Do Candidates' Policy Positions Matter in Regional Elections? Evidence from the 2021 Elections to the Welsh Senedd

Online Appendix

Contents

1 Policy Statements

Table A1: MyVoteChoice issue statements

#	Issue statement	Theme	Less collin.
1	There should be no private sector involvement in the NHS in Wales	HSC	
2	The Welsh Government should scrap minimum unit price regulations for the sale of alcohol	HSC	X
3	Accident and emergency care should be centralised in larger hospitals even if some smaller hospitals lose these services	HSC	X
4	The restrictions on personal freedom imposed to combat COVID-19 were an over-reaction to the threat posed by the virus	HSC	
5	The Welsh government should seek close alignment with the UK government in its response to COVID-19	WLS	X
6	Individuals who refuse the COVID-19 vaccine should be subject to restrictions on their movements	HSC	X
7	The Welsh government should use its power to set income tax rates at a different level from England during the next term of Parliament	WLS	X
8	The state should actively redistribute from the richest people to the poorest	ECO	X
9	Welsh businesses should have lower tax rates	ECO	X
10	It should be easier for trans people to have their gender identity legally recognised in Wales	LCT	
11	Free school meals in Wales should be extended to all pupils whose families receive Universal Credit	EDU	X
12	Recreational use of cannabis by individuals should be decriminalised	LCT	X
13	Wales should become an independent country	WLS	X
14	Senedd Cymru (the Welsh Parliament) should be abolished	WLS	
15	More powers should be devolved from the UK government to Senedd Cymru (the Welsh Parliament)	WLS	
16	Brexit is good for Wales	LCT	
17	Wales should develop alternatives to the comprehensive school model, such as academies, free schools, and grammar schools	EDU	X
18	Welsh residents should receive a subsidy to undertake their higher education in Wales	EDU	X
19	Moving away from exam-based assessment for primary and secondary students will lower the standard of education in Wales	EDU	X
20	The Welsh Government should bring back the Right to Buy for housing association tenants in Wales	ECO	
21	Upper limits should be placed by the Welsh Government on rents charged by private landlords	ECO	X
22	Wales should continue to hold its rail operations in public ownership	ECO	
23	Cardiff airport should be returned to commercial (private) ownership	ECO	X
24	Welsh transport strategy should prioritise public transport over private vehicles	ENV	
25	The effects of human-made global warming have been exaggerated	ENV	
26	The Welsh government should allow the extraction of gas and petroleum by hydraulic fracturing (fracking)	ENV	
27	The Welsh government's promotion of the Welsh language has gone too far	LCT	
28	Immigration undermines the cultural values of Wales	LCT	

Note: This table provides the full list of issue statements that were included in the *MyVoteChoice* application. The theme provides the broad policy areas to which policy issues belong. The final column flags the 15 statements we retain in robustness check two, where we focus on statements for which we observe the highest levels of disagreement between candidates and their parties. ECO = economy; EDU = education; ENV = environment; HSC = health and social care; LCT = language and culture; WLS = Wales' relationship with the rest of the UK.

2 Data Cleaning

The aim of data cleaning is to ensure that our results are not unduly influenced by inputs from users who either fill in the questionnaire very rapidly without paying attention to the items or users who completed the questionnaire more than once. Therefore, we remove all of the following from our analyses: 1) all observations in which the time taken to complete the 28 issue statements of the VAA was less than 120 seconds; 2) all observations in which the time taken to respond to any one issue statements was one second or less; 3) all observations in which the time taken to respond to three or more issue statements was two seconds or less; 4) all observations in which the respondent answered ten successive issue statements in the same way; 5) all observations in which there are twenty or more "no opinion" responses to issue statements; 6) all observations that appear to be repeat attempts made on the same device (to identify these a cookie was installed). Overall, the cleaning leads us to drop 1,190 out of the totally 15,807 user records collected, or about 7.5%.

3 Robustness Checks

Table A2: Weighted regressions

	All issues			Personally salient issues		
	(1)	(2)	(3)	(4)	(5)	(6)
Proximity to candidate	0.53***	-0.03	-0.24	0.26***	0.11**	-0.20
	(0.03)	(0.14)	(0.30)	(0.02)	(0.04)	(0.16)
Proximity to party		0.59***	0.61***		0.40***	0.38***
		(0.14)	(0.15)		(0.10)	(0.10)
Party identification		0.54***	0.54***		0.55***	0.55***
		(0.03)	(0.03)		(0.03)	(0.03)
Previous vote choice		0.08***	0.08***		0.08***	0.07***
		(0.02)	(0.02)		(0.02)	(0.02)
p(candidate win)		0.02***	0.02***		0.02***	0.02***
		(0.00)	(0.00)		(0.00)	(0.00)
Proximity to candidate * interest in election			0.06			0.10*
			(0.09)			(0.05)
Candidate fixed effects	No	Yes	Yes	No	Yes	Yes
Voter fixed effects	No	Yes	Yes	No	Yes	Yes
Candidates	59	59	59	59	59	59
Voters	10481	7911	7772	9180	7099	6979
Observations	19035	14091	13843	16733	12683	12468

Note: The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with weighted linear regression adjusting the sample means of age, sex, and education to the population values. Standard errors clustered by voter in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A1: Marginal effects plots [weighted regressions]

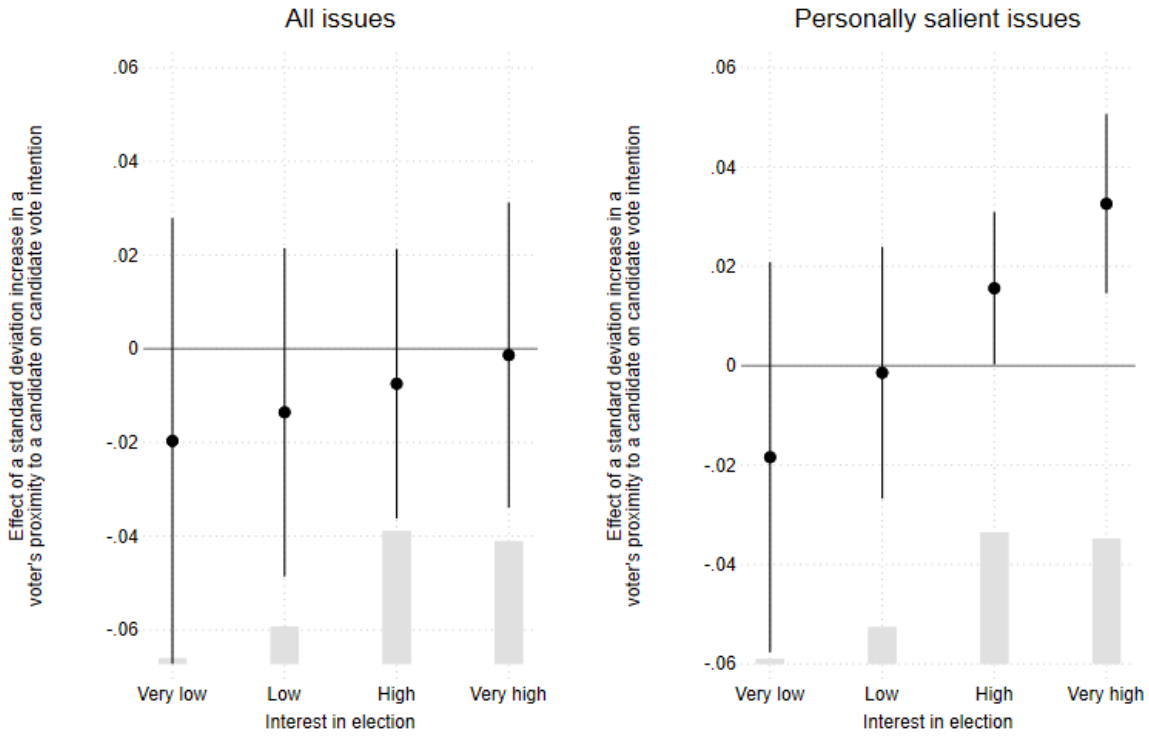


Table A3: Additional control variables

	All issues			Personally salient issues		
	(1)	(2)	(3)	(4)	(5)	(6)
Proximity to candidate	0.55***	0.13	0.02	0.26***	0.08*	-0.19
	(0.02)	(0.10)	(0.19)	(0.01)	(0.03)	(0.11)
Proximity to party (all issues)		0.32**	0.33**		0.39***	0.37***
		(0.11)	(0.11)		(0.07)	(0.07)
Proximity to party (personally salient)		0.08*	0.08*		0.03	0.03
		(0.03)	(0.03)		(0.04)	(0.04)
Party identification		0.55***	0.55***		0.55***	0.54***
		(0.02)	(0.02)		(0.02)	(0.02)
Previous vote choice		0.07***	0.07***		0.07***	0.07***
		(0.02)	(0.02)		(0.02)	(0.02)
p(candidate win)		0.02***	0.02***		0.02***	0.02***
		(0.00)	(0.00)		(0.00)	(0.00)
Proximity to candidate * interest in election			0.02			0.08**
			(0.04)			(0.03)
Interactions with key demographics	No	Yes	Yes	No	Yes	Yes
Candidate fixed effects	No	Yes	Yes	No	Yes	Yes
Voter fixed effects	No	Yes	Yes	No	Yes	Yes
Candidates	59	59	59	59	59	59
Voters	11580	7099	6979	10047	7099	6979
Observations	20997	12683	12468	18305	12683	12468

Note: In addition to the controls shown above, models 2, 3, 5, and 6 include multiplicative interactions between the candidate fixed effects and voters' age, gender, and educational attainment. The interactions account for the candidate-specific effects of key demographics. The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with linear regression. Standard errors clustered by voter in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A2: Marginal effects plots [additional controls]

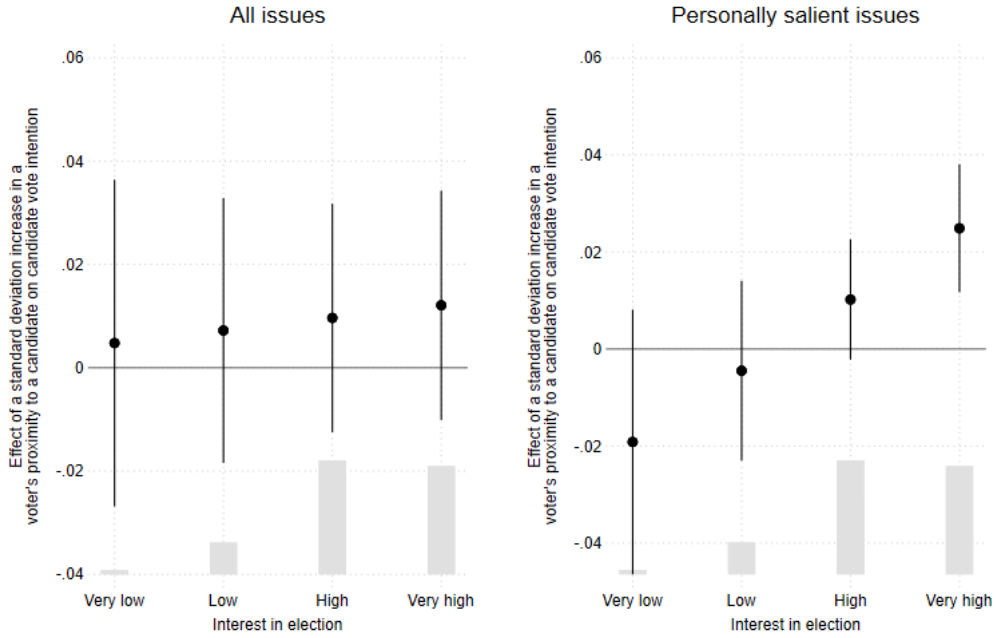


Table A4: Less collinear proximity measures

	All issues			Personally salient issues		
	(1)	(2)	(3)	(4)	(5)	(6)
Proximity to candidate	0.58***	0.11	-0.29	0.25***	0.08**	-0.27**
	(0.02)	(0.07)	(0.15)	(0.01)	(0.03)	(0.09)
Proximity to party		0.36***	0.38***		0.40***	0.39***
		(0.06)	(0.06)		(0.05)	(0.05)
Party identification		0.54***	0.54***		0.55***	0.54***
		(0.02)	(0.02)		(0.02)	(0.02)
Previous vote choice		0.06***	0.06***		0.05**	0.05**
		(0.01)	(0.01)		(0.02)	(0.02)
p(candidate win)		0.02***	0.02***		0.02***	0.02***
		(0.00)	(0.00)		(0.00)	(0.00)
Proximity to candidate * interest in election			0.12**			0.11***
			(0.04)			(0.03)
Candidate fixed effects	No	Yes	Yes	No	Yes	Yes
Voter fixed effects	No	Yes	Yes	No	Yes	Yes
Candidates	59	59	59	59	59	59
Voters	11580	8513	8288	8466	6408	6240
Observations	20997	15163	14758	15388	11421	11122

Note: The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with linear regression. Standard errors clustered by voter in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A3: Marginal effects plots [less collinear proximity measures]

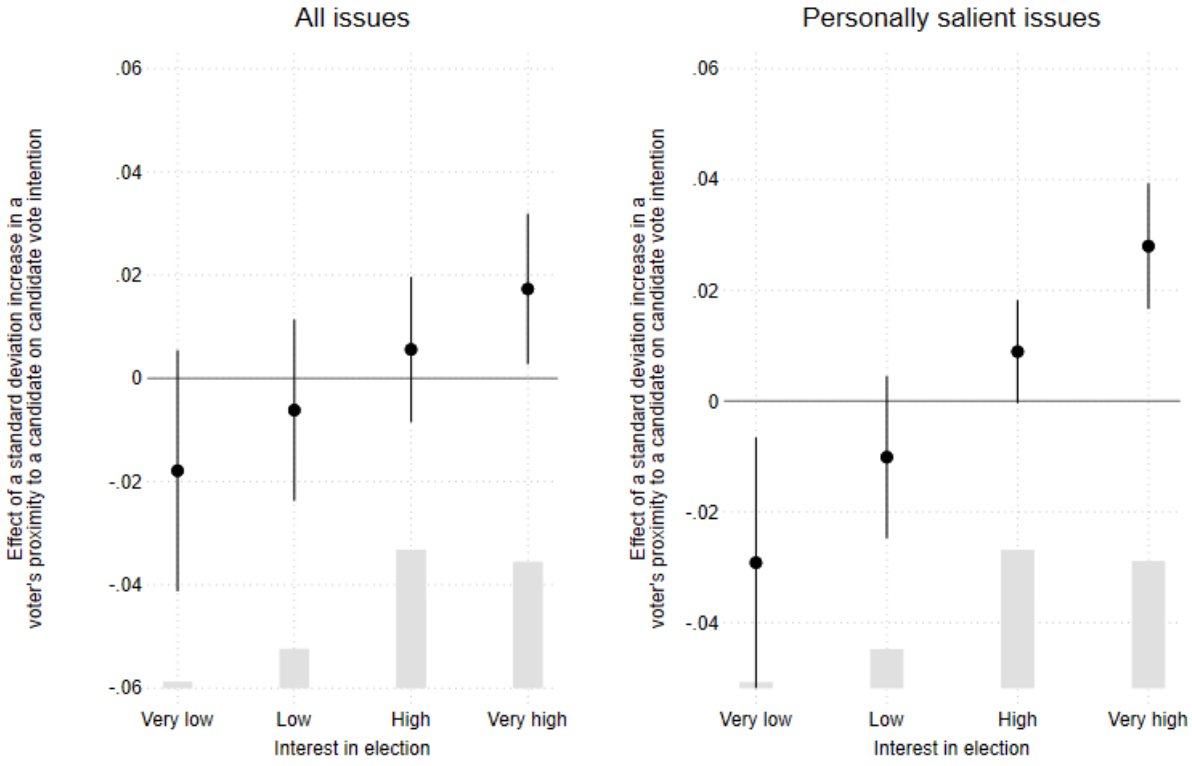


Table A5: Two-way clustered standard errors

	All issues			Personally salient issues		
	(1)	(2)	(3)	(4)	(5)	(6)
Proximity to candidate	0.55***	0.03	-0.05	0.26***	0.09*	-0.20
	(0.09)	(0.12)	(0.19)	(0.04)	(0.04)	(0.12)
Proximity to party		0.53***	0.55***		0.42***	0.40***
		(0.11)	(0.11)		(0.07)	(0.08)
Party identification		0.55***	0.54***		0.55***	0.55***
		(0.02)	(0.02)		(0.02)	(0.02)
Previous vote choice		0.06**	0.06**		0.06**	0.06**
		(0.02)	(0.02)		(0.02)	(0.02)
p(candidate win)		0.02***	0.02***		0.02***	0.02***
		(0.00)	(0.00)		(0.00)	(0.00)
Proximity to candidate * interest in election			0.02			0.09*
			(0.05)			(0.04)
Candidate fixed effects	No	Yes	Yes	No	Yes	Yes
Voter fixed effects	No	Yes	Yes	No	Yes	Yes
Candidates	59	59	59	59	59	59
Voters	59	59	59	59	59	59
Observations	20997	15163	14758	18305	13624	13273

Note: The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with linear regression. Standard errors clustered by voter **and** candidate in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A4: Marginal effects plots [two-way clustered standard errors]

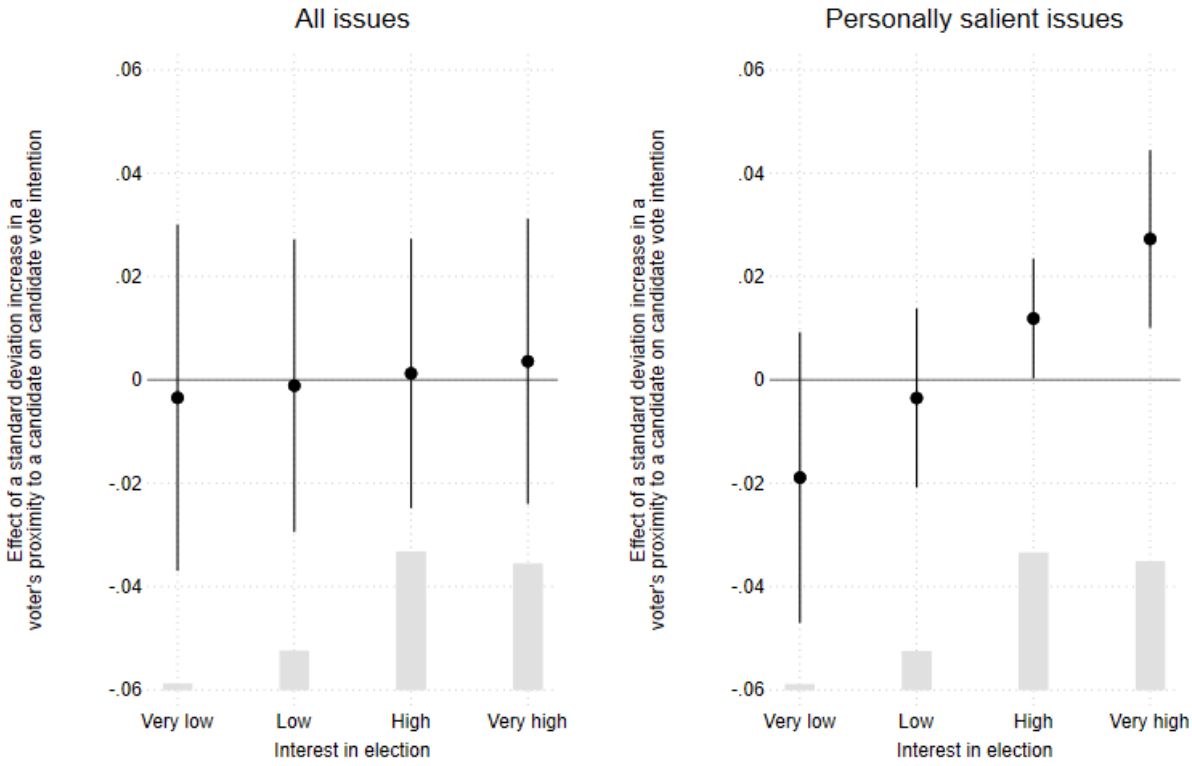


Table A6: Block-bootstrapped standard errors

	All issues			Personally salient issues		
	(1)	(2)	(3)	(4)	(5)	(6)
Proximity to candidate	0.55***	0.03	-0.05	0.26***	0.09**	-0.20*
	(0.02)	(0.09)	(0.17)	(0.01)	(0.03)	(0.10)
Proximity to party		0.53***	0.55***		0.42***	0.40***
		(0.09)	(0.09)		(0.06)	(0.06)
Party identification		0.55***	0.54***		0.55***	0.55***
		(0.02)	(0.02)		(0.02)	(0.02)
Previous vote choice		0.06***	0.06***		0.06***	0.06***
		(0.01)	(0.01)		(0.01)	(0.02)
p(candidate win)		0.02***	0.02***		0.02***	0.02***
		(0.00)	(0.00)		(0.00)	(0.00)
Proximity to candidate * interest in election			0.02			0.09**
			(0.04)			(0.03)
Candidate fixed effects	No	Yes	Yes	No	Yes	Yes
Voter fixed effects	No	Yes	Yes	No	Yes	Yes
Candidates	59	59	59	59	59	59
Voters	11580	8513	8288	10047	7622	7428
Observations	20997	15163	14758	18305	13624	13273

Note: The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with linear regression. Standard errors (in parentheses) are estimated using block-bootstrap (by voter) with 1,000 replications. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A5: Marginal effects plots [block-bootstrapped standard errors]

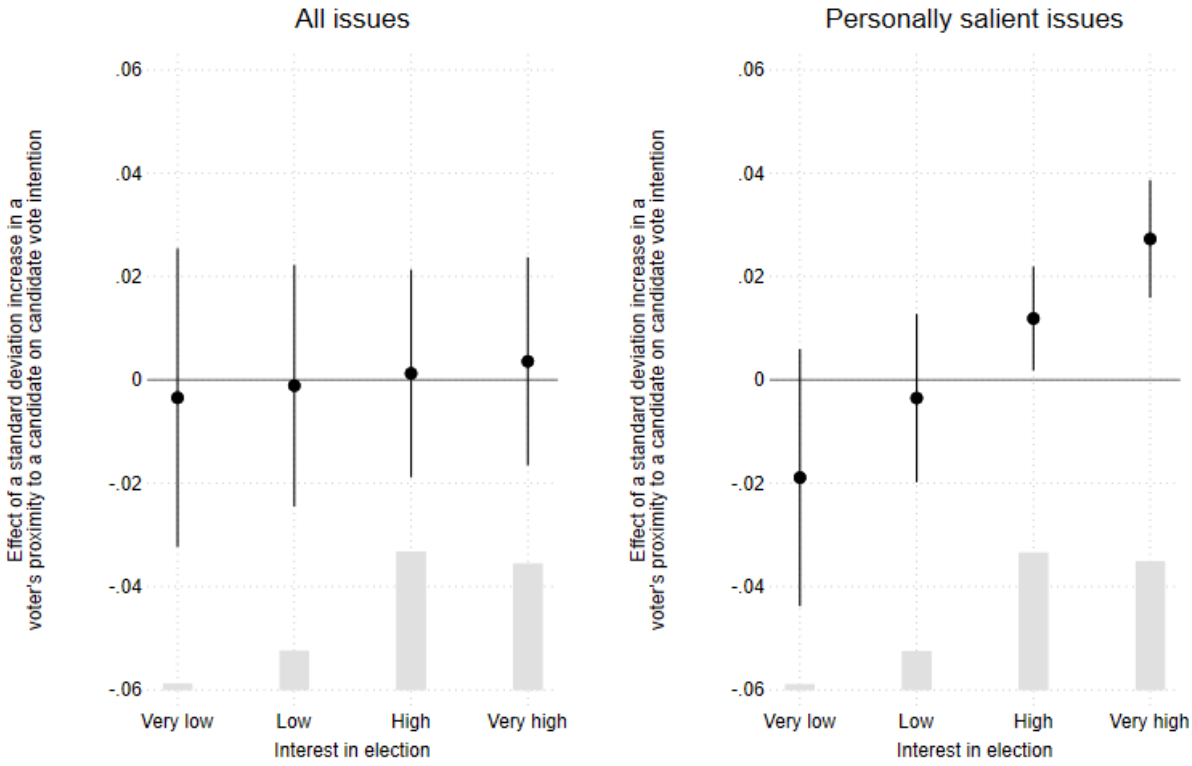
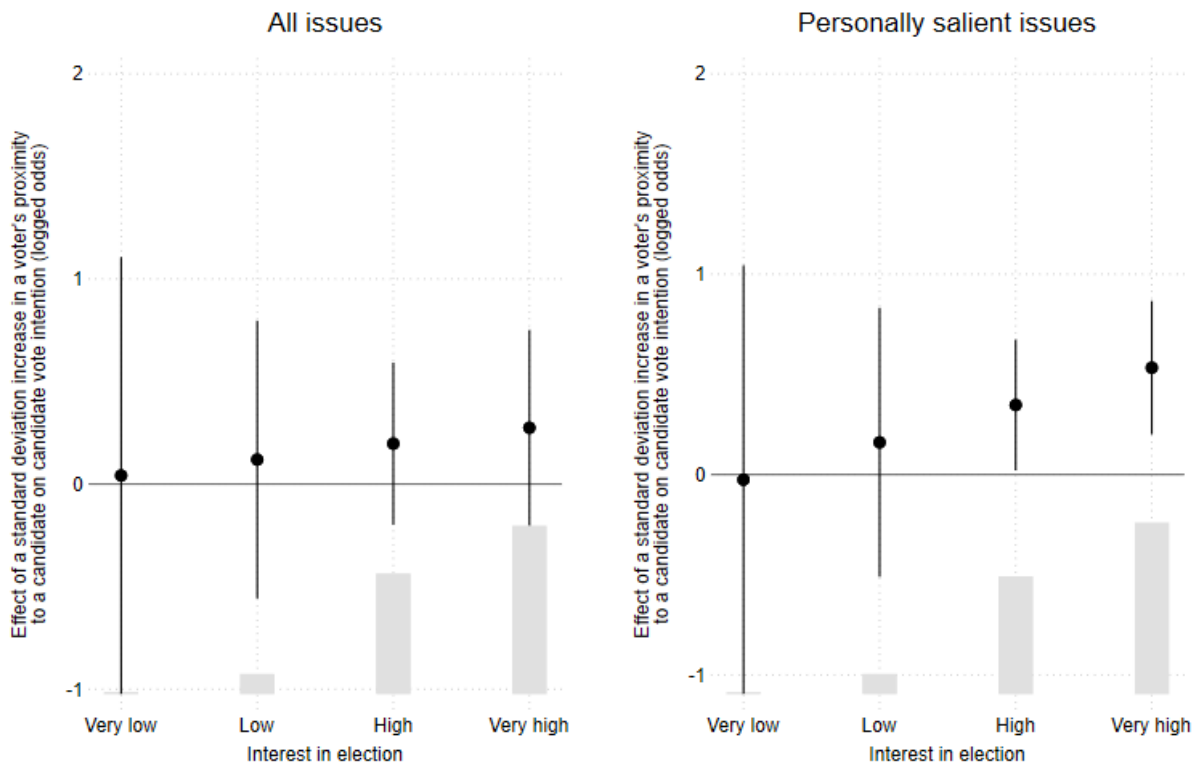


Table A7: (Conditional) logit regressions

	All issues			Personally salient issues		
	(1)	(2)	(3)	(4)	(5)	(6)
Proximity to candidate	8.26***	1.73	-0.34	3.17***	2.29**	-1.21
	(0.31)	(1.76)	(7.07)	(0.14)	(0.75)	(4.30)
Proximity to party		11.95***	11.57***		10.71***	10.51***
		(1.75)	(1.76)		(1.51)	(1.54)
Party identification		2.23***	2.23***		2.28***	2.28***
		(0.16)	(0.16)		(0.16)	(0.17)
Previous vote choice		0.52***	0.52***		0.46**	0.47**
		(0.15)	(0.15)		(0.15)	(0.15)
p(candidate win)		0.40***	0.39***		0.42***	0.42***
		(0.03)	(0.03)		(0.03)	(0.03)
Female candidate		-0.13	-0.12		-0.06	-0.05
		(0.16)	(0.17)		(0.15)	(0.16)
Incumbent		-0.00	-0.03		-0.24	-0.25
		(0.21)	(0.22)		(0.22)	(0.22)
Local candidate		-0.12	-0.10		-0.01	0.02
		(0.18)	(0.18)		(0.18)	(0.18)
Proximity to candidate * interest in election			0.73			1.06
			(2.07)			(1.21)
Candidate fixed effects	No	No	No	No	No	No
Voter fixed effects	No	Yes	Yes	No	Yes	Yes
Candidates	59	45	45	59	45	45
Voters	11580	1470	1436	10047	1354	1324
Observations	20997	3921	3828	18305	3602	3521

Note: The dependent variable in all models is whether a voter intends to vote for a candidate. Models 1 and 3 are estimated using logistic regression. All other models are estimated with conditional logistic regression accounting for unobserved heterogeneity at the voter level. Adding candidate fixed effects on top of the voter fixed effects would cause conditional logit coefficient estimates to be biased due to the incidental parameters problem (Timoneda 2021). To account for candidate-level heterogeneity, we instead introduce three candidate-level variables in models 2, 3, 5, and 6: candidate gender, incumbency, and whether candidates live in the constituency (all measured as binaries). Note that conditional logistic regression drops all clusters of voters with consistently positive or negative outcomes. Standard errors clustered by voter in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A6: Marginal effects plots [conditional logit regressions]



Note: In conditional logit regressions, predicted probabilities can only be calculated under the (non-sensical) assumption that all voter fixed effects are zero (Timoneda 2021). Therefore, we instead show conditional effects in terms of logged odds, which are unaffected by the fixed effects.

Table A8: Double de-meaned interactions

	All issues	Personally salient issues
	(1)	(2)
Proximity to candidate	0.02	0.09**
	(0.09)	(0.03)
Proximity to party	0.55***	0.41***
	(0.09)	(0.06)
Party identification	0.54***	0.55***
	(0.02)	(0.02)
Previous vote choice	0.06***	0.06***
	(0.01)	(0.01)
p(candidate win)	0.02***	0.02***
	(0.00)	(0.00)
Proximity to candidate * interest in election (dd)	0.06	0.10**
	(0.07)	(0.04)
Candidate fixed effects	Yes	Yes
Voter fixed effects	Yes	Yes
Candidates	59	59
Voters	8288	7428
Observations	14758	13273

Note: The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with linear regression. Standard errors clustered by voter in parentheses. dd = double de-meaned. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A7: Marginal effects plots [double de-meaned interactions]

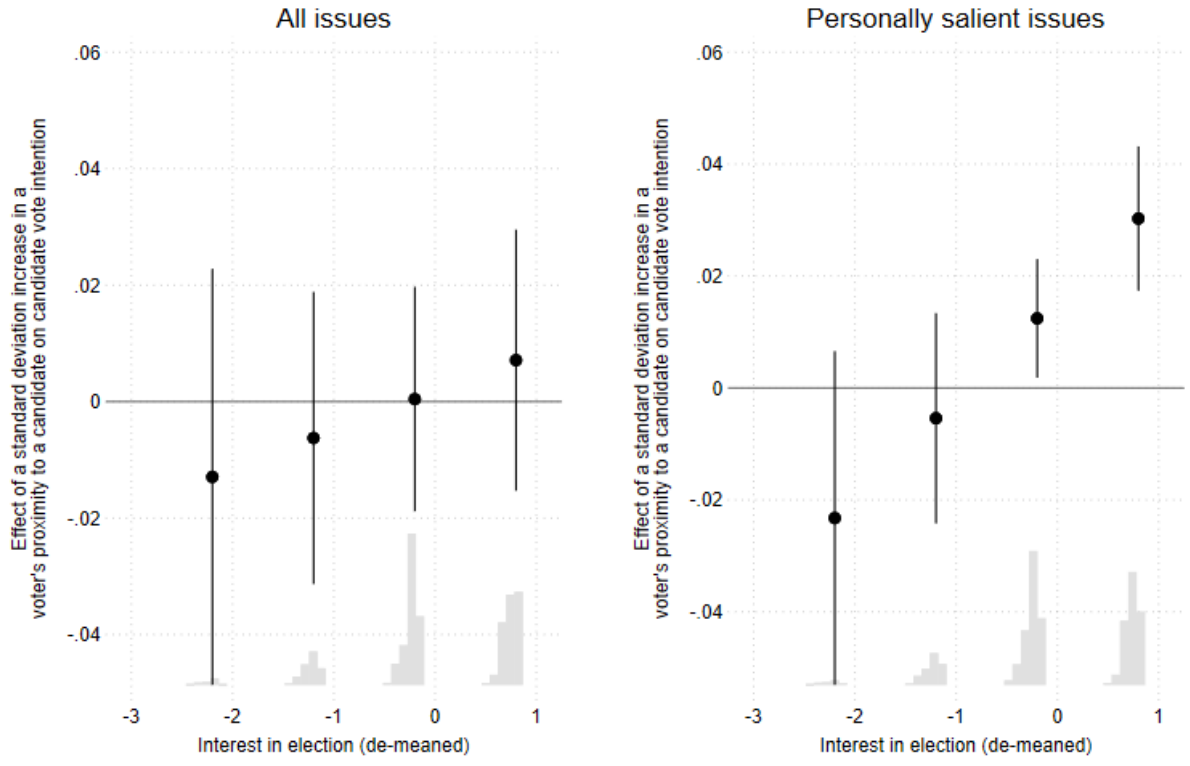


Table A9: Interactions with self-reported familiarity with candidates

	All issues	Personally salient issues
	(1)	(2)
Proximity to candidate	-0.10	-0.04
	(0.12)	(0.06)
Proximity to party	0.53***	0.40***
	(0.09)	(0.06)
Party identification	0.54***	0.55***
	(0.02)	(0.02)
Previous vote choice	0.06***	0.06***
	(0.01)	(0.01)
p(candidate win)	0.02***	0.02***
	(0.00)	(0.00)
Proximity to candidate * familiarity with candidates	0.05	0.06**
	(0.03)	(0.02)
Candidate fixed effects	Yes	Yes
Voter fixed effects	Yes	Yes
Candidates	59	59
Voters	8439	7553
Observations	15029	13499

Note: The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with linear regression. Standard errors clustered by voter in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A8: Marginal effects plots [interactions with self-reported familiarity with candidates]

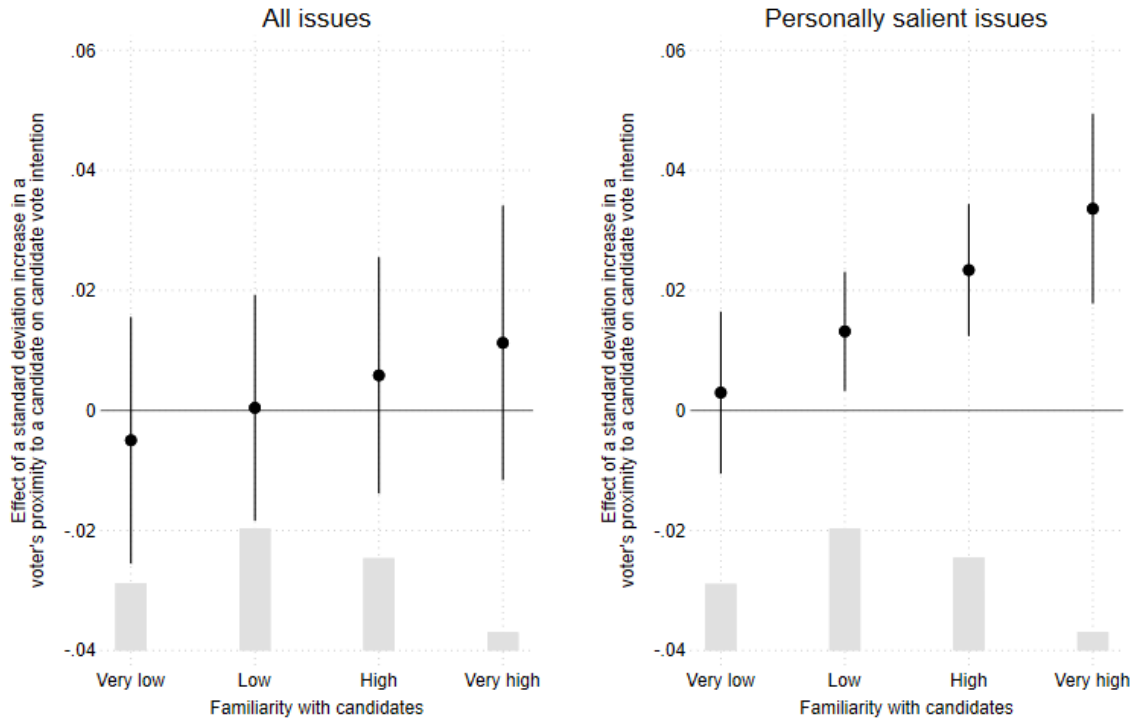
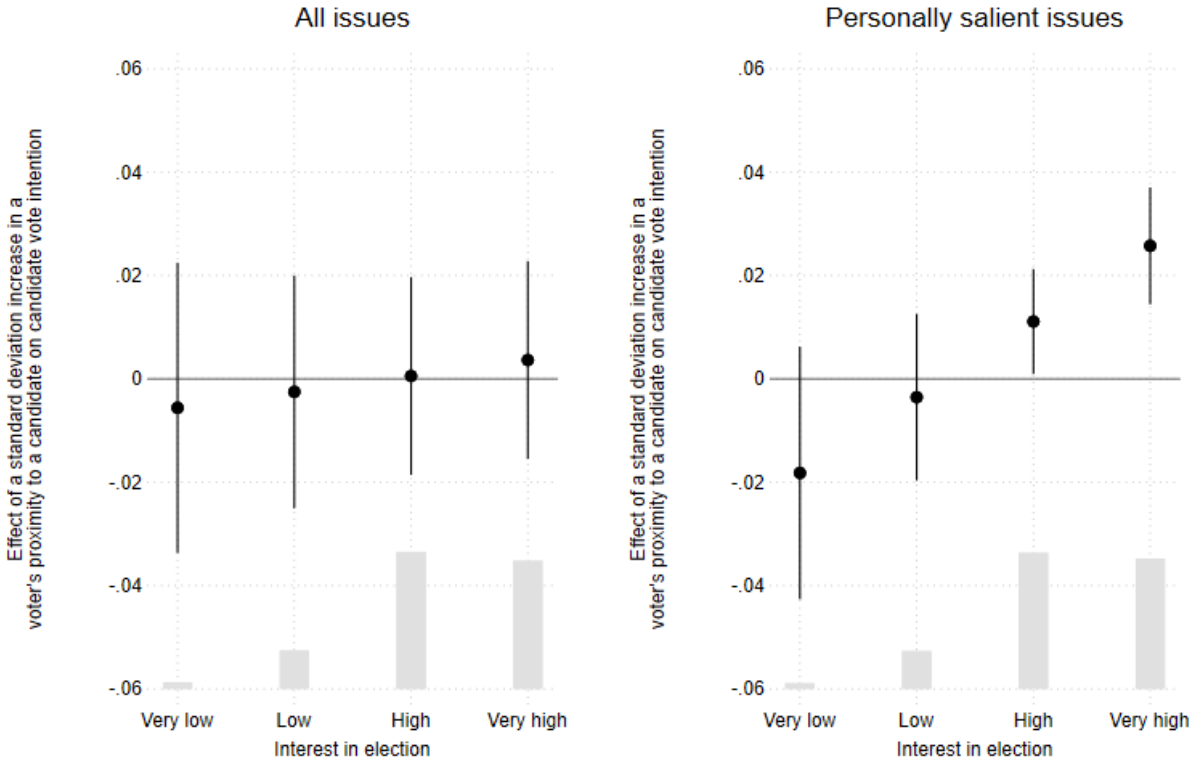


Table A10: No sample restrictions/data cleaning

	All issues			Personally salient issues		
	(1)	(2)	(3)	(4)	(5)	(6)
Proximity to candidate	0.56***	0.03	-0.08	0.27***	0.09**	-0.19
	(0.02)	(0.09)	(0.16)	(0.01)	(0.03)	(0.10)
Proximity to party		0.54***	0.55***		0.44***	0.42***
		(0.09)	(0.09)		(0.06)	(0.06)
Party identification		0.54***	0.54***		0.54***	0.54***
		(0.02)	(0.02)		(0.02)	(0.02)
Previous vote choice		0.06***	0.07***		0.06***	0.06***
		(0.01)	(0.01)		(0.01)	(0.01)
p(candidate win)		0.02***	0.02***		0.02***	0.02***
		(0.00)	(0.00)		(0.00)	(0.00)
Proximity to candidate * interest in election			0.03			0.08**
			(0.04)			(0.03)
Candidate fixed effects	No	Yes	Yes	No	Yes	Yes
Voter fixed effects	No	Yes	Yes	No	Yes	Yes
Candidates	59	59	59	59	59	59
Voters	12391	9139	8894	10716	8162	7954
Observations	22443	16254	15820	19499	14565	14192

Note: The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with linear regression. Standard errors clustered by voter in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A9: Marginal effects plots [no sample restrictions/data cleaning]



4 Additional Analyses

Table A11: Results by issue area

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PTC - economy	-0.07						-0.04
	(0.04)						(0.04)
PTC - health & social care		-0.05					-0.04
		(0.03)					(0.03)
PTC - environment			-0.02				0.00
			(0.03)				(0.03)
PTC - education				-0.04			-0.04
				(0.03)			(0.03)
PTC - language & culture					-0.08*		-0.11***
					(0.03)		(0.03)
PTC - Wales vs UK						0.19***	0.19***
						(0.03)	(0.03)
Proximity to party	0.62***	0.59***	0.57***	0.59***	0.62***	0.30***	0.46***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.09)
Party identification	0.54***	0.55***	0.55***	0.54***	0.55***	0.54***	0.53***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Previous vote choice	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
p(candidate win)	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Candidate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Voter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Candidates	59	59	59	59	59	59	59
Voters	8509	8513	8501	8506	8513	8511	8489
Observations	15154	15163	15141	15150	15163	15161	15121

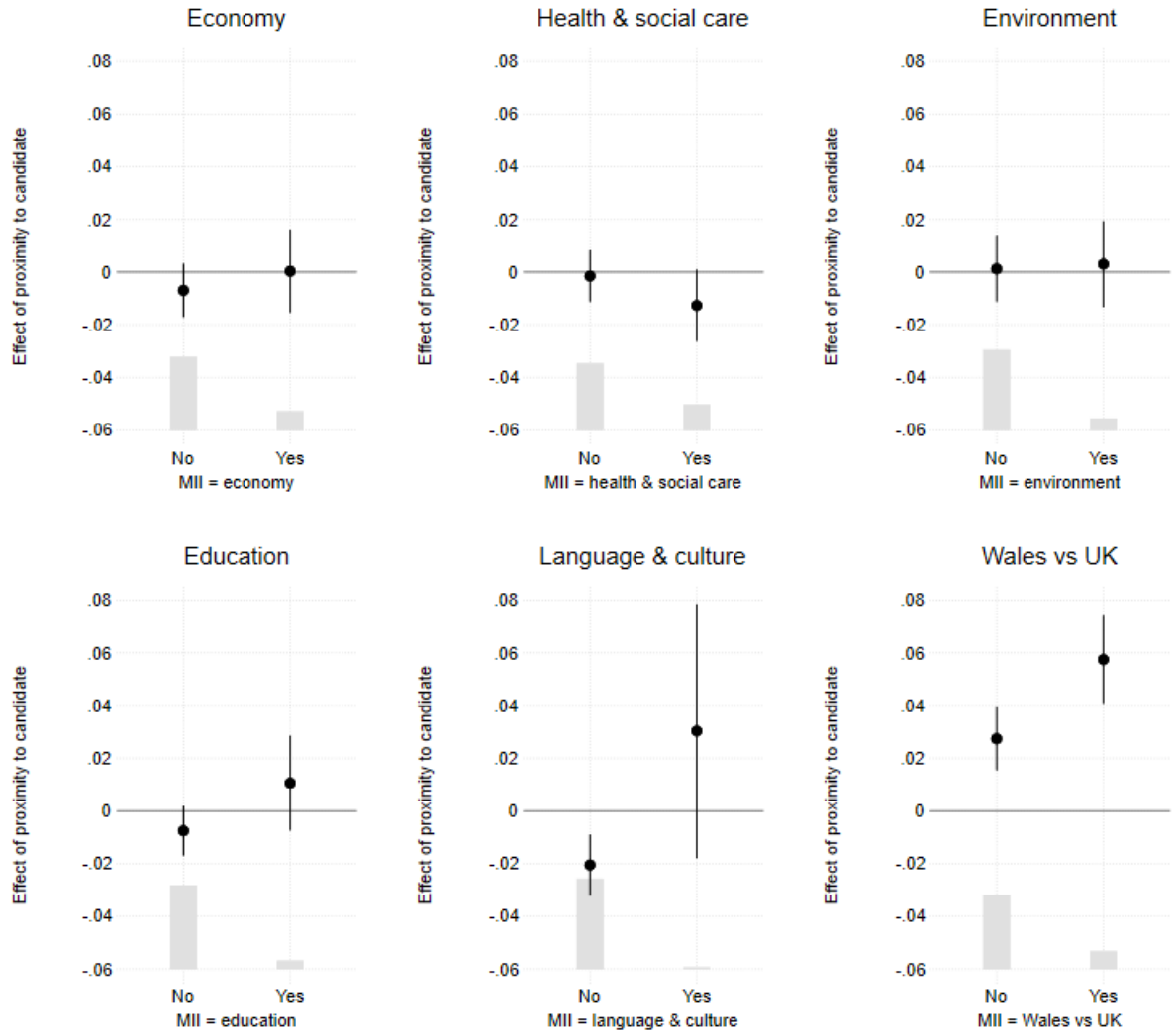
Note: The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with linear regression. Standard errors clustered by voter in parentheses. PTC = proximity to candidate. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A12: Results by issue area including interactions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PTC - economy	-0.08						-0.05
	(0.04)						(0.04)
PTC - economy * MII = economy	0.03						0.06
	(0.06)						(0.06)
PTC - health & social care		-0.01					-0.01
		(0.04)					(0.04)
PTC - health & social care * MII = health & social care		-0.13*					-0.08
		(0.06)					(0.06)
PTC - environment			-0.01				0.01
			(0.03)				(0.03)
PTC - environment * MII = environment			-0.03				0.01
			(0.04)				(0.04)
PTC - education				-0.05			-0.05
				(0.03)			(0.03)
PTC - education * MII = education				0.09			0.12
				(0.06)			(0.06)
PTC - language & culture					-0.08*		-0.11***
					(0.03)		(0.03)
PTC - language & culture * MII = language & culture					0.24		0.28*
					(0.13)		(0.14)
PTC - Wales vs UK						0.15***	0.15***
						(0.03)	(0.03)
PTC - Wales vs UK * MII = Wales vs UK						0.15***	0.16***
						(0.04)	(0.04)
Proximity to party	0.62***	0.58***	0.57***	0.59***	0.62***	0.28***	0.45***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)	(0.09)
Party identification	0.55***	0.55***	0.55***	0.55***	0.55***	0.54***	0.53***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Previous vote choice	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
p(candidate win)	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Candidate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Voter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Candidates	59	59	59	59	59	59	59
Voters	8320	8324	8312	8317	8324	8322	8300
Observations	14837	14846	14824	14833	14846	14844	14804

Note: The dependent variable in all models is whether a voter intends to vote for a candidate. All models are estimated with linear regression. Standard errors clustered by voter in parentheses. PTC = proximity to candidate; MII = most important issue area. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A10: Marginal effects plots [by issue area, model 7 in Table A12]



References

Timoneda, J.C. (2021) “Estimating Group Fixed Effects in Panel Data with a Binary Dependent Variable: How the LPM Outperforms Logistic Regression in Rare Events Data.” *Social Science Research* 93: 102486.