ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Kelly, Edmund; Odermatt, Angela; Metson, Lennard

Working Paper

Computational Reproduction and Robustness Replication for Hollyer, Klašnja and Titiunik (2022): A Replication Report from the Nottingham Replication Games

I4R Discussion Paper Series, No. 54

Provided in Cooperation with:
The Institute for Replication (I4R)

Suggested Citation: Kelly, Edmund; Odermatt, Angela; Metson, Lennard (2023) : Computational Reproduction and Robustness Replication for Hollyer, Klašnja and Titiunik (2022): A Replication Report from the Nottingham Replication Games, I4R Discussion Paper Series, No. 54, Institute for Replication (I4R), s.l.

This Version is available at:
https://hdl.handle.net/10419/274553

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

INSTITUTE for REPLICATION

No. 54
I4R DISCUSSION PAPER SERIES

Computational Reproduction and Robustness Replication for Hollyer, Klašnja and Titiunik (2022): A Replication Report from the Nottingham Replication Games

Edmund Kelly

Angela Odermatt

Lennard Metson

August 2023

INSTITUTE for REPLICATION

I4R DISCUSSION PAPER SERIES

I4R DP No. 54

Computational Reproduction and Robustness Replication for Hollyer, Klašnja and Titiunik (2022): A Replication Report from the Nottingham Replication Games

Edmund Kelly[1], Angela Odermatt[1], Lennard Metson[1]

[1]University of Oxford/United Kingdom

AUGUST 2023

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the ZBW – Leibniz Information Centre for Economics, and RWI – Leibniz Institute for Economic Research, I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

Computational reproduction and robustness replication for Hollyer, Klašnja and Titiunik (2022): A replication report from the Nottingham Replication Games

Edmund Kelly* Angela Odermatt† Lennard Metson‡

August 17, 2023

Abstract

Hollyer, Klašnja, and Titiunik (2022) analyse the trade-off that political parties face between running programmatic campaigns and fielding charismatic candidates, whose electoral appeal may come at the cost of undermining the party brand. They argue that higher electoral volatility prompts parties to rely on charismatic candidates, even though they might not be as loyal to the party's programmatic stance. They substantiate their argument with a cross-national dataset and a quantitative case study in Brazil. We computationally reproduced and conducted further robustness tests for their cross-national study by translating the Stata code to R. Next, we conducted a computational reproduction and some additional robustness tests for the quantitative case study. We find that their cross-national analysis is reproducible, albeit with some minor discrepancies. The quantitative case study is also largely reproducible and both are robust in several ways. We conclude by making some suggestions about data dissemination and robustness checks for authors of regression discontinuity designs.

*University of Oxford. E-mail: edmund.kelly@politics.ox.ac.uk.
†University of Oxford. E-mail: angela.odermatt@politics.ox.ac.uk.
‡University of Oxford. E-mail: lennard.metson@politics.ox.ac.uk.

1 Introduction

Hollyer, Klašnja, and Titiunik (2022), henceforth HKT, examine how political parties balance the competing strategies of fielding charismatic candidates in elections and maintaining a unified, programmatic party brand. While charismatic candidates are appealing to the electorate, they are also more likely to threaten the long-term party platform. HKT present a formal model and then test its predictions with two empirical studies. The first, a cross-national analysis, tests the correlates of two dependent variables: the programmaticness and personalism of a party's campaigning efforts. The second analysis, a regression discontinuity design (henceforth RDD), tests whether winning one election means a party is less likely to nominate a charismatic candidate in the next election.

We found that both analyses were largely reproducible and robust to several additional checks. This being said, we were unable to reproduce some of the coefficients in the cross-national analysis. Most importantly, we were unable to reproduce the coefficients for extreme parties, used to operationalize internal party cohesion. We also found minor discrepancies in their underlying data. Further, while the RDD was robust in several ways, we were not able to conduct some robustness checks because the replication packages did not contain the raw data used to construct all variables included in the analysis. These caveats aside, we do not dispute the headline results or substantive interpretation of the analysis presented by HKT. The results are largely reproducible and are robust in several ways in addition to those discussed by HKT.

In what follows, we first demonstrate the computational reproducibility of the cross-national analysis after converting the original code into R. We then demonstrate the computational reproducibility of the RDD after re-writing the original R code. In the second section, we detail the results of our robustness replication of the cross-national analysis and then of the RDD. We conclude by making some suggestions about data dissemination and robustness checks for authors of RDDs.

2 Computational reproduction

2.1 Cross-national analysis

To test the corollaries of their formal model, HKT perform a cross-national analysis over 75 countries. Their main hypothesis is that higher electoral volatility is correlated with lower programmaticness and higher personalism. This effect is mitigated by intra-party group cohesion, which HKT operationalise through ethnic parties or parties with extreme ideology. Parties with high group-cohesion are expected to have higher personalism scores in contexts with low electoral volatility and higher programmaticness scores in contexts with high electoral volatility (compared to parties with lower group cohesion).

HKT present their results by first plotting the raw correlations of their personalism and programmaticness scores with the electoral volatility index. They then repeat the same analyses but subset the data by ethnic and non-ethnic parties, and separately by extreme and moderate parties. Conducting the same analyses in R we reproduced the raw correlation plots, which can be found in Figures 1, 2, and 3. While the linear fit is similar to that in the paper, the underlying scatterplot does not precisely match the figures shown by HKT in Figures 2 and especially in Figure 3. Despite this discrepancy, the overall findings hold. Further, the available party-level data does not include either information on the which party or country the observations belong to. This makes it difficult to assess what might be underlying reasons for the found discrepancies.

HKT then conduct parametric analyses and calculate predicted values. Here, we found a minor coding error in the summary statistics in Table C3 in the appendix of the paper. The summary statistics for the variables for closed and open list systems are reported the wrong way around. Because the type of PR system used is included as one categorical variable in the analyses, this does not affect them. The reproduced summary statistics can be found in Table 1.

2.2 Quantitative case study

To test whether the effect of electoral volatility on candidate type selection is causal, HKT use an RDD to test whether winning an election leads to parties choosing a less charis-

matic candidate in the next election. HKT re-analyse data on Brazilian local elections from Klašnja and Titiunik (2017). They compare candidates who barely won election to those who barely lost, arguing that electoral victory proxies for the electoral volatility faced by their party. Their argument is that an electoral win in the previous election (representing lower electoral volatility) will discourage the party from selecting a charismatic candidate in the subsequent election.

HKT operationalise charisma using the age of the candidates. Candidates aged over 35 are coded as non-charismatic. HKT argue that younger candidates are likely to be more charismatic; that they were able to become candidates at a younger age implies this (they do not have the experience of their older counterparts). They also operationalise charisma with a binary variable recording whether the candidate has ever held a government job or public office (insider status). Again, the argument is that those who become candidates without these advantages are likely to be more charismatic. They demonstrate that parties whose candidate just won election are 8.4% less likely to select a young candidate, and 9.2% less likely to select a young and outsider candidate.

We did not have any issues with reproducibility for the quantitative case study. After re-writing the code we were able to reproduce the original results in full. Figure 6 is a computational reproduction of Figure 6 from the original paper. This plots the effect of winning election in the previous period on the probability of selecting a charismatic candidate in the subsequent period. This draws on the results in Table 5, which is a computational reproduction of Table 2 in the original paper.

3 Robustness replication

Not having found any substantive coding errors, we now directly move on to our replication. We again divided this section into the cross-national and the Brazilian RD analysis.

3.1 Cross-national analysis

For the cross-national analysis, we conducted a computational reproduction using the data provided and operationalisations used by HKT. However, we translated the code originally written in Stata to R. Instead of translating the code from the Stata script,

we used the empirical strategy as described in the paper as our basis for the analyses conducted with R. Not using the Stata-code to conduct the analyses in R is a further robustness check as we thereby also tested whether the description and actual analyses performed by HKT match.

Turning first to the non-parametric analyses of the correlations between electoral volatility and personalism respectively programmaticness, we were able to replicate the magnitude and direction of the raw correlations. Replications of Figures 3, 4, and 5 from the original paper which illustrate these results can be found in the appendix in Figures 1, 2, and 3. As in the original study, we find that personalism and volatility are positively correlated while the opposite is the case for programmaticness. Additionally, we also find that this relationship is mitigated by ethnicity as well as ideological extremism, both of which HKT use as an indicator of intraparty group attachments. While our replicated figures report the same raw correlations, the individual data points in the scatterplots differ from the ones reported in the original study. Additionally, we also included confidence intervals in the plots. This shows that especially for the extreme and ethnic parties the slopes are associated with higher uncertainty. On one hand, this plays to the argument made by HKT that volatility matters less for these parties. On the other, it could also just be a function of there being many less observations for these two party groups.

Results of the replicated parametric analyses found in the Appendix in Table D1 are displayed in Table 3. We were able to reproduce all the coefficients except for the Gini variable. For this variable we get effect sizes of 0 across all models. For all other variables we find the same effect sizes and directions and only minor deviations in the reported standard errors which might be due to different rounding rules used. HKT calculate predicted values based on these models and compare them across different configurations. That way, they show, for instance, that a nonethnic party moving form a lowly to a highly volatile context is predicted to have a higher personalism score.

Using the `marginaleffects` package we are able to replicate all predicted values except the ones for extremist parties. The largest difference is for the predicted value on the programmaticness score for an extremist party in a highly volatile context. The

reported value by HKT is 3.20 while we find 3.13. The difference of 0.07 is larger than the standard error of 0.05 respectively the replicated one of 0.04. While the differences thus are not large, there appears to be something amiss with the reported values as we were able to replicate all other values. Standard errors associated with the predicted values differ only slightly between the reported and replicated models. Except for the extremist parties we were thus also able to replicate the overall size, direction and significance of the differences in predicted values reported by HKT.

Finally, we conducted a further robustness check to try and figure out why the values for extremist parties might be different. HKT use an arbitrary cut-off point of an absolute difference of 3 from the mean ideology level (on a 0-6 scale) in a country to classify an extremist party. In Figures 4 and 5 we show the predicted values of the personalism respectively programmaticness scores at the 1^{st}, 2^{nd}, and 3^{rd} quantiles of electoral volatility across the whole range of the extremism variable. We see that, overall, the hypothesised direction of the effects hold. However, by including confidence intervals, we also see that the changes are associated with large uncertainty, especially for the personalism score. While we can thus confirm the overall direction of the findings by HKT using the threshold, we need to be aware that the findings might not hold to the same extent for personalism and programmaticness, i.e. that they are not inversely related.

3.2 Quantitative case study

For the RDD, HKT did not conduct any direct robustness checks. They argue that the required robustness checks were conducted in Klašnja and Titiunik (2017). We disagree that this is sufficient because the dependent variable in the original study differs (the original authors consider incumbency, whereas HKT consider the characteristics of elected candidates). Only the density tests and placebo outcome tests conducted on the running variable in the original study are independent of the dependent variable and as such can be carried through. We therefore conducted a robustness replication in four stages.

First, we demonstrate the robustness of HKT's results to different bandwidths. Figure 9 plots the point estimates and 95 per cent confidence intervals from the RDD with bandwidths between 5 and 100 per cent for young candidates, and Figure 10 does the same

for young outsider candidates. The point estimates are consistent across bandwidths, ranging between -5 and -10 per cent.

Second, we demonstrate robustness to different functional forms. We re-ran the model with a series of kernel estimators and with a quadratic regression function. In no case did this significantly affect the results. See Table 6 for full results and Figure 7 for an illustration of the first order polynomial result (which is given numerically but not presented graphically in the original paper). Third, we demonstrate robustness to different binning procedures. See Figure 8 for the full results.

Fourth, we demonstrate robustness to placebo cut-offs. We re-ran the original model with placebo cut-offs at one per cent increments between -10% and +10%. We plot the results in Figure 11 for young candidates, and in Figure 12 for young and outsider candidates. At none of the placebo cut-offs did we obtain a statistically significant result.

Although the required data to reproduce the original finding is contained in the replication package, there are no additional contextual variables. The replication file contains the running variable (election performance), the two dependent variables recording candidate charisma, and a variable to exclude cases in which a lame duck mayor ran for re-election. We reached out to HKT to ask if they still had a copy of the continuous age variable from which the dichotomous outcome variable was constructed. Although they did not, they noted that the continuous variable could be constructed from the Brazilian municipal elections database.[1] Given the time constraints we faced during the Replication Games, this limited the range of robustness tests we could conduct. In particular, we suggest that interested researchers could use the raw data to re-run the analysis with a continuous age variable instead of the binary under/over 35 variable.

4 Conclusion

Overall, aside from some minor discrepancies, we were able to reproduce the analysis reported by HKT and to demonstrate its robustness in several ways. Most importantly, we do not challenge the substantive interpretation of the original results. We do nonetheless make two suggestions for future replications and one suggestion for authors of RDDs.

[1] https://www.tse.jus.br/eleicoes/estatisticas/estatisticas

First, we found that attempting to reproduce the analysis from the description in the body text and appendices to be fruitful. We feel that this approach not only tests whether the methods described in the paper are sufficiently detailed, but that it also ensures that reproducibility analysis does not carry over any original coding errors. While we took this approach in order to write new code in R (as the original was in Stata) we believe that this is a useful strategy even if the replication is conducted in the same coding language. Second, we stress the importance of including raw data and the code to reconstruct it in replication packages where this is practically and ethically appropriate. In our case, this would have allowed us to more feasibly conduct further robustness checks on the operationalisation of the running variable in the RDD within a reasonable time frame. While authors may, understandably, not wish to overwhelm their replication packages with raw data and early-stage processing code, we suggest that including as much of this as possible maximises the usefulness of their replication package. We suggest that such data and code be kept separate from the 'headline' replication material to maintain user friendliness while improving usefulness for future replicators. Finally, we suggest that authors of RDDs conduct a separate set of robustness tests if they change the dependent variable in their model. We suggest that robustness to different functional forms, bandwidths and placebo cut-offs should not be assumed when changing dependent variables, even if the underlying data is the same. These thoughts aside, in conclusion, it appears that HKT's results are largely reproducible and robust in several ways.

References

Hollyer, James R, Marko Klašnja, and Rocío Titiunik (2022). "Parties as Disciplinarians: Charisma and Commitment Problems in Programmatic Campaigning". In: *American Journal of Political Science* 66.1, pp. 75–92.

Klašnja, Marko and Rocío Titiunik (2017). "The incumbency curse: Weak parties, term limits, and unfulfilled accountability". In: *American Political Science Review* 111.1, pp. 129–148.

Appendix A: Tables.

Table 1: Reproduced Summary Statistics

Variables	mean	sd	max	min	no.values
partysize	0.15	0.15	0.79	0.00	506
unions	0.35	0.35	1.00	0.00	504
business	0.47	0.33	1.00	0.00	504
religious	0.26	0.32	1.00	0.00	504
ev_total	0.29	0.15	0.69	0.02	464
ctot	-0.17	0.44	3.80	-1.62	484
mdmh	1.94	1.46	6.11	-0.33	500
demsys	3.09	0.93	4.36	0.00	473
ethfrac	0.38	0.23	0.86	0.00	499
polity	7.91	2.89	10.00	-6.00	506
gdppc	9.58	0.96	11.09	6.71	496
gini	38.33	9.31	64.73	23.72	478
personal	2.84	0.71	4.00	1.00	506
programmatic	3.03	0.44	3.96	1.42	506
programmatic2	3.15	0.51	4.00	1.56	506
extremist2	1.73	1.20	6.28	0.01	505
extremist2_2d	3.13	1.92	9.83	0.11	505
extremist2_3d	4.50	2.79	14.76	0.26	505
pluralty	0.23	0.42	1.00	0.00	503
pr	0.57	0.49	1.00	0.00	503
erule	4.69	1.84	7.00	1.00	504
ethnic	0.18	0.39	1.00	0.00	504
presidential	0.42	0.49	1.00	0.00	502
parliamentary	0.52	0.50	1.00	0.00	502
assembly	0.06	0.24	1.00	0.00	502
closed.list	0.57	0.50	1.00	0.00	506
open.list	0.25	0.43	1.00	0.00	506
non.list	0.18	0.38	1.00	0.00	506

Table 2: Replication of Table D1 in the Appendix

	Baseline Model		Extended Model			
	Personalism	Programmaticness	Personalism	Programmaticness	Personalism	Programmaticness
(Intercept)	2.83***	2.76***	2.84***	2.72***	2.71***	2.73***
	(0.55)	(0.55)	(0.57)	(0.54)	(0.55)	(0.51)
ev_total	1.08***	−0.41*	1.94***	−1.09***	1.29***	−0.54**
	(0.26)	(0.22)	(0.48)	(0.33)	(0.25)	(0.24)
partysize	1.42***	0.99***	1.36***	1.14***	1.50***	0.98***
	(0.22)	(0.19)	(0.24)	(0.18)	(0.23)	(0.19)
unions	−0.13	0.36***	−0.08	0.29***	−0.12	0.38***
	(0.10)	(0.07)	(0.10)	(0.06)	(0.10)	(0.07)
business	0.19	−0.13**	0.20*	−0.11**	0.19	−0.11**
	(0.12)	(0.06)	(0.12)	(0.05)	(0.12)	(0.05)
religious	0.11	0.29***	0.12	0.30***	0.11	0.26***
	(0.11)	(0.07)	(0.11)	(0.07)	(0.11)	(0.07)
pluralty	0.36**	0.23*	0.33**	0.28**	0.30*	0.19*
	(0.14)	(0.11)	(0.13)	(0.12)	(0.15)	(0.10)
pr	0.05	0.18**	0.07	0.16**	0.05	0.18**
	(0.08)	(0.08)	(0.08)	(0.07)	(0.08)	(0.07)
factor(cl)closed list	0.51***	0.21*	0.47***	0.26**	0.46***	0.18*
	(0.15)	(0.10)	(0.15)	(0.11)	(0.15)	(0.09)
factor(cl)open list	0.26	0.18*	0.23	0.22*	0.23	0.14
	(0.14)	(0.10)	(0.14)	(0.11)	(0.14)	(0.08)
mdmh	0.00	0.02	−0.00	0.02	0.00	0.02
	(0.03)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)
factor(system)parliamentary	0.20	−0.16	0.16	−0.13	0.23	−0.21
	(0.16)	(0.10)	(0.17)	(0.10)	(0.16)	(0.11)
factor(system)presidential	0.04	−0.33**	0.02	−0.31*	0.09	−0.33**
	(0.17)	(0.11)	(0.18)	(0.11)	(0.17)	(0.11)
demsys	−0.01	0.02	−0.00	0.00	−0.01	0.02
	(0.05)	(0.03)	(0.05)	(0.04)	(0.06)	(0.03)
polity	0.02	0.00	0.01	0.01	0.02	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
gdppc	−0.15***	−0.02	−0.16***	−0.02	−0.15***	−0.01
	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)
gini	0.00	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
ethfrac	0.13	0.10	0.08	0.14	0.08	0.07
	(0.22)	(0.11)	(0.22)	(0.10)	(0.23)	(0.11)
extremist2			0.11	−0.04		
			(0.07)	(0.03)		
ev_total:extremist2			−0.50**	0.39***		
			(0.23)	(0.12)		
ethnic					0.47**	−0.06
					(0.19)	(0.11)
ev_total:ethnic					−1.15**	0.64*
					(0.52)	(0.33)
R^2	0.24	0.40	0.26	0.45	0.25	0.42
Adj. R^2	0.21	0.38	0.22	0.43	0.22	0.39
Num. obs.	431	431	430	430	431	431
RMSE	0.63	0.35	0.63	0.34	0.63	0.35
N Clusters	75	75	75	75	75	75

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

	Baseline Model		Extended Model			
	Personalism	Programmaticness	Personalism	Programmaticness	Personalism	Programmaticness
(Intercept)	2.8295	2.7624	2.8417	2.7168	2.7094	2.7350
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0000)
ev_total	1.0830	−0.4078	1.9359	−1.0907	1.2914	−0.5411
	(0.0002)	(0.0770)	(0.0003)	(0.0023)	(0.0000)	(0.0344)
partysize	1.4211	0.9888	1.3641	1.1383	1.5023	0.9821
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
unions	−0.1284	0.3637	−0.0829	0.2927	−0.1207	0.3753
	(0.1944)	(0.0000)	(0.4079)	(0.0000)	(0.2204)	(0.0000)
business	0.1921	−0.1317	0.1990	−0.1103	0.1901	−0.1107
	(0.1091)	(0.0226)	(0.0963)	(0.0386)	(0.1179)	(0.0418)
religious	0.1102	0.2911	0.1171	0.2972	0.1051	0.2596
	(0.3206)	(0.0003)	(0.2785)	(0.0001)	(0.3654)	(0.0011)
pluralty	0.3583	0.2335	0.3259	0.2789	0.2963	0.1875
	(0.0254)	(0.0514)	(0.0307)	(0.0311)	(0.0748)	(0.0691)
pr	0.0497	0.1817	0.0726	0.1642	0.0533	0.1795
	(0.5307)	(0.0284)	(0.3837)	(0.0353)	(0.5169)	(0.0258)
factor(cl)closed list	0.5059	0.2148	0.4713	0.2583	0.4569	0.1817
	(0.0046)	(0.0523)	(0.0061)	(0.0329)	(0.0080)	(0.0636)
factor(cl)open list	0.2586	0.1841	0.2326	0.2220	0.2336	0.1449
	(0.1023)	(0.0855)	(0.1229)	(0.0705)	(0.1311)	(0.1125)
mdmh	0.0039	0.0217	−0.0002	0.0238	0.0019	0.0175
	(0.9103)	(0.3729)	(0.9945)	(0.2763)	(0.9574)	(0.4356)
factor(system)parliamentary	0.2047	−0.1568	0.1642	−0.1308	0.2311	−0.2068
	(0.2799)	(0.2042)	(0.3936)	(0.2567)	(0.2232)	(0.1281)
factor(system)presidential	0.0390	−0.3275	0.0187	−0.3105	0.0881	−0.3289
	(0.8290)	(0.0430)	(0.9214)	(0.0517)	(0.6267)	(0.0423)
demsys	−0.0126	0.0187	−0.0046	0.0018	−0.0133	0.0204
	(0.8126)	(0.5923)	(0.9317)	(0.9616)	(0.8144)	(0.5248)
polity	0.0206	0.0039	0.0126	0.0069	0.0210	0.0010
	(0.3876)	(0.8518)	(0.5859)	(0.7214)	(0.3766)	(0.9623)
gdppc	−0.1518	−0.0250	−0.1622	−0.0169	−0.1456	−0.0073
	(0.0046)	(0.5634)	(0.0034)	(0.7016)	(0.0072)	(0.8540)
gini	0.0017	0.0016	0.0014	0.0011	0.0017	0.0008
	(0.7593)	(0.6682)	(0.7998)	(0.7450)	(0.7576)	(0.8122)
ethfrac	0.1302	0.0985	0.0838	0.1437	0.0757	0.0742
	(0.5662)	(0.3961)	(0.7122)	(0.1873)	(0.7447)	(0.4977)
extremist2			0.1150	−0.0397		
			(0.1230)	(0.2649)		
ev_total:extremist2			−0.4975	0.3889		
			(0.0361)	(0.0035)		
ethnic					0.4731	−0.0636
					(0.0193)	(0.5749)
ev_total:ethnic					−1.1454	0.6354
					(0.0384)	(0.0644)
R^2	0.2419	0.4017	0.2554	0.4542	0.2540	0.4172
Adj. R^2	0.2107	0.3771	0.2209	0.4289	0.2195	0.3902
Num. obs.	431	431	430	430	431	431
RMSE	0.6337	0.3537	0.6289	0.3385	0.6302	0.3500
N Clusters	75	75	75	75	75	75

Table 3: Replication of Table D1 in the Appendix with exact p-values in parentheses

Table 4: Replication of Table 1

	Low.Volatility.Person	High.Volatility.Person	Differences.Person	Low.Volatility.Program	High.Volatility.Program	Differences.Program
Overall	2.67	2.97	0.29	3.08	2.97	-0.11
Overall.SE	0.05	0.05	0.07	0.04	0.04	0.06
Nonethnic	2.62	2.97	0.35	3.07	2.93	-0.15
Nonethnic.SE	0.05	0.05	0.07	0.04	0.04	0.07
Ethnic	2.92	2.96	0.04	3.11	3.13	0.03
Ethnic.SE	0.11	0.11	0.14	0.06	0.05	0.07
Differences.Ethnic	0.30	-0.01		0.03	0.20	
Differences.Ethnic.SE	0.13	0.12		0.07	0.06	
Moderate	2.60	3.12	0.52	3.05	2.76	-0.29
Moderate.SE	0.10	0.10	0.13	0.06	0.07	0.09
Extremist	2.72	2.84	0.12	3.11	3.13	0.02
Extremist.SE	0.07	0.08	0.10	0.04	0.04	0.06
Differences.Ideo	0.12	0.12		0.06	0.37	
Differences.Ideo.SE	0.14	0.14		0.06	0.08	

Table 5: Computational reproduction of original Table 2

Outcome	τ	p value	95 per cent CI	h	b	N+	N-	95 per cent CI (h = b)
Young	-0.084	0.040	[-0.185; -0.005]	13.491	24.643	527	559	[-0.172; 0.043]
Young and Outsider	-0.092	0.015	[-0.185; -0.020]	12.280	22.333	501	523	[-0.185; -0.020]

Table 6: Robustness to different polynomials

Fit	Outcome	τ	p value	95 per cent CI	h	b	N+	N-	95 per cent CI (h = b)
Original	Young	-0.084	0.040	[-0.185; -0.005]	13.491	24.643	527	559	[-0.172; 0.043]
Original	Young and Outsider	-0.092	0.015	[-0.185; -0.020]	12.280	22.333	501	523	[-0.185; -0.020]
Quadratic	Young	-0.095	0.047	[-0.194; -0.001]	24.150	37.942	699	756	[-0.178; 0.041]
Quadratic	Young and Outsider	-0.105	0.041	[-0.197; -0.020]	21.157	33.185	667	725	[-0.189; 0.010]
Uniform Kernel	Young	-0.078	0.036	[-0.168; -0.013]	14.219	35.219	540	578	[-0.196; 0.009]
Uniform Kernel	Young and Outsider	-0.100	0.034	[-0.186; -0.038]	12.662	29.492	509	539	[-0.177; 0.013]
Triangular Kernel	Young	-0.084	0.040	[-0.185; -0.005]	13.491	24.463	527	559	[-0.172; 0.043]
Triangular Kernel	Young and Outsider	-0.092	0.037	[-0.185; -0.020]	12.280	22.333	501	523	[-0.194; -0.001]
Epanechnikov Kernel	Young	-0.085	0.040	[-0.187; -0.006]	12.420	24.038	504	530	[-0.177; 0.043]
Epanechnikov Kernel	Young and Outsider	-0.092	0.037	[-0.201; -0.002]	11.521	11.521	475	496	[-0.194; -0.001]

Appendix B. Figures.

Figure 1: Replication Original Figure 3: Electoral Volatility, Personalism, and Programmaticness - Raw Correlations

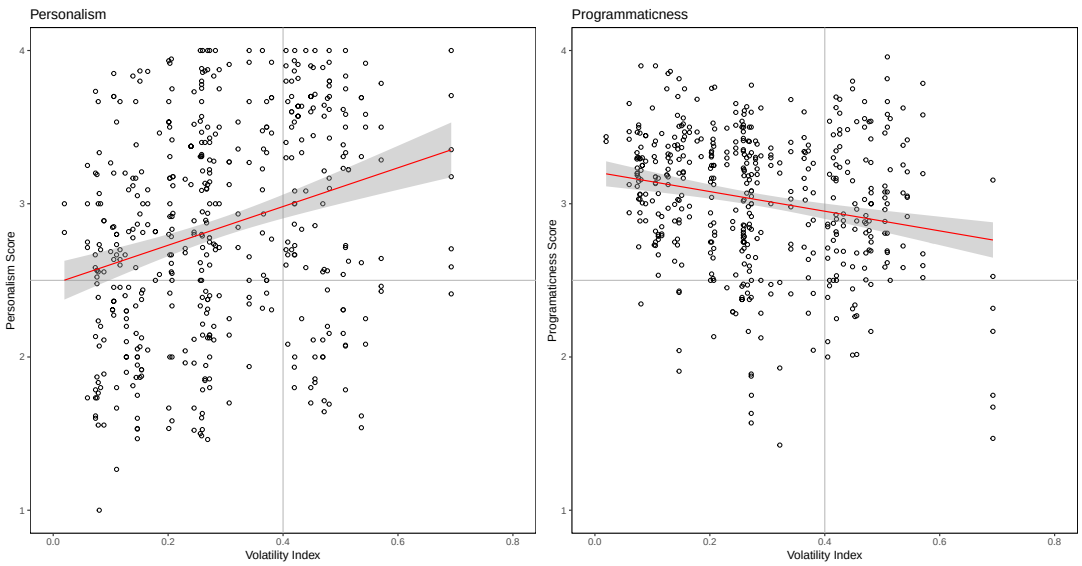

Figure 2: Replication Original Figure 4: Electoral Volatility, Ethnic Parties, and Electoral Strategies - Raw Correlations

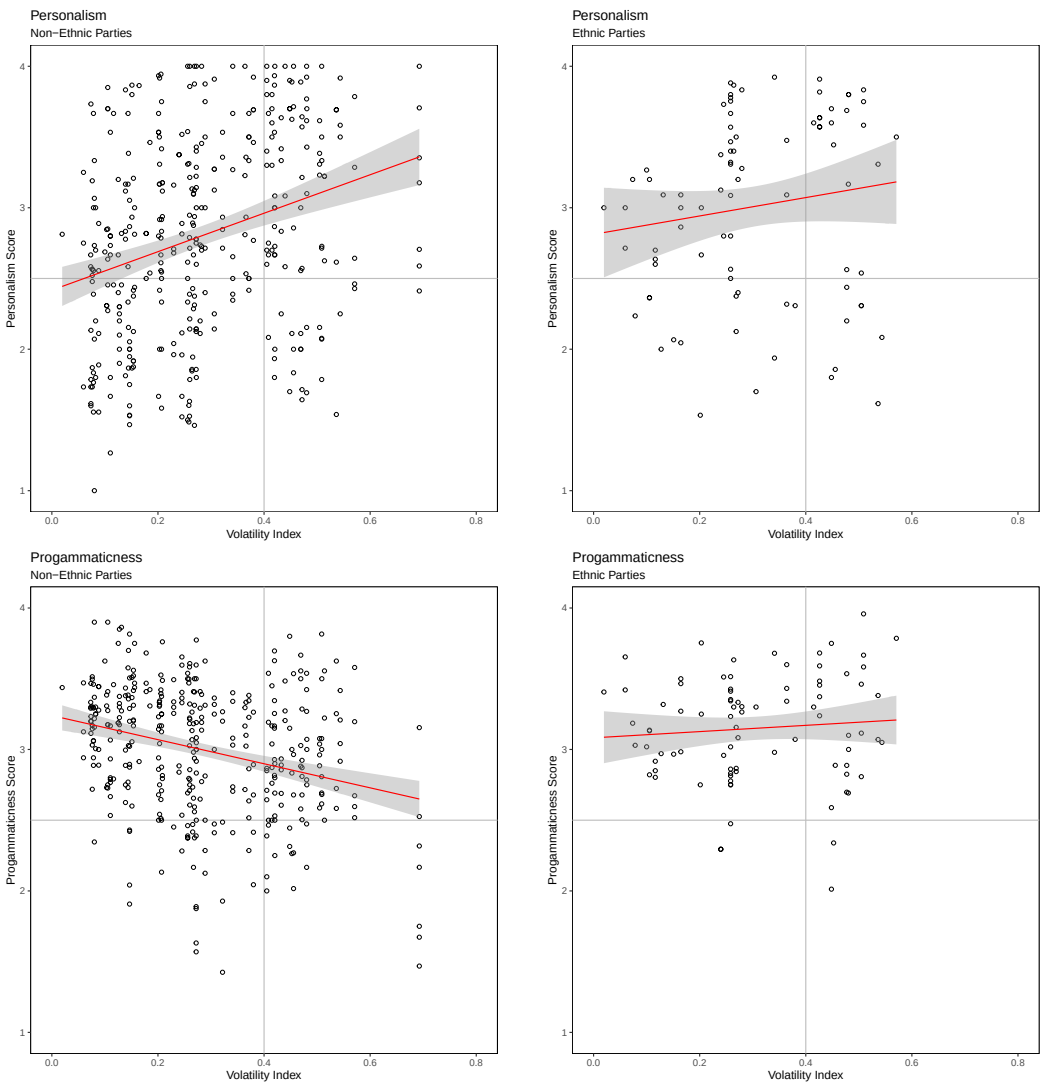

Figure 3: Replication Original Figure 5: Electoral Volatility, Ideological Extremism, and Electoral Strategies - Raw Correlations

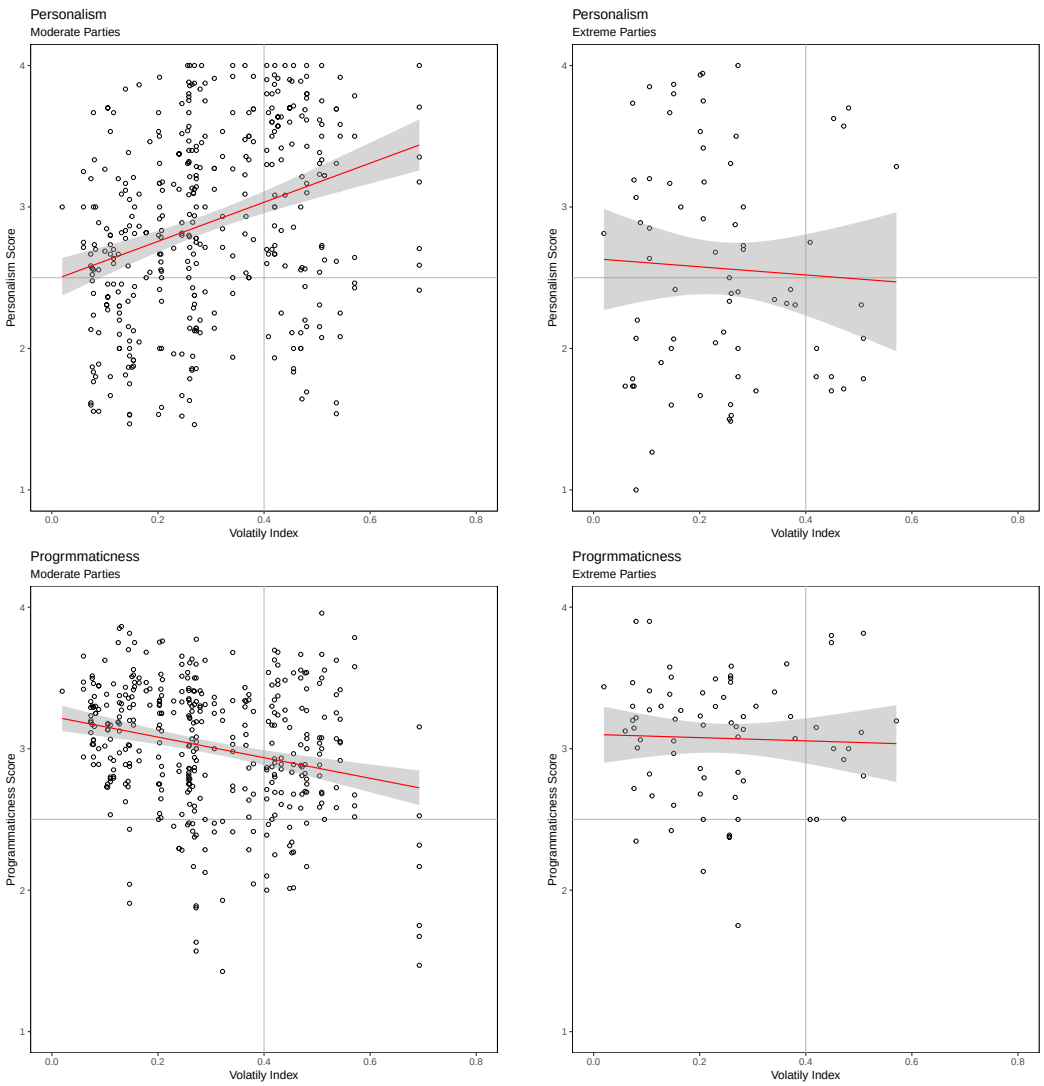

Figure 4: Predicted Personalism at Volatility Quantiles across Ideology

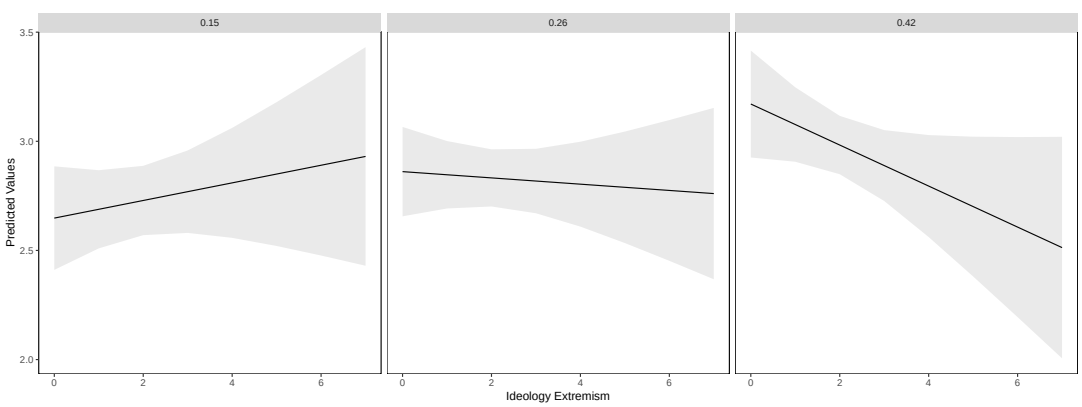

Figure 5: Predicted Programmaticness at Volatility Quantiles across Ideology

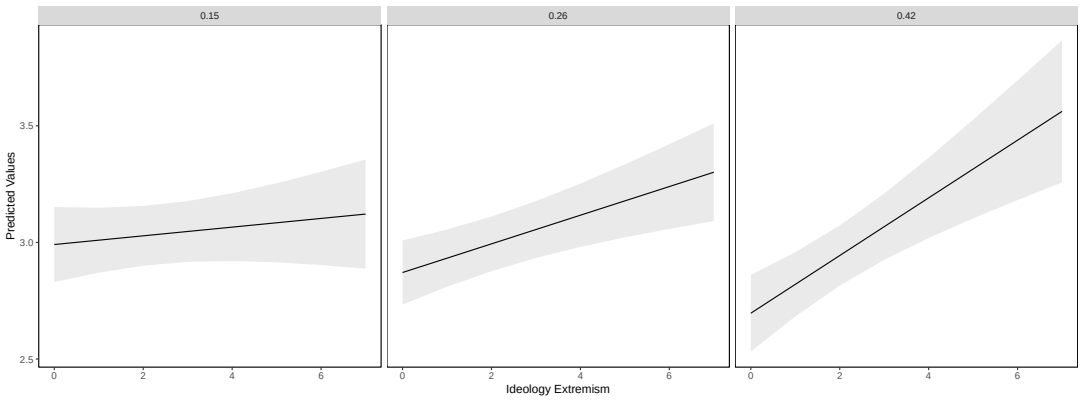

Figure 6: Replication Original Figure 6: Quantitative case study in Brazil

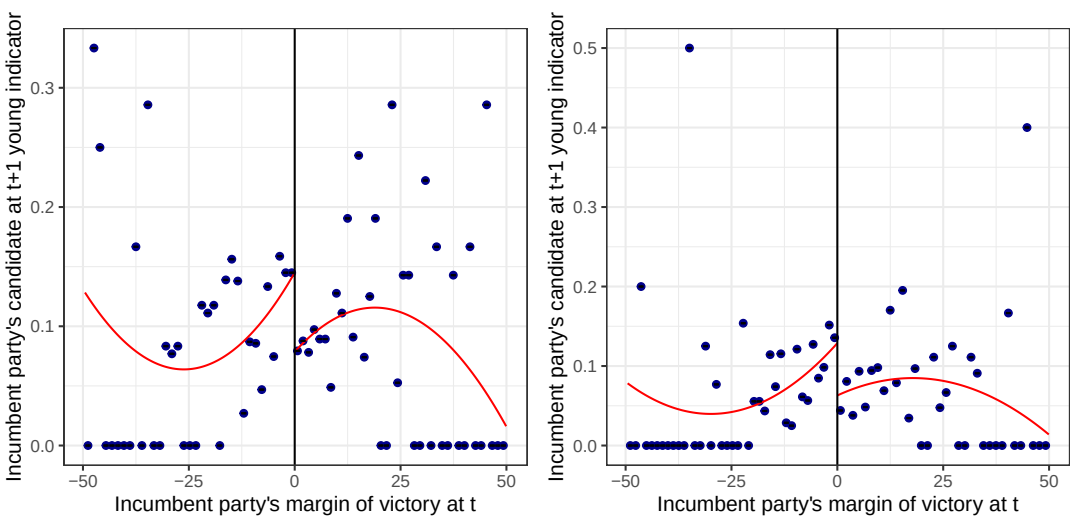

Figure 7: Replication Figure 6: Robustness to first order polynomial

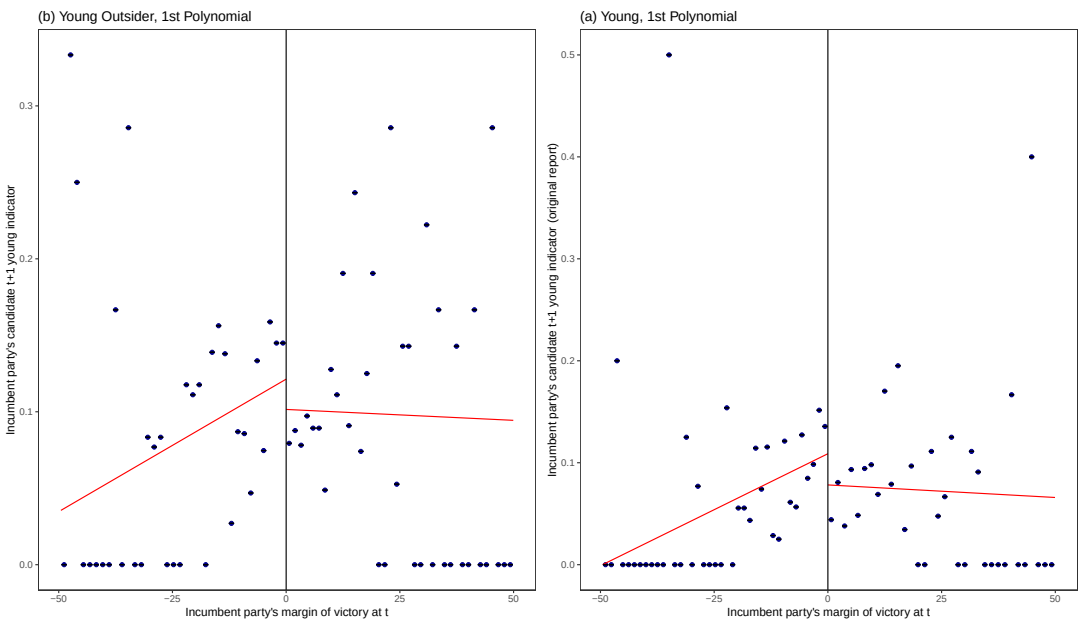

Figure 8: Replication Figure 6: Robustness to different binning procedures

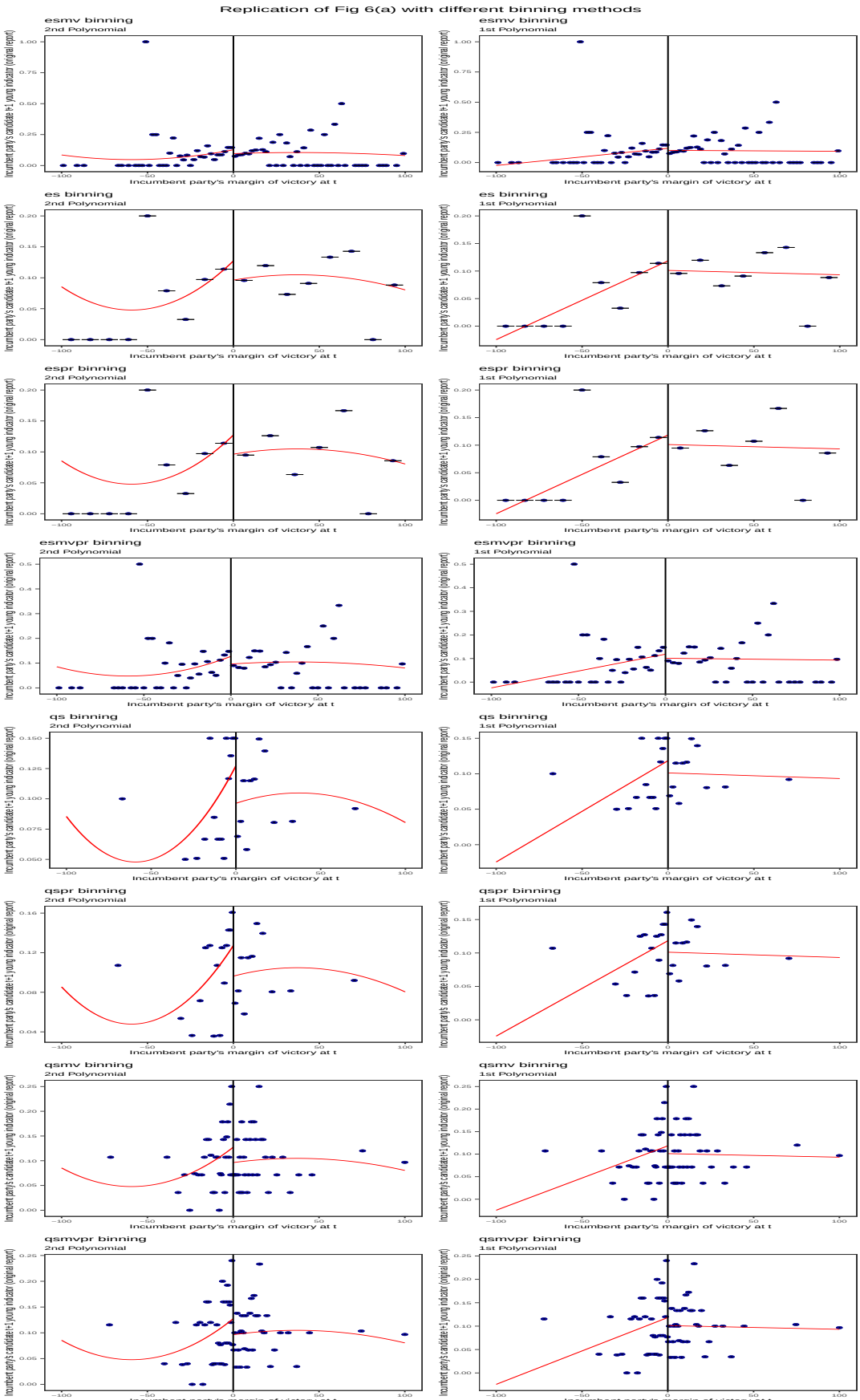

Figure 9: Replication Figure 6: Bandwidth selection, young candidates

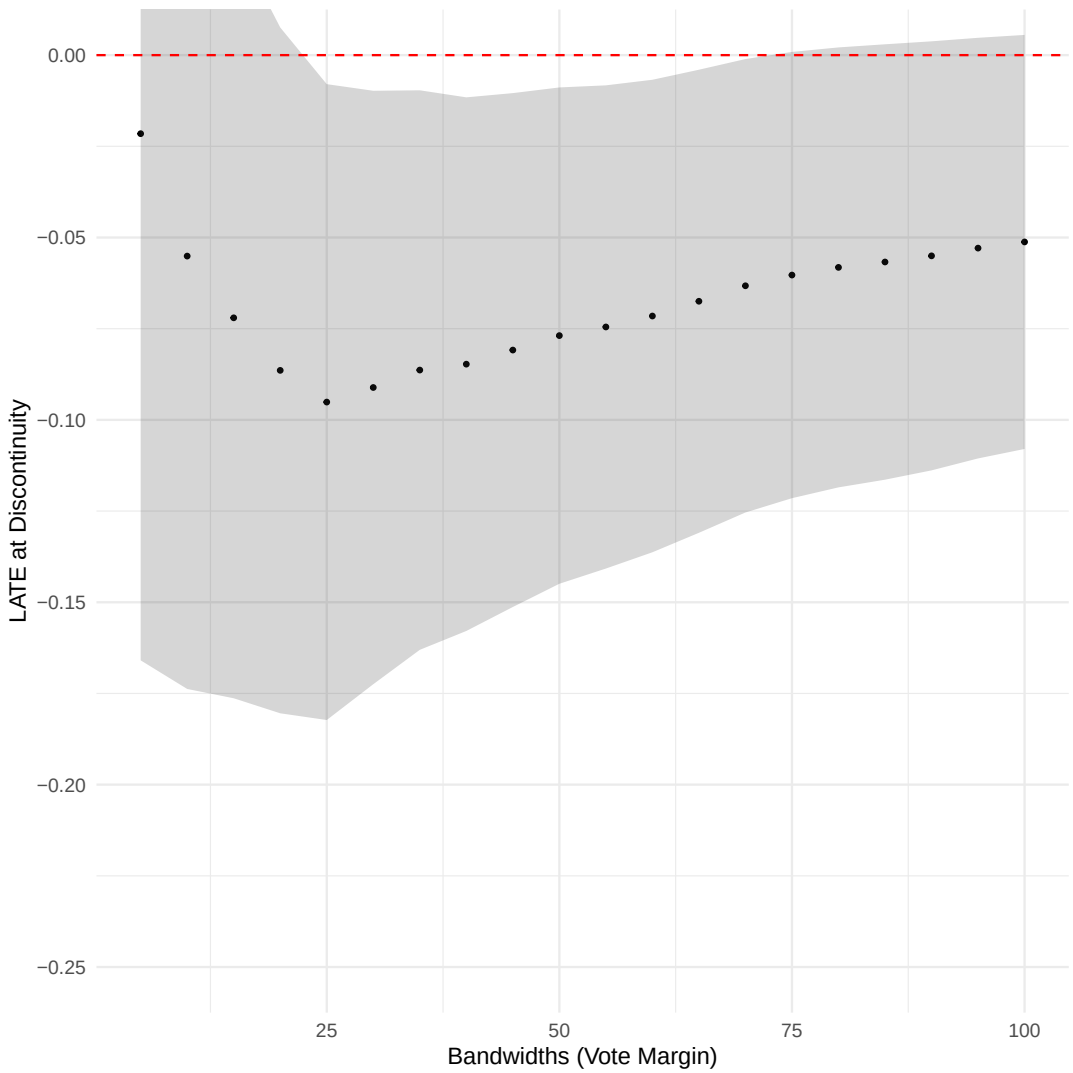

Figure 10: Replication Figure 6: Bandwidth selection, young outsider candidates

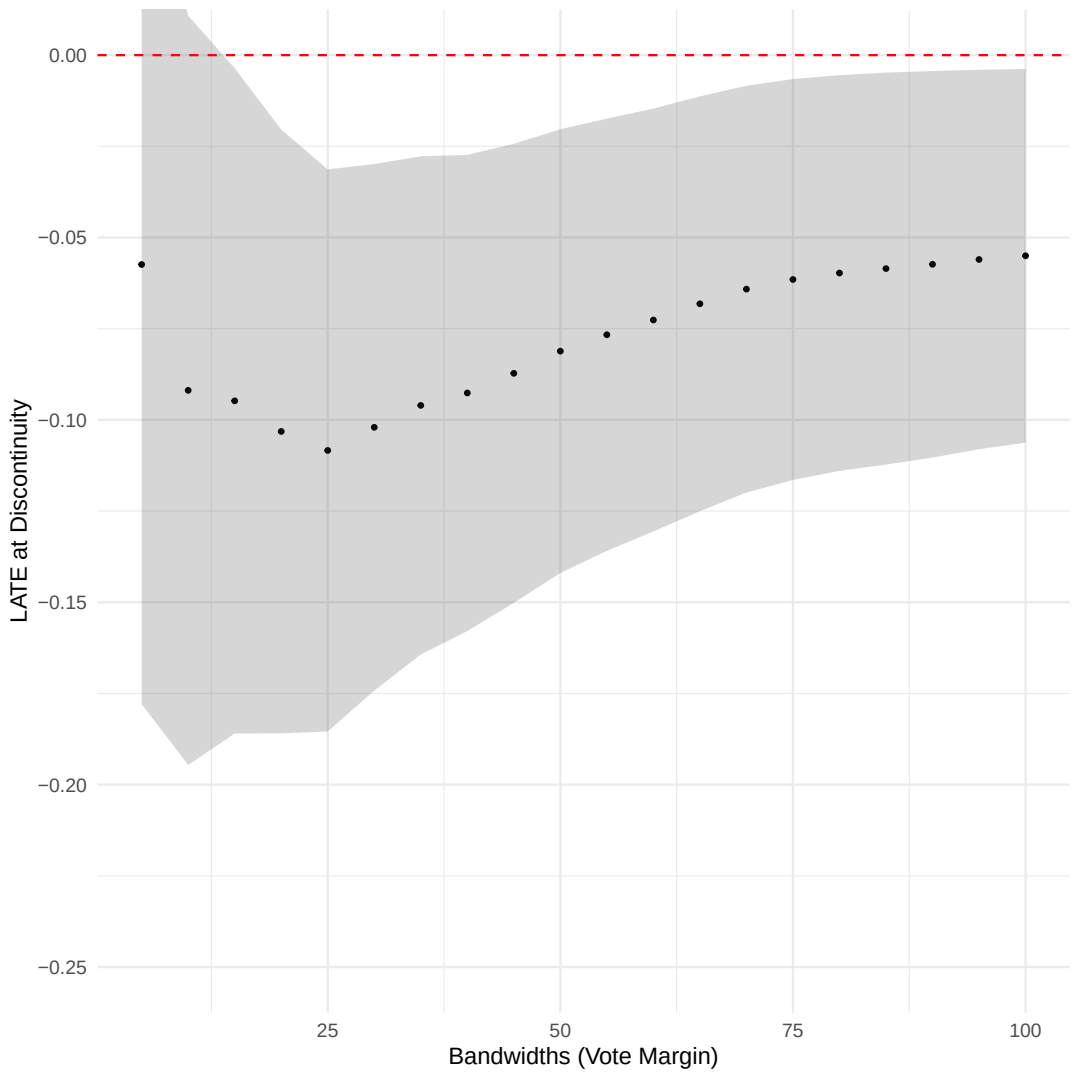

Figure 11: Replication Figure 6: Placebo cut-offs, young candidates

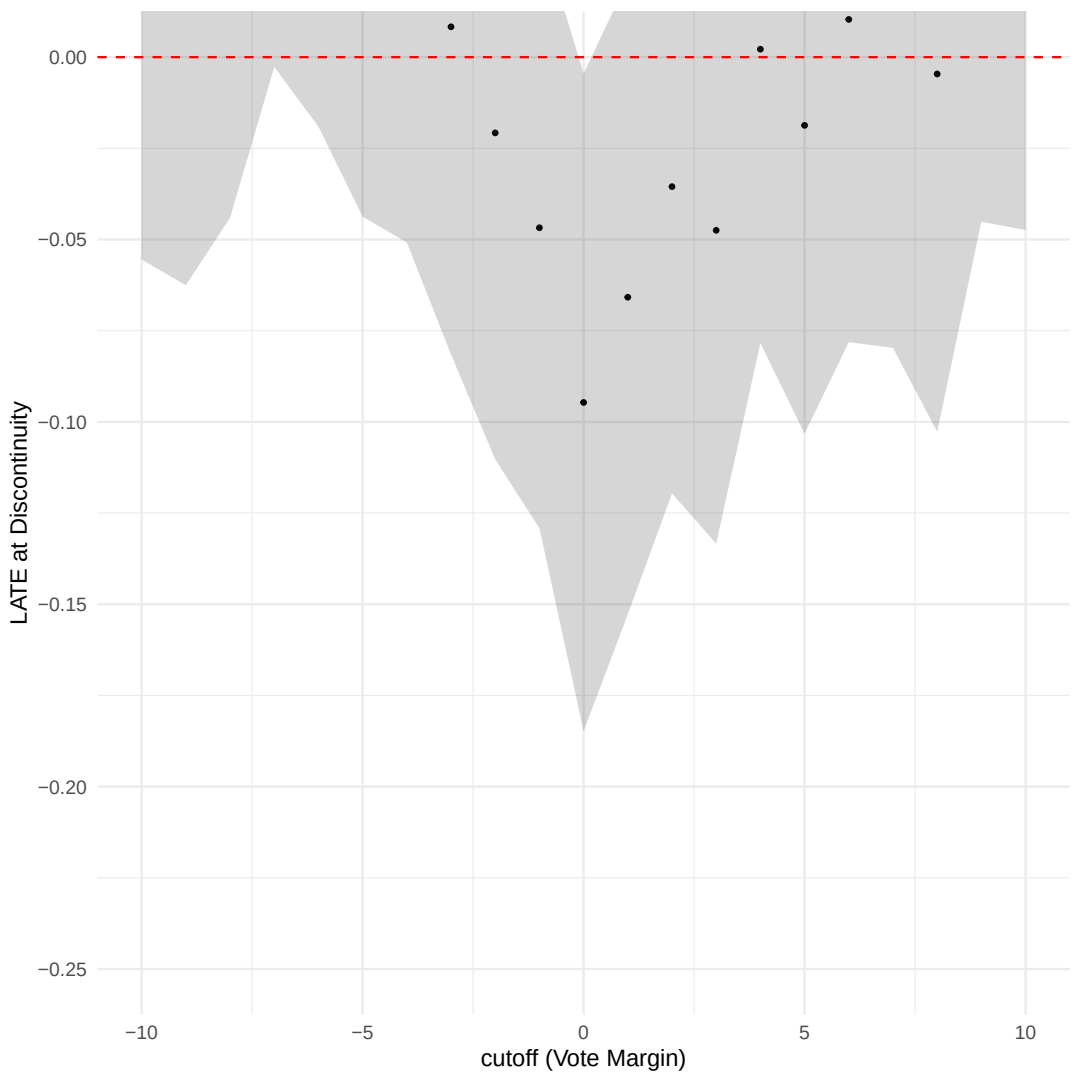
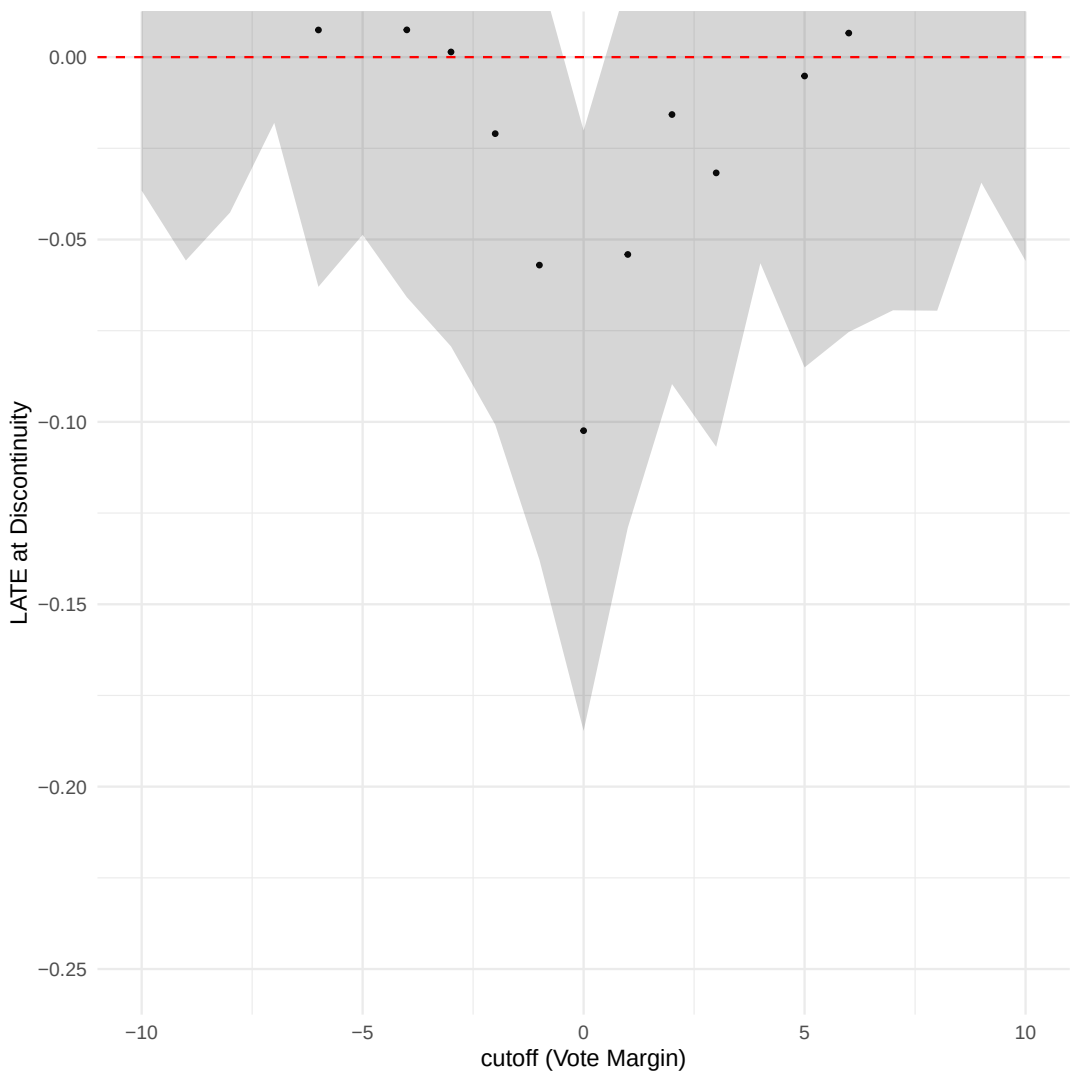
Figure 12: Replication Figure 6: Placebo cut-offs, young and outsider candidates