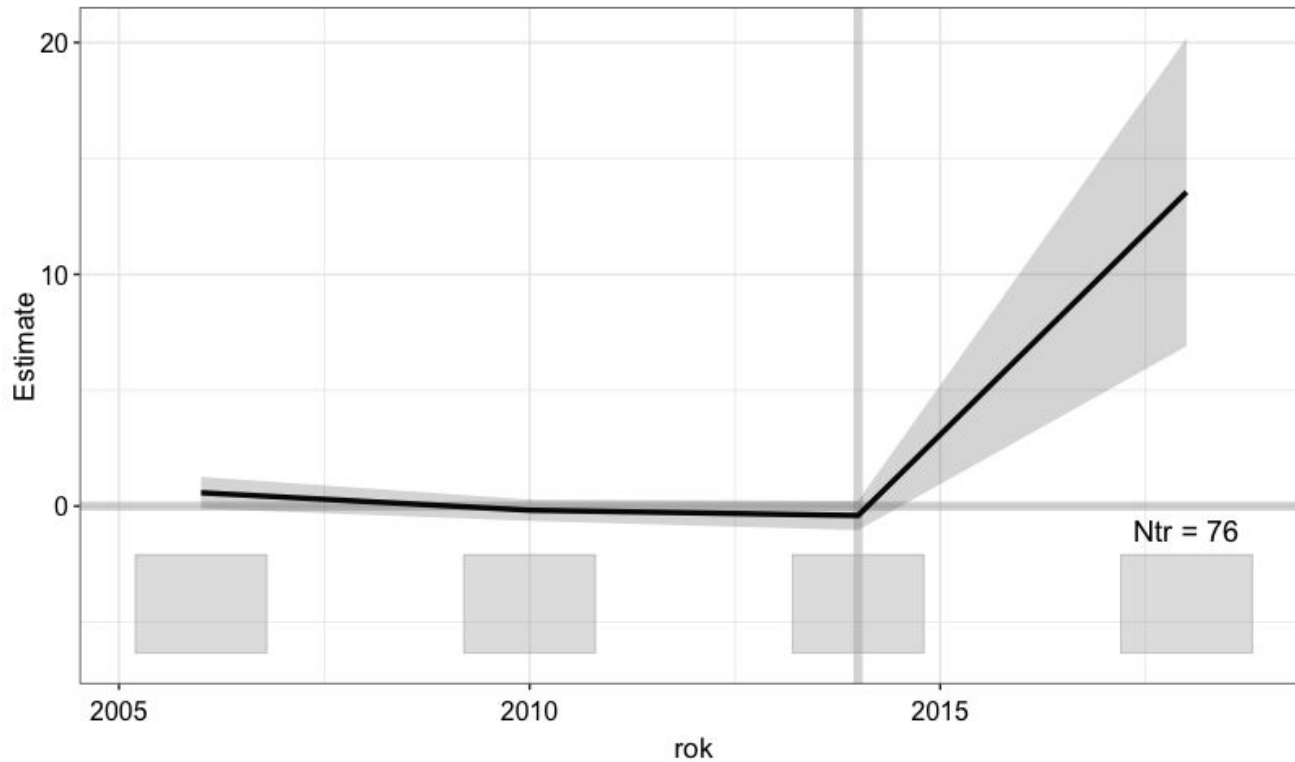
	ATT	S.E.	z-score	CI.lower	CI.upper	p.value	n.Treated
2006	0.3272	0.3038	1.0770	-0.2682	0.9227	0.2815	78
2010	-0.1324	0.2661	-0.4975	-0.6538	0.3891	0.6188	78
2014	-0.1948	0.2752	-0.7081	-0.7342	0.3445	0.4789	78
2018	3.5878	2.9204	1.2285	-2.1360	9.3116	0.2192	78

Average Treatment Effect on the Treated:

	ATT	S.E.	z-score	CI.lower	CI.upper	p.value
[1,]	3.588	2.92	1.228	-2.136	9.312	0.2192

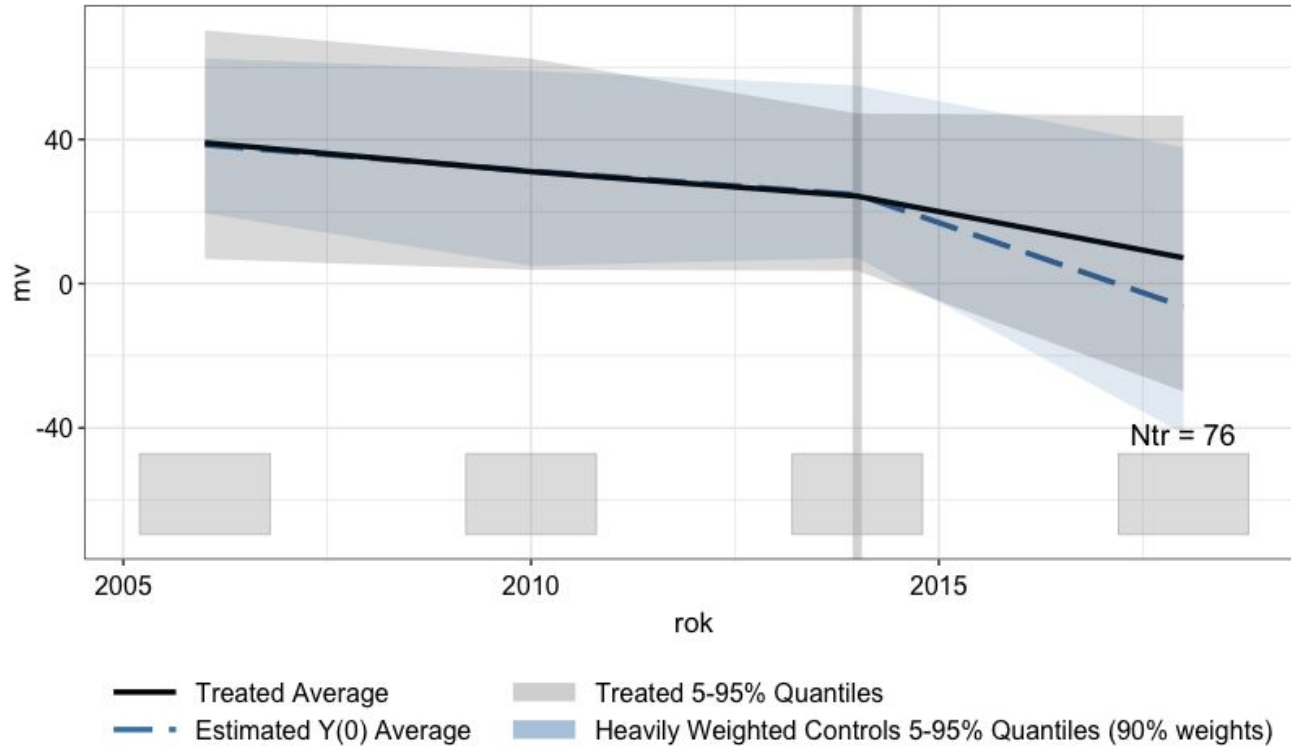
# MODEL II (exp x kont 2)

## Average Treatment Effect on the Treated



# MODEL II (exp x kont 2)

## Treated and Counterfactual Averages



# MODEL III (exp x kont 1)

## Model 3: NIEW ~ D

Call:

```
tjbal.default(data = data_test, Y = "niew2", D = "D", X = c("bez",  
  "P2627_ogolem", "konk", "wmipc", "publwog", "kw"), X.avg.time = list(c(2006,  
  2010, 2014), c(2006, 2010, 2014), c(2006, 2010, 2014), c(2006,  
  2010, 2014), c(2006, 2010, 2014), c(2006, 2010, 2014)), index = c("id_cand",  
  "rok"), Y.match.time = c(2006, 2010, 2014), demean = T, estimator = "kernel",  
  vce = "jackknife")
```

~ by Period (including Pre-treatment Periods):

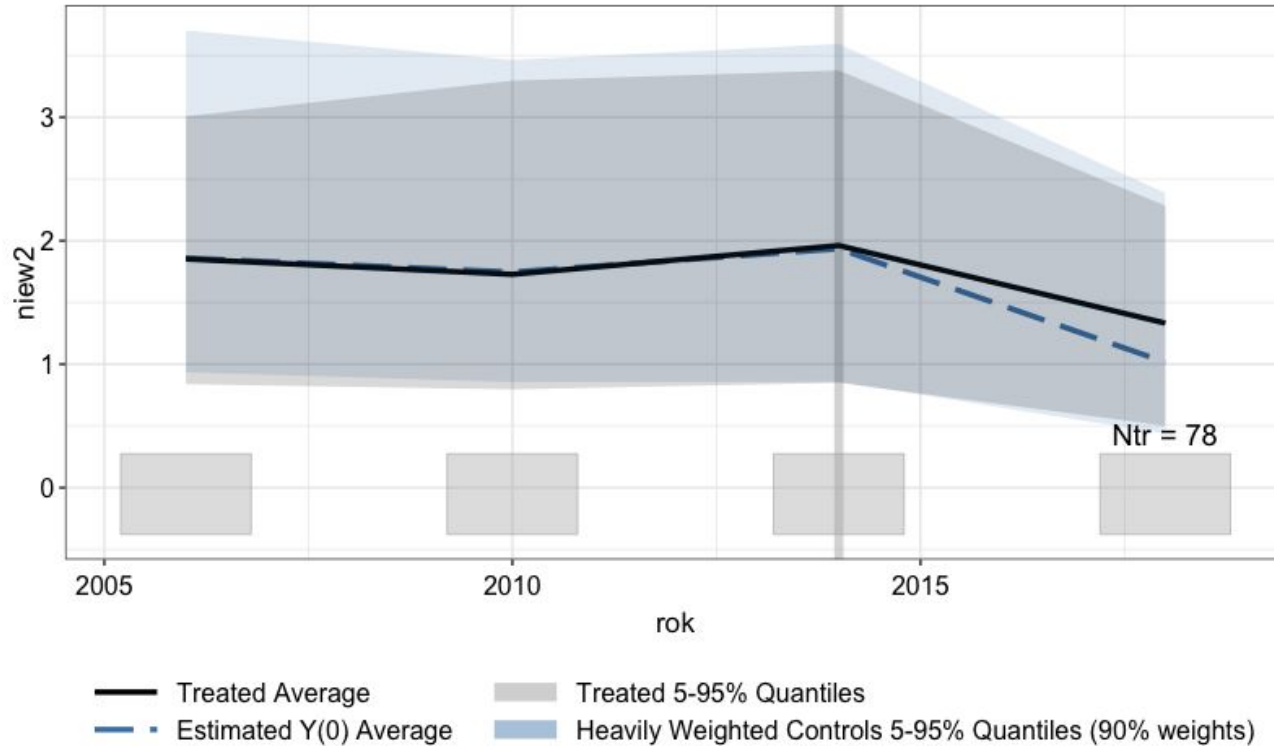
	ATT	S.E.	z-score	CI.lower	CI.upper	p.value	n.Treated
2006	-0.0075	0.0147	-0.5072	-0.0363	0.0214	0.6120	78
2010	-0.0192	0.0155	-1.2381	-0.0496	0.0112	0.2157	78
2014	0.0266	0.0205	1.2992	-0.0136	0.0668	0.1939	78
2018	0.3165	0.0783	4.0426	0.1631	0.4700	0.0001	78

Average Treatment Effect on the Treated:

	ATT	S.E.	z-score	CI.lower	CI.upper	p.value
[1,]	0.3165	0.0783	4.043	0.1631	0.47	1e-04

# MODEL III (exp x kont 1)

## Treated and Counterfactual Averages





# MODEL IV (exp x kont 2)

## Model 4: NIEW ~ D

Call:

```
tjbal.default(data = data_test, Y = "niew2", D = "D", X = c("bez",  
  "P2627_ogolem", "konk", "wmipc", "publwog", "kw"), X.avg.time = list(c(2006,  
  2010, 2014), c(2006, 2010, 2014), c(2006, 2010, 2014), c(2006,  
  2010, 2014), c(2006, 2010, 2014), c(2006, 2010, 2014)), index = c("id_cand",  
  "rok"), Y.match.time = c(2006, 2010, 2014), demean = T, estimator = "kernel",  
  vce = "jackknife")
```

~ by Period (including Pre-treatment Periods):

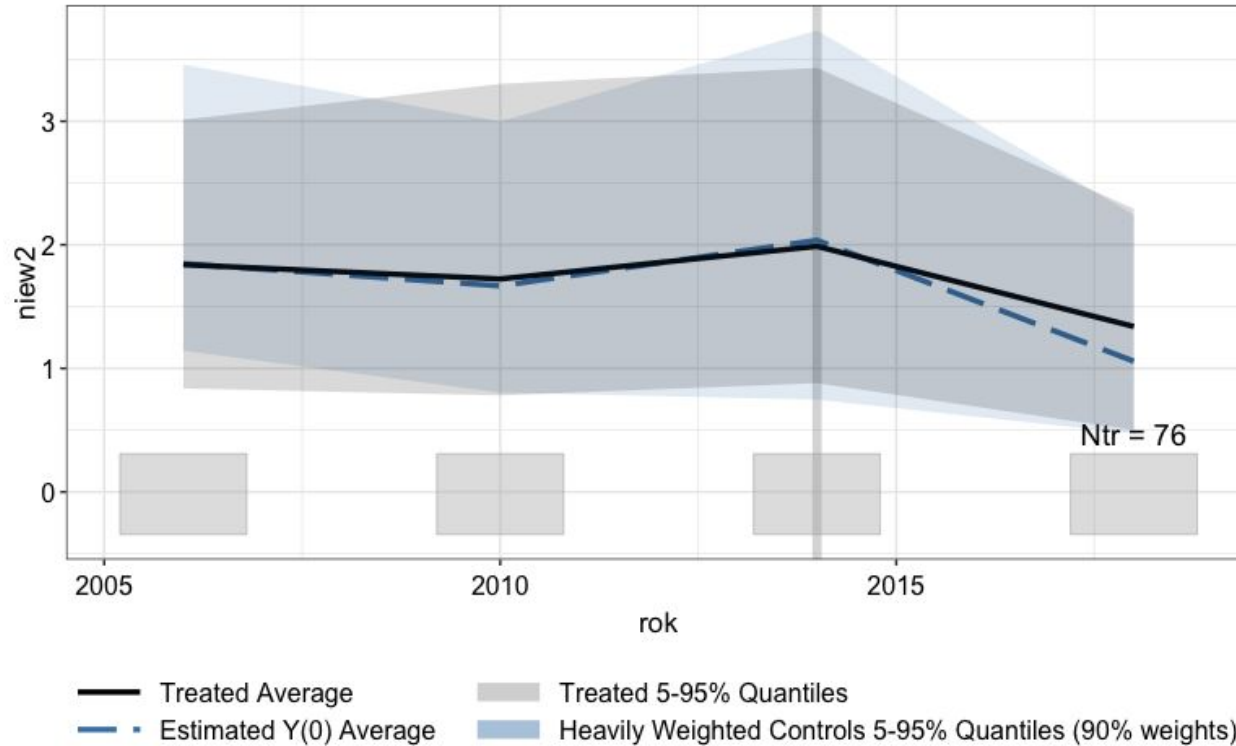
	ATT	S.E.	z-score	CI.lower	CI.upper	p.value	n.Treated
2006	-0.0071	0.0148	-0.4807	-0.0360	0.0218	0.6308	76
2010	0.0560	0.0162	3.4637	0.0243	0.0877	0.0005	76
2014	-0.0489	0.0209	-2.3377	-0.0899	-0.0079	0.0194	76
2018	0.2816	0.0979	2.8755	0.0897	0.4735	0.0040	76

Average Treatment Effect on the Treated:

	ATT	S.E.	z-score	CI.lower	CI.upper	p.value
[1,]	0.2816	0.0979	2.876	0.0897	0.4735	0.004

# MODEL IV (exp x kont 2)

## Treated and Counterfactual Averages



# Wnioski

- potwierdzenie wniosków RDD w zakresie występowania anomalii w 2018 wśród inkumbentów (inc) z doświadczeniem z lat 90. w małych gminach
- mod I: w małych gminach inc z doświadczeniem z lat 90. mają istotnie; wyższą przewagę wyborczą w porównaniu do inc bez doświadczenia z lat 90.;
- mod III: w małych gminach inc z doświadczeniem z lat 90. mają istotnie więcej gł. niew. w porównaniu do inc bez doświadczenia z lat 90.
- w mod II efekt nie jest istotny statystycznie (duże gminy).
- w mod IV efekt jest istotny statystycznie, lecz dopasowanie syntetycznej grupy kontrolnej jest słabe (duże gminy).

# APPENDIX

