The Economic Roots of Cross-National Similarity in Voter Preferences

Supplemental Information

Contents

1 Descriptive Statistics

In Table A.1 we provide the descriptive statistics for the variables based on the estimation sample of 8,554 country pair-years.

Table A.1: Descriptive statistics of the unimputed data

	N	**Mean**	**S.D.**	**Min.**	**Max.**
Voter Congruence	8,554	1.37	0.23	0.05	1.79
Similarity	7,142	0.27	0.22	0.05	4.60
Average Export	8,119	3.40	4.28	0.005	31.10
Maastricht	8,554	0.78	0.41	0	1
Eurozone	8,554	0.27	0.44	0	1
Party Landscape Similarity	8,554	1.85	1.52	0	8.3
Education Similarity	5,583	9.63	6.99	0	39.80
GDP Per Capita Similarity	8,453	21,106.19	21,322.56	0.17	124,360
Income Similarity	7,192	15,145.01	13,478.09	0.66	76,845.18
National Election	8,554	0.06	0.23	0	1
European Election	8,554	0.19	0.39	0	1

2 Non-Response Bias in Self-Placement

In the manuscript we use multiple imputation to address possible non-response in self-placement, which can bias inferences if non-response is correlated with other observable (or unobservable) characteristics rather than missing completely at random. At the least, listwise deletion of non-responding participants will skew the survey's weighting scheme, which is calibrated to all participants in the survey, whether they answer all questions or not. Past studies have nearly always ignored missinginess despite research from (Homola 2019) documenting that across 40 years of Eurobarometer survey waves in 14 countries, women are significantly less likely to record an ideological self-placement. This means that ignoring missingness will lead to an estimate in which the preferences of men are given more weight than the preferences of women. Further, if these patterns of missingness vary across countries—and Homola shows that they do, with the gender gap in non-response varying from 1.7% in Sweden to 8.4% in Luxembourg—listwise deletion will lead to systematic, but *inconsistent* biases in mean estimates across countries, compounding inferential problems. In other words, listwise deletion may cause the sample means to be biased from population means *and* that the bias may vary systematically across countries and over time. Of course, this is likely to have systematic relationships across more than just gender.

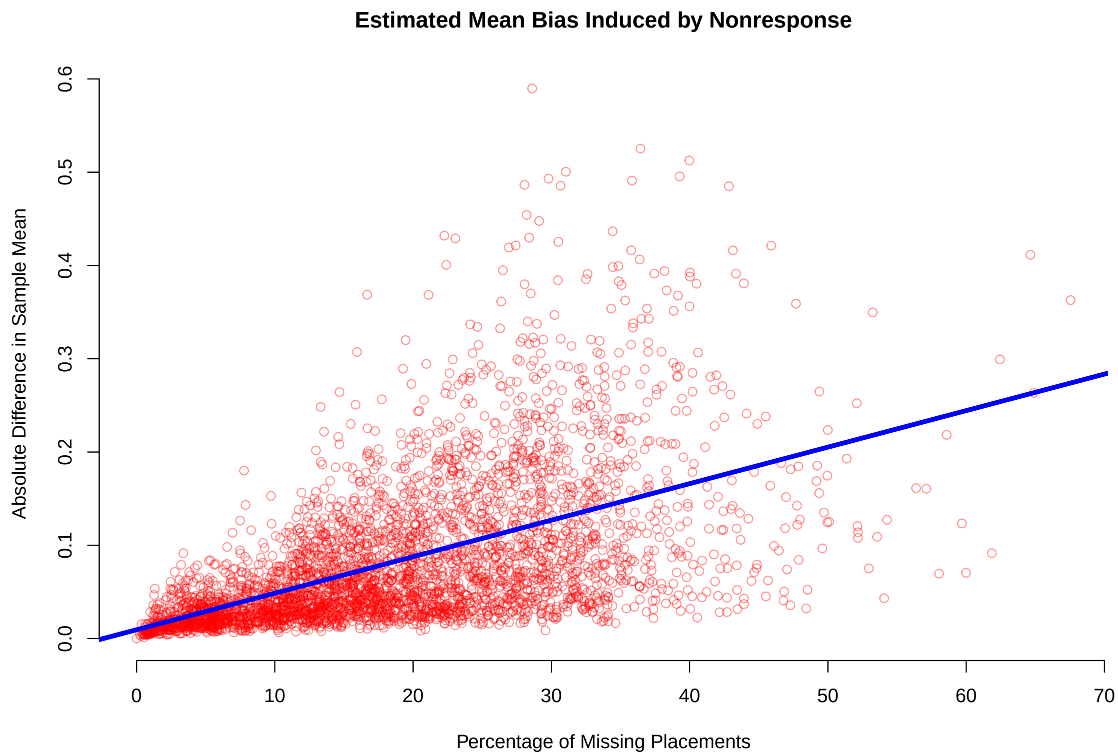
Figure A.1 plots the percentage of missing left-right placement data for each country-survey in our sample on the x-axis against the absolute difference between sample means when missingness is ignored (listwise deletion) and treated with multiple imputation—we may think of this as listwise deletion-induced bias in the mean. The plot shows that country-surveys with greater levels of item non-response on the self-placement question tend to have larger absolute differences between sample means, but also that the relationship is noisy, or far from deterministic. There are surveys with small levels of missingness but large estimated mean bias. Of course the inverse is also manifest. This means that while, in general, the choice of approach—either listwise deletion or multiple imputation—becomes more consequential as the percentage of missingness increases in the sample (as indicated by the positively-sloped OLS regression line), a low level of missingness *does not* ensure little or no mean bias.

2.1 Imputation Model

The analysis above suggests that imputation is important to recover a more accurate unweighted distribution of voter preferences. Of course, we know that imputation is necessary to utilize the survey's population weights. To address this point, we impute left-right self-placements with a predictive model (from the Amelia package in R) based on individual demographics (gender, age, education, urban, church attendance, income quartiles, marital status, socioeconomic class, and household union membership) along with country and year fixed effects (King, et al. 2001).

We want to show here, however, that the key findings in the manuscript are not a re-

Figure A.1: Percentage of missing placement data and the absolute difference in sample means between listwise deletion and multiple imputation



sult of the imputation of self-right placement data. Table A.2 shows, the results from the unimputed data (listwise deletion) are nearly identical in magnitude and efficiency to those in the manuscript. While we prefer to present the imputed data analyzed in the manuscript and believe that that is the proper approach to measurement and estimation, it is reassuring to know that the relationship between political and economic integration and preference congruence is not an artifact of those measurement choices.

We also use multiple imputation to deal with missing values at the dyadic level in the aggregated data frame. This imputation model includes information on education rates, income, GDP, and GDP per capita from the OECD and World Bank. To improve the predictive accuracy of the model, we include country and year fixed effects, two temporal polynomials, and lags and leads of all the independent variables. We provide the descriptive statistics for the imputed dyadic data in Table A.3.

Table A.2: OLS regression results for the relationship between voter congruence and economic similarity, bilateral trade and political integration: listwise deletion

	β 95% CI
Similarity	0.040**
	[0.015, 0.064]
Average Export	0.011**
	[0.006, 0.016]
Maastricht (Both)	0.035**
	[0.012, 0.057]
Eurozone (Both)	0.045**
	[0.029, 0.060]
Year Fixed Effects	$\checkmark$
Dyad Fixed Effects	$\checkmark$
N	8,554
Adjusted R^2	0.482
RMSE	0.168

Note: Dependent variable is the reversed EMD measure where higher values indicate country pairs with more congruent distributions of left-right preferences. Brackets show 95% confidence intervals. Estimates are from listwise deletion.

* p-value < 0.10; ** p-value < 0.05

Table A.3: Descriptive statistics of the imputed data ($N = 8,854$). Variables not tabulated here have no missing values.

	Mean	**S.D.**	**Min.**	**Max.**
Similarity	0.29	0.21	0.0005	4.60
Average Export	3.36	4.24	0.002	31.10
Education Similarity	9.72	6.74	0	39.80
GDP Per Capita Similarity	21,262.82	21,323.8	0.17	124,360
Income Similarity	15,599.58	13,605.45	0.66	76,845.18

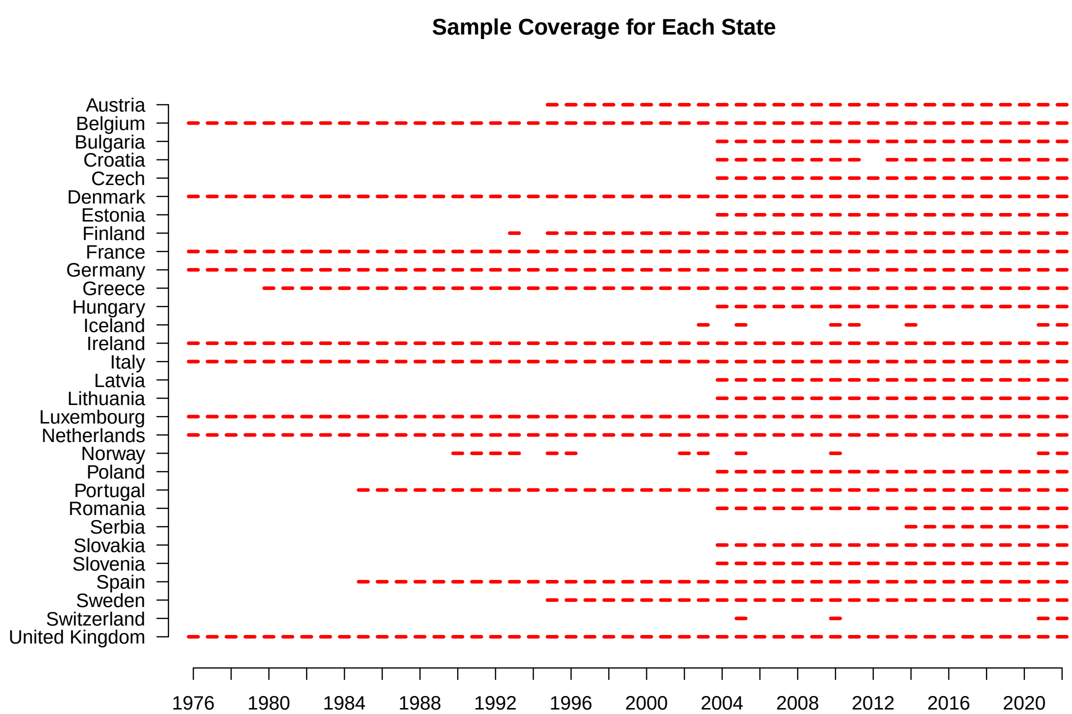
3 Robustness Checks

In this section we probe the robustness of our central analysis with differing specifications, many of which were discussed (but not presented) in the manuscript.

3.1 Frequency Weights

The frequency that dyads appear in the data varies based upon their how often Eurobarometer conducted surveys in the each country. Coverage for each country is given in Figure A.2, which is also in the main text. The figure reveals that the primary divide in coverage is between older, richer, western democracies like Denmark, Germany, the UK, which enter the sample early and are survey each year, and younger, less developed democracies, like Bulgaria, Hungary, and Poland, which enter the sample much later. This means that the effective weights exerted by dyads on the parameter estimates will be large for dyads composed of western countries relative to dyads composed of their central and eastern counterparts. Here we consider the potential impact of these variable weights on the analysis.

Figure A.2: Survey coverage for each sample country.



To that end, we assess the robustness of our main results to estimation where each dyad exerts equal weights on the parameter estimates. The results are given in Table A.4. The left model has the original estimates from the main text while the right model has been reweighted such that all dyads contribute equal weights to the final estimates: $w_{id} = \frac{1}{n_d}$, where w_{id} is the weight placed on any observation i in dyad d, and n_d is the total number of observations within dyad d. The results are fairly consistent across specifications, with slightly larger coefficients for both political integration variables.

Table A.4: Comparing results from the main-text model using equal weights on all observations, to estimates using dyad-equalizing weights

	Main text	Dyad equalizing
Similarity	0.040^{**}	0.046^{**}
	[0.014, 0.064]	[0.042, 0.050]
Average Export	0.011^{**}	0.007^{**}
	[0.006, 0.016]	[0.006, 0.007]
Maastricht (Both)	0.032^{**}	0.047^{**}
	[0.007, 0.055]	[0.040, 0.054]
Eurozone (Both)	0.044^{**}	0.056^{**}
	[0.027, 0.059]	[0.051, 0.058]
Year Fixed Effects	√	√
Dyad Fixed Effects	√	√
Observations	8,554	8,554
Adjusted R^2	0.474	0.506
RMSE	0.170	0.170
$\bar{y}$ (σ^2_y)	1.369 (0.055)	1.363 (0.054)

3.2 Over-Time Specification

Here we first present two autoregressive models containing lagged dependent variables and lags of each of the independent variables. The model takes on the following form:

$$y_{dt} = \alpha y_{dt-1} + \beta_0 \mathbf{X}_{dt} + \beta_1 \mathbf{X}_{dt-1} + \gamma_d + \phi_t + \epsilon_{dt} \tag{1}$$

y is our dependent variable, congruence, indexed by dyads d and years t, $\mathbf{X}$ are our key covariates, α estimates the persistence, or the level of the lagged dependent variable carrying through to the next time period, β_0 are the weights on concurrent values of the covariates, β_1 are the weights on lagged values of the covariates, γ and ϕ are vectors of fixed effects for dyad and years, and of course ϵ_{dt} is the residual error.

With minimal rearrangement, this model specification allows us extract all values of interest necessary to understand the persistence of congruence and the short- and long-term effects of our key covariates. As discussed, the degree of persistence (or the carry-over of congruence from one period to the next) is α, the short-run effect of concurrent values of covariates $\mathbf{X}_{dt}$ is captured by β_0, the short-run effect of past values of covariates $\mathbf{X}_{dt-1}$ is β_1, and the total effect of these covariates, given their persistence on past and present realizations of the dependent variable, is called the "long-run multiplier" of $\mathbf{X}$ is $\frac{\beta_0+\beta_1}{(1-\alpha)}$.

Table A.5 shows the ADL (1;1) and Table A.6 shows the ADL (2;2) model with dyad fixed effects and dyad equalizing weights. Instead of focusing solely on within-year variation – as in the model presented in the manuscript – the ADL models allow us to observe the over-time variation. Since they contain lags of the dependent and independent variables, the effects of the covariates bleed over into subsequent time periods. As such, it is important to evaluate the covariates in terms of both short- and long-term effects. Similar to the manuscript's models, the short- and long-term effects depict the effects of a within-unit one-standard deviation (one-unit) increase in the continuous (binary) variables. 95% confidence intervals are calculated via the percentile method and are shown in brackets.

The long-term effects for both autoregressive models are supportive of our three hypotheses: economic similarity, average exports and political integration clearly and positively increase voter congruence in the long-term. Furthermore, the long-term effects are generally comparable to those found in the main model in the manuscript. The one exception is the long-term effect of the *Maastricht* variable which is over twice the size of the coefficient in the manuscript (0.072 in the ADL (1;1) compared to 0.032). We equalize dyads here because the persistence estimate for the short-run dyads (primarily those containing central and eastern countries) is much smaller than the persistence estimate for the long-run dyads, which exert much more influence in the unweighted models. As a result, the unweighted model yields long-term effects estimates that are unusually large for some parameters—nearly 4x the size in the case economic similarity.

3.3 Similarity of Party Landscapes

A possible confounder in our theory of voter congruence is the ideological landscape of the systems. Voters across countries may appear to be more or less similar because the parties that operate within those systems are more or less similar. This is particularly troubling if one subscribes to a model of public opinion in which parties drive the positions of voters, rather than the inverse. To control for this confounding variable, we measure the absolute difference in the effective number of electoral parties (Laakso & Taagepera 1979).

Table A.7 shows the results of the OLS model when we control for this confounding variable. The inclusion of this variable does not meaningfully influence the conclusions from our model and all three hypotheses are still supported. Moreover, the effect sizes are nearly identical. The coefficient for *party landscape similarity* is positive and statistically different from zero, which suggests that systems where the difference in number of electoral parties is smaller have less congruent electorates. Since we do not have a theoretical expectation to explain this result, we leave it to other scholars to effectively explore this relationship.

3.4 Development

Our main empirical model omits covariates that are central to classic theories of societal change that are likely correlated with the focal variables in our model. In particular, the chief confounding factors we derived are from Lipset and Rokkan (the party system similarity covariate discussed above) and factors discussed by Inglehart and others. With respect to Inglehart (1977, 1997), note that our argument's central prediction is different from his—we are studying convergence (two distributions more closely resembling one another), whereas he made predictions regarding parallel change (two distributions moving in the same direction). Also, we believe our economic production similarity covariate may be the best single covariate to test his predictions, which would imply 0 convergence as co-moving distributions do not become more similar (they remain equally dissimilar, say, one unit to the left). Our next best covariate to capture Inglehart's argument would be development, so we have modeled similarity in GDP and income. Finally, the other dominant, classic theories of socio-political change with which we are familiar (e.g., Braverman, Maslow) similarly focus on development, income in particular, but also on the role of education in reorienting patterns of conflict, consumption, and priorities.

To account for these confounders, we add three variables to our main model: GDP per capita and income per capita (both from the World Bank's World Development Indicators), and the percentage of the population ages 25-34 with a tertiary education. To create *similarity* measures we take the absolute difference between each country in the country-pair. We expect that all three measures will be negative since greater similarities in economic development and education should correspond to greater voter congruence. We present an OLS model (with year and dyad fixed effects) including these confounders in Table A.8.

The inclusion of these covariates does not affect the conclusions from the main text model. Moreover, the effect sizes are similar to those in the manuscript model. The results are mixed among the three possible confounders: *GDP per capita similarity* is negative and statistically significant (as expected), but *education similarity* is not statistically different from zero, and *income per capita similarity* is not in the expected direction. We leave it to other scholars to speculate about the causes of these correlations, but we are more confident that the manuscript's key findings related to economic similarity, complementarity and political integration are robust to including measures of wealth and development.

3.5 Proximity to Elections

It is possible that proximity to elections – either national or European elections – might confound the relationship between political and economic integration and ideological congruence. We add three variables to our model: *national elections*, which is coded 1 if both countries have national elections that year, *European elections*, which is coded 1 if European Parliament elections take place in both countries that year, and the interaction of *European elections* with *Eurozone*, to account for different effects of EU elections based on Eurozone membership.

The results in Table A.9 illustrates the robustness of our findings in the manuscript: the conclusion of all the hypothesis tests remain unchanged and the effect sizes do not change in a meaningful manner. The results of the control variables suggest that proximity to national elections leads to a slight divergence in preferences, whereas proximity to European elections generates a sizeable increase in the congruence of preferences. Eurozone co-participation does not exacerbate this effect.

3.6 Shock Similarity

Below we have estimated the long-term effects of China shock dosage similarity on the congruence of political preferences. Recall that (Colantone & Stanig 2018*a*, Colantone & Stanig 2018*b*) show that the china shock produces a parallel movement across countries, increasing the support for Brexit voting and nationalist parties across countries. Their estimates are based on changes to Chinese imports over 3 or 6 years—that is, a 2 year change, lagged to outcome, and a 5 year change, lagged to outcome. Below, we estimate the effect of covariance in China shock dosage on the congruence of political preferences across dyads at current values and then lags of up to 6 years. The effects are negative at concurrent value and 1 year lag, null at 2 year lag, and positive thereafter, comporting with (Colantone & Stanig 2018*a*, Colantone & Stanig 2018*b*).

We cannot say for certain why positive effects manifest at concurrent value and 1 year lag, but our best guess is that early reaction to China shock are consumption based—driven by the novel availability of lower priced goods. Practical effects of Chinese imports manifest in wages pressures and labor force displacement, which (Colantone & Stanig 2018*a*, Colantone & Stanig 2018*b*) argue are the central drivers to political preference change.

Table A.5: Autoregressive distributed lag models (ADL 1;1) for the relationship between voter congruence and economic similarity, bilateral trade and political integration

	β 95% CI	STE 95% CI	LTE 95% CI
Congruence_{t-1}	0.339^{**}		
	[0.316, 0.360]		
Similarity_{t}	-0.013	-0.002	0.009^{**}
	[-0.055, 0.024]	[-0.008, 0.003]	[0.002, 0.016]
Similarity_{t-1}	0.053^{**}	0.008^{**}	
	[0.013, 0.092]	[0.002, 0.013]	
Average Export_{t}	-0.002	-0.001	0.008^{**}
	[-0.013, 0.009]	[-0.009, 0.007]	[0.001, 0.015]
Average Export_{t-1}	0.009^{*}	0.007^{*}	
	[-0.002, 0.019]	[-0.001, 0.015]	
Maastricht $(\text{Both})_{t}$	0.082^{**}		0.072^{**}
	[0.053, 0.112]		[0.047, 0.097]
Maastricht $(\text{Both})_{t-1}$	-0.034^{**}		
	[-0.064, -0.007]		
Eurozone $(\text{Both})_{t}$	0.023		0.032^{**}
	[-0.008, 0.058]		[0.010, 0.054]
Eurozone $(\text{Both})_{t-1}$	-0.003		
	[-0.033, 0.028]		
Constant	0.786^{**}		
	[0.622, 0.918]		
Dyad Equalizing Weights	√		
Dyad Fixed Effects	√		
N	8,119		
Adjusted R^2	0.552		
RMSE	0.157		

Note: Coefficients reflect median value of 1,000 simulations. 95% confidence intervals via percentile method. * p-value < 0.10; ** p-value < 0.05
STE and LTE are for 1-standard deviation increases (for continuous variables).

Table A.6: Autoregressive distributed lag models (ADL 2;2) for the relationship between voter congruence and economic similarity, bilateral trade and political integration

	β 95% CI	STE 95% CI	LTE 95% CI
Congruence_{t-1}	0.308**		
	[0.282, 0.334]		
Congruence_{t-2}	0.081**		
	[0.055, 0.105]		
Similarity_{t}	-0.021	-0.003	0.010**
	[-0.063, 0.017]	[-0.009, 0.003]	[0.002, 0.018]
Similarity_{t-1}	0.034	0.005	
	[-0.013, 0.087]	[-0.002, 0.013]	
Similarity_{t-2}	0.028	0.004	
	[-0.012, 0.067]	[-0.002, 0.010]	
Average Export_{t}	0.001	0.0008	0.010**
	[-0.011, 0.014]	[-.008, 0.010]	[0.003, 0.018]
Average Export_{t-1}	0.010	0.008	
	[-0.007, 0.027]	[-0.005, 0.020]	
Average Export_{t-2}	-0.002	-0.002	
	[-0.016, 0.011]	[-0.012, 0.008]	
Maastricht (Both)_{t}	0.085**		0.076**
	[0.054, 0.118]		[0.047, 0.10]
Maastricht (Both)_{t-1}	0.001		
	[-0.039, 0.042]		
Maastricht (Both)_{t-2}	-0.040**		
	[-0.071, -0.010]		
Eurozone (Both)_{t}	0.027		0.037**
	[-0.007, 0.060]		[0.011, 0.065]
Eurozone (Both)_{t-1}	0.002		
	[-0.042, 0.045]		
Eurozone (Both)_{t-2}	-0.008		
	[-0.041, 0.028]		
Constant	0.681**		
	[0.522, 0.834]		
Dyad Equalizing Weights	√		
Dyad Fixed Effects	√		
N	7,684		
Adjusted R^2	0.570		
RMSE	0.154		

Note: Coefficients reflect median value of 1,000 simulations. 95% confidence intervals via percentile method. * p-value < 0.10; ** p-value < 0.05
STE and LTE are for 1-standard deviation increases (for continuous variables).

Table A.7: OLS regression results for the relationship between voter congruence and economic similarity, bilateral trade and political integration: controlling for similarity of party landscapes

	β 95% CI	SD($\tilde{x}$) Δ 95% CI
Similarity	0.040**	0.006**
	[0.015, 0.064]	[0.002, 0.010]
Average Export	0.011**	0.008**
	[0.006, 0.015]	[0.005, 0.011]
Maastricht (Both)	0.032**	
	[0.007, 0.055]	
Eurozone (Both)	0.046**	
	[0.030, 0.062]	
Party Landscape Similarity	0.005**	0.005**
	[0.0009, 0.009]	[0.0008, 0.008]
Constant	1.064**	
	[0.889, 1.213]	
Year Fixed Effects	√	
Dyad Fixed Effects	√	
N	8,554	
Adjusted R^2	0.475	
RMSE	0.170	
$\bar{y}$ (σ_y^2)	1.369 (0.055)	

Note: Coefficients reflect median value of 1,000 simulations. 95% confidence intervals via percentile method. * p-value < 0.10; ** p-value < 0.05

Table A.8: OLS regression results for the relationship between voter congruence and economic similarity, bilateral trade and political integration: controlling for similarity of economic development and education

	β 95% CI	SD($\tilde{x}$) Δ 95% CI
Similarity	0.030**	0.004*
	[0.004, 0.057]	[0.006, 0.016]
Average Export	0.011**	0.009**
	[0.007, 0.016]	[0.005, 0.012]
Maastricht (Both)	0.032**	
	[0.008, 0.056]	
Eurozone (Both)	0.046**	
	[0.030, 0.062]	
Education Similarity	-0.0006	-0.003
	[-0.001, 0.0001]	[-0.007, 0.0006]
GDP Per Capita Similarity	-0.000001**	-0.010**
	[-0.00002, -0.000006]	[-0.016, -0.004]
Income Per Capita Similarity	0.000001**	0.007**
	[0.0000003, 0.000003]	[0.001, 0.013]
Year Fixed Effects	✓	
Dyad Fixed Effects	✓	
N	8,554	
Adjusted R^2	0.475	
RMSE	0.170	
$\bar{y}$ (σ_y^2)	1.369 (0.055)	

Note: Coefficients reflect median value of 1,000 simulations. 95% confidence intervals via percentile method. * p-value < 0.10; ** p-value < 0.05

Table A.9: OLS regression results for the relationship between voter congruence and economic similarity, bilateral trade and political integration: controlling for proximity of elections

	β 95% CI	SD($\tilde{x}$) Δ 95% CI
Similarity	0.040**	0.006**
	[0.013, 0.066]	[0.002, 0.009]
Average Export	0.011**	0.008**
	[0.006, 0.016]	[0.005, 0.012]
Maastricht (Both)	0.033**	
	[0.007, 0.060]	
Eurozone (Both)	0.044**	
	[0.026, 0.059]	
National Election (Both)	-0.015*	
	[-0.030, 0.003]	
European Election (Both)	0.285**	
	[0.202, 0.360]	
Eurozone × European Election	-0.005	
	[-0.039, 0.027]	
Constant	1.058**	
	[0.894, 1.215]	
Year Fixed Effects	✓	
Dyad Fixed Effects	✓	
N	8,554	
Adjusted R^2	0.475	
RMSE	0.170	
$\bar{y}$ (σ_y^2)	1.369 (0.055)	

Note: Coefficients reflect median value of 1,000 simulations. 95% confidence intervals via percentile method. * p-value < 0.10; ** p-value < 0.05

Table A.10:

	Concurrent	1 year	2 years	3 years	4 years	5 years	6 years
Shock similarity	-0.119^*	-0.122^*	-0.026	0.081^*	0.082^*	0.044^*	0.093^*
	(0.015)	(0.015)	(0.016)	(0.016)	(0.017)	(0.018)	(0.018)
All covariates	Y	Y	Y	Y	Y	Y	Y
Dyad FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y
Observations	5,773	5,307	5,332	5,315	5,005	4,738	4,421
R^2	0.553	0.553	0.542	0.554	0.558	0.556	0.554

95% confidence intervals via percentile method. * p-value < 0.05

References

Colantone, Italo & Piero Stanig. 2018*a*. "Global competition and Brexit." *American political science review* 112(2):201–218.

Colantone, Italo & Piero Stanig. 2018*b*. "The trade origins of economic nationalism: Import competition and voting behavior in Western Europe." *American Journal of Political Science* 62(4):936–953.

Homola, Jonathan. 2019. "Are parties equally responsive to women and men?" *British Journal of Political Science* 49(3):957–975.

Laakso, Markku & Rein Taagepera. 1979. "Effective Number of Parties: A Measure with Application to Western Europe." *Comparative Political Studies* 12(1):3–27.