Matthew Diana
PPPA 8022
Replication Project
April 22, 2022

I did **not** use the data posted by the authors

Replication and Extension: Kim and Urpelainen (2017) "The Polarization of American Environmental Policy: A Regression Discontinuity Analysis of Senate and House Votes, 1971-2013"

Section 1: Introduction

Environmental policy in the United States has risen to the forefront of political discourse in recent years. Debates surrounding America's commitments to the Paris Climate Accords and construction of the Keystone XL Pipeline have driven a wedge between the two major parties, while consecutive Trump and Biden Administrations have taken drastically different positions surrounding the United States' approach to addressing manmade climate change. Kim and Urpelainen (2017) use a regression discontinuity model to assess how much environmental-policy polarization in Congress is caused by partisan ideology; this design attempts to hold voter preferences constant to isolate the effect of partisan elite polarization.

In this paper, I will attempt to replicate the primary specifications of Kim and Urpelainen's study; generally, these specifications are regression discontinuity designs that assess the average pro-environmental voting record of members of Congress based on their party, with a discontinuity at the 50-percent Democratic vote threshold (i.e., the threshold where a Democrat wins the election). I included minor changes in the research design, primarily by omitting third-party and independent candidates from the regression discontinuity to assess the polarization between the two major parties, which is the central focus of the original paper. This change served both practical considerations due to limitations in my data sources, as well as this conceptual aim. To extend the authors' framework, I will include data from 2014-2020 in a separate model, as Kim and Urpelainen's original timeframe includes only the years between

1971-2013 and will not capture recent shifts in how the two parties approach environmental policy and politics.

The results of the replication that includes only the years from 1971-2013 show substantively similar results to Kim and Urpelainen, while some specifications that limit the sample size show identical results. Overall, both the authors' and my own results reveal that electing a Democrat to Congress causes a roughly 40-percentage point increase in pro-environmental voting record. Many of the discrepancies may be explained by the omission of third-party and independent candidates, but this should have little effect on the sample of elections with results near the 50-percent vote threshold. By expanding the timeframe of the study through 2020, my extension reveals a starker divide between the parties in the average environmental voting record, indicating that polarization on environmental issues has increased in that period relative to the 1971-2013 timeframe.

Section 2: Research Design

Kim and Urpelainen assess partisan polarization surrounding environmental politics in the United States Congress using a regression discontinuity design (RDD). They argue that since the establishment of the Environmental Protection Agency in 1970 by the Nixon Administration, environmental issues have lost bipartisan support, with federal environmental policies now coming primarily from Democratic executives. The literature provides two explanations for such polarization; the first relies on median-voter theorem, suggesting that the "Democrat-Republican wedge reflects the preferences of this median voter in different electoral conditions" (Kim & Urpelainen, 2017, p. 457). The second explanation argues that a growing divide between party elites, reflecting ideological differences among partisans in Congress, better explains partisanship in Congress, rather than a cleavage between Republican and Democratic *voters* (Kim & Urpelainen, 2017). Kim and Urpelainen note that while this latter explanation is accepted among researchers, it is difficult to empirically research how much polarization is due to shifting electoral conditions as opposed to ideological rifts among partisan elites.

An RDD is a useful tool overcome some of these difficulties. Kim and Urpelainen (2017) argue that the design is supported by the literature as a means to "consider the election of Republican and Democratic officials in close elections as essentially random" (p. 458). In the context of this study, the RDD is meant to isolate the effect of *partisan elite* polarization, rather than polarization among the *electorate*: the authors rationalize that the RDD will hold median voter preferences constant, allowing them to isolate the effect of partisanship on average voting decisions. Therefore, the causal question of this study focuses on "the causal effect of electing a Democrat instead of a Republican in close elections on pro-environmental voting" (Kim & Urpelainen, 2017, p. 456).

To operationalize environmental voting records, the authors' use the League of Conservation Voters (LCV) yearly scorecard for legislators' voting decisions (League of Conservation Voters, 2021). The LCV assesses all votes taken by Congress, classifies those that are "environmentally" focused, and then determines whether a "yea" or "nay" vote is the "pro-environmental" or "anti-environmental" position.

Voting decision serves as the unit of observation for this RDD; between 1971-2020, the sample includes 501,393 total observations, with 61,822 votes in the Senate and 439,571 votes in the House of Representatives. Kim and Urpelainen (2017) utilize several different specifications; first, in a parametric design, the authors use an RDD with the full sample in each chamber of Congress, where treatment is a binary indicator for Democratic victory and the running variable is the proportion of the vote won by the Democratic candidate. In separate specifications, the authors also include the second, third, and fourth powers of the Democratic vote share. This model is displayed in the following equation:

$$Pro\_Env_{i,t,r} = \beta_0 + \beta_1 Democrat_{i,t} + f\left(Vote\ Share_{i,t,-\alpha}\right) + \epsilon_{i,t,r}$$

where "i denotes legislator (House Representative or Senator), *t* denotes year, and *r* denotes a roll call in that year" (Kim & Urpelainen, 2017, p. 467). $Pro\_Env_{i,t,r}$ is a dummy variable indicating whether legislator *i* voted in the pro-environmental position for roll-call vote *r.* $Democrat_{i,t}$ is a dummy variable indicating whether legislator *i* was elected as a Democrat. The term $Vote\ Share_{i,t,-\alpha}$ is specified such that it denotes the vote-share in the most recent election for member i, and $\epsilon_{i,t,r}$ is the error term. Standard errors for each specification are clustered by legislator. The objective of this study is to yield an unbiased estimate of β_1, which indicates the percentage point difference in the pro-environmental voting record for Democratic legislators over Republican legislators.

The authors also use a nonparametric approach with a limited sample size, assessing the difference between electing a Democratic and Republican candidate in races with a 3, 2, 1, and 0.5 percentage point margin of victory. This specification uses the same logic as the parametric equation shown above by limiting the sample size to only close races. Finally, in both the parametric and nonparametric models, the authors include two further specifications with state fixed-effects and state-and-year fixed-effects.

2.1: Data: The congressional voting scorecard from the League of Conservation Voters (2021) is the most important set of data included in this research, as it provides both the total number of observations (roll-call votes) and the dependent variable (Pro-Environmental Position). I collected this dataset from the LCV's official website for every vote cast between 1971-2020. Using these data, I merged the roll-call votes with data from CQ Press (2022) on the results of federal elections for all relevant years, such that a roll-call vote cast by a member of Congress is merged to the results of the election that put that legislator in office for the relevant term. Members elected by a special election are included in the RDD; however, members who are appointed to fill a vacancy are not included while they are serving the term that they were appointed to, as they were not elected to that position and thus have no vote share.

2.2: Extension: To extend Kim and Urpelainen's work, I will include the years 2014-2020 in a separate model, expanding the total timeframe of the study to include the years 1971-2020. The authors and I are limited by the LCV scorecard, which does not include any years prior to 1971.

I argue that the addition of the years 2014-2020 will include interesting new data because environmental policy rose to the forefront of the congressional agenda several times in that period. In 2015, the Obama Administration entered the United States into the Paris Climate Agreement, which included pledges to reduce carbon dioxide emissions; the agreement was opposed by many Republicans. Beginning in 2017, the Trump Administration took actions that the League of Conservation Voters (2021) classified as anti-environmental, including several Executive Branch appointments to lead of the Environmental Protection Agency, Department of the Interior, and Department of Energy. These actions pushed environmental issues to the forefront of American political discourse and congressional attention, so I believe that the addition of these years will provide interesting data to expand upon Kim and Urpelainen's original work on partisan polarization.

2.3: Adjustments Made for this Replication: In replicating this study, I made the strategic and substantive choice to limit the analysis to members and candidates aligned with one of the two major parties (i.e., Democratic or Republican), while excluding third party or independent candidates. This means that roll-call votes are limited to members who run for office as either a Democrat or Republican, and that Democratic vote share is measured as a percentage of the total major-party votes. I made this decision partly due to data availability; members of Congress were identified by the party identification that they ran under (CQ Press, 2022), so it would have been difficult and subjective to code individual members who run as independents but caucus with a major party as having partisan alignment. Conceptually, the focus of this paper is on the polarization in Congress between the two major parties, so I argue that I am justified in examining only the effect of Democratic voting patterns against Republican voting patterns in Congress. This similarly applies to assessing the discontinuity at the vote-share threshold of 50-percent among major party votes, as that is the threshold where voters make the choice to send a Democrat or Republican to Congress over the other party, rather than over an independent or third-party candidate. In sum, this adjustment means that my model will not assess the tendency of Democrats to take the pro-environmental position against all other members, but rather it will assess Democrats against only Republicans. Kim and Urpelainen provide justification for this approach in their report, noting that in their dataset, a small percentage of all votes were cast by legislators unaffiliated with a major party, and that the results of their RDD were not substantively changed by the exclusion of those legislators. This is similarly true for my replication dataset; around 0.3 percent of the roll-call votes between 1971-2020 were cast by members unaffiliated with the Democratic or Republican party.

Section 3: Results

Figure 1 in Appendix II displays a replication of the same figure from Kim and Urpelainen, albeit with my addition of years 2014 – 2020. Replications of Figure 1 using only the years 1971 – 2013 are available in Appendix I, along with the author's original chart for comparison. These charts show results that are very similar to the authors' in both the trend lines and the variation in binned results. The differences in the dispersion of some notable outlier data may be due to different bin sizes used by the authors, which is unspecified.

The Table 1 Replication in Appendix II displays the results for the various models that the authors' employed, with the addition of the years 2014-2020. Replication results of the timeframe of 1971-2013 are again available in Appendix I, with the authors' original results for reference. With the additional years, these results indicate a consistent effect of around a 50-perecentage point increase in pro-environmental voting position when a Democrat is elected. These results show a clear increase over Kim and Urpelainen's original findings, indicating that in those seven years, polarization on environmental issues increased above the average effect in the years 1971-2013.

Kim and Urpelainen include an ordinary least squares (OLS) estimation in their table with little explanation; ostensibly, it serves as a comparison for the effect of examining this question through an RDD. The estimated effect in each RDD specification is larger than the one provided by an OLS, suggesting that examining pro-environmental voting records as a difference of averages between the parties will underestimate its effect.

3.1: Potential Reasons for Discrepancies: I found a stark difference in the trend line of the RDD chart for Senators near the leftward extreme where Democratic vote-share is close to zero. Kim and Urpelainen's charts show a clear upwards shift in the trend line at the left extreme of the graph, whereas my replication displays a downward trend. This may be due to my decision to limit this analysis to only major-party candidates, excluding independents and third-party candidates. To use an example to display this hypothesis, there are currently two independent Senators, Bernie Sanders (I-VT) and Angus King (I-ME). While classified as independents, they each caucus with Senate Democrats and vote largely in-line with the Democratic party. Because of this, they do not face serious challengers from the Democratic party in their home-states; in 2018, for example, Sanders ran as an independent against a Republican and other independent candidates, with no Democratic candidate officially on the ballot. This has the effect of classifying Sanders as a candidate who won an election with a Democratic vote share of zero percent; in reality, Sanders acts as a *de facto* Democratic Senator. Including Sanders and other *de facto* partisans in the RDD may to have this effect at the tail-end of these figures, as they are grouped in with heavily Republican leaning victories along the vote-share axis but vote in-tandem with the Democratic party. Regardless, this effect is not present in the vicinity of the discontinuity, which is the area of interest for this study; as such, their inclusion or omission should have minimal effect on the results.

Turning to the replication of Table 1 using only the years between 1971-2013, found in Appendix I, I hypothesize several potential reasons for the slight discrepancy in results. The most obvious is that I removed all legislators and candidates unaffiliated with the Democratic or Republican party from the sample. While those legislators accounted for just above 0.3 percent of all votes cast, their exclusion is nonetheless a change that may have influenced results, particularly in the parametric specifications. In the non-parametric specifications using only races with a 0.5 percent margin of victory, there is no discrepancy between Kim and

Urpelainen's original findings and the replication results; this is intuitively sound, as the sample sizes are reduced to only 915 observations. It follows that no legislators elected as an independent or third-party candidates were elected with such a slim margin in this time frame, so their exclusion did not affect that specification. Further, limiting the data to those races should remove candidates that run as independents or third-party candidates but act as *de facto* partisans, as I hypothesize that they would typically fall towards the extreme ends of the vote-share spectrum.

Additionally, there are some areas where our data sources were potentially unaligned. While the LCV vote data was gathered from the same source, the authors did not specify where they obtained data on congressional election results. For this replication, I utilized the CQ Press (2022) database on congressional general and special election results, but there were cases of missing or unmatched data when merging with the LCV votes. This data loss was minimal but may still potentially account for the discrepancy in sample sizes and RDD coefficient outcomes. Regardless, the replication table for the timeframe between 1973-2013 is substantively the same as the authors' findings, indicating that a Democratic victory in close congressional races resulted in around a 40-percentage point increase in the likelihood that the legislator will vote for the pro-environmental position.

Section 4: Discussion and Conclusion

The results of this RDD indicate a strong causal effect of partisanship on polarization in congressional voting patterns. Further, the use of an RDD suggests that we can "rule out public opinion as an explanation" in favor of polarization among partisan elites (Kim & Urpelainen, 2017, p. 481). In both Kim and Urpelainen's original findings, as well as my replication of their original timeframe of 1971-2013, the results indicate about a 40-percentage point effect on pro-environmental voting of electing a Democrat over a Republican. The authors did not identify any of their specifications as their primary model; rather, they provided all of their specifications in Table 1 and concluded that Democratic victories results in *about* a 40-percentage point increase in pro-environmental voting. I am confident that my findings display the same rigor as theirs because, although they do not display the same exact findings in many specifications, they are in line with an estimated effect of about 40-percentage points. Similarly, the differences in estimations across individual specifications generally trend in the same direction, both across and down the table; for example, in the RDD specification with a second-order polynomial using votes from the Senate, the estimated effect of electing a Democrat decreased in specifications as State and year fixed effects were added for both the authors' original table and my replication.

By extending the Kim and Urpelainen's work through 2020, I show that partisan polarization surround environmental issue has increased in that time frame. The results in Table 1 of Appendix II show that electing a Democrat has a positive effect of around 50-percentage points on the tendency to vote in the pro-environmental position. These results are a significant increase from the timeframe of 1971-2013 and support the hypothesis that polarization on environmental issues has increased in the period between 2014-2020. However, the findings cannot be used to conclude that any specific variable caused the increase in polarization in that time. Kim and Urpelainen (2017) suggest that further research should "focus on understanding the origins of elite partisan polarization" (p. 481); my results indicate that the years since their publication may provide further opportunities to do so.

By using the LCV data to measure the voting records of legislators, Kim and Urpelainen introduce a key limitation to the findings of this study. The results are limited to the issues that

LCV identifies as relevant to environmental policy, so we must rely on their assessment of the pro-environmental position, each of which are subjective judgements that may change over time. Further, given that the LCV is an environmental advocacy organization, their classifications may reflect the priorities and opinions of *advocates* in the environmental space; these priorities may exclude issues that would reasonably be considered environmental at some times, or include some which would not be considered environmental, depending on the demands of their membership or biases held by advocates in the field of environmental advocacy. If these concerns are founded, it may introduce bias to the estimates of polarization on environmental issues.

Appendix I

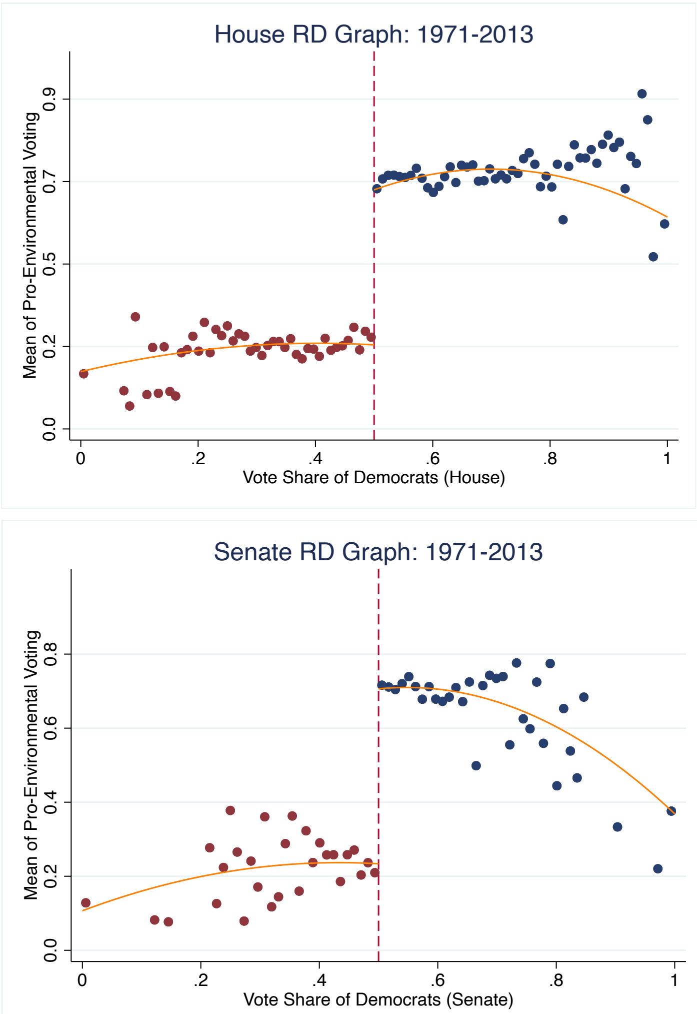


Figure 1: Replication of Kim and Urpelainen's graphical representation of RDD results for the House and Senate, for the timeframe of 1971-2013. Vertical line represents the 50-percent threshold for the Democratic candidate, the threshold for victory

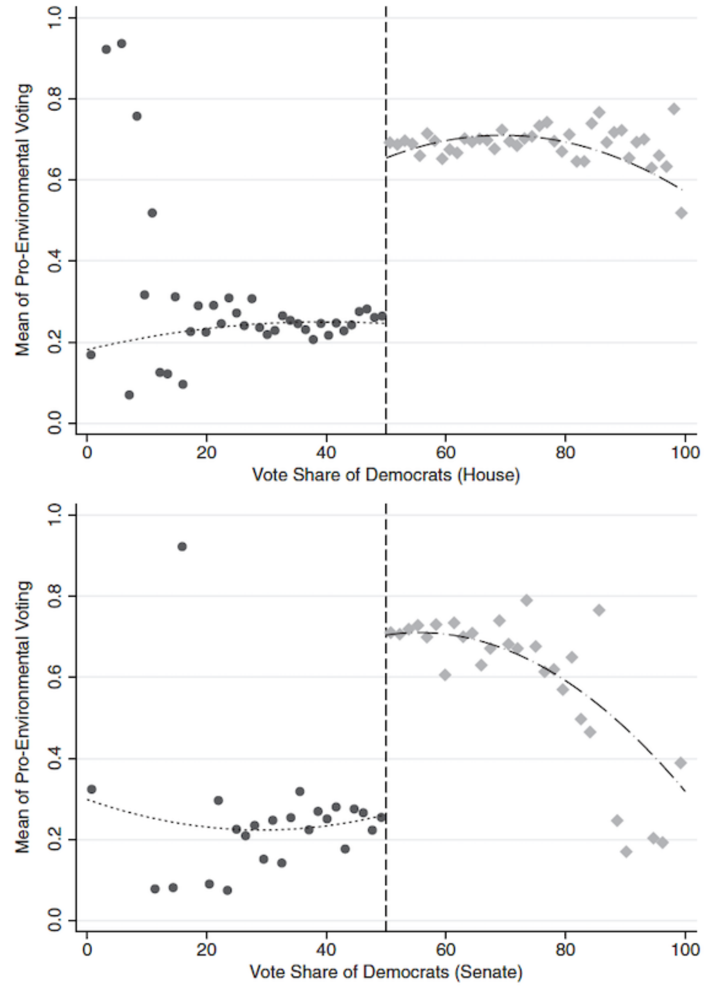


Figure 1 Original
(Kim & Urpelainen, 2017, p. 473)

Table 1 Replication: 1971-2013

Table 1 Replication: Voting on Environmental Bills by Senators and House Representatives, 1971 - 2013	(1)	(2)	(3)	(4)	(5)	(6)
		Senate			House	
OLS	0.451***	0.382***	0.387***	0.452***	0.435***	0.443***
	(0.026)	(0.017)	(0.017)	(0.012)	(0.010)	(0.010)
	50520	50520	50520	336190	336190	336190
RDD, second-order polynomial	0.497***	0.423***	0.418***	0.465***	0.464***	0.469***
	(0.030)	(0.019)	(0.019)	(0.014)	(0.012)	(0.012)
	50423	50423	50423	336190	336190	336190
RDD, third-order polynomial	0.475***	0.442***	0.440***	0.422***	0.438***	0.445***
	(0.033)	(0.021)	(0.020)	(0.017)	(0.015)	(0.015)
	50423	50423	50423	336190	336190	336190
RDD, fourth-order polynomial	0.475***	0.442***	0.439***	0.435***	0.445***	0.451***
	(0.033)	(0.021)	(0.020)	(0.017)	(0.015)	(0.015)
	50423	50423	50423	336190	336190	336190
3% Margin	0.481***	0.480***	0.494***	0.411***	0.401***	0.407***
	(0.045)	(0.033)	(0.033)	(0.025)	(0.023)	(0.022)
	4996	4996	4996	14438	14438	14438
2% Margin	0.554***	0.502***	0.517***	0.407***	0.403***	0.413***
	(0.046)	(0.039)	(0.033)	(0.029)	(0.027)	(0.026)
	3453	3453	3453	9496	9496	9496
1% Margin	0.529***	0.526***	0.741***	0.428***	0.443***	0.476***
	(0.076)	(0.065)	(0.139)	(0.040)	(0.032)	(0.032)
	1771	1771	1771	4988	4988	4988
0.5% Margin	0.493***	0.612***	0.785**	0.407***	0.448***	0.497***
	(0.010)	(0.044)	(0.224)	(0.062)	(0.054)	(0.076)
	915	915	915	2460	2460	2460
State FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	No	Yes	No	No	Yes

Standard errors in parentheses, clustered by legislator
* p<0.05, ** p<0.01, *** p<0.001

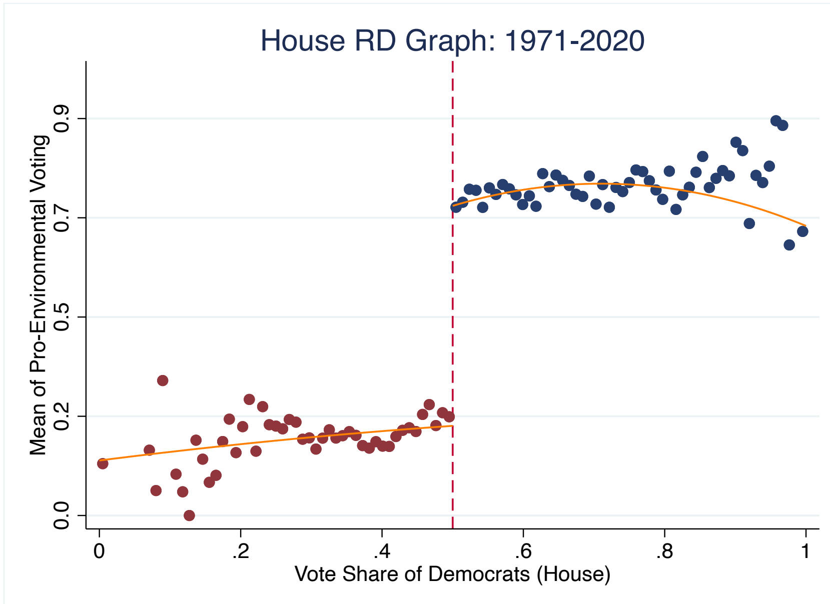
Table 1 Original: 1971-2013

(Kim & Urpelainen, 2017, p. 472)

Table 1. Voting on Environmental Bills by Senators and House Representatives, 1971–2013	(1)	(2)	(3)	(4)	(5)	(6)
		Senate			House	
OLS	.436***	.369***	.373***	.458***	.438***	.446***
	(.027)	(.018)	(.018)	(.012)	(.010)	(.010)
	49716	49716	49716	319258	319258	319258
RDD, second-order polynomial	.485***	.401***	.394***	.473***	.462***	.466***
	(.037)	(.023)	(.023)	(.014)	(.012)	(.012)
	49311	49311	49311	313604	313604	313604
RDD, third-order polynomial	.429***	.399***	.395***	.431***	.424***	.431***
	(.043)	(.027)	(.027)	(.016)	(.014)	(.014)
	49311	49311	49311	313184	313184	313184
RDD, fourth-order Polynomial	.429***	.399***	.395***	.443***	.432***	.438***
	(.043)	(.027)	(.027)	(.015)	(.014)	(.014)
	49311	49311	49311	313184	313184	313184
3% Margin	.453***	.467***	.467***	.422***	.411***	.422***
	(.048)	(.031)	(.029)	(.025)	(.023)	(.022)
	4933	4933	4933	13479	13479	13479
2% Margin	.524***	.499***	.515***	.416***	.407***	.421***
	(.054)	(.040)	(.034)	(.029)	(.028)	(.028)
	3367	3367	3367	8764	8764	8764
1% Margin	.462***	.523***	.739***	.446***	.441***	.497***
	(.093)	(.066)	(.139)	(.040)	(.034)	(.031)
	1724	1724	1724	4546	4546	4546
0.5% Margin	.493***	.612***	.785***	.407***	.467***	.520***
	(.100)	(.044)	(.224)	(.064)	(.051)	(.078)
	915	915	915	2259	2259	2259
State FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	No	Yes	No	No	Yes

Notes. As the table shows, the effect of electing a Democrat over a Republican in narrow races on pro-environmental voting is large and positive: all coefficients are positive and statistically significant, while most of them are larger than 40% points.
Standard errors clustered by legislators in parentheses.
$^{***}p < .01$.

Appendix II



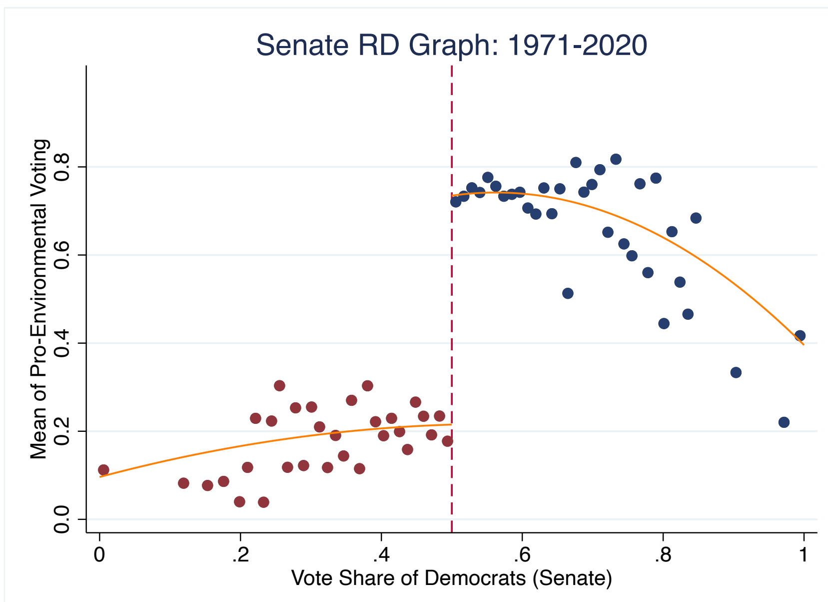


Figure 1: Replication of Kim and Urpelainen's graphical representation of RDD results for the House and Senate for the timeframe of 1971-2020. Vertical line represents the 50-percent threshold for the Democratic candidate, the threshold for victory

Table 1 Replication: 1971-2020

Table 1 Replication: Voting on Environmental Bills by Senators and House Representatives, 1971 - 2020

	(1)	(2)	(3)	(4)	(5)	(6)
	Senate			House		
OLS	0.514***	0.421***	0.428***	0.545***	0.510***	0.519***
	(0.024)	(0.017)	(0.017)	(0.011)	(0.010)	(0.010)
	60758	60758	60758	438107	438107	438107
RDD, second-order polynomial	0.555***	0.465***	0.461***	0.544***	0.530***	0.538***
	(0.027)	(0.019)	(0.019)	(0.013)	(0.012)	(0.012)
	60661	60661	60661	438107	438107	438107
RDD, third-order polynomial	0.526***	0.487***	0.485***	0.504***	0.507***	0.517***
	(0.030)	(0.021)	(0.020)	(0.017)	(0.015)	(0.015)
	60661	60661	60661	438107	438107	438107
RDD, fourth-order polynomial	0.526***	0.487***	0.484***	0.512***	0.513***	0.522***
	(0.030)	(0.021)	(0.020)	(0.016)	(0.014)	(0.0145)
	60661	60661	60661	438107	438107	438107
3% Margin	0.512***	0.517***	0.536***	0.476***	0.453***	0.459***
	(0.041)	(0.027)	(0.029)	(0.024)	(0.022)	(0.022)
	6081	6081	6081	18100	18100	18100
2% Margin	0.566***	0.505***	0.521***	0.480***	0.455***	0.458***
	(0.043)	(0.037)	(0.039)	(0.027)	(0.025)	(0.026)
	4113	4113	4113	12078	12078	12078
1% Margin	0.554***	0.529***	0.791***	0.467***	0.465***	0.516***
	(0.068)	(0.052)	(0.133)	(0.037)	(0.033)	(0.033)
	2039	2039	2039	6073	6073	6073
0.5% Margin	0.533***	0.612***	0.785**	0.444***	0.438***	0.498***
	(0.095)	(0.044)	(0.223)	(0.059)	(0.057)	(0.071)
	1003	1003	1003	2985	2985	2985
State FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	No	Yes	No	No	Yes

Standard errors in parentheses, clustered by legislator
* p<0.05, ** p<0.01, *** p<0.001

Works Cited

CQ Press. (2022). *Election Results*. Retrieved from CQ Press Voting and Elections Collection: https://library.cqpress.com/elections/index.php

Kim, S. E., & Urpelainen, J. (2017). The Polarization of American Environmental Policy: A Regression Discontinuity Analysis of Senate and House Votes, 1971–2013. *Review of Policy Research*, 456-484.

League of Conservation Voters. (2021). *National Environmental Scorecard*. Retrieved from LCV: https://scorecard.lcv.org/scorecard?year=2021