Online Appendix for "Proportional Representation and Right-Wing Populism: Evidence from Electoral System Change in Europe"

Michael Becher* Irene Menéndez González† Daniel Stegmueller‡

November 29, 2021

Contents

*School of Global and Public Affairs, IE University; IAST. michael.becher@ie.edu
†School of Global and Public Affairs, IE University.Irene.Menendez@ie.edu
‡Duke University. daniel.stegmueller@duke.edu

A. Data details

Parties and votes Table A.1 lists the right-wing populist parties included in the study. Recall that our study covers the nine EU member states, including the UK, that participated in all European elections between 1979 and 2014. As noted in main text, Luxembourg is excluded due to its size. The Synthetic control analyses furthermore exclude France, since no pre-reform data on immigration inflows is available. Subsection B.4 below provides synthetic control results when including France (and excluding immigration).

The terminology in the literature varies and, perhaps unavoidably, concepts remain somewhat contested (Golder 2016; Kitschelt 2007; Mudde 2007; Müller 2017; Muis and Immerzeel 2017; Zulianello 2020). But there is considerable conceptual overlap between different approaches, and empirically the literature broadly agrees in terms of identifying the universe of cases that fall under the umbrella of right-wing populism in the countries under study. Following Müller (2017) and similar to several other scholars, *right-wing populist* parties are defined as political parties that are populist (i.e., anti-elitist and anti-pluralist) and culturally conservative or exclusionary (Golder 2016). The core claim of populist parties is that "only some of the people are really the people. Think of Nigel Farage celebrating the Brexit vote by claiming it had been a 'victory for real people' " (Müller 2017: 22). This definition of right-wing populism closely resembles what Norris and Inglehart (2019) call authoritarian populism and Mudde (2007) calls radical populist right. It is intentionally broad rather than narrow (Kitschelt 2007) and identifies core elements present in all expressions of right-wing populism.

Empirically, right-wing populist parties are identified based on the agreement of two sources: (1) the Comparative Political Data Set (Armingeon et al. 2018: A3-4), which identifies right-wing populist parties from 1960 to 2016; (2) the list of authoritarian-populist parties based on the analysis of expert surveys in a two-dimensional space by Norris and Inglehart (Norris and Inglehart 2019: Table 7.2). The resulting list is also identical to the recent compilation of Zulianello (2020), except that the latter does not cover the earlier years of our study. Data on vote shares in European elections is from the Parliaments and Governments database (Döring and Manow 2019). Their dataset is also used to calculate the effective number of electoral parties used as an additional outcome variable in OA Table C.4.

Table A.1 shows that in Ireland and in the UK before 1994 the sources identify no viable RWP party competing in elections. This is substantively meaningful information that we use in the analysis. In the pre-reform UK, possible entrants trying to compete in the single-member district system faced a high electoral threshold. Similarly, Ireland is a case where the effective electoral threshold under its STV system—17.4 on average in the 2004 election (Farrell and Scully 2007: 75)—is higher than in most other countries represented in the European Parliament because it uses four electoral districts to elect a total of 15 MEPs. This imposes high hurdles for the entry of any new party. Theoretically, the effect of electoral system reform on vote shares may be driven by the extensive margin (i.e., party entry) and/or the intensive margin (i.e., the

Table A.1
Right-Wing Populist Parties and EU countries identified by various sources, 1979-2014.

	Party Name (English)	Abbreviation
Belgium	Flemish Bock	VB
Belgium	National Front	FN-NF
Denmark	Danish People's Party	DF
France	National Front	FN
Germany	Alternative for Germany	AfD
Germany	National Democratic Party	NPD
Germany	Republicans	REP
Germany	German People's Union	DVU
Greece	Independent Greeks	AE
Greece	People's Association – Golden Dawn	XO
Greece	Popular Orthodox Rally	LAOS
Ireland	none	
Italy	Northern League	LN
Italy	National Alliance (formerly Social Movement, MSI-DN)	AN
Italy	Brothers of Italy - National Centre-right	FdI-CN
Netherlands	Centre Democrats	CD
Netherlands	List Pim Fortuyn	LPF
Netherlands	Party for Freedom	PVV
UK	UK Independence Party	UKIP
UK	British National Party	BNP
UK	Anti-Federalist League (only Westminster 1992)	
UK	Referendum Party (only Westminster 1997)	

Note: See Online Appendix section A for the definition of right-wing populist parties and related literature. Empirically, parties are identified based on the agreement of two sources: Armingeon et al. (2018) and Norris and Inglehart (2019). While there are different conceptual treatments, there is broad agreement on the set of core RWP parties for the elections under study (1979 to 2014). Data on vote shares in European elections is from the Parliaments and Governments database (Döring and Manow 2019).

growth of an initially small or niche party). In game theoretic models, a party's decision to enter the electoral arena is endogenous to beliefs about its vote-winning potential given the existing party system, electoral rules and voting behavior (Cox 1997: ch. 8-9). That is, having no RWP party competing in a given election is theoretically and substantively meaningful. In election-years without a RWP party competing, their vote share is zero.

Other variables Information on electoral institutions for the European Parliament are from European Parliament Directorate General for Research (1997, 1999). In addition, Farrell and Scully (2007: ch.4) provide an excellent overview. Sources for covariates are as follows:

- Bi-annual survey data on Euroscepticism and Satisfaction with national democracy are from the European Commission's online database Eurobarometer Interactive (https://ec.europa.eu/commfrontoffice/publicopinion).
- The measure of Chinese import competition is from the Atlas of Economic Complexity (The Growth Lab at Harvard University 2019).
- The regional authority index (Marks et al. 2008) used in our DiD models is a summary measure of the authority of regional governments (for subnational units with a population of at least 150,000). It is obtained from the updated database of Hooghe et al. (2016).
- Mainstream party position on European integration are calculated from the Chapel Hill Expert Survey trend file (Bakker et al. 2020), for 1999-2014, and the earlier Ray and Marks/Steenbergen Party Dataset, for 1979-1994 (both retrieved from https://www.chesdata.eu/). Each dataset contains a variable measuring the overall orientation of the party leadership towards European integration, with responses ranging from 1 (strongly opposed) to 7 (strongly in favor). We calculate the mean for mainstream parties in a given country and year, using years close to the next EP election.
- Data on immigration inflows in each EP election period are from the United Nations Department of Economic and Social Affairs (UNDESA) immigration flow database (United Nations Populations Division 2015). Inflows are defined as the number of individuals (irrespective of their citizenship status) seeking to establish residence in the destination country in a given year. We aggregate yearly flow data to EP election periods. While this database provides us with the most comprehensive coverage, it lacks information for France in the pre-reform years. The consequence of this is that our synthetic control analysis reported in the main text excludes France. We provide an alternative analysis with France included in appendix B.4 and find no substantive difference in results. The panel difference-in-difference models reported in the main text indicate instances where France is excluded.
- The generosity of unemployment insurance is from the Comparative Welfare Entitlements Data Set (Scruggs et al. 2013).
- Our remaining economic and political controls are obtained from the Comparative Political Data Set (Armingeon et al. 2018).

If a covariate time series in a specific country has missing observations in a given year, we interpolate (predict) the time series using a Kalman filter from a flexible, country-specific local linear trend time series model (Harvey 1990).

B. Synthetic control analysis

Abadie et al. (2010) provides an excellent introduction to the synthetic control method (SCM). This exposition follows Becker and Kloessner (2018) and highlights that the SCM can be implemented using predictor variables at a more disaggregated time scale (i.e., we have covariates that vary within election periods) and employing an improved algorithm for calculating the required weights.

B.1. Method

Denote by Y_{tj} the right-wing populist vote share for country j out of $J+1$ countries at election t ($t = 1, \ldots, T$). The treated country (United Kingdom) is denoted by $j = 1$, the remaining $J = 8$ countries are possible donors for a synthetic UK. Denote the number of pre-reform elections by T^{pre} so that $Y_{1j} \ldots Y_{T^{pre}j}$ are observed prior to the reform, and $Y_{T^{pre}+1,j} \ldots Y_{Tj}$ are observed subsequently.

A SCM estimator models the effect of the reform using a weighted linear combination of optimally chosen donor countries representing counterfactual outcomes that would have obtained absent the change of electoral rules. More precisely, it generates $\hat{Y}_{T^{pre}+1,j} \ldots \hat{Y}_{Tj}$ which approximate the unobserved counterfactual post-reform outcomes $\tilde{Y}_{T^{pre}+1,j} \ldots \tilde{Y}_{Tj}$. In the post-reform period the effect of the reform on the tth observation (with $t > T^{pre}$) is given by $Y_{t1} - \tilde{Y}_{t1}$ approximated empirically by $Y_{t1} - \hat{Y}_{t1}$.

This empirical approximation is achieved by weighting donor countries by a vector of weights $W = (w_2, \ldots, w_{J+1})'$ (with weights constrained to be non-negative and sum to unity). The approximated outcome for $\tilde{Y}_{t1}$ is obtained as

$$\hat{Y}_{t1}(W) = \sum_{j=2}^{J+1} w_j Y_{tj} \tag{1}$$

The effect of the reform δ_t can then be estimated by the difference between actual and synthetic post-reform outcomes

$$\hat{\delta}_t(W) := Y_{t1} - \hat{Y}_{t1}(W). \tag{2}$$

Optimal weights would minimize the approximation error $\tilde{Y}_{t1} - \hat{Y}_{t1}(W)$. As the first term is an (unobservable) counterfactual, Abadie et al. propose to pursue two objectives. First, pursue the best possible pre-reform approximation $Y_{t1} - \hat{Y}_{t1}(W)$ by minimizing the root mean squared error

$$e_Y(W) := \sqrt{\frac{1}{T^{pre}} \sum_{t=1}^{T^{pre}} \left(Y_{t1} - \sum_{j=2}^{J+1} Y_{tj} w_j \right)^2}. \tag{3}$$

Second, in order for the post-reform counterfactual values to be approximated well, ensure that the synthetic controls also approximate a set of K variables that are predictive of right-wing vote shares . Denote by X_{ksj} the observed value of variable $k (k = 1, \ldots, K)$ for country $j (j =$

$1, \ldots, J+1)$ at time $s(s = 1, \ldots, S_k)$. We use a different time index to signify that these variables may be available at a more disaggregated time scale. The difference between observed values for the United Kingdom and its synthetic control approximation are given by $X_{ks1} - \sum_{j=2}^{J+1} X_{ksj} w_j$ for each covariate. Then the second quantity to be minimized is given by

$$e_X(V, W) := \sqrt{\sum_{k=1}^{K} v_k \frac{1}{S_k} \sum_{s=1}^{S_k} \left(X_{ks1} - \sum_{j=2}^{J+1} X_{ksj} w_j \right)^2}. \tag{4}$$

Here v_k are non-negative weights (collected in $V = (v_1, \ldots, v_K)'$ which capture the relative importance of each variable in predicting vote outcomes. Note that this criterion is purely *predictive*, not a reflection of the causal role of these variables. In our application, X includes changing political and economic conditions (such as, among others, support for the EU, unemployment rates, and Chinese import exposure) as well as lagged outcomes. Below, we present sensitivity analyses showing that results are not sensitive to the choice of pre-reform characteristics or the sequential exclusion of countries in the donor pool. As the lagged outcome variable, we include RWP vote shares in 1994 in order to capture the general upward trend in populist votes.

In terms of identifying the effect of interest, the smaller $e_X(V, W)$ and $e_Y(W)$ the smaller is the potential bias of the estimated $\hat{\delta}_t(W)$. See Abadie et al. (2010) for a detailed discussion of how $e_X(V, W) = e_Y(W) = 0$ enables the estimation of the effect of interest. To estimate this model (see Becker and Kloessner (2018) for full details), define the function W^* which maps covariate weights v_k onto weights for donor units minimizing the approximation error of covariates: $W^*(V) := \arg\min_W e_X(V, W)$. Then use W^* to define the corresponding approximation error for right-wing votes:

$$e_Y^*(V) := e_Y(W^*(V)) \tag{5}$$

To optimally determine V minimize (5) above w.r.t. $v_1, \ldots, v_K$ which produces optimal covariate weights $v_1^*, \ldots, v_K^*$ and country weights $W^*(v_1^*, \ldots, v_K^*)$.

B.2. Synthetic control details and results

Table B.1 shows the weights used in the SCM. Panel (a) displays weights attached to potential donor countries, panel (b) displays the weights assigned to covariates. Synthetic UK is predominantly composed of the Netherlands and Ireland. This deserves a couple of comments. First, we note that it is not unusual for the SCM to choose only a limited set of units from the pool of possible donors. For example, the study of Fowler (2013) uses SCM to examine the impact of the adoption of compulsory voting in Australia on turnout and social policy: In the analysis of pension spending, 2 out of 20 OECD countries receive most weight; in the analysis of turnout, 4 out of 20 countries receive most weight.[1] This is a feature, not a bug, of the approach that differs from traditional regression-based methods.

Secondly, the selection of Ireland and the Netherlands as key components of a synthetic UK is plausible. They are highly open economies with historical trade and political links to the UK. Ireland was the second Anglo-Saxon country in the EU at the time of the reform. As already noted, its electoral system—STV in small multi-member electoral districts—entails a high effective electoral threshold for new parties, similar to Britain's pre-reform system. While the effective threshold was lower in the Netherlands than in Ireland and the UK, it is politically relevant and high compared to national elections because of the the lower number of seats available to be allocated in the European electoral arena. Thus, the electoral weakness of RWP challengers in these countries in the early 1990s is consistent with institutional explanations. Similar to the UK, both countries were highly exposed to demand factors such as Chinese import competition and, subsequently, labor market competition from the EU's Eastern enlargement. In the Netherlands, a popular vote in 2005 soundly rejected the treaty to establish a Constitution for the EU. The country also saw a considerable rise of RWP parties, in particular the Party for Freedom (PVV). Moreover, the Netherlands does not use an open-list PR system but an ordered system in which there is only "limited scope for candidates to improve their list placement through personal votes" (Farrell and Scully 2007: 77). Recall that among the countries under study, only two (Ireland and Italy) have an open-list PR system (Farrell and Scully 2007: ch. 4), while the remaining ones employ a closed-list or ordered list system with limited scope for voters to change candidates' ranking on the party list.

That said, thirdly, the fact that some countries receive a zero weight in our baseline SCM does *not* mean that they are not a useful comparison case. It simply means that the method has found another combination of country weights that approximates pre-reform UK outcomes better. In line with this, we show below that our results are robust to sequentially excluding each possible donor country in turn. Our results are not driven by Netherlands or Ireland, though they get most of the weight because they provide the best fit for the pre-reform UK.

Panel (b) shows that the key predictor in X is the RWP vote share in 1994. This makes sense given the divergent trend in RWP that had emerged at the time. In previous applications of SCM, the lagged outcome variable often receives most weight (Becker and Kloessner 2018). It is important to note that this stage of the analysis is not a horse race between competing supply side explanations. Weights on covariates are determined solely by their potential to decrease a predictive criterion (cf. equation (4)). See Botosaru and Ferman (2019) for a detailed discussion of the role of covariates in synthetic control estimation. Substantively, these results simply indicate that beyond their impact on the strength of RWP parties in 1994, additional variables are not needed to construct a counterfactual UK. Below we show that alternative specifications that force different weight combinations on covariates produce comparable substantive results.

[1] Similarly, on country weights in other application, see for instance analyses of German reunification (Abadie et al. 2015) or terrorism in the Basque country (Abadie and Gardeazabal 2003; Becker and Kloessner 2018).

Table B.1
Entries of $W^*(V)$ and V matrices: estimated final weights of countries and additional predictor variables. Root mean squared error of approximation

(a) $W^*(V)$

Country	w_j
Belgium	0.018
Denmark	0.000
Germany	0.000
Greece	0.000
Ireland	0.229
Italy	0.000
Netherlands	0.753
RMSPE	0.0072

(b) V

Variable	v_k
Dissatisfaction with EU	0.000
Satisfaction with national democracy	0.000
Unemployment rate	0.000
Unemployment insurance generosity	0.000
Capital openess	0.000
Chinese imports	0.001
Right government share	0.000
Avg. party position EU policy	0.000
Migration inflow	0.000
RWP vote share in 1994	0.999
RMSPE	0.0383

Table B.2 provides estimates for the reform effect for each post-reform election with corresponding one-sided p-values based on placebo reform assignments to EU8 countries as shown in Figure 2 in the main text. Inference from synthetic controls has to be based on placebo tests, and when interpreting our p values one should keep in mind the small size of the sample. Nonetheless, Table B.2 signifies that the magnitude of the RWP vote share in the United Kingdom exceeds that of placebo-reform EU8 countries in every post-reform election.

Table B.2
Difference between United Kingdom and synthetic control for each post-reform election. One-sided placebo p-values

	European Election				
	1999	2004	2009	2014	1999-2014
$\Delta(Y_{1t}, Y_{0t})$	6.9	18.9	9.7	18.5	13.5
p-value	0.000	0.000	0.000	0.000	0.000

B.3. Specification tests

In the following two figures we present two sets of sensitivity analyses. In Figure B.1 we examine the influence of the choice of specific covariates. As our previous discussion indicated, the aim of SCM estimation is to minimize an objective criterion (RMSPE) not to produce a subjectively "meaningful" set of covariate weights. But a critical reader might nonetheless worry about the inclusion or exclusion of specific covariates, which changes the calculation of optimal weights, affects our conclusion (see Ferman et al. 2020 for a discussion of "cherry picking" of SCM covariates). We address this issue by forming *all* possible combinations of covariates in X and re-estimating the model. Figure B.1 plots estimates from 116 models representing various possible covariate combinations. Note that we excluded estimates from models that were decidedly worse in terms of their match between the synthetic control group and the pre-reform data (where the ratio of $e_X(V, W)$ relative to the main model is > 20).

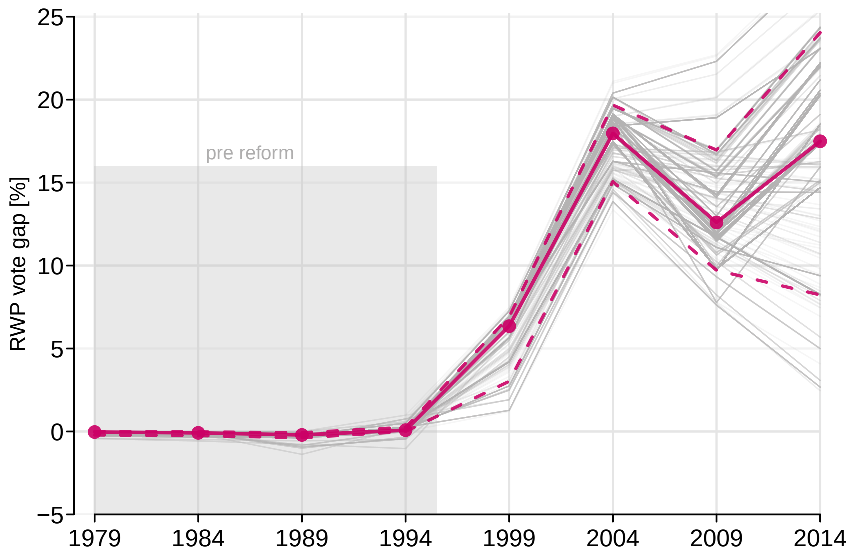

Figure B.1

Sensitivity analysis of synthetic control predictors. Average estimate and 95% bounds.

This figure plots RWP vote share gaps estimated from a series of models using all possible combination of covariates in the SCM estimator. The bold magenta line represents the average of all estimated models; dashed lines represent 95% bounds based on 116 estimates of possible covariate combinations. The bounds and average estimate signify that the key SCM result does not depend on the specific combinations of predictors added to the model. They show a pattern similar to the result in the main text: a clear increase in the RWP vote with the introduction of PR with an even greater increase in 2004.

Figure B.2 studies the impact of removing a country from the synthetic control donor pool. Removing a country induces different country weights as well as different weights on covariates. As a result, the approximation of the synthetic UK case to its observed counterpart in the pre-reform period can vary considerably (for example, it is worse when excluding the

Netherlands from the donor pool). Despite variation in the closeness of the approximation, we find the core pattern of our results generally confirmed. All panels of Figure B.2 show a close match between the United Kingdom and its synthetic counterpart in the pre-reform period.

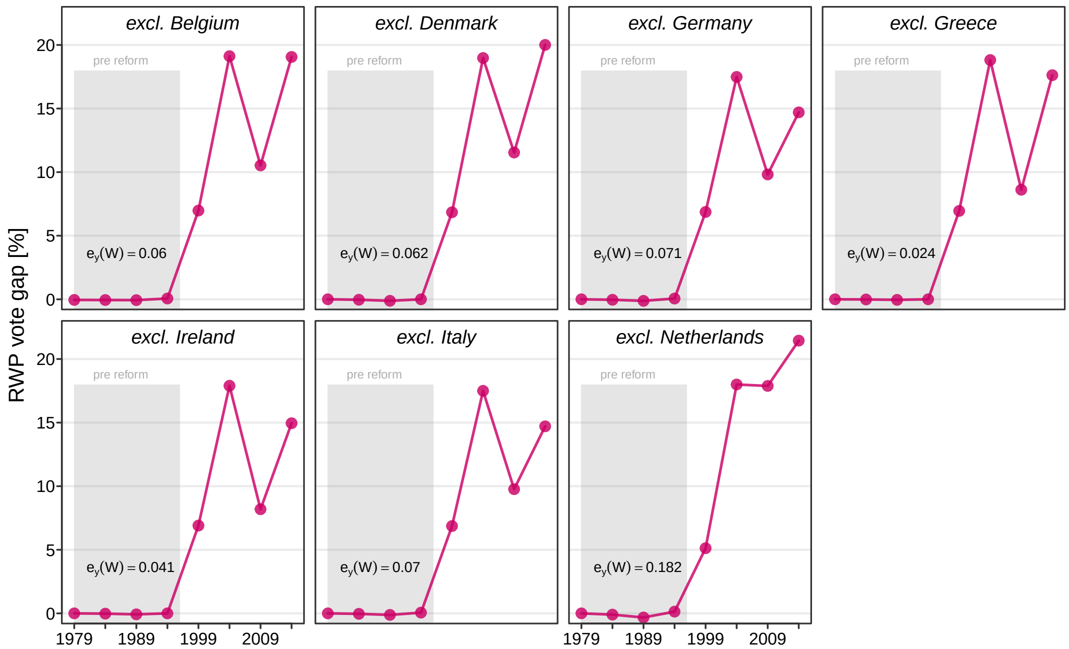

Figure B.2

Resulting SCM estimates when excluding a country from donor pool

This figure plots estimates from models excluding countries from the set of possible synthetic control donors. The magenta line shows resulting estimates. $e_Y(W)$ shows the root mean squared error of approximation in the pre-reform period.

B.4. Extended sample including France

Our SCM analyses reported in the main text excludes France due to missing pre-reform observations of immigration inflows. In this subsection, we present an alternative analysis that includes France (and consequently excluded immigration as a covariate). The right panel of Figure B.3 shows synthetic control estimates when including France while the left panel simply replicated the synthetic control used in the main text. A quick visual comparison of both panels reveals that the broad pattern of results remains unchanged. More specifically, results for the first two post-reform elections in 1999 and 2004 are virtually identical. In later elections (2009 and 2014) we obtain lower synthetic control values when excluding immigration inflows. Thus, the implied gap between the UK and its synthetic counterpart is somewhat larger. In this sense, the specification including immigrations is slightly more conservative and we thus chose it as the one presented in the main text.

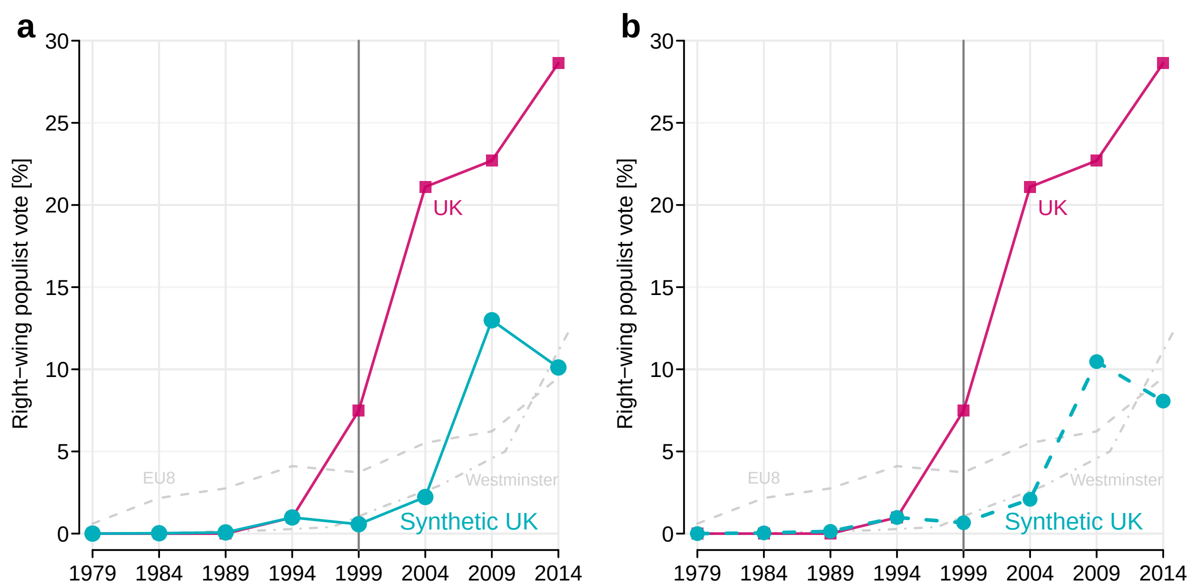

Figure B.3

Comparison of synthetic control UK estimates when including France.

Panel (a) shows the synthetic UK used in the main text, which includes immigration inflows as a covariate (which necessitates the exclusion of France). Panel (b) shows the synthetic UK estimate obtained when including France (and excluding immigration inflows as a covariate).

C. Difference-in-difference analyses

We conduct a set of additional analyses using difference-in-difference (DiD) specifications to complement our SCM results. These provide estimates of the electoral reform impact using different identifying assumptions. The following section discusses the difference in assumptions made between the synthetic control approach and DiD estimates.

Before doing so, we present a simple difference in means comparison between the UK and the EU8. Table C.1 shows mean support for populist parties [in percent] in the United Kingdom and the EU8 average before and after the reform in the UK, as well as the pre-post difference and the difference-in-differences. This calculation on the raw data shows the relatively larger magnitude of the change in the UK compared to the EU where the electoral rules stay constant. The difference of the differences (shown in the bottom right of Table C.1) provides an estimate close to the the baseline two-period, two-group DiD regression reported in the main text.

Table C.1
Support for right-wing populist parties (in percent) in the United Kingdom and EU8 before and after UK reform. Means and differences with standard errors in parentheses.

	Before reform	After reform	Difference
EU8	3.1	8.5	5.4
	(1.1)	(1.1)	(1.6)
United Kingdom	0.2	20.0	19.7
	(3.1)	(3.1)	(4.4)
Difference	−2.9	11.4	*14.3*
	(3.3)	(3.3)	*(4.7)*

C.1. Identifying assumptions in SCM and DiD approaches

This subsection discussed differences in identifying assumptions between synthetic control models and difference-in-difference analyses. We use a simple (parametric) model to focus on the core ideas. For each country i ($i = 1, \dots, n$) we observe T time periods, of which $t_1, \dots T_0$ are pre-reform periods and $T_0 + 1, \dots, T$ are post-reform. Denote by Y_{it}^1 and Y_{it}^0 the respective potential outcomes (vote shares) of country i at time t under either electoral reform or no-reform conditions. Reform status is denoted by an indicator variable, D_{it}, which is equal to 1 if country i at time t is exposed to the electoral reform and 0 otherwise. The potential outcome absent an electoral reform can be written as (Angrist and Pischke 2008):

$$Y_{it}^0 = X_{it}\beta + \lambda_t \mu_i + \delta_t + \epsilon_{it}. \tag{6}$$

Denoting by τ_{it} the (additive) effect of the reform, the potential outcome in the reform state is:

$$Y_{it}^1 = X_{it}\beta + \lambda_t \mu_i + \delta_t + \tau_{it} + \epsilon_{it}. \tag{7}$$

Here, X_{it} is a vector of observed covariates, μ_i are unobserved country characteristics which are time-constant but can have time-specific impacts captured by their associated coefficients λ_t; δ_t captures common time shocks; and ϵ_{it} are unobserved idiosyncratic shocks.

If we assume that the reform only affects subsequent elections in the country implementing it, we can write observed vote shares by the well-known switching equation:

$$Y_{it} = D_{it} Y_{it}^1 + (1 - D_{it}) Y_{it}^0. \tag{8}$$

We are interested in the average treatment effect on the UK in each post-reform period:

$$\tau_t = E(Y_{it}^1 - Y_{it}^0 | D_{it} = 1). \tag{9}$$

If assignment to 'reform status' and vote shares are both influenced by μ_i then unobserved country characteristics are a confounder that may bias the estimated effect of the reform. More precisely, bias occurs when μ_i is imbalanced between the UK and other countries and $\lambda \neq 0$. Estimating τ_t necessitates making an assumption about Y_{it}^0—the outcome that would have occurred absent the reform. Since this is an *unobservable* quantity, the validity of this identifying assumption cannot be verified. Thus, our aim here is to employ two different assumptions and examine to what extent the resulting effect estimates agree. The two assumptions discussed next differ in what has to be conditioned upon in order to ensure that potential outcomes in the control condition Y_{it}^0 are independent of reform assignment.

Parallel trends assumption Assume that the change in Y^0 from time period t to t' is independent of the assignment to the reform country (UK) or unchanged countries (EU8), after conditioning on observable variables Abadie (2005): $E(Y_{it}^0 - Y_{it'}^0 | D_{it} = 1, X_{it}) = E(Y_{it}^0 - Y_{it'}^0 | D_{it} = 0, X_{it})$. In terms of our model above (see eq.6), unobserved country characteristics μ_i maybe be imbalanced between reform and non-reform groups, but their effect has to be constant over time ($\lambda_t = \lambda$). More compactly, we write

$$Y_{it}^0 \perp D_{it} | (X_{it}, t, \lambda\mu_i) \tag{10}$$

which encodes the assumption that the outcome in the control condition is independent from reform assignment after conditioning on observed covariates and both time and country fixed effects. The difference-in-difference analysis presented in the main text relies on this assumption (although it can be somewhat weakened with more sophisticated specifications).

Conditional independence (or ignorability) assumption A different assumption is that the expected values for the potential control outcome Y_{it}^0 for both reform and non-reform countries is the same in expectation after conditioning on observed covariates and *past outcomes* (Angrist and Pischke 2008). More formally,

$$Y_{it}^0 \perp D_{it} | (X_{it}, Y_{ih}^0). \tag{11}$$

Here, Y_{ih}^0 is a vector of potential outcomes in h time periods before T_0. This assumption implies that countries with similar outcomes before the reform ($Y_{it}^0, t = 1, \ldots, T_0$) are expected to have similar potential control outcomes after it ($Y_{it}^0, t = T_0 + 1, \ldots, T$) after conditioning on covariates. The synthetic control approach used in our main text relies on this assumption.

C.2. Test for pre-reform non-parallel trends

Table C.2 tests for non-parallel trends in the pre-reform period. The test is based on two-way fixed effects models (with country and time fixed effects) including reform-time interactions in linear and quadratic form. We calculate cluster-robust p-values for these interactions specified without and with covariates. Overall, we find no evidence to reject the null hypothesis of parallel trends, especially when conditioning on the set of covariates we use in Table I in the main text.

Table C.2
Tests for pre-reform period non-parallel trends

	Linear trend		Quadratic trend	
	p_{rob}	p_{wild}	p_{rob}	p_{wild}
Test without covariates	0.102	0.034	0.429	0.324
Test with covariates	0.483	0.171	0.911	0.672

Note: Test based on generalized DiD model with country and time fixed effects. Entries are p values from tests of reform × linear time-trend and reform × quadratic time-trend interactions calculated using robust (HC3; p_{rob}) and cluster-wild-bootstrapped (p_{wild}) standard errors. Covariates include a regional authority index, capital openness index, Chinese import penetration, the generosity of unemployment insurance, Euroskepticism among citizens, and average party position on EU (of non-RWP parties).

C.3. Placebo DiD analysis

Figure C.1 plots the distribution of estimates from a placebo difference-in-difference analysis. We use specification M3 of Table I including devolution, economic integration, welfare generosity, and Euroscepticism and successively assign reform status to EU8 countries that did not actually experience a switch to PR. The median of placebo estimate is −4.3 percentage points, which is of the opposite sign of the UK estimate and more than three times smaller in terms of absolute size. Furthermore, the estimate of for the UK is clearly far in the tails of the distribution of placebo estimates. Thus, our placebo analysis suggests that the estimated size of the impact of the UK reform is unlikely to be found by chance.

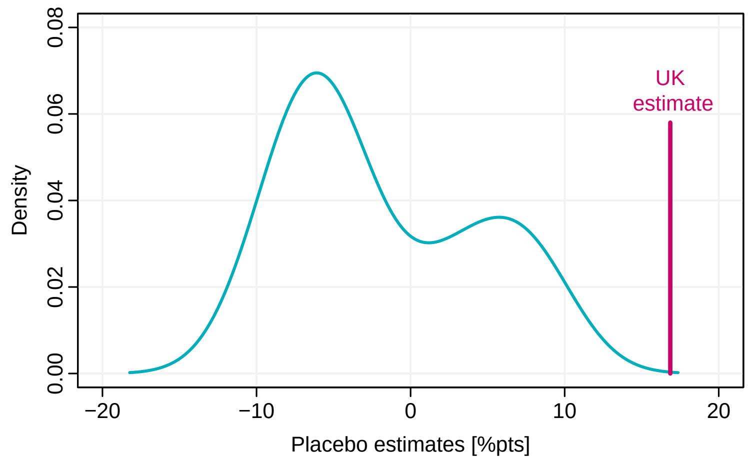

Figure C.1
Placebo difference-in-difference analysis.

Placebo estimates are generated by successively assigning reform status to non-reform countries. The median of placebo estimates is −4.3 percentage points. The figure visually illustrates the distribution of placebo estimates (using a Gaussian kernel density smoother) compared to the UK estimate. It suggests that the estimated size of the effect of the UK reform is unlikely to be found by chance.

D. Electoral reform and Euroscepticism

Table D.1 tests whether the reform had an effect on public attitudes toward the EU, measured as the share of respondents in a Eurobarometer survey who say that EU is a bad thing for their country, using the same difference-in-difference estimators employed for the analysis of RWP votes in the main text. Canonical theories of strategic electoral coordination under alternative electoral institutions imply that the reform made it much easier for voters to switch their support to new political parties critical of the political establishment without wasting their vote. These theories assume that political preferences are relatively stable and highlight that changing the rules of the game affects the strategic behavior of parties and voters. In reality, preferences may be affected as well. While a full analysis of the mechanisms requires individual-level panel data that are not available, the additional analyses using aggregate data summarized in Table D.1 do not reveal evidence that the reform changed public preferences toward the EU. We find no strong or statistically significant effect of the reform on aggregate public opinion toward the EU. This suggests that the reform did not work mainly through altering mass preferences on the European Union and is consistent with the strategic coordination channel of institutional theories.

Table D.1
Difference-in-difference estimates of electoral reform and Euroscepticism.

M1	Panel DiD (assuming parallel trends)	0.322	(0.196)
M2	Panel DiD, parallel trends conditional on covariates		
	Devolution (regional authority index)	0.335	(0.244)
	+ Economic Integration + welfare generosity	0.115	(0.189)
	+ Immigration inflows	0.137	(0.362)

Notes: Estimates with standard errors in parentheses. *M1:* Time-average ATT from multi-period DiD panel estimator. N=72. *M2* uses an estimator relaxing the parallel trends assumption conditional on covariates (Callaway and Sant'Anna 2020). Covariates include a regional authority index, capital openness index, Chinese import penetration, generosity of unemployment insurance, and immigration inflows (the latter limits the analysis period to 1984–2014 and excludes France due to missing immigration inflow data). See appendix for more details. Bootstrapped standard errors using 5000 replicates.

E. Electoral reform and the effective number of parties

In additional difference-in-difference analyses displayed in Table E.1, we use the effective number of electoral parties (ENEP) in EP elections as the dependent variable. It is calculated based on vote shares from the Parliaments and Governments database (Döring and Manow 2019). ENEP has been widely studied in the literature and does not require analysts to classify parties as populist or not. As noted in the main text, theories of electoral coordination imply that the electoral reform under study makes the electoral system more permissive to the entry and growth of new parties. Under the uncontroversial assumption that party competition under the original first-past-the-post system involves meaningful national issues and party labels, for instance, Gary Cox's work implies that the reform increases the upper limit on the viable number of parties (Cox 1997). Indeed, this is what we find using the same flexible model specification, which relaxes the parallel trends assumption, used to study RWP votes. The estimates in Table E.1 suggest that the switch to PR increased the effective number of electoral parties by approximately one. This holds whether covariates are included or not. Consistent with the theory, this finding indicates that impact of the reform on RWP votes reflects an increase in the size of the party system, not the substitution of one party by another.

Table E.1
Difference-in-difference estimates of effective number of electoral parties in post-reform election.

M1 Panel DiD (assuming parallel trends)	0.946	(0.336)
M2 Panel DiD, parallel trends conditional on covariates		
Devolution (regional authority index)	1.036	(0.320)
+ Economic Integration + welfare generosity	0.991	(0.327)
+ Immigration inflows	1.083	(0.219)

Notes: Estimates with standard errors in parentheses. *M1:* First period ATT from multi-period DiD panel estimator. N=72. *M2* uses an estimator relaxing the parallel trends assumption conditional on covariates (Callaway and Sant'Anna 2020). Covariates include a regional authority index, capital openness index, Chinese import penetration, generosity of unemployment insurance, and immigration inflows (the latter limits the analysis period to 1984–2014 and excludes France due to missing immigration inflow data). See appendix for more details. Bootstrapped standard errors using 5000 replicates.

Online appendix references

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies 72*(1), 1–19.

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association 105*(490), 493–505.

Abadie, A., A. Diamond, and J. Hainmueller (2015). Comparative politics and the synthetic control method. *American Journal of Political Science 59*(2), 495–510.

Abadie, A. and J. Gardeazabal (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review 93*(1), 113–132.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly Harmless Econometrics.* An Empiricist's Companion. Princeton University Press.

Armingeon, K., V. Wenger, F. Wiedemeier, C. Isler, L. Knöpfel, D. Weisstanner, and S. Engler (2018). Comparative political data set 1960-2016. Technical report, Institute of Political Science, University of Berne.

Bakker, R., L. Hooghe, S. Jolly, G. Marks, J. Polk, J. Rovny, M. Steenbergen, and M. A. Vachudov (2020). 1999-2019 chapel hill expert survey trend file. (version 1.2). Technical report, Chapel Hill.

Becker, M. and S. Kloessner (2018). Fast and reliable computation of generalized synthetic contols. *Econometrics & Statistics 5*, 1–19.

Botosaru, I. and B. Ferman (2019). On the role of covariates in the synthetic control method. *The Econometrics Journal 22*(2), 117–130.

Callaway, B. and P. H. Sant'Anna (2020). Difference-in-differences with multiple time periods and an application on the minimum wage and employment. *Journal of Econometrics in press.* arXiv:1803.09015.

Cox, G. W. (1997). *Making Votes Count.* New York: Cambridge University Press.

Döring, H. and P. Manow (2019). Parliaments and governments database (parlgov): Information on parties, elections and cabinets in modern democracies. http://www.parlgov.org/.

European Parliament Directorate General for Research (1997). Legislation governing elections to the european parliament. http://www.europarl.europa.eu/workingpapers/poli/w13/country_en.htm. European Parliament, Luxembourg: Working Paper Political series Working Paper Political series W 13.

European Parliament Directorate General for Research (1999). Legislation governing elections to the european parliament. http://www.europarl.europa.eu/RegData/etudes/etudes/join/1999/

166617/IPOL-AFCO_ET(1999)166617_EN.pdf. European Parliament, Luxembourg. Working Document Political Series 01.1999.

Farrell, D. M. and R. Scully (2007). *Representing Europe's Citizens? Electoral Institutions and the Failure of Parliamentary Representation.* Oxford: Oxford University Press.

Ferman, B., C. Pinto, and V. Possebom (2020). Cherry picking with synthetic controls. *Journal of Policy Analysis and Management 39*(2), 510–532.

Fowler, A. (2013). Electoral and policy consequences of voter turnout: Evidence from compulsory voting in australia. *Quarterly Journal of Political Science 8*(2), 159–182.

Golder, M. (2016). Far right parties in europe. *Annual Review of Political Science 19*, 477–497.

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter.* Cambridge: Cambridge University Press.

Hooghe, L., G. Marks, A. H. Schakel, S. Chapman-Osterkatz, S. Niedzwiecki, and S. Shair-Rosenfield (2016). *Measuring regional authority. Volume I: a postfunctionalist theory of governance.* Oxford: Oxford University Press.

Kitschelt, H. (2007). Growth and persistence of the radical right in postindustrial democracies: Advances and challenges in comparative research. *West European Politics 30*(5), 1176–1206.

Marks, G., L. Hooghe, and A. H. Schakel (2008). Measuring regional authority. *Regional & Federal Studies 18*(2-3), 111–121.

Mudde, C. (2007). *Populist Radical Right Parties in Europe.* Cambridge: Cambridge University Press.

Muis, J. and T. Immerzeel (2017). Causes and consequences of the rise of populist radical right parties and movements in europe. *Current Sociology 65*(6), 909–930.

Müller, J.-W. (2017). *What is Populism?* (2nd with new Afterword ed.). Penguin Books.

Norris, P. and R. Inglehart (2019). *Cultural Backlash: Trump, Bexit, and Authoritarian Populism.* Cambridge, U.K.: Cambridge University Press.

Scruggs, L., D. Jahn, and K. Kuitto (2013). Comparative welfare entitlements data set. Technical Report V2, University of Connecticut.

The Growth Lab at Harvard University (2019). International trade data (sitc, rev. 2). Technical Report V3, Harvard Dataverse.

United Nations Population Division (2015). International migration flows to and from selected countries: The 2015 revision (POP/DB/MIG/Flow/Rev.2015). Department of Economic and Social Affairs.

Zulianello, M. (2020). Varieties of populist parties and party systems in europe: From state-of-the-art to the application of a novel classification scheme to 66 parties in 33 countries. *Government and Opposition 55*(2), 327–347.