The Role of Majority Status in Close Election Studies

Matteo Alpino[1] and Marta Crispino[2]

[1] Structural Economic Analysis Directorate, Bank of Italy, Rome, Italy. E-mail: alpino.mtt@gmail.com
[2] Statistical Analysis Directorate, Bank of Italy, Rome, Italy. E-mail: crispino.marta8@gmail.com

Abstract

Many studies exploit close elections in a regression discontinuity framework to identify partisan effects, that is, the effect of having a given party in office on some outcome. We argue that, when conducted on single-member districts, such design may identify a compound effect: the partisan effect, plus the majority status effect, that is, the effect of being represented by a member of the legislative majority. We provide a simple strategy to disentangle the two, and test it with simulations. Finally, we show the empirical relevance of this issue using real data.

Keywords: partisan effect, single-member districts, regression discontinuity design

1 Introduction

Since Lee (2008), Lee, Moretti, and Butler (2004), and Pettersson-Lidbom (2008), many papers use a regression discontinuity design (RDD) that exploits close elections (CEs) to estimate the effect of a given party being in office on some outcome (e.g., public spending).

We argue that when the data are made of first-past-the-post districts to elect members of a parliament, the treatment effect cannot be interpreted as a pure *partisan effect* (PE), because it is potentially compounded with the effect of being represented by a member of the majority, that is, a *majority status effect*. Consider one term when the democrats have conquered the majority of seats. In this case, all districts are either won by a democrat in the majority or by a republican in the opposition. Instead, if republicans have won the majority of seats, all districts are either won by a democrat in the opposition or by a republican in the majority. In other words, representatives differ not only in their party affiliation, but also in their majority status. Since most applications combine data pooled from several election-years, the estimated effect is a weighted average of these two different joint effects, making its interpretation complicated.

Note that the bundling of these two effects naturally occurs in this electoral system due to institutional features that make party affiliation mechanically correlated with majority status. This issue is therefore distinct from the fact that party identity is sometimes correlated with politician's characteristics such as gender or ethnicity due to complicated patterns of representation, a problem analyzed by Marshall (2022).

Majority status is a characterizing feature of all members of parliament, and has the potential to have an effect on the outcome in many applications that aim at estimating the PE: pork barrel spending, party incumbency advantage, roll call voting, campaign financing, etc. In fact, majority members are likely to have greater agenda setting power and to serve in key positions in legislative committees, or in the cabinet; together they can pass legislation without relying on the support of members of different parties; in some countries, the majority in the parliament elects the executive. Finally, there is evidence that majority status matters for the ability to secure federal transfers and campaign contributions (Albouy 2013; Cox and Magar 1999).

Political Analysis (2024)
vol. 32: 148–155
DOI: 10.1017/pan.2023.14

Published
18 April 2023

Corresponding author
Matteo Alpino

Edited by
Jeff Gill

2 The Compounded Effect

Consider an electoral system, where representatives are elected in n single-member first-past-the-post districts. Each of two parties fields one candidate in every district. Define D_{it} as a dummy equal to one if the democratic party (D) wins the election in district i, in election-year t, and M_{it} as a dummy equal to one if the district i in year t belongs to the majority, that is, to the party whose candidates won in the majority of districts. Thus, D_{it} captures the party affiliation and M_{it} the majority status. Note that, by definition, D_{it} and M_{it} are mechanically related: when party D holds the majority then $D_{it} = M_{it}$; when D is in the opposition, then $D_{it} = 1 - M_{it}$.

We are interested in estimating the PE, that is, the causal effect of party D being in office on some outcome Y_{it}. Assume that in true data-generating process Y_{it} is a function of both D_{it} and M_{it} (e.g., the level of federal funding of a district may depend on the party affiliation of its representative, and on its majority status)[1] and that electoral outcomes in all districts are randomized.[2]

Consider regressing Y_{it} on D_{it} using cross-sectional data from one election-year t when the democrats have the majority. Using this dataset the coefficient on D_{it} corresponds to the compound effect of being represented by a democrat, and by a majority member, because $D_{it} = M_{it}, \forall i$. If instead at t republicans have won the majority, the same coefficient would capture the compound effect of being represented by a democrat, and by an opposition member, because $D_{it} = 1 - M_{it}, \forall i$. Finally, when data include several election-years, the estimated coefficient is a weighted average of these two joint effects. In particular, it identifies the pure PE only if majority status has no effect on the outcome (ruled out by assumption), or if the covariance between D_{it} and M_{it} is zero,[3] which is not true in general. In fact, such covariance crucially depends on the relative number of democratic-controlled (when $D_{it} = M_{it}$, positive covariance) versus republican-controlled (when $D_{it} = 1 - M_{it}$, negative covariance) years. Specifically, it decreases (in absolute value) as the dataset is more balanced in terms of democratic-controlled and republican-controlled years; it becomes negligible in case of perfect balance, because for each observation such that $D_{it} = M_{it}$, there is one such that $D_{it} = 1 - M_{it}$. Starting from perfect balance, the covariance increases (decreases) as the fraction of democratic-controlled years increases (decreases).[4] Note that typically studies that estimate a (local) regression of Y_{it} on D_{it} use datasets with an unbalanced number of republican-controlled and democratic-controlled years, and thus they do not necessarily identify the pure PE.

2.1 Identification of the PE

To identify the PE, formally defined in Section C of the Supplementary Material, the data must include more than one election-year and exhibit variation in the party who controls the assembly.[5] Assume that D_{it} is randomized; our main strategy is to simply control for M_{it} in the regression of Y_{it} on D_{it}. Note that M_{it} depends only on D_{it} and on which party has the majority in the assembly. It is therefore sufficient to assume that the overall majority is determined at the national level (and not at the district level) and to control for time fixed effects to safely include M_{it} in the regression without introducing a selection bias. The assumption is more likely to hold (i) when the number of districts n is large, and thus small is the probability that the outcome in one district determines the overall majority, and (ii) the smaller the fraction of districts that never changes political color, because in that case the control of the assembly would be determined only by the outcome in the few contestable districts. Both (i) and (ii) are testable. Finally, note that Albouy (2013) already

1 See Albouy (2013) for evidence in this respect.
2 Indeed, the issue under discussion is not limited to RDD CE, but to all research designs.
3 This follows from the omitted variable bias formula.
4 See Section B of the Supplementary Material for a proof.
5 Note that it is not possible to identify heterogeneous effects, such as the PE on majority members. In fact, we cannot credibly compare democratic districts in years when democrats have the majority to republican districts when republicans have the majority due to year-level confounders. See Section C of the Supplementary Material.

makes the same assumption with the aim to identify M_{it}, but he does not discuss the importance of controlling for M_{it} in order to identify the PE, which is our focus.

In reality D_{it} is not randomized and thus researchers rely on the RDD CE. In this design, Calonico et al. (2019) recommend including controls, which is crucial in our identification strategy, only to improve precision and after checking that such controls are balanced at the threshold. This recommendation is based on the presumption that covariates imbalance might suggest that the potential outcome function is not continuous at the threshold, so that the crucial identifying assumption is violated. Furthermore, the authors add that covariates can be included to restore identification if the researchers are willing to impose additional assumptions. In our case, we are aware that M_{it} might not be balanced at the threshold, and that the outcome might be a function of it. In fact, as elaborated above, we propose to include M_{it} in the regression under the additional assumption that assembly control is determined at the national level.

Finally, note that if our argument does not convince the reader on the viability of controlling for M_{it}, it is always possible to balance the sample in terms of years with democratic/republican control, so that the correlation between M_{it} and D_{it} is negligible and is not necessary to include majority status. In practice, one may selectively drop years or, more efficiently, use post-stratification (Miratrix, Sekhon, and Yu 2013), that is, re-weight the sample such that observations under the two types of years have equal weight.

3 Simulations

We simulate elections in 601 single-member districts to elect representatives of a parliament in a two-party system for 100 election years.[6] The outcome Y_{it} is a function of majority status, party identity, the vote share X_{it} for the democratic party, and random components at the year and district level.[7]

We estimate two models: (A) the standard one with a constant and D_{it}, and (B) our specification augmented with M_{it} and year fixed effects. Both include a linear function in the margin of victory estimated separately on each side of the threshold. Figure 1 plots the point estimate of the coefficient on D_{it} for the two models together with the 95% confidence intervals (CIs), as a function of the bandwidth. Crucially, the estimates are performed separately in nine different samples of 50 election years, each characterized by a different ratio of democratic to republican years, corresponding to the panels of Figure 1.

Model A (black) provides an unbiased estimate of the PE (i.e., 0.3) only when the sample is composed by the same number of democratic and republican years (central panel). In all other cases, the estimate is either upward biased (with more democratic years) or downward biased (with more republican years). The sign and size of the bias is thus consistent with what predicted in Section 2. On the contrary, model B (red) always estimates a coefficient centered on the true effect.

4 Evidence from Real Data on the U.S. House

We perform similar analyses on real data, aiming at showing that controlling for majority status can affect estimates of the PE in the predicted direction. Throughout the section, we present results from models A and B, as well as a third specification with both D_{it} and M_{it} but without fixed effects. For more details on data and estimation, see Sections F–H of the Supplementary Material.

6 Replication material for this section and the next one is available at Alpino and Crispino (2023) at https://doi.org/10.7910/DVN/GAK3QS.

7 See Section D of the Supplementary Material for details.

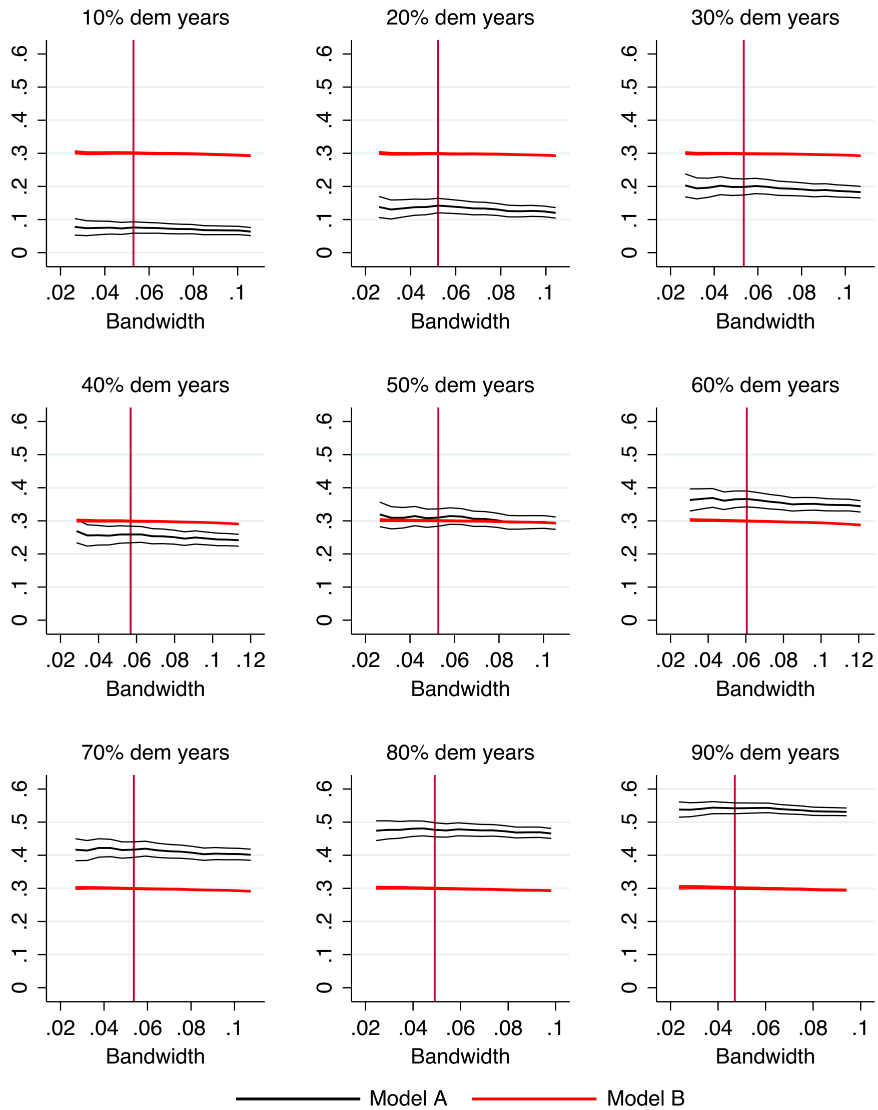

Figure 1. Estimates of PE in simulated data. True PE is 0.3. The vertical red lines indicate the optimal bandwidth by Calonico, Cattaneo, and Titiunik (2014). Linear model estimated with OLS with standard errors adjusted for heteroskedasticity.

4.1 Roll-Call Voting and Incumbency Advantage 1946–1994

We replicate the analysis in Lee, Moretti, and Butler (2004) using the original dataset, which includes results for the U.S. House in the period 1946–1994, and voting scores of representatives on a right–left scale 0–100 based on roll-call votes. In this sample, there is only one republican-controlled year. The authors use a RDD CE to estimate the PE on three outcomes: contemporaneous policy stance RC_{it}, policy stance in the next term RC_{it+1}, and the treatment in the next term D_{it+1} (incumbency advantage). Results, reported in Table 1, show that including majority status considerably affects the estimate of the coefficient on D_{it} for all outcomes.

Despite some differences, the qualitative conclusion in Lee, Moretti, and Butler (2004) is robust to this replication. Nevertheless this exercise shows that the PE changes more than one would expect in a valid RDD CE when we control for majority status.

4.2 Roll-Call Voting 1947–2008

We extend the dataset in the previous section until 2008, obtaining a sample with 23 terms under democratic control and 8 under republican control. The estimation is conducted separately

Table 1. Replication of Lee, Moretti, and Butler (2004)

	Outcome: RC_{it+1}			Outcome: RC_{it}			Outcome: D_{it}		
	1	2	3	4	5	6	7	8	9
D_{it}	20.75	13.15	17.63	48.28	60.99	57.91	0.530	0.337	0.389
	(1.98)	(2.84)	(2.94)	(1.30)	(1.87)	(1.93)	(0.058)	(0.069)	(0.064)
M_{it}		10.31	7.17		−14.18	−11.36		0.262	0.182
		(2.84)	(2.94)		(1.82)	(1.93)		(0.050)	(0.048)
Time-FE	No	No	Yes	No	No	Yes	No	No	Yes
observations	887	887	887	955	955	955	887	887	887

Note: Linear model estimated with OLS without controlling for the margin of victory. Robust standard errors in parenthesis. Bandwidth = 2 percentage points.

on subsamples that feature a different ratio of observations from democratic- and republican-controlled years, resulting in different covariance between D_{it} and M_{it}. For simplicity, we only focus on the PE on contemporaneous roll-call voting RC_{it}. Table 2 reports the results. In the most balanced period 1982–2004, the correlation between D_{it} and M_{it} is close to zero. As expected, the coefficient on D_{it} is the same (approximately 56) irrespective of whether we control for majority status. The coefficient on M_{it} is approximately −5, suggesting that majority members have on average a less liberal stance compared to opposition members, holding party constant. Results from the other subsamples are broadly consistent with what predicted theoretically in Section 2: relative to 1982–2004, the coefficient on D_{it} in the model without M_{it} is lower the more democratic years (positive covariance), and higher the more republican years (negative covariance). Furthermore, in all partially unbalanced subsamples controlling for majority status yields a coefficient on D_{it} closer to 56, relative to the model without M_{it}. Introducing time fixed effects makes little difference. The results confirm our theoretical insights which, however, has a limited quantitative relevance in this application, due to the moderate effect of majority status on roll-call voting.

4.3 Electoral Financing 1979–2006

We estimate the effect of a victory of the democratic party in a district on the campaign funds raised by the incumbent party in the next election.[8] Since most incumbents seek reelection, this is almost equivalent to testing whether democratic members raise more funds than their republican colleagues to finance their reelection campaign. This could happen if members of one party are on average more able to attract funds, or if donors have a partisan bias. The analysis is interesting in light of Cox and Magar (1999), who find that majority status yields an advantage in terms of campaign financing. The outcome is the amount of campaign funds (in thousands of 1990 dollars) raised in a district from non-investor donors by the party that won the previous election.

As before, in the balanced subsample (1978–2004) the coefficient on D_{it} is the same (approximately −133) irrespective of whether we control for majority status (see Table 3). Moreover, here the coefficient on M_{it} is sizable (80), and thus its omission makes for very large difference in the estimate of the coefficient on D_{it} in unbalanced subsamples: −51 in 1978–1992 versus −205 in 1994–2004. As before, controlling for majority status makes the estimate of the coefficient on D_{it} more similar across subsamples.

8 Data are from Fouirnaies and Hall (2014) but our analysis is different and it is not a replication.

https://doi.org/10.1017/pan.2023.14 Published online by Cambridge University Press

Table 2. Roll-call voting.

	1946–2006			Dem. control: 1978–1992		Rep. control: 1994–2004		1978–1994			1990–2004		
	1	2	3	4	5	6	7	8	9	10	11	12	13
D	49.48	53.72	53.59	48.32	48.43	61.80	61.79	50.35	53.49	53.43	58.93	57.54	57.47
	(1.54)	(1.51)	(1.51)	(3.19)	(3.25)	(2.90)	(2.97)	(2.83)	(2.66)	(2.71)	(2.48)	(2.51)	(2.53)
M		−6.73	−6.29						−4.83	−4.74		−3.80	−3.62
		(0.79)	(0.79)						(1.42)	(1.38)		(1.31)	(1.25)
Electoral cycle FE	No	No	Yes	No	Yes	No	Yes	No	No	Yes	No	No	Yes
Mean Y if D=1	67	67	67	68	68	75	75	69	69	69	73	73	73
Mean Y if D=0	15	15	15	16	16	11	11	15	15	15	12	12	12
No. of obs. in dem years (%)	78	78	78	100	100	0	0	86	86	86	32	32	32
Corr(D,M)	0.56	0.56	0.56	1.00	1.00	−1.00	−1.00	0.72	0.72	0.72	−0.36	−0.36	−0.36
No. of obs.	3,699	3,682	3,682	843	843	531	531	980	969	969	781	777	777
	1982–2004			**Dem. control: 1954–1976**		**Rep. control: 1946+1952**		**1946–1976**			**1946–1958**		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
D	55.90	56.36	56.31	44.17	44.55	57.08	56.95	47.70	52.48	51.63	49.22	52.02	51.28
	(2.53)	(2.46)	(2.46)	(2.44)	(2.40)	(3.44)	(3.31)	(1.89)	(2.02)	(1.96)	(2.31)	(2.37)	(2.31)
M		−5.01	−4.82						−5.88	−4.53		−4.82	−3.46
		(1.20)	(1.18)						(1.05)	(1.07)		(1.00)	(1.05)
Electoral cycle FE	No	No	Yes	No	Yes	No	Yes	No	No	Yes	No	No	Yes
Mean Y if D=1	71	71	71	63	63	63	63	64	64	64	64	64	64
Mean Y if D=0	13	13	13	19	19	7	7	16	16	16	14	14	14
No. of obs. in dem years (%)	54	53	53	100	100	0	0	88	88	88	74	74	74
Corr(D,M)	0.07	0.07	0.07	1.00	1.00	−1.00	−1.00	0.75	0.75	0.75	0.47	0.47	0.47
No. of obs.	1,145	1,135	1,135	1,677	1,677	279	279	2,269	2,264	2,264	1,067	1,063	1,063

Note: Linear model estimated with OLS controlling linearly for the margin of victory on each side of the threshold. Standard errors clustered at the electoral district. Bandwidth = 0.183 selected using the method by Calonico, Cattaneo, and Titiunik (2014).

Table 3. Campaign financing.

	1978–2004			Dem. control: 1978–1992		1978–1994			Rep. control: 1994–2004		1990–2004		
	1	2	3	4	5	6	7	8	9	10	11	12	13
D	−132.83	−145.06	−133.02	−51.13	−33.76	−95.17	−127.84	−114.25	−205.34	−219.08	−165.61	−95.01	−108.75
	(45.92)	(47.76)	(41.57)	(37.92)	(30.51)	(38.75)	(43.57)	(31.63)	(84.42)	(79.65)	(75.32)	(73.49)	(66.90)
M		82.80	78.61				57.34	64.63				109.01	120.66
		(27.20)	(22.49)				(27.74)	(21.19)				(39.45)	(35.41)
Electoral cycle FE	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes
Mean Y if *D*=1	327	327	327	220	220	256	256	256	461	461	442	442	442
Mean Y if *D*=0	467	467	467	258	258	324	324	324	669	669	622	622	622
No. of obs. in dem years (%)	52	52	52	100	100	80	80	80	0	0	16	16	16
Corr(*D*,*M*)	0.05	0.05	0.05	1.00	1.00	0.60	0.60	0.60	−1.00	−1.00	-0.68	−0.68	−0.68
No. of obs.	1,056	1,056	1,056	554	554	690	690	690	502	502	599	599	599

Note: Linear model estimated with OLS controlling linearly for the margin of victory on each side of the threshold. Standard errors clustered at the electoral district in parenthesis. Bandwidth = 0.09 selected using the method by Calonico, Cattaneo, and Titiunik (2014).

5 Conclusion

We show how and when majority status can affect the interpretation of the PE in RDD CE studies. We propose an identification strategy based on controlling for majority status and validate it with simulated and real data, including those used in Lee, Moretti, and Butler (2004). In the latter case, our specification does not alter the qualitative conclusion of the study, but in other applications, the empirical relevance of our point is significant.

Despite our focus on first-past-the-post systems, where party and majority status are realized simultaneously, our argument is more broadly relevant to contexts where the alignment between different layers (local versus national) or branches (president versus parliament) of government is expected to matter. Furthermore, our paper is relevant not only for RDD CE studies, but also for other research designs aimed at estimating the PE, since our argument is not about failure of specific identification assumptions.

Acknowledgment

We thank the Editor Jeff Gill, three anonymous referees, Jon Fiva, Andreas Kotsadam, Eliana La Ferrara, Edwin Leuven, Halvor Mehlum, Johanna Rickne, and Rocìo Titiunik for insightful comments. The views in this paper do not necessarily represent those of the Bank of Italy.

Data Availability Statement

Replication code for this article is available at Alpino and Crispino (2023) at https://doi.org/10.7910/DVN/GAK3QS.

Funding

We acknowledge support from ESOP, University of Oslo, funded by the Research Council of Norway (227072/F10).

Supplementary Material

For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2023.14.

References

Albouy, D. 2013. "Partisan Representation in Congress and the Geographic Distribution of Federal Funds." *Review of Economics and Statistics* 95 (1): 127–141. https://doi.org/10.1162/REST_a_00343.

Alpino, M., and M. Crispino. 2023. "Replication Data for: '*The Role of Majority Status in Close Election Studies*'." Harvard Dataverse, V1, UNF:6:SORyg9FB6zzSav1MbRXJoQ== [fileUNF]. https://doi.org/10.7910/DVN/GAK3QS.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik. 2019. "Regression Discontinuity Designs Using Covariates." *Review of Economics and Statistics* 101 (3): 442–451. https://ideas.repec.org/a/tpr/restat/v101y2019i3p442-451.html.

Calonico, S., M. D. Cattaneo, and R. Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82 (6): 2295–2326. https://doi.org/10.3982/ECTA11757.

Cox, G. W., and E. Magar. 1999. "How Much Is Majority Status in the U.S. Congress Worth?" *American Political Science Review* 93 (2): 299–309. https://doi.org/10.2307/2585397.

Fouirnaies, A., and A. B. Hall. 2014. "The Financial Incumbency Advantage: Causes and Consequences." *Journal of Politics* 76 (3): 1–14. https://doi.org/10.1017/S0022381614000139.

Lee, D. S. 2008. "Randomized Experiments from Non-random Selection in U.S. House Elections." *Journal of Econometrics* 142 (2): 675–697. https://doi.org/10.1016/j.jeconom.2007.05.004.

Lee, D. S., E. Moretti, and M. J. Butler. 2004. "Do Voters Affect or Elect Policies? Evidence from the U.S. House." *Quarterly Journal of Economics* 119 (3): 807–859. https://doi.org/10.1162/0033553041502153.

Marshall, J. 2022. " *Can Close Election Regression Discontinuity Designs Identify Effects of Winning Political Characteristics?*" *American Journal of Political Science*. https://doi.org/10.1111/ajps.12741.

Miratrix, L., J. Sekhon, and B. Yu. 2013. "Adjusting Treatment Effect Estimates by Post-Stratification in Randomized Experiments." *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 75: 369–396. https://doi.org/10.2307/23360930.

Pettersson-Lidbom, P. 2008. "Do Parties Matter for Economic Outcomes? A Regression-Discontinuity Approach." *Journal of the European Economic Association* 6 (September): 1037–1056. https://doi.org/10.1162/JEEA.2008.6.5.1037.