

Appendix for "Young People Punish Undemocratic Behavior Less Than Older People"

Kristian Vrede Skaaning Frederiksen[‡]

October 24, 2023

Contents

Appendix A: Undemocratic Behaviors, Treatment Examples, and Ethical Considerations	2
Appendix B: Illustration of Pooled Fixed Effects Estimate	6
Appendix C: Attentiveness and Effects of Other Factors	7
Appendix D: Additional Analysis Using Other Data Sources	23
Appendix E: Robustness to External Validity Bias	27

[‡]PhD Candidate, Department of Political Science, Aarhus University. ORCID: 0000-0003-1845-3297.
Email: ksf@ps.au.dk

Appendix A: Undemocratic Behaviors in Included Studies

Table A1 in this appendix shows all of the undemocratic behaviors employed in the included studies and thus in the analysis of the letter. As mentioned in the letter, these behaviors are violations of key democratic principles such as free and fair elections (e.g., reducing the number of polling stations in areas of opposition support), civil liberties (e.g., legitimizing physical attacks on political opponents or harassing journalists), and checks and balances (e.g., packing the courts or ignoring court rulings). Studies 1, 2, and 5 set up behaviors in undemocratic and democratic almost-symmetric pairs, whereas Studies 3-4 did not. Figures A1-A2 show examples of how the entire treatment scenarios look in the included conjoint (A1) and vignette (A2) experiments. For Figure A1, I use Study 4 as the example, because it exemplifies the voting likelihood scales as well as the forced-choice outcomes. For Figure A2, it is the last three lines of the vignette text which manipulates undemocratic behavior and the second outcome which is the relevant voting intentions outcome.

As these treatment scenarios are what the experiments presented in the letter are about, I also describe ethical practices here. Regarding deception, harm, and impact, it is clearly stated in the surveys that the candidates and scenarios—in all studies—are fictitious and hypothetical. I chose to present fictitious candidates and scenarios for the respondents in order not to let impressions about real-world candidates or contexts bias the results or deceive respondents about political candidates that actually exist. Consent was obtained and the respondents were compensated economically for participating in all studies.

Table A1, Part 1: Undemocratic behaviors in the included studies.

Study	Democratic behaviors	Undemocratic behaviors
1	Supported a proposal to preserve existing polling stations in all areas	Supported a proposal to reduce polling stations in areas that support opposing parties
	Said court rulings by judges appointed by opposing parties should be adhered to	Said court rulings by judges appointed by opposing parties should be ignored
	Said it is unacceptable to fight political opponents in the streets even though one feels provoked	Said it is legitimate to fight political opponents in the streets if one feels provoked
2	Said it is unacceptable to harass journalists even though they do not reveal sources	Said it is acceptable to harass journalists that do not reveal sources
	Supported a proposal to preserve existing polling stations in all areas	Supported a proposal to reduce polling stations in areas that support opposing parties
	Said court rulings by judges appointed by opposing parties should be adhered to	Said court rulings by judges appointed by opposing parties should be ignored
	Said it is unacceptable to fight political opponents in the streets even though one feels provoked	Said it is legitimate to fight political opponents in the streets if one feels provoked
	Said it is unacceptable to harass journalists even though they do not reveal sources	Said it is acceptable to harass journalists that do not reveal sources
	Said that we must trust election officials	Said that election officials cannot be trusted
	Said elected officials must obey the courts even when they think that the decisions are wrong	Said elected officials should not be bound by court decisions they regard as politicized
Said that use of force against certain groups should be avoided even if they retain information about crimes	Said that use of force against certain groups can be necessary if they retain information about crimes	
3	Make changes to the tax scheme in the country	[...]permanently close polling stations in areas not supporting the party
	Make changes to the level of welfare spending in the country	[...]put permanent restrictions on media channels critical of the party
		[...]purge existing public officials and replace them with party loyalists
4		[...]replace existing supreme court judges with party loyalists
	Supported a proposal to preserve existing polling stations in all areas	Supported a proposal to reduce polling stations in areas that support opposing parties
	Said the rights of media platforms and journalists should be protected	Said it is acceptable to harass journalists who do not reveal sources
		Proposed to permanently close polling stations in areas that support the opposing party
5		Proposed to put permanent restrictions on media channels critical of the party
	Proposed to preserve existing polling stations in all areas	Proposed to reduce polling stations in areas that support [Labour/the Conservatives]
	Said that vote buying is never acceptable	Said that vote buying is acceptable under some circumstances
	Said the rights of media platforms should be protected	Said we should put restrictions on media platforms supporting [Labour/the Conservatives]
	Said it is unacceptable to harass journalists even though they do not reveal sources	Said it is acceptable to harass journalists who do not reveal sources
	Said it is unacceptable to ght political opponents in the streets even though one feels provoked	Said it is legitimate to ght political opponents in the streets if one feels provoked
Proposed legislation protecting the Supreme Court from partisan inuence	Proposed to appoint Supreme Court justices loyal to [the Conservatives/Labour]	

Figure A1: Conjoint treatment example for Study 4 (US).

Imagine that the following two candidates run for the 2028 presidential election.

	CANDIDATE A	CANDIDATE B
Age	60	71
Gender	Female	Female
Party	Republican	Republican
Positions	<ul style="list-style-type: none"> - Increase income tax on 10 percent richest - Increase efforts to arrest and eventually deport illegal immigrants 	<ul style="list-style-type: none"> - Decrease power of labor unions - Make it harder for women to get an abortion
Recent Behavior	Proposed to permanently close polling stations in areas that support the opposing party	Said it is acceptable to harass journalists who do not reveal sources

Which candidate do you prefer?

Candidate A

Candidate B

Don't know

How likely would you be to vote for each of the candidates?

	Very unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewhat likely	Very likely	Don't know
Candidate A	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Candidate B	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure A2: Vignette treatment example for Study 3 (US).

Imagine that a candidate from your preferred party runs for the 2028 presidential election. The years before the election were characterized by normal economic conditions. However, both the public and experts voiced their fears of democratic breakdown, and national newspapers published articles with headlines like **"OUR DEMOCRACY IS VULNERABLE"** frequently.

One month before the election, the candidate from your preferred party proposes to purge existing public officials and replace them with party loyalists. Would you support this proposal?

Definitely not <input type="radio"/>	Probably not <input type="radio"/>	Might or might not <input type="radio"/>	Probably yes <input type="radio"/>	Definitely yes <input type="radio"/>	Don't know <input type="radio"/>
---	---------------------------------------	---	---------------------------------------	---	-------------------------------------

Would you vote for the candidate from your preferred party in this election?

Definitely not <input type="radio"/>	Probably not <input type="radio"/>	Might or might not <input type="radio"/>	Probably yes <input type="radio"/>	Definitely yes <input type="radio"/>	Don't know <input type="radio"/>
---	---------------------------------------	---	---------------------------------------	---	-------------------------------------

Would you vote for an other candidate in this election?

Definitely not <input type="radio"/>	Probably not <input type="radio"/>	Might or might not <input type="radio"/>	Probably yes <input type="radio"/>	Definitely yes <input type="radio"/>	Don't know <input type="radio"/>
---	---------------------------------------	---	---------------------------------------	---	-------------------------------------

Would you vote at all in this election?

Definitely not <input type="radio"/>	Probably not <input type="radio"/>	Might or might not <input type="radio"/>	Probably yes <input type="radio"/>	Definitely yes <input type="radio"/>	Don't know <input type="radio"/>
---	---------------------------------------	---	---------------------------------------	---	-------------------------------------

Considering the candidate's proposal to purge existing public officials and replace them with party loyalists, how would you feel about your preferred party afterwards?

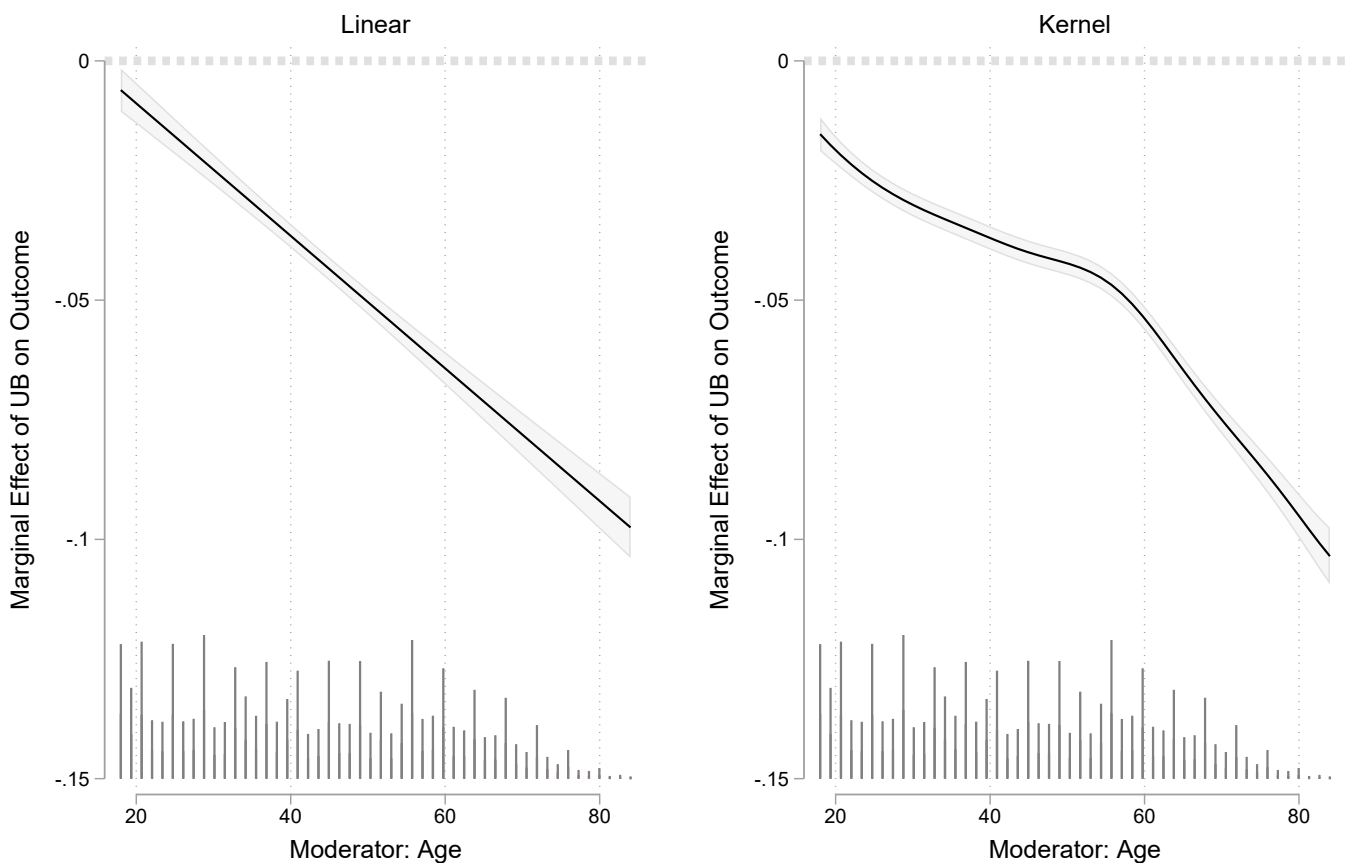
Much worse <input type="radio"/>	Somewhat worse <input type="radio"/>	About the same <input type="radio"/>	Somewhat better <input type="radio"/>	Much better <input type="radio"/>	Don't know <input type="radio"/>
-------------------------------------	---	---	--	--------------------------------------	-------------------------------------



Appendix B: Illustration of Pooled Fixed Effects Estimate

In this appendix, I provide the illustration of the pooled fixed effects estimate, paralleling the random effects estimates shown in Figure 2. Thus, Figure B1 illustrates the pooled linear effects and kernel estimate with country-year fixed effects. As Figure B1 shows, the fixed effects estimates are very similar to the random effects estimates shown in Figure 2. Punishment of undemocratic behavior increases with age, and the kernel estimate shows that this interaction is approximately linear.

Figure B1: Illustration of interaction between undemocratic behavior (UB) and age on voting intentions. Fixed effects.



Appendix C: Attentiveness and Effects of Other Factors

In this appendix, I test the possibility that mere survey attentiveness explains the age-differences in the effects of undemocratic behavior. Specifically, inattentiveness may attenuate the effects toward zero for young people (Ternovski and Orr 2022). First, I show the results when controlling for or conditioning on attentiveness. In case the results remain similar to the original findings, we should be less worried about such attenuation. Second, I show results by different response times and reduce the number of tasks in the included conjoint experiments (Study 1, 2, 4, and 5) to check how sensitive the findings are to excluding respondents possibly showing satisficing behavior or inattentiveness. Third, I show the effects of other candidate attributes and factors interacted by respondent age. We should also be less worried about attenuation of effects among young people if these other factors follow distinct patterns from that related to undemocratic behavior (e.g., if the effects of some factors decrease rather than increase with age).

Controlling for or Conditioning on Attentiveness

Figures C1-C4 show the interaction when controlling for or conditioning on attentiveness (i.e., showing the results for inattentive and attentive respondents separately). As mentioned in the letter, I measured attentiveness in all studies. The attentiveness measures consist of factual manipulation checks, which measure attentiveness to treatment specifically (Kane and Barabas 2019). Study 2, 4, and 5 included a question asking which job the candidates ran for with the correct answer being president (see also Figure A1), whereas the measures in Study 1 and 3 asked about other treatment features. Respondents were coded as attentive if they answered correctly and coded as inattentive if they answered incorrectly.

Figure C1: Interaction between undemocratic behavior and age across country-year samples and studies when controlling for attentiveness.

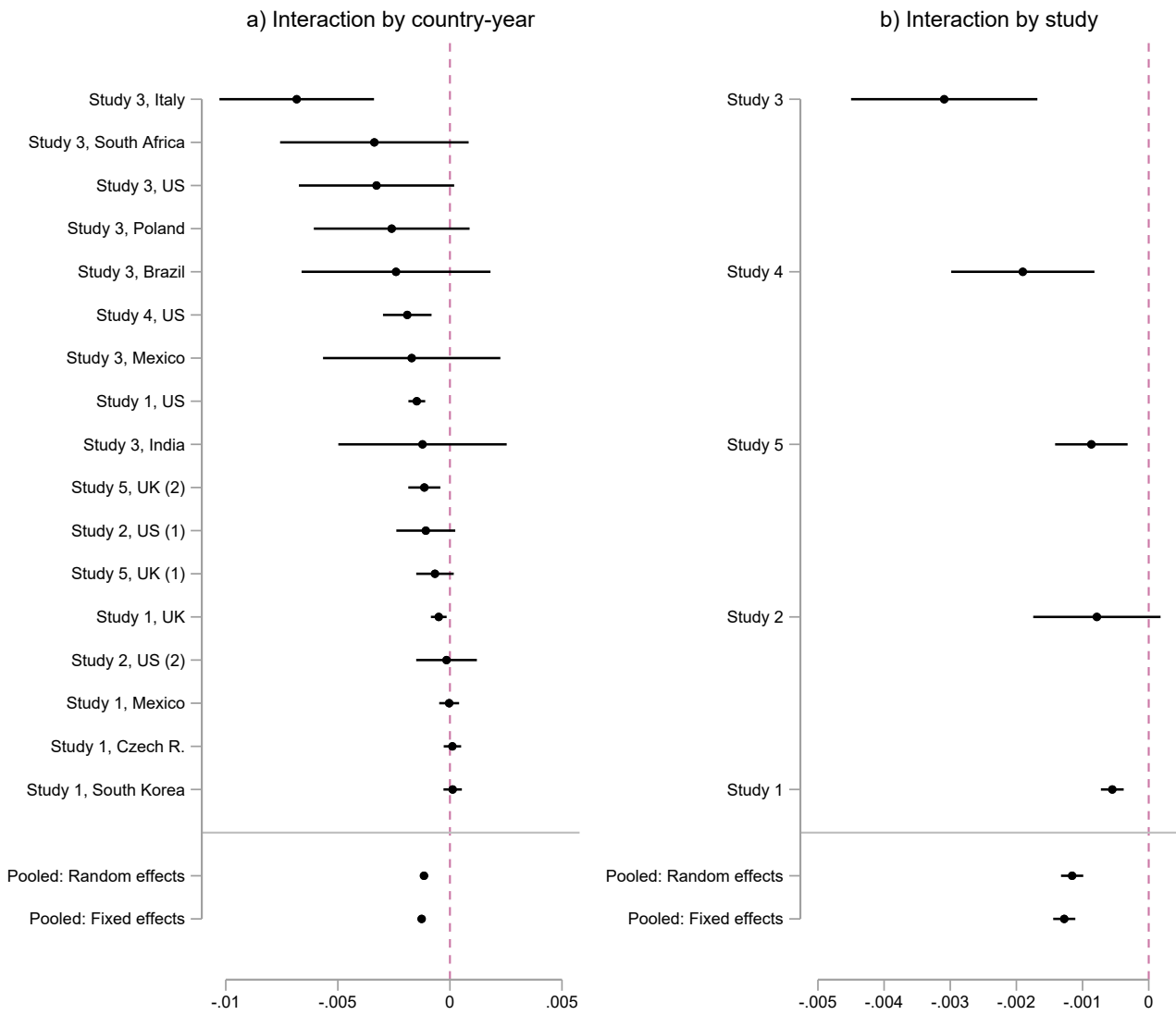


Figure C2: Interaction between undemocratic behavior and age across country-year samples and studies by attentiveness.

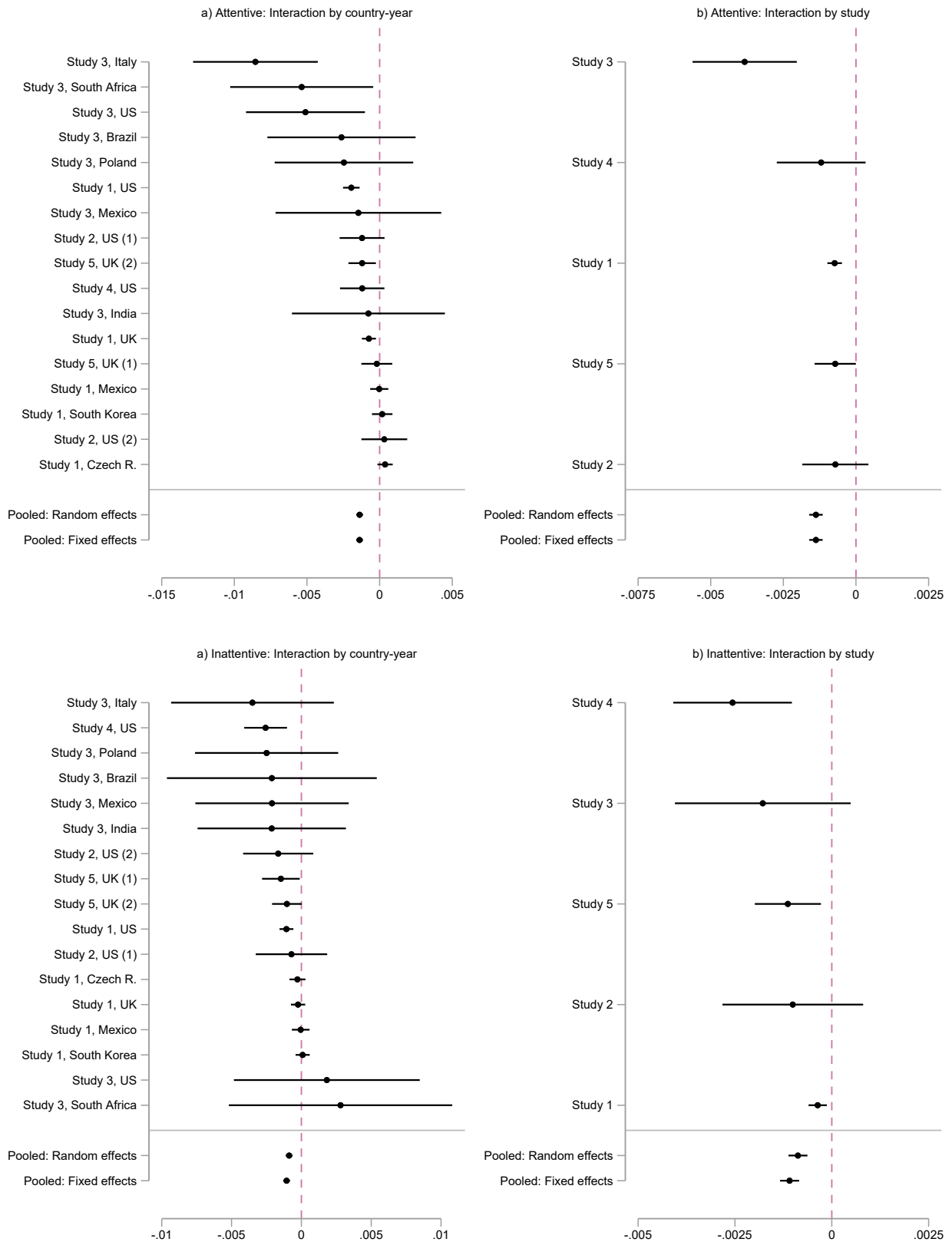


Figure C3: Illustration of interaction between undemocratic behavior and age when controlling for attentiveness.

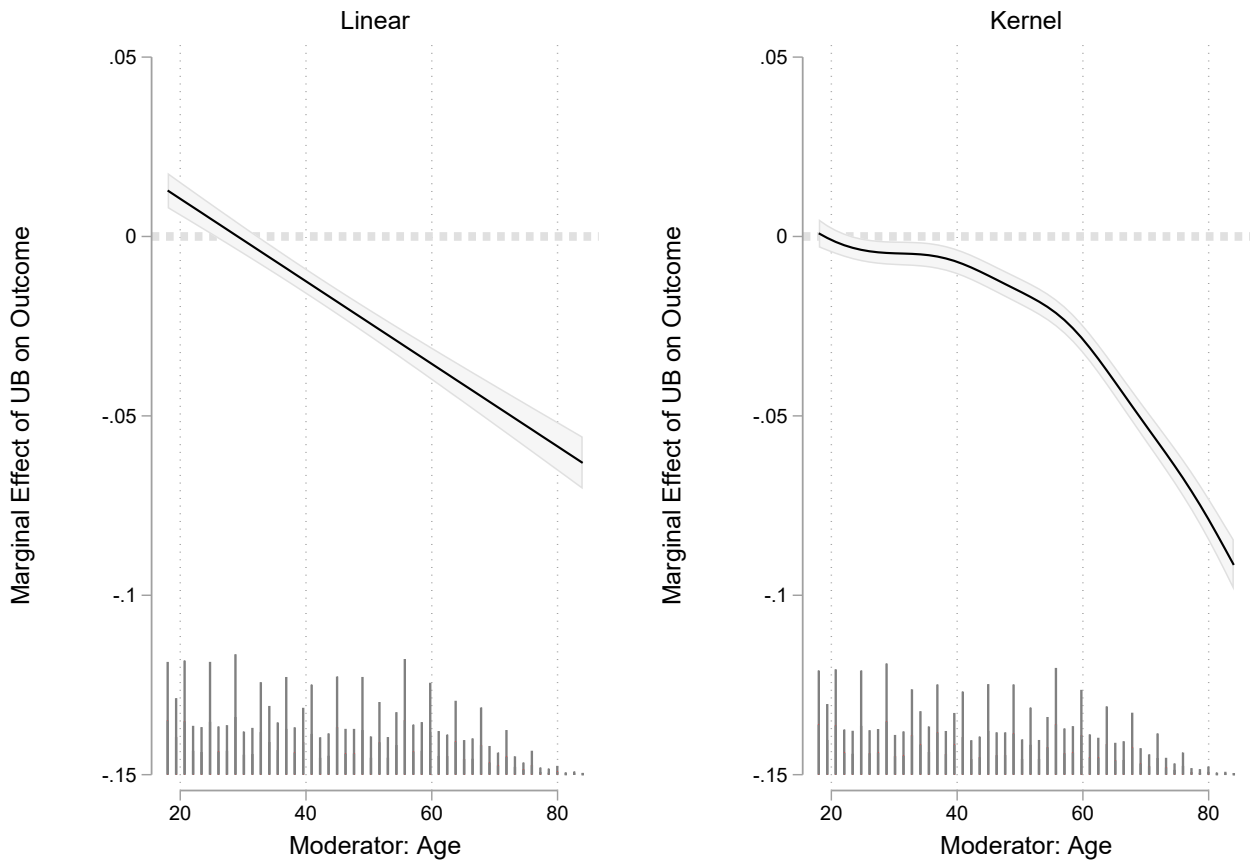


Figure C4: Illustration of interaction between undemocratic behavior and age by attentiveness.

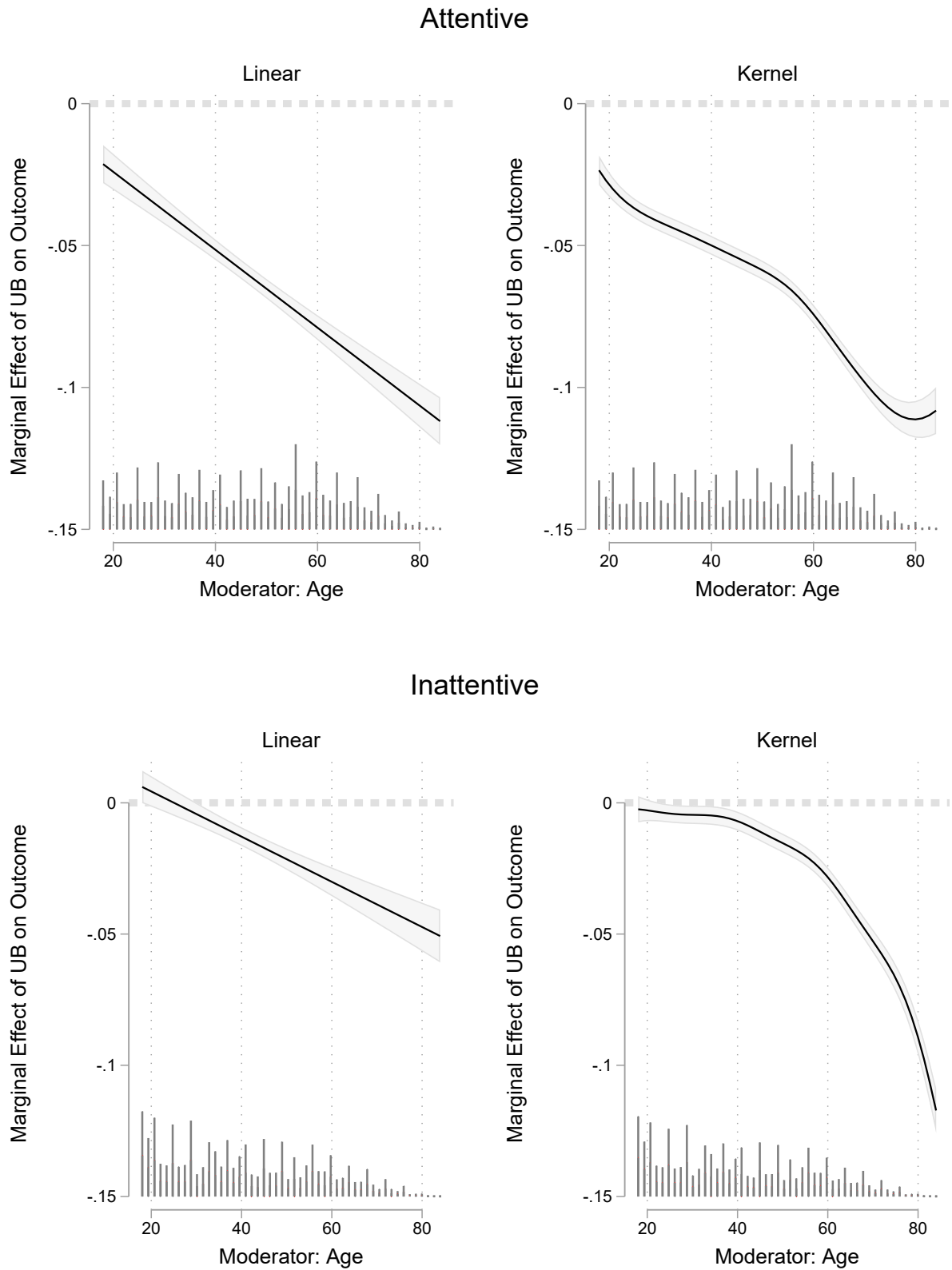
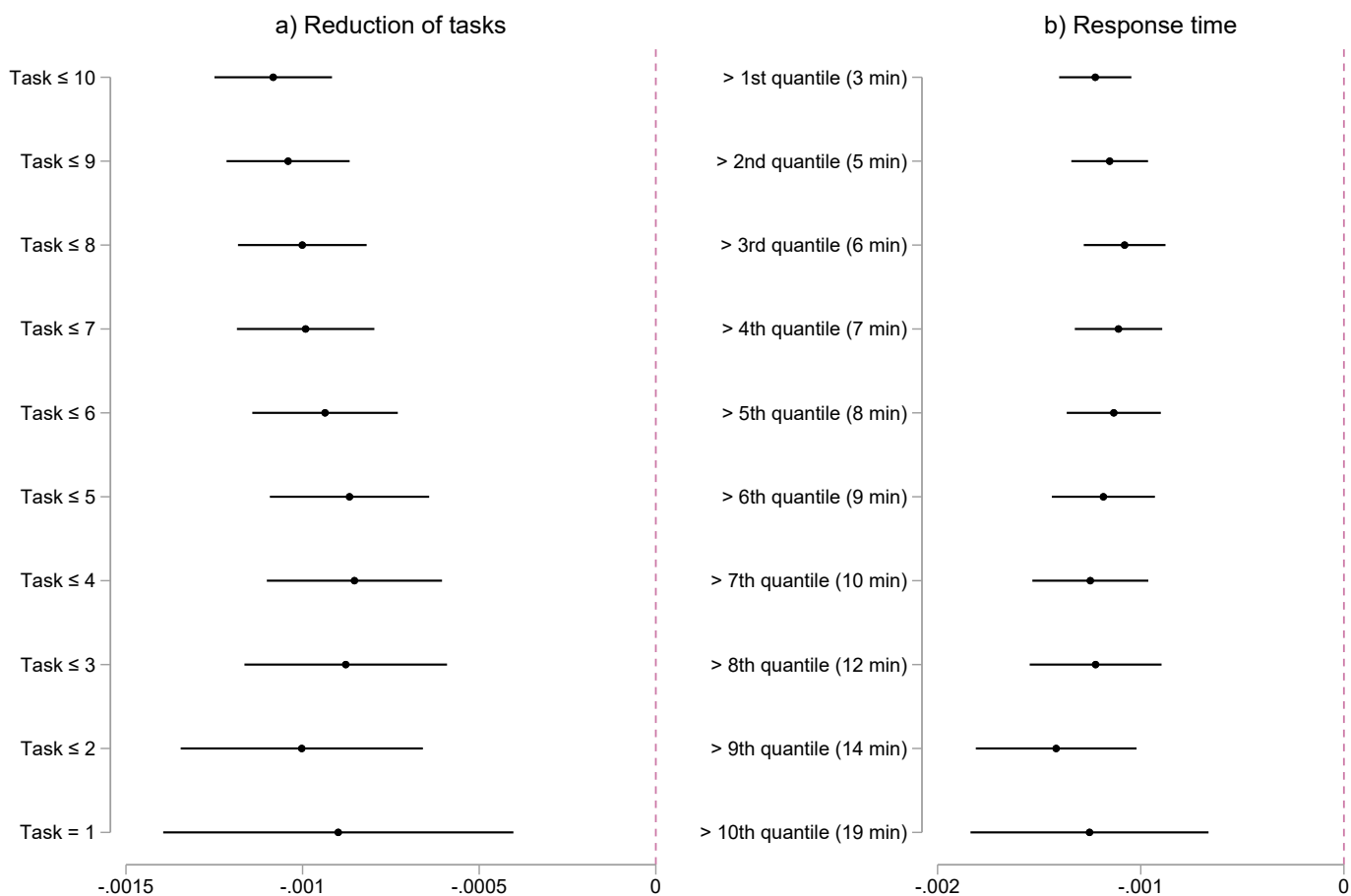


Figure C1 and C3 show that the results hold when controlling for attentiveness. As in the original results, the interaction between undemocratic behavior and age is statistically significant in 4 out of 5 studies and in the pooled estimates (Figure C1). Moreover, the graphical illustration remains similar to in the original results (Figure C3). Figure C2 and C4 show that the findings remain largely the same when conditioning on attentiveness as well. Three studies remain statistically significant regardless of whether we focus on attentive or inattentive respondents, and the pooled estimates remain significant statistically in all cases. The estimates are, of course, more uncertain when focusing on inattentive respondents, who are fewer in number, whereas the results for attentive respondents look like the original results the most. Accordingly, the graphical illustration for attentive respondents is more similar to the original results than the illustration for inattentive respondents (Figure C4), although both illustrations show that the effect of undemocratic behavior increases with age.

Reduction of Tasks and Response Time

Figure C5 shows the pooled results for Study 1, 2, 4, and 5 when the number of tasks is reduced—that is, shows results for the first task, the first two tasks, and so on—and by different response times. Study 3 is excluded, as this study only includes 1 task—which in itself supports the argument that a large amount of tasks did not produce the age-differences observed, as the tendency of young people to punish undemocratic behavior less was very strong in Study 3 (see also the original Figure 1).

Figure C5: Interaction between undemocratic behavior and age when limiting the number of tasks or restricting on response time in the pooled sample for Study 1, 2, 4, and 5.



I reduce the number of tasks gradually from 10 to 1 in panel a) of Figure C5. I start from 10 tasks, because this is the minimum total number of tasks in the included conjoint experiments (Studies 1-10 include 10 tasks, Study 4 consists of 15 tasks, and Study 5 includes 18 tasks). Panel a) shows that the findings are robust to limiting the number of tasks, as the interaction between undemocratic behavior and age remain strongly significant regardless of how many tasks are included. In other words, the results for the first couple of tasks are similar to the results including all tasks.

I split response times in 10 quantiles and gradually exclude quick-responding quantiles of respondents from the regressions in panel b) of Figure C5. For example, the first/upper estimate excludes only respondents using less than 3 minutes to complete the survey, whereas the last estimate only includes respondents using more than 19 minutes

to complete the survey. We see that excluding quick-responding respondents does not change the results, as the interaction between undemocratic behavior and age is strongly significant for all ten estimates. In sum, the findings are very robust to limiting the number of tasks or excluding respondents based on response times and—judging from these tests—therefore do not seem to be biased by satisficing behavior or inattentiveness.

Anti-Minority and -Redistribution Behavior

This section investigates age-differences in effects of other controversial behaviors such as anti-redistributive and anti-minority/nativist behaviors. As mentioned in the letter, I assigned such candidate behaviors in Study 1, 2, and 4. The anti-egalitarian behaviors consist of limiting redistribution of resources, whereas the anti-minority/nativist behaviors consist of rights-restrictions on immigrants, abortion, and gay marriage. The behaviors are assigned and coded similarly to undemocratic behaviors; that is, the measures are coded as 1 if the candidate respectively displays anti-redistributive or anti-minority/nativist behavior and coded as 0 if the candidate displays pro-redistributive or minority friendly behavior.

Figures C6-C7 show the interaction between anti-redistributive behavior and age, which is largely insignificant. Figure C6 shows that the interaction is significantly positive—such that the negative effect of anti-redistributive behavior decreases with age—in the two samples provided by Study 2, but otherwise insignificant. Figure C7 illustrates that the interaction is insignificant in the pooled sample, with the linear interaction being rather flat and the kernel estimator showing that the interaction fluctuates unsystematically as age increases.

Figures C8-C9 show the interaction between anti-minority or nativist behavior and age. Figure C8 shows that this interaction is significantly positive in *all* samples and studies, whereas Figure C9 corroborates this finding, as the effect is strongly negative among young people and goes toward zero as age increases. Punishment of anti-minority or nativist behavior thus runs counter to punishment of violations of core democratic principles; whereas punishment of the former decreases with age, punishment of the

latter increases. The distinct patterns related to the effects of anti-minority and anti-redistributive behaviors—as compared to undemocratic behavior—suggest that effects of candidate behaviors are not generally attenuated among young people.

Figure C6: Interaction between anti-redistribution and age across country-year samples and studies.

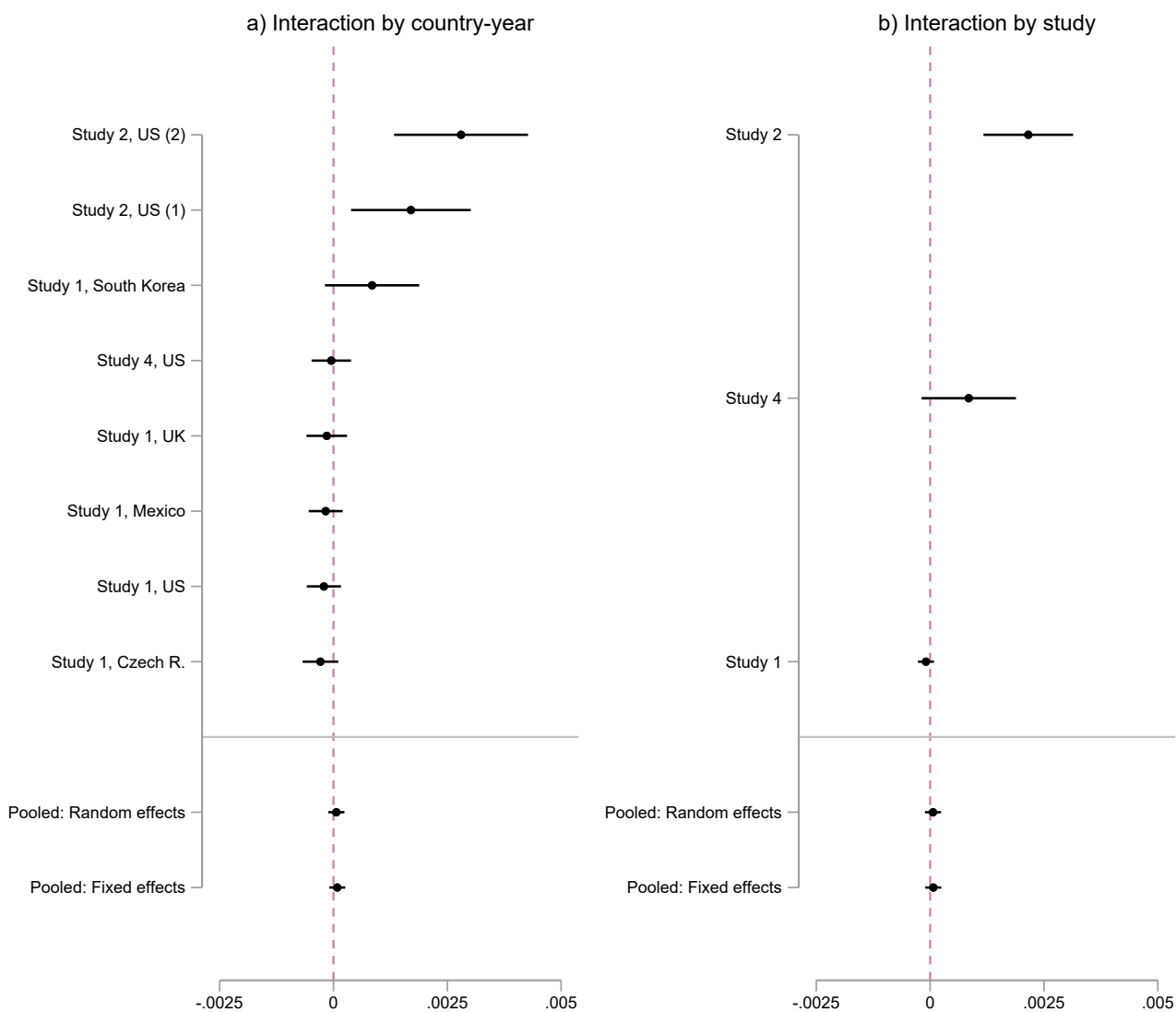


Figure C7: Illustration of interaction between anti-redistribution and age.

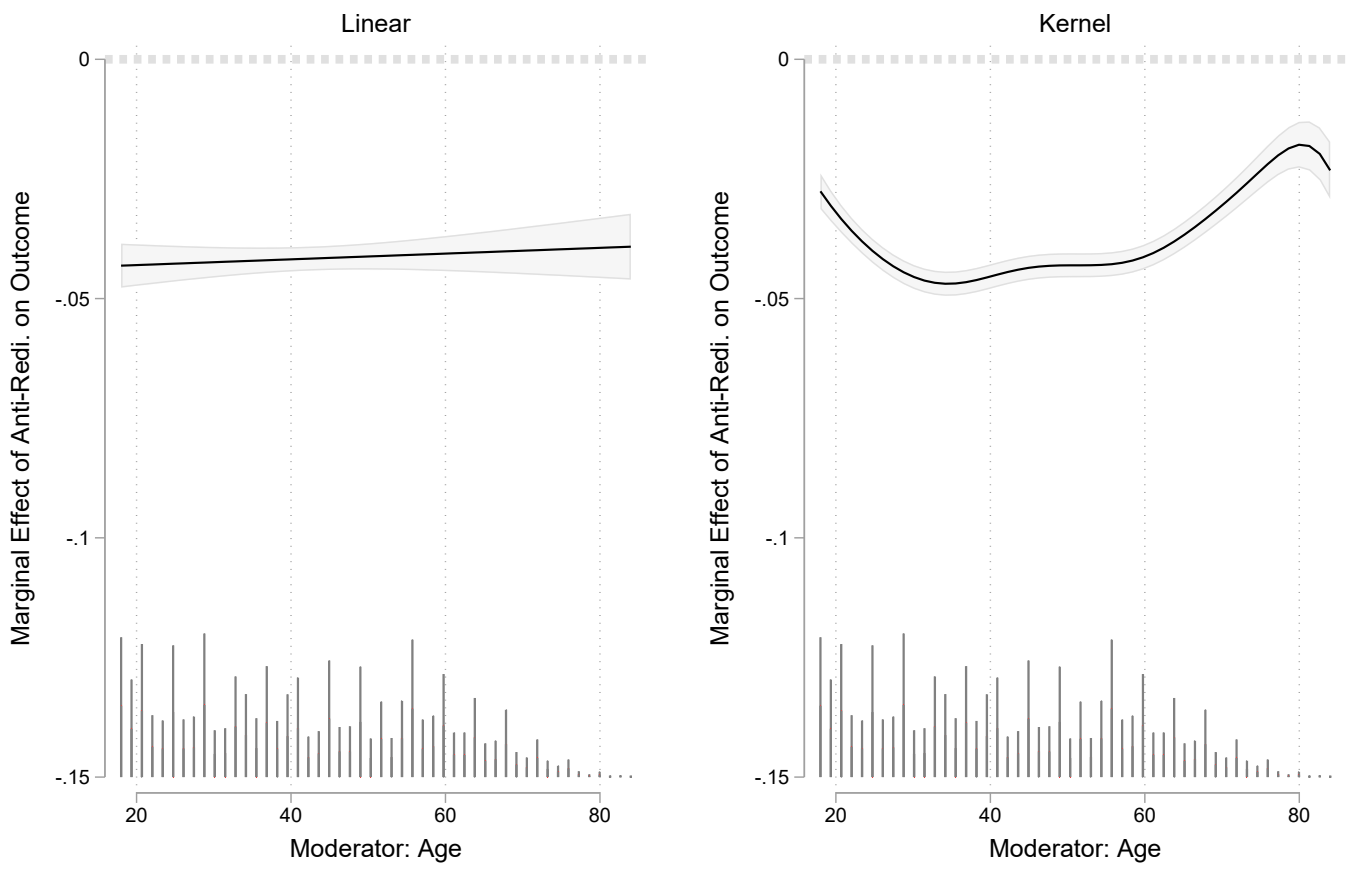


Figure C8: Interaction between anti-minority or nativist behavior and age across country-year samples and studies.

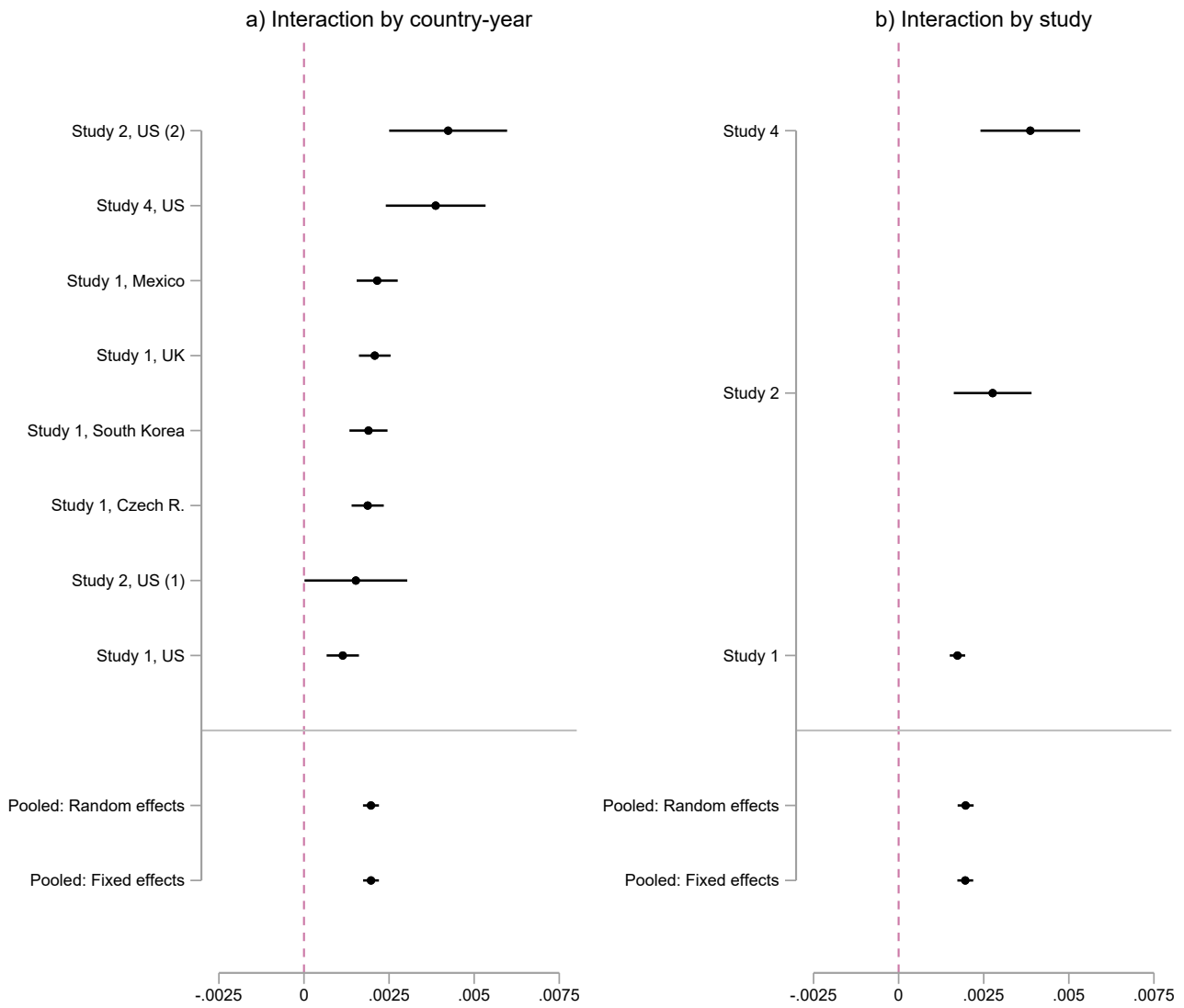
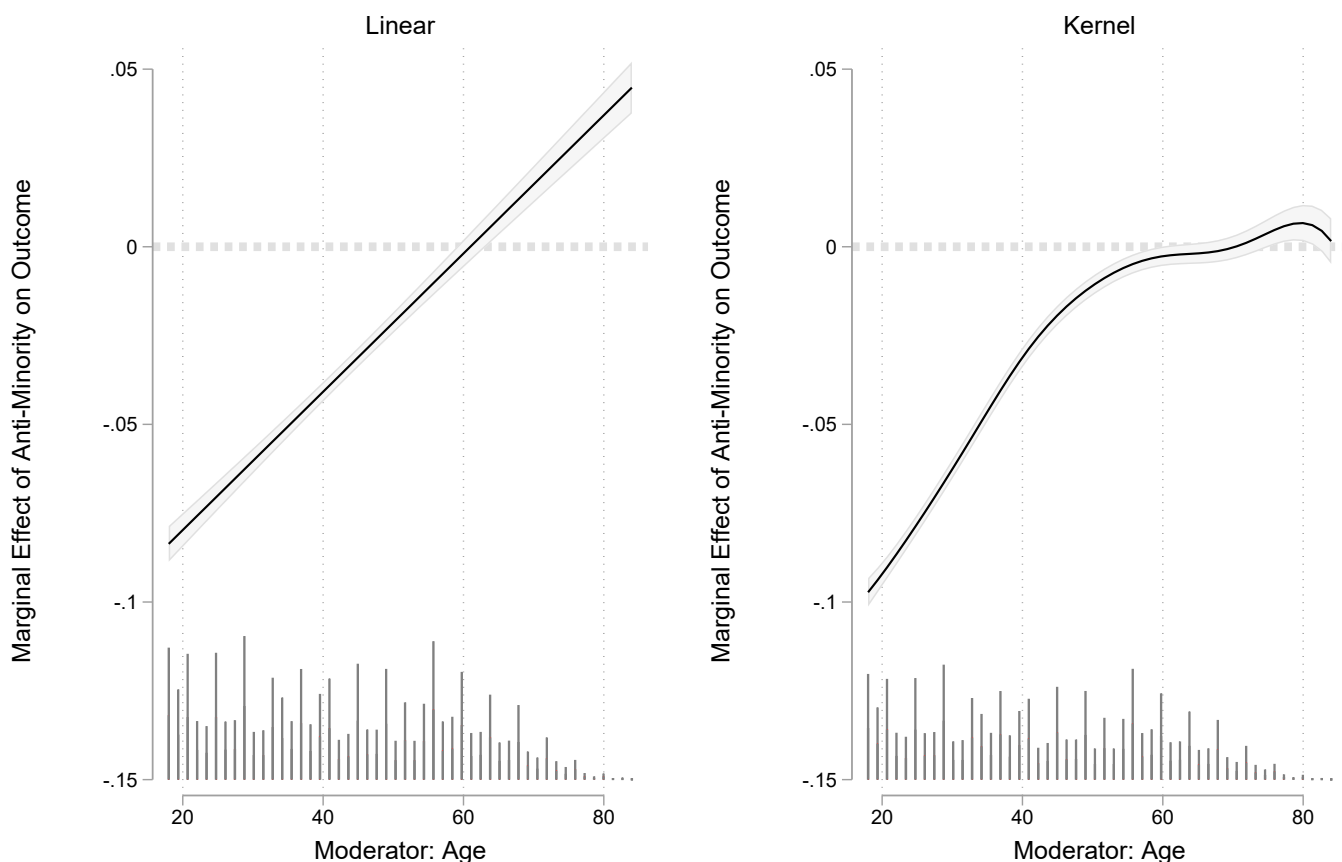


Figure C9: Illustration of interaction between anti-minority or nativist behavior and age on voting intentions.



Partisanship and Policy Distance

In this section, I estimate the effects of policy agreement and copartisanship between respondents and candidates interacted by age. I employ measures of policy distance between respondent and conjoint candidate profiles from Studies 1, 2, and 4, and measures of partisanship from Studies 1, 2, 4, and 5. The measures are scaled 0-1 so that 1 signals strong copartisanship and strong policy distance/disagreement. All measures except partisanship in Study 5, which is binary, are based on scales.

Figures C10-C13 show the interactions between copartisanship and age and distance and age, respectively. Figures C10 and C12 confirm the finding described in the letter that the effects of copartisanship increase with age. The tendency is more mixed for policy distance: Figure C11 and C13 show that the interaction related to policy distance

is heterogeneous across studies, with Study 2 yielding a significantly positive interaction and Study 1 and 4 yielding significantly negative interactions, and the pooled estimate being rather non-linear. The pattern for partisanship is therefore the only one following that of undemocratic behavior. In sum, this appendix shows that the results are robust to accounting for attentiveness and to putting restrictions on the number of tasks and response times and that the effects of several other factors follow patterns that are distinct from the age-differences in sanctioning of undemocratic behavior. Thus, differences in survey attentiveness do not seem to explain the results.

Figure C10: Interaction between copartisanship and age across country-year samples and studies.

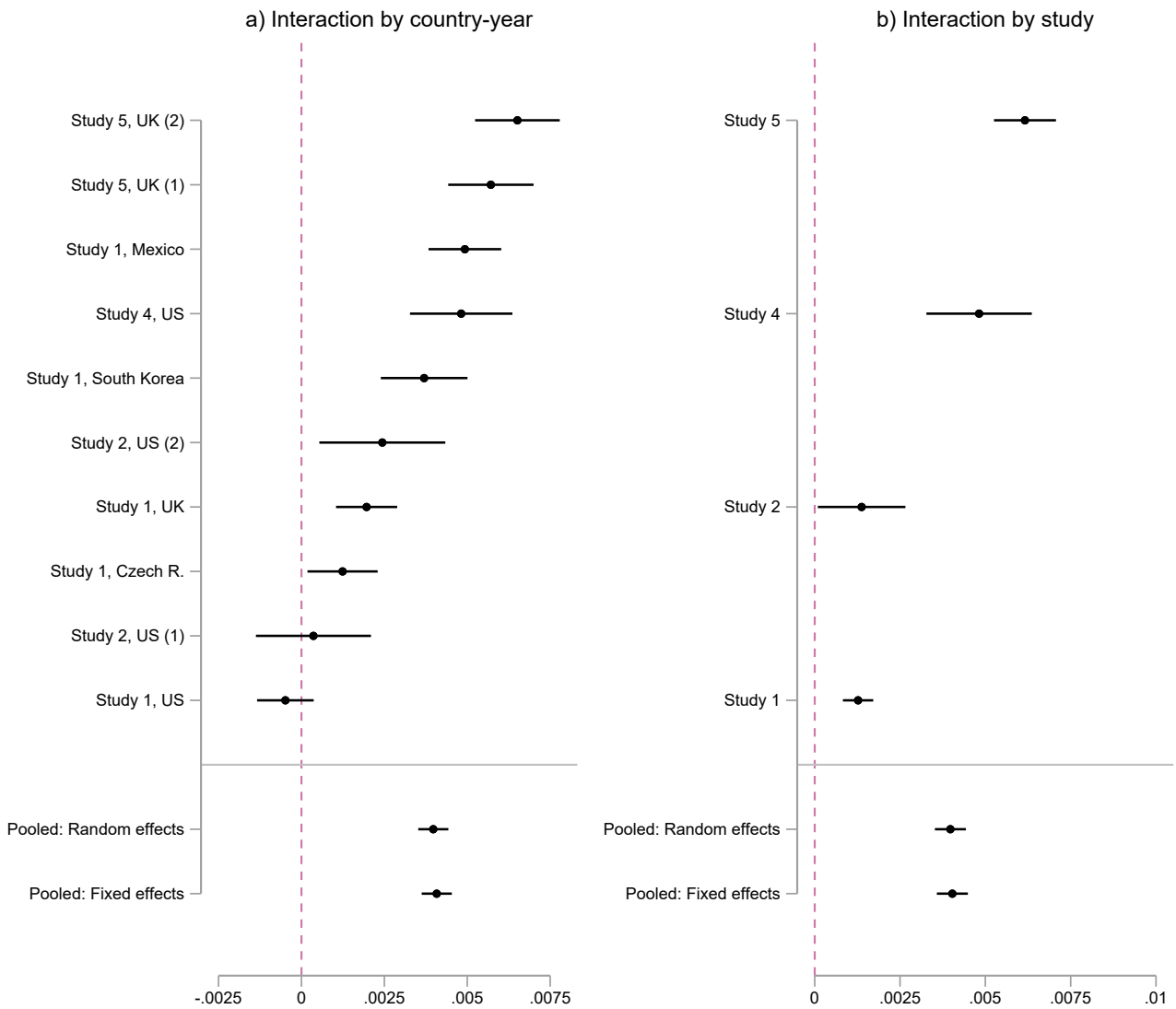


Figure C11: Interaction between policy distance and age across country-year samples and studies.

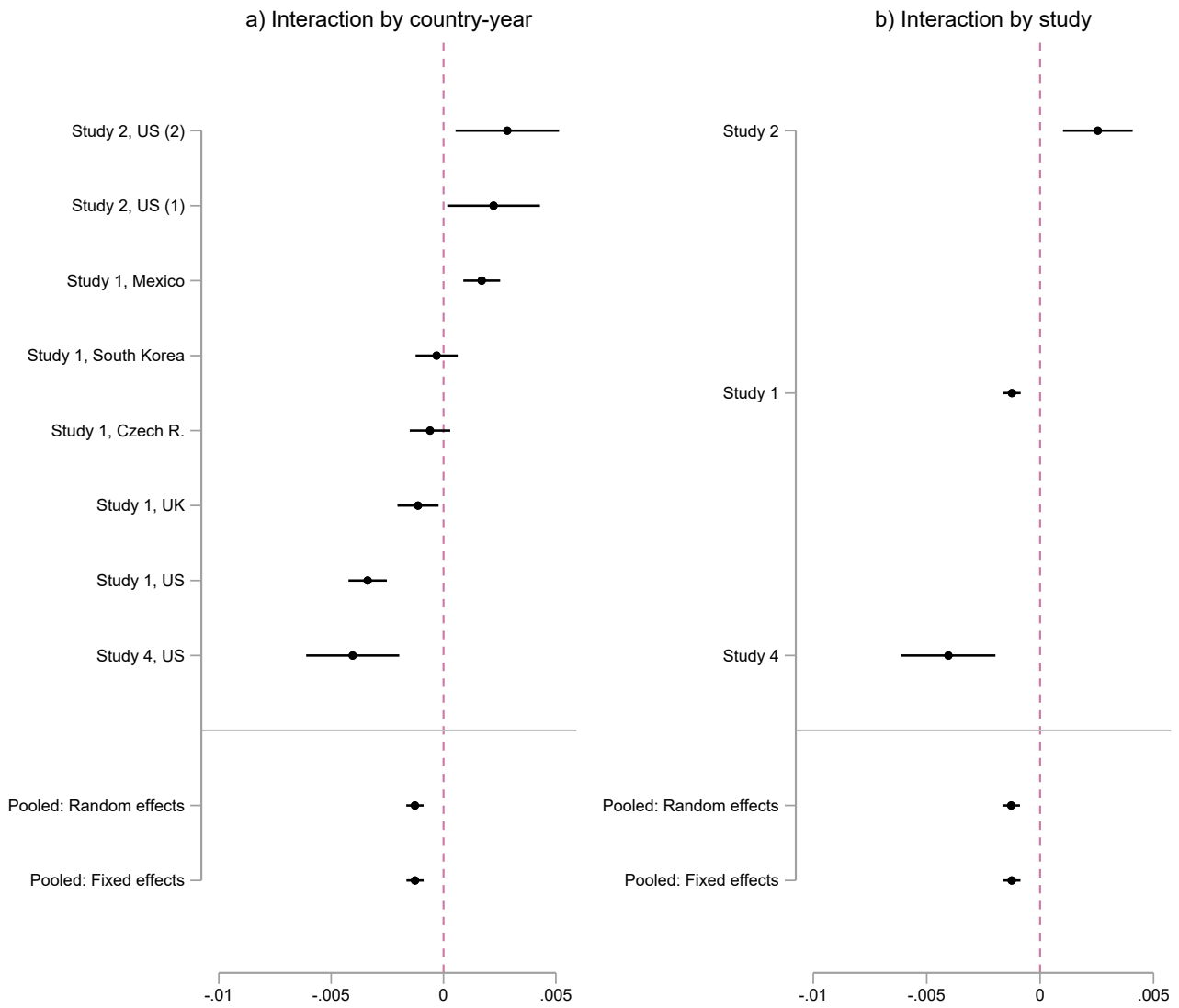


Figure C12: Illustration of interaction between copartisanship and age across country-year samples and studies.

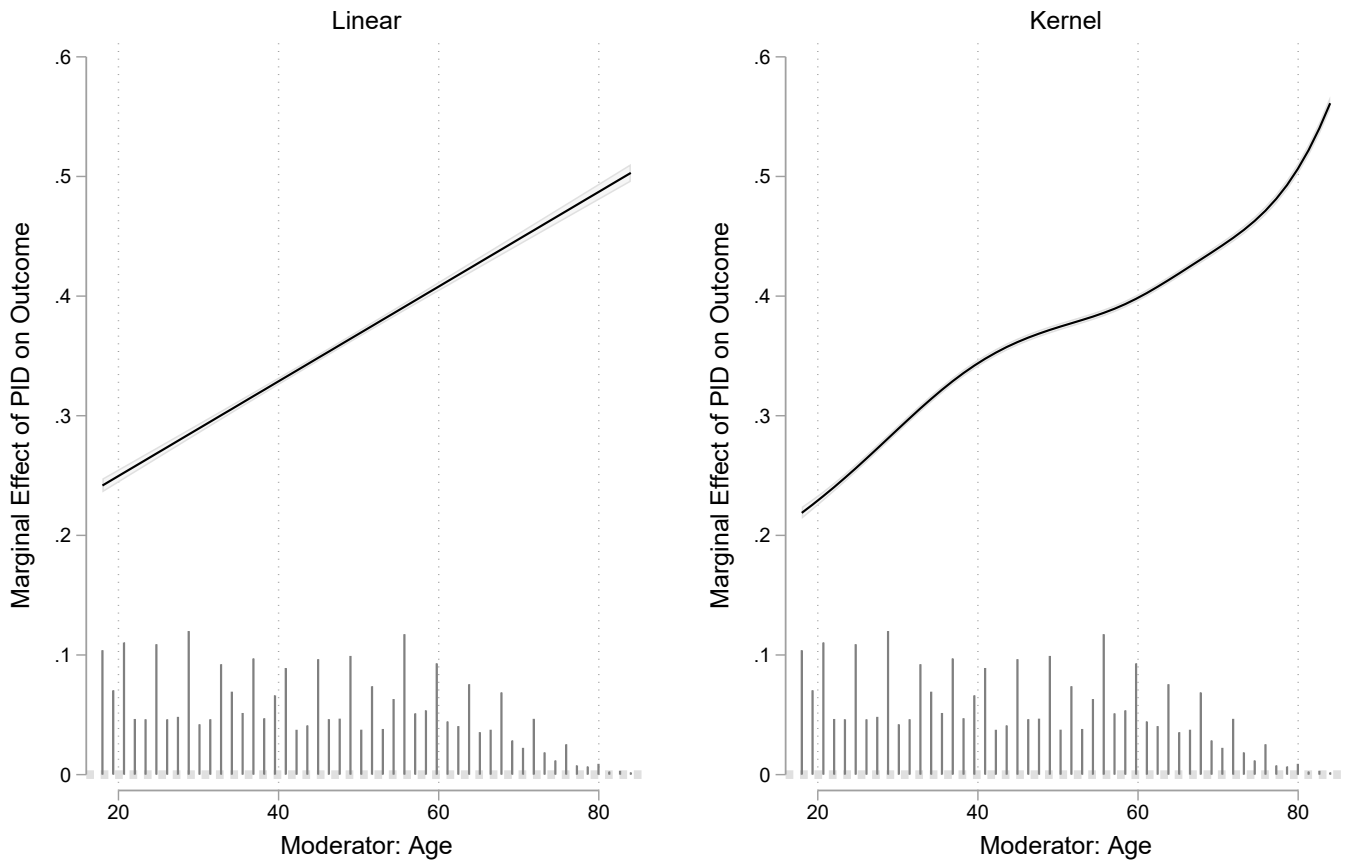
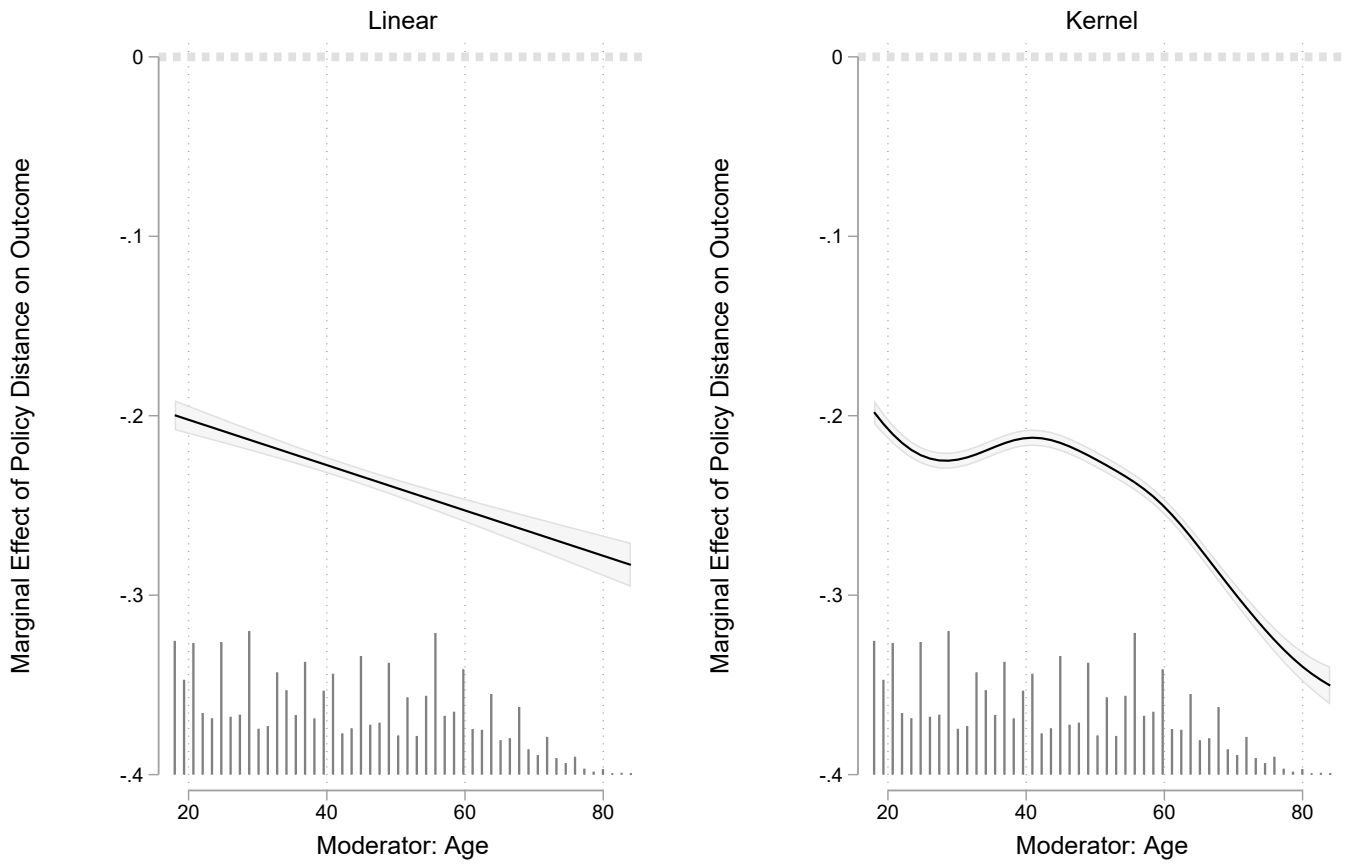


Figure C13: Illustration of interaction between policy distance and age across country-year samples and studies.



Appendix D: Additional Analyses Using Other Data Sources

In this appendix, I draw on alternative data sources and studies to explore any sensitivity of the findings related to the choice of source. As discussed in the main manuscript, Lucid and Amazon Mechanical Turk are challenged by distinct problems related to inattentiveness and professional survey taking. The selection criteria for these additional studies—beyond being published and having publicly available data—are that they contain random assignment of undemocratic behavior (still defined as violations of the core democratic principles of free elections, civil liberties, and checks and balances), a measure of respondent age, and an outcome measuring an evaluation of the undemocratic (or democratic) actor. This leads to the inclusion of Graham and Svulik (2020), Carey et al. (2022), Saikkonen and Christensen (2022), Lewandowsky and Jankowski (2023), Reuter and Szakonyi (2021), Mares and Visconti (2020), Krishnarajan (2023), and Aarslew (2023).

As Table D1 shows, these studies were conducted using different sources: YouGov (Carey et al. 2022; Aarslew 2023; Krishnarajan 2023), Qualtrics (Saikkonen and Christensen 2022), pre-pandemic Lucid (Graham and Svulik 2020)—which is crucial as inattentiveness on Lucid rose and remained high during 2020/the pandemic (Ternovski and Orr 2022)—the Russian Election Study (Reuter and Szakonyi 2021), Respondi (Lewandowsky and Jankowski 2023), and a random walk approach (Mares and Visconti 2020). Reuter and Szakonyi (2021) is thus based on a probability-based, representative survey, whereas the remaining studies—except Lewandowsky and Jankowski (2023), who focus on extreme voters—match nationally representative quotas on one or more demographic variables (as did the original Lucid samples). Additionally, these other platforms may be less challenged by inattentiveness than Lucid and less challenged by fraudulent/professional respondent behavior than MTurk (Ternovski and Orr 2022). Table D1 also sums up these studies in terms of countries included, year of fielding, design, number of tasks, and number of respondents.

Krishnarajan (2023) and Aarslew (2023) differ from the remaining studies by not having voting intentions as dependent variable. Instead, Krishnarajan (2023) and Aarslew (2023) measure perceived undemocraticness of policy proposals and post-election eval-

Table D1: Included studies using other data sources.

Study	Source	Countries	Years	Design	Tasks	Resp.
Graham and Svolic (2020)	Lucid	US	2018	Conjoint	16	1,691
Carey et al. (2022)	YouGov	US	2019	Conjoint	10	954
Aarslew (2023)	YouGov	Mexico and Denmark	2020/2021	Vignette	1	2,500
Krishnarajan (2023)	YouGov	22 countries	2021	Vignette	4	30,073
Saikkonen and Christensen (2022)	Qualtrics	Finland	2020	Conjoint	6	1,030
Lewandowsky and Jankowski (2023)	Respondi	Germany	2021	Conjoint	4	999
Mares and Visconti (2020)	Random walk approach	Romania	2016	Conjoint	5	502
Reuter and Szakonyi (2021)	Russian Election Study	Russia	2016	Vignette	1	2,010

uations of hypothetical governments, respectively. I included these studies due to their resemblance to the remaining experiments—to the best of my knowledge, no other studies with similar resemblance and publicly available data exist—and, in particular regarding Krishnarajan (2023), the cross-country scope of the data. Aarslew (2023) includes four separate outcomes, which I pool to a sum scale. Krishnarajan (2023) contains a categorical measure of age, which I recode by assigning the mean value of each interval.¹

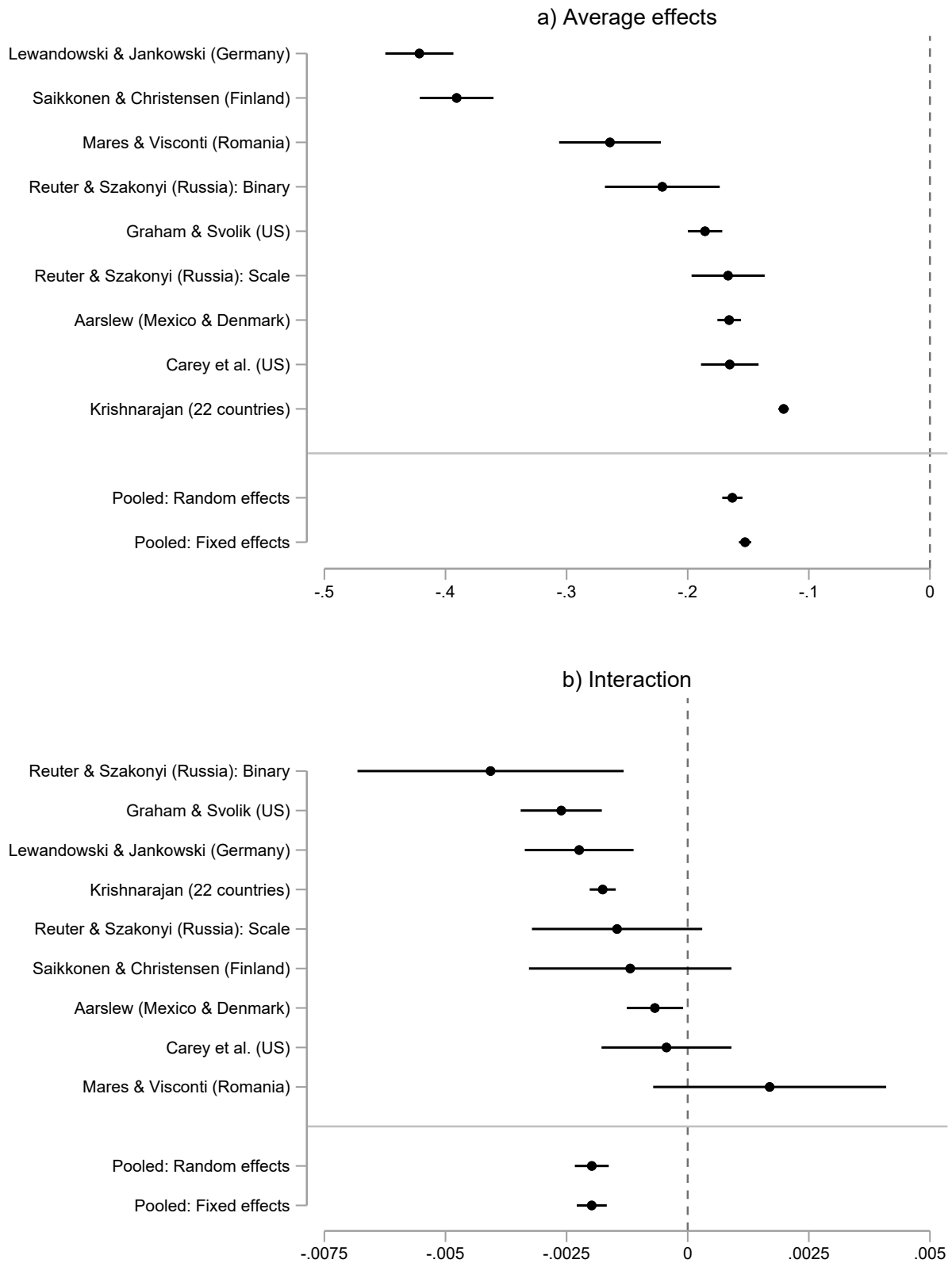
Carey et al. (2022), Saikkonen and Christensen (2022), and Mares and Visconti (2020) differ from the remaining studies by randomly assigning multiple undemocratic or democratic behaviors to each candidate. I conducted sum scales of undemocratic behavior from least (0) to most (1) undemocratic for these studies. Lewandowsky and Jankowski (2023) also assign multiple undemocratic behaviors but always jointly and always contrasted with a fully democratic candidate. Lewandowsky and Jankowski (2023) then measure the outcome as whether one votes for the undemocratic candidate (i.e., the data is structured on the task rather than the candidate level), wherefore I estimate the average effects of respondent age on this outcome for this study. Reuter and Szakonyi (2021) contains a binary as well as a scale outcome, and I include both below.² Each study resembles the original studies—which consisted of a binary indicator of undemocratic behavior, a measure of respondent age in years, and voting intentions as outcome—on aspects not mentioned above. I scale all outcomes and undemocratic behavior measures to 0-1 as in the original analysis.

Figure D1 shows the average effects of undemocratic behavior (upper panel) and interaction between undemocratic behavior and respondent age (lower panel) using these alternative data sources. The upper panel shows that the effects of undemocratic behavior, unsurprisingly, are negative in all studies. The lower panel shows that the negative interaction between undemocratic behavior and age is not sensitive to using Lucid or Me-

¹The intervals are 18-34, 35-54, and 55+. I code the latter interval as 65 years, because that is the mean age among respondents who are 55 years old or older in the two remaining YouGov studies (Aarslew 2023; Carey et al. 2022). Saikkonen and Christensen (2022) contains a similar, five-category measure of age based on decades, but as none of the categories are open-ended, I simply assign the mean value of each interval.

²I do not pool these, as they are both based on the scale outcome, and only the scale outcome adds to the pooled estimate. This provides a smaller estimate of the interaction between undemocratic behavior and age than if I only included the binary outcome or both in the pooled estimate.

Figure D1: Average effects of undemocratic behavior and interaction between undemocratic behavior and age across studies using other data sources (see also Table D1 above).



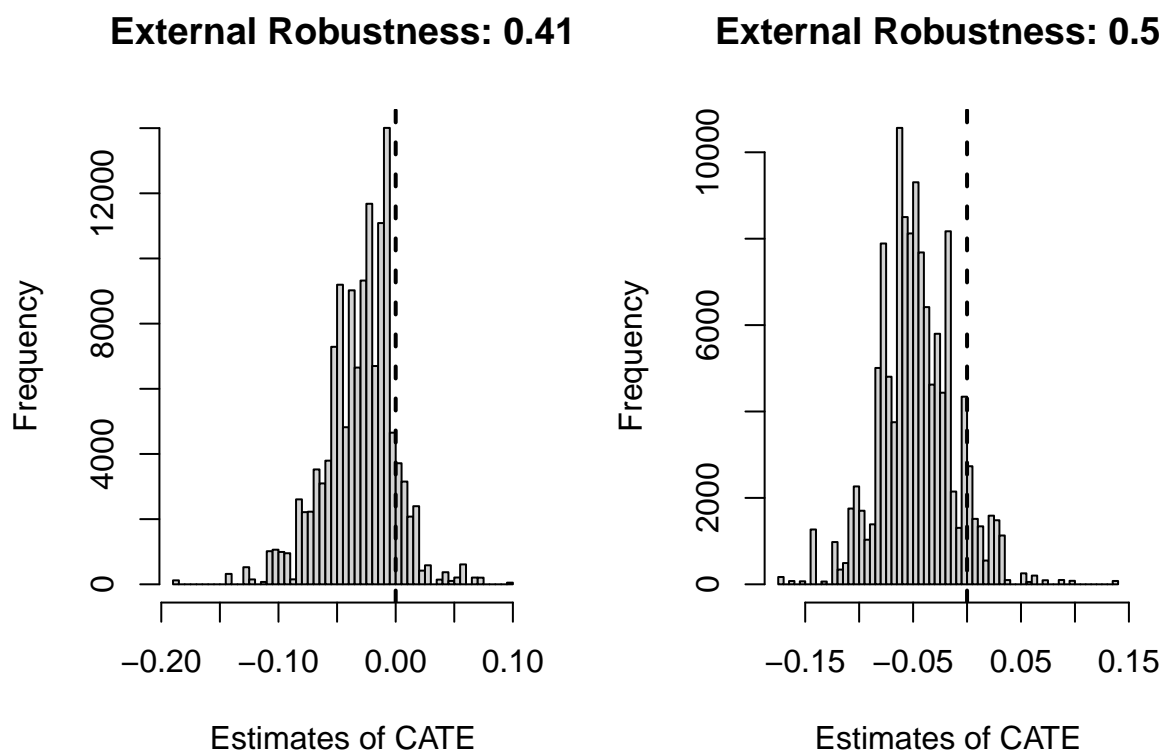
chanical Turk as survey platform. Recalling that six out of eight studies consist of single country-year samples (which often fail in reaching statistical significance, as the original Figure 1 also shows), all but one yield negative interaction coefficients. Five out of nine coefficients reach statistical significance, and the pooled estimates (-.002) are similar to or larger than the original pooled estimates (-.0017 to -.002). The two studies resembling the original studies the least (Aarslew 2023; Krishnarajan 2023) both yield coefficients smaller than the pooled estimate. The findings are therefore robust to using other data sources among which some are challenged less by fraudulent respondent behavior than MTurk and challenged less by inattentiveness than Lucid (Ternovski and Orr 2022).

Appendix E: Robustness to External Validity Bias

In this appendix, I quantify possible external validity bias arising from differences between the samples included and the populations of interest using the approach developed by Devaux and Egami (2022). I do this to address concerns related to lack of sample representativeness, which may lead to low external validity with regards to the participants of the study. By re-weighting the experimental samples according to observed covariates and showing how much conditional average treatment effects (CATEs) vary, the approach quantifies how different the populations of interest should be from the samples to explain away the population average treatment effect. This approach therefore does not require population data.

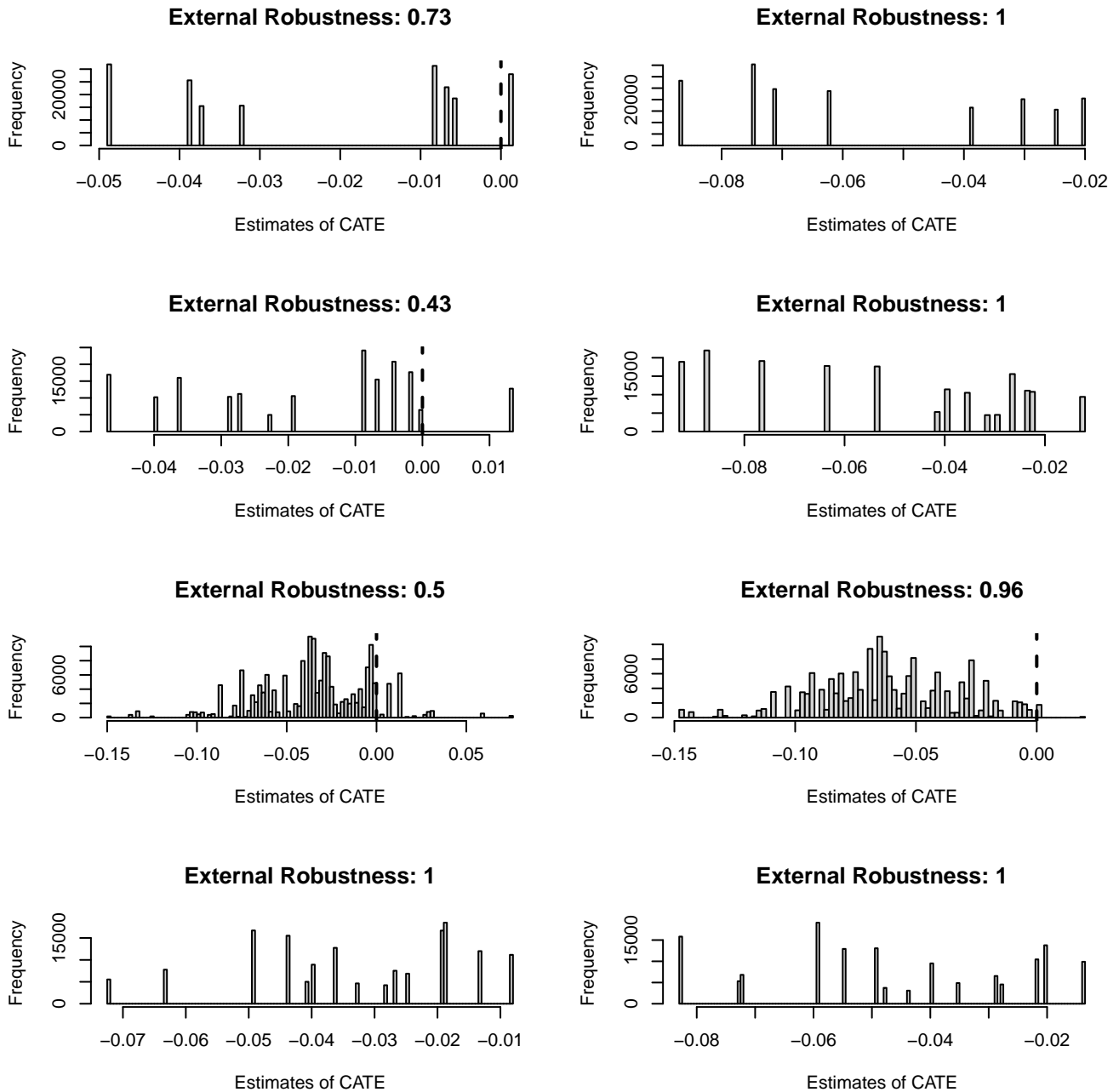
Across Studies 1-5, six covariates were each measured in at least four out of all five studies. These covariates are education [equivalent to high school or lower; equivalent to college or higher], partisanship [non-partisans; partisans], gender [female; male], residence [not living in a big city; living in a big city], attentiveness [attentive; inattentive], and political attitudes on a left/right scale. In the pooled sample, 58% have a college education or higher, 72% are partisans, 51% are female, 34% live in big cities, 54% are attentive, and the mean political attitude is 2.7 on a scale from 1 to 5 (i.e., slightly leftist). Residence was not measured in Study 2, partisanship was not measured in Study 5, and political attitudes were not measured in Study 3.

Figure E1: External robustness of the effects of undemocratic behavior for young people (left panel) and older people (right panel) when employing all six covariates (education, gender, residence, attentiveness, partisanship, and attitudes).



In Figures E1-E2, I show how externally robust the results are to differences between sample and population. As the statistical software (i.e., the *exr* package in R) requires treatments to be binary and does not permit interactions, I split the sample in two and estimate the effects of undemocratic behavior for young respondents below the sample mean of 44 years and older respondents above the sample mean. If the effects of undemocratic behavior in each of these groups are externally robust, our confidence in the main findings should also increase. Figure E1 plots external robustness for young people (left panel) and older people (right panel) when including all six covariates and thus excludes Studies 2, 3, and 5. Figure E2 plots external robustness when the number of covariates is limited to three to four and thus includes more studies. The first row includes the covariates attentiveness, education, and gender, the second row adds residence to these three covariates, the third row adds attitudes, and the fourth row adds partisanship. External robustness estimates below 0.14 are categorized as low robustness, estimates

Figure E2: External robustness of the effects of undemocratic behavior for young people (left column) and older people (right column) when employing three to four covariates. First row includes attention, education, and gender, whereas second row adds residence, third row adds attitudes, and fourth row adds partisanship.



above 0.57 are categorized as high robustness, and estimates in between 0.14 and 0.57 are categorized as moderate robustness. This is because 0.14 is the average amount of re-weighting required for national probability samples to resemble populations and 0.57 is the average amount of re-weighting required for MTurk samples to resemble populations (Devaux and Egami 2022).

We see that external robustness is quite high. When all six covariates are included, and Study 2, 3, and 5 thus are excluded (Figure E1), external robustness is moderate but quite close to the upper threshold of 0.57. When all studies and a limited number of covariates are included (first row of Figure E2), external robustness is very high (0.73-1).³ When one study is excluded and four covariates are included, external robustness varies between 0.43 and 1, and only falls below the upper threshold of 0.57 in two of six instances. The findings are therefore quite robust to populations that are rather different from the actual samples.

References

- Aarslew, Laurits F. 2023. “Why Don’t Partisans Sanction Electoral Malpractice?” *British Journal of Political Science* 53 (2): 407–423. doi:10.1017/S0007123422000126.
- Carey, John M., et al. 2019. “Searching for Bright Lines in the Trump Presidency”. *Perspectives on Politics* 17 (3): 699–718.
- Carey, John, et al. 2022. “Who will defend democracy? Evaluating tradeoffs in candidate support among partisan donors and voters”. *Journal of Elections, Public Opinion and Parties* 32 (1): 230–245. doi:10.1080/17457289.2020.1790577.
- Devaux, Martin, and Naoki Egami. 2022. “Quantifying Robustness to External Validity Bias”. Available at SSRN. <https://ssrn.com/abstract=4213753>.

³Part of the reason why external robustness reaches the maximum level of 1 several times—the other part of the reason plausibly being that external robustness is indeed high—is that some covariates contain only few categories, and few simulations therefore are performed.

- Graham, Matthew H., and Milan W. Svobik. 2020. "Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States". *American Political Science Review* 114 (2): 392–409.
- Kane, John V., and Jason Barabas. 2019. "No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments". *American Journal of Political Science* 63 (1): 234–249.
- Krishnarajan, Suthan. 2023. "Rationalizing Democracy: The Perceptual Bias and (Un)Democratic Behavior". *American Political Science Review* 117 (2): 474–496. doi:10.1017/S0003055422000806.
- Lewandowsky, Marcel, and Michael Jankowski. 2023. "Sympathy for the devil? Voter support for illiberal politicians". *European Political Science Review* 15 (1).
- Mares, Isabela, and Giancarlo Visconti. 2020. "Voting for the lesser evil: evidence from a conjoint experiment in Romania". *Political Science Research and Methods* 8 (2): 315–328.
- Reuter, Ora John, and David Szakonyi. 2021. "Electoral Manipulation and Regime Support: Survey Evidence from Russia". *World Politics* 73 (2): 275–314.
- Saikkonen, Inga A.-L., and Henrik Serup Christensen. 2022. "Guardians of Democracy or Passive Bystanders? A Conjoint Experiment on Elite Transgressions of Democratic Norms". *Political Research Quarterly* 0 (0): 10659129211073592. doi:10.1177/10659129211073592.
- Ternovski, John, and Lilla Orr. 2022. "A Note on Increases in Inattentive Online Survey-Takers Since 2020". *Journal of Quantitative Description: Digital Media* 2.