How Ideology, Economics and Institutions Shape Affective Polarization in Democratic Polities

Noam Gidron,[1] James Adams,[2] and Will Horne[3]

Abstract

While many studies analyze affective polarization in the US public, i.e., partisan voters' tendencies to dislike the out-party, we lack comparative knowledge about the factors that intensify affective polarization cross-nationally. We analyze national election survey data from 20 polities between 1996 and 2015 to understand how mass-level affective polarization is related to elite-level ideological polarization, to economic conditions, and to political institutions. We conclude that affective polarization is more intense where unemployment and income inequality are high, and in countries with majoritarian political institutions. We find only inconsistent evidence that elite-level ideological polarization is a source of mass-level affective polarization. Our findings illuminate the political consequences of the recent world-wide recession, and speak to debates about political institutions and mass-elite linkages.

[1] noam.gidron@mail.huji.ac.il
[2] jfadams@ucdavis.edu
[3] rhorne@princeton.edu

Concerns over the health of democratic norms and institutions have intensified in recent years (e.g., Levitsky and Ziblatt 2018; Lieberman et al. 2018; Mounk 2018), with political polarization emerging as a key driver of democratic dysfunction (Somer and McCoy 2018). The rise of *affective polarization* in the mass public, defined as dislike and hostility across partisan lines (Iyengar et al. 2012; Lelkes 2016), is especially disconcerting since it fractures the social fabric by damaging cooperation and trust across partisan lines. In this regard Levitsky and Ziblatt (2018, 220) observe that "the fundamental problem facing American democracy remains extreme partisan division – one fueled not just by policy differences but by deeper sources of resentment." Although research on affective polarization focuses heavily on the United States, there is abundant evidence of this phenomenon across western democracies. As German President (and former Vice-Chancellor) Frank-Walter Steinmeier noted in his 2018 Christmas address:

> "Wherever you look – especially on social media – we see hate; there is shouting and daily outrage. I feel that we Germans are spending less and less time talking to each other…What happens when societies drift apart, and when one side can barely talk to the other without it turning into an all-out argument, is all too evident in the world around us: burning barricades in Paris, deep political rifts in the United States and anxiety in the United Kingdom ahead of Brexit."[4]

To date there is relatively little cross-national research on mass-level affective polarization (for exceptions see Carlin and Love 2018; Huddy et al. 2018; Lauka et al. 2018; Westwood et al. 2017; Reiljan forthcoming; Wagner 2017). However, a comparative framework can illuminate factors that drive affective polarization cross-nationally, which may in turn inform policy-makers in the US (and elsewhere) who seek remedies for affective polarization. That is what we provide here. Based on analyses of 76 national election surveys across 20 countries between 1996 and 2015, we analyze three possible drivers of affective polarization across western polities: ideological polarization

---

[4] https://www.dw.com/en/frank-walter-steinmeier-our-democracy-is-as-strong-as-we-make-it/a-46849802

among political elites; adverse economic conditions (income inequality and unemployment), and majoritarian political institutions that concentrate policy-making authority in the hands of a single party that may lack majority public support. Based on comparisons across countries, and also within countries over time, we report four findings.

First, we find that *affective polarization in the United States is not unduly high in comparative perspective.* Several countries (including Spain, Greece and France) display markedly more intense mass-level affective polarization than the US, while many additional countries (including Austria, Britain, Canada, and Portugal) display affective polarization levels similar to the US.

Second, we find that adverse economic conditions intensify affective polarization in that – in comparisons between countries – countries with higher unemployment rates display more intense affective polarization, and – in comparisons within countries over time – affective polarization intensifies as unemployment increases. We also find that affective polarization is more intense in countries with greater income inequality, although we cannot precisely estimate income inequality effects in temporal analyses within countries, because inequality tends to change slowly over time.

Third, we find that countries with majoritarian institutions that concentrate policy-making authority in the hands of a single party (or multiple parties that lack majority popular support) display sharply more intense affective polarization than countries featuring consensual institutions that tend to disperse power more broadly. In particular, the countries in our study that feature single-member districts that tend to concentrate political power (France, Canada, Great Britain, Australia, and the US) all display above-average levels of affective polarization. We also find that other features of majoritarianism (including a stronger executive and a more unitary, less federal system) are associated with affective polarization. While we hesitate to conclude that majoritarian institutions cause affective polarization (due to endogeneity issues discussed below), this relationship is consistent

with Lijphart's argument that "kinder, gentler" consensual institutions promote positive democratic outcomes (1999, chapter 12).

Fourth, we engage with debates in American politics about whether elite ideological polarization drives mass affective polarization (see Abramowitz and Webster 2017; Lelkes 2018; Mason 2015; Mason 2018) and report relatively weak and inconsistent findings, in both analyses within countries over time or in analyses that additionally account for comparisons between countries. Thus, our analyses do not imply that elite-level polarization – at least on the overarching Left-Right dimension – is a major driver of mass-level affective polarization.

Our findings carry implications for mass-elite linkages, and the consequences of economic conditions and political institutions. Specifically, our results speak to the effects of the global economic crisis beginning in 2008-2009 on affective polarization across western publics. Our analyses imply that the unemployment spikes across this period have largely driven the intense affective polarization we observe in Greece, Portugal, and Spain, three of the countries hit hardest by the financial crisis. And while unemployment has subsequently declined in most western democracies, these recoveries have in some cases been accompanied by increased income inequality (e.g., on the American recovery, see Saez 2012), which may also intensify affective polarization.

In addition, our work comparatively extends previous theorizing on American mass-level polarization. The US represents the most widely studied example of mass- and elite-level polarization over the past 40 years – yet it is not the only case. It is important to analyze affective polarization across western publics, first because non-US cases are intrinsically important, and second because the comparative patterns we identify may illuminate causes of affective polarization in the United States. In particular, our cross-national analyses of the effects of economic conditions and elite polarization speak to arguments American politics scholars have advanced to explain intensify-

4

ing mass-level polarization (Fiorina and Abrams 2008; Hetherington 2009; Lelkes 2016; McCarty et al. 2006). Our study follows the call for more comparative research on affective polarization (Iyenger et al. 2019; on comparative polarization more generally, see Lupu 2015), which we hope illuminates the causes of affective polarization both inside and outside the US.

What Drives Affective Polarization? Hypotheses

Affective polarization in the United States, defined as mass-level hostility across partisan lines (Lelkes 2016), has sharply intensified since the 1970s. For instance, Iyengar et al. (2012) document that the proportion of Americans who state that they would be displeased if their child married someone from the other party has increased from about 5% in the 1960s to nearly 50% by 2010. Iyengar and Westwood (2014) note that by the mid-2010s, negative affect based on partisanship was as strong as negative affect based on race. One leading scholar even argues that "partyism [prejudice toward partisan opponents] is now worse than racism" (Sunstein 2015, 2).

Affective polarization erodes the civic foundations of democratic institutions. It prompts preferential treatment of co-partisans (Lelkes and Westwood 2017), and polarized partisans are more likely to discriminate against partisan opponents in economic transactions (McConnell et al. 2018). Polarization also diminishes trust in government, especially when the opposing party holds power (Hetherington and Rudolph 2015), which speaks to the deadlock of US politics since the mid-2000s. In cross-national experiments Carlin and Love (2018) find that perceived partisan ideological polarization is strongly associated with lack of cooperation across party lines in economic decision-making games. And a June 2018 Rasmussen public opinion poll, showing that thirty-one

percent of American likely voters believe another civil war is likely in the next five years, plausibly reflects this intense mass-level affective polarization.[5]

The rise of mass-level affective polarization in the US may be driven in part by idiosyncratic features, including the rise of American partisan media – notably the consolidation of a conservative media universe led by Fox News – that intensify partisans' negative affect toward partisan opponents (e.g., Levendusky 2013; Lau et al. 2017). However American scholars identify additional factors that may generalize across western democracies. Here we review these studies and formulate testable hypotheses about the possible drivers of affective polarization.

Elite ideological polarization

The first factor that may intensify affective polarization in the mass public is the growing *elite-level ideological polarization* in the United States beginning in the 1970s, whereby Democratic and Republican elites have sharply diverged from each other with respect to their policy behavior, and have also clearly sorted themselves into opposing ideological camps (Hetherington 2009). This elite polarization, which is documented in analyses of congressional roll-call votes (McCarty et al. 2006) and in studies of the patterns of financial donations to political campaigns (Bonica 2014), has been visible to the US public: Hetherington (2009) documents that survey respondents' party placements have increasingly diverged since the 1970s, with citizens ascribing increasingly conservative positions to the Republican Party and more liberal positions to the Democrats.

Some scholars argue that elite ideological polarization and mass-level affective polarization are closely interlinked (Abramowitz and Webster 2017). Rogowski and Sutherland (2015), for instance, report experimental findings that stronger ideological differences between American candi-

---

[5] http://www.rasmussenreports.com/public_content/politics/general_politics/june_2018/31_think_u_s_civil_war_likely_soon

dates (and officeholders) drive affective polarization among citizens. Carlin and Love (2018) draw on social psychology research to explain the mechanism: when partisans view their opponents as opposing their core beliefs, they tend to attribute negative traits to the other party and experience negative emotions toward them (Fiske et al. 2007, 81). In this regard, Zakharova and Warwick (2015) report cross-national analyses of surveys from 14 democracies demonstrating that citizens ascribe more negative character traits to parties with whom they have ideological disagreements.

Other scholars, however, question the relationship between ideological and affective polarization (Lelkes 2018 reviews this debate). Mason (2018) suggests that affective polarization in the US is primarily driven by the growing overlap of religious, racial, and partisan identities, not ideology. Thus, rival parties' partisans may share similar policy stands, or lack coherent policy preferences (Kinder and Kalmoe 2017), but still dislike their political opponents.

The studies summarized above motivate our first hypothesis:

H1 (The elite ideological polarization hypothesis). Elite ideological polarization intensifies mass-level affective polarization.

National economic conditions

Scholars postulate that economic inequality may intensify affective polarization. Levitsky and Ziblatt (2018, 228-229) suggest, for instance, that reducing inequality could decrease affective polarization in the US. Hitlin and Harkness (2017, chapter 6) outline the causal mechanism, arguing that economic inequality "begets negative moral emotions," and is socially divisive because it prompts envy towards the top and scorn toward those at the bottom. Since envy and scorn are associated with anger, we might expect economic inequality to generate affective polarization in politics.

7

Moreover, other scholars argue that income inequality intensifies elite ideological polarization, which as discussed above may intensify affective polarization. McCarty et al. (2006) link rising inequality in the United States to increased elite ideological polarization (see also Mickey et al. 2017). By contrast, Fenzi (2018) argues that economic inequality prompts elite ideological <u>moderation</u>, because inequality decreases participation among the poor, which motivates center-left parties in Europe to court middle-class, centrist voters by adopting centrist positions (see also Iversen and Soskice 2015).

The Hitlin and Harkness (2017) study summarized above implies that income inequality directly intensifies mass-level affective polarization, while the studies discussed in the preceding paragraph suggest that income inequality may indirectly influence affective polarization via its effect on elite ideological polarization. These studies motivate the following hypothesis about the direct effects of income inequality, when controlling for elite ideological polarization:

H2a (<u>The income inequality hypothesis</u>). Income inequality intensifies affective polarization, in analyses that control for elite-level ideological polarization.

Scholars highlight additional economic conditions that may influence affective polarization. Research documents a direct relationship between weaker national economic conditions and declining political trust and satisfaction with democracy (Clarke et al. 1993; Gilley 2006) – attitudes that may be related to partisans' dislike for out-parties, and hence to affective polarization. With respect to indirect effects, Lopez and Ramirez (2004) argue that economic downturns – especially rising unemployment – prompt elite ideological polarization, as right-wing parties call for stimulus via *lower* taxation and spending, while left-wing parties demand a stimulus via *higher* taxes and spending. Finally, Funke et al. (2016) and Eichengreen (2018) find that economic downturns prompt the rise of radical populist parties, which exacerbates elite ideological polarization. These considera-

tions prompt our hypothesis about the direct effects of economic downturns, in analyses that control for elite ideological polarization:

H2b (The unemployment hypothesis). Unemployment intensifies affective polarization, in analyses that control for elite-level ideological polarization.

Political institutions

Lijphart (1999) distinguishes between majoritarian political institutions such as disproportional, plurality-based voting systems that tend to concentrate policy-making into the hands of a single party – one that often lacks majority popular support – versus consensual institutions, including proportional voting systems that disperse policy-making authority among multiple parties that collectively enjoy broader support.[6] Lijphart (1999, chapter 16) argues and presents empirical evidence that the publics in the "kinder, gentler" consensual systems are more satisfied with democracy, while Anderson and Guillory (1997) show that the degrees of democratic satisfaction expressed by supporters of "winning" versus "losing" parties in democratic competition are more similar in consensual democracies. Along similar lines, McCoy and Somer's (2019, 261) comparative study of eleven polarized countries concludes that the most extreme cases of polarization "emerge in contexts of majoritarian electoral systems that produce a disproportionate representation of the majority or plurality party." These considerations suggest that inter-party hostility and partisans' resentment towards out-parties may be most intense in majoritarian systems.

---

[6] As discussed below the distinction between majoritarian/consensual institutions rests on additional factors besides electoral systems including federalism, executive-legislative relations and interest group pluralism. The British "Westminster" system is the archetype majoritarian democracy, while the Netherlands and Switzerland are archetypical consensual democracies.

In addition, political institutions may also indirectly influence affective polarization via their effects on income inequality and on elite ideological polarization. Lijphart (1999, chapter 16) finds that more consensual systems display significantly less income inequality than majoritarian systems (see also Iversen and Soskice 2006), while Dow (2011) finds that rival parties tend to be more ideologically polarized in more proportional, consensual systems (but see Ezrow 2008).

The above considerations motivate our hypothesis about the relationship between political institutions and affective polarization:

H3 (The majoritarian institutions hypothesis). Affective polarization is more intense in majoritarian political systems, in analyses controlling for economic conditions and for elite polarization.

We note that some scholars argue that institutional design is endogenous to existing power balances between political actors (e.g., Boix, 1999, North 1990), in particular that countries with less power sharing and more intergroup hostility at the time institutions are chosen – which may reflect social cleavages related to ethnicity, religion, or geography – tend to adopt more majoritarian institutions (for an overview see Colomer 2018). This is why H3 posits a statistical – not a causal – relationship between majoritarian institutions and affective polarization. At the same time, the political institutions in the countries we analyze were largely chosen before 1950, which weakens the likelihood they are endogenous to current affective polarization levels. To account for possible endogeneity issues, we will estimate models both with and without institutional variables.

Data and Measures

We analyze whether affective polarization in western publics is more pronounced when party elites are more ideologically polarized (the elite ideological polarization hypothesis), when income ine-

quality and unemployment are higher (the income inequality and the unemployment hypotheses), and where political institutions are more majoritarian (the majoritarian institutions hypothesis).

Dependent variable: Affective polarization

We measure affective polarization in mass publics by analyzing survey data from the Comparative Study of Electoral Systems (CSES), which has compiled national election studies since 1996. The CSES surveys include a common module asking respondents to rate the political parties in their country on a 0-10 thermometer scale, where higher numbers denote more positive evaluations. We reverse this scale so that 10 denotes the most negative party evaluation and 0 the most positive, which simplifies the interpretation of our results. We analyze all western countries in Modules 1 to 4 in the CSES that feature at least two election surveys (the minimum number required to estimate country fixed-effects models we analyze below), and for which we have reliable data on income inequality and unemployment which we require to evaluate our economic effects hypotheses. Our study encompasses 76 election surveys from 20 Western democracies between 1996 and 2015; Table S1 in the Supplementary Information memo lists the countries and elections in the data. The countries with the most CSES election studies during this period are Germany, Iceland, New Zealand, and Norway (five surveys each), while those with the fewest surveys are Austria and Greece (two surveys each). We do not include in our analyses post-communist Eastern European countries, which are characterized by very different political traditions (Pop-Eleches and Tucker 2011).

In measuring affective polarization, US scholars have often emphasized partisans' <u>utility differential</u> for their preferred party versus its opponents, defined as the difference in the thermometer scores partisans assign to their in-party versus those they assign to the out-party(ies) (e.g., Iyengar et al. 2012). Empirically, growing affective polarization in the US is primarily driven by

partisans' growing out-party dislike, not by improved in-party evaluations (Iyengar et al. 2012).

Moreover, short-term fluctuations in feelings toward one's in-party may obscure long-term affective polarization trends.[7] We therefore begin with a measure of affective polarization based on out-party dislike, and then consider how our results generalize using a difference-based measure.

In the American two-party system, each partisan's out-party dislike score is simply the thermometer rating that a Democratic (Republican) partisan assigns to the rival Republican (Democratic) Party. Matters are more complicated in multiparty systems outside the US, where partisans rate multiple out-parties, that are moreover of different sizes. To address this issue, we follow Reiljan (forthcoming) and Wagner (2017) in constructing a *weighted out-party evaluation* measure, which we compute in two steps. First, we compute each partisan constituency's weighted mean evaluation of all out-parties in the system. Thus for the partisans of party A, competing in a system that additionally includes parties B, C,…, K, the mean evaluation that A's partisans assign to out-parties is:

$$\text{Affective Polarization Party A} = \sum_{i=B}^{K} (\text{Thermometer Score Party}_i) *$$
$$((\text{Vote Share Party}_i) * (1 - \text{Vote Share Party A}))$$

$$(1)$$

where the thermometer score denotes the thermometer ratings that party A's partisans assign to parties B, C,…, K, and each out-party's vote share is its national vote in the current election for the

---

[7] For instance, we find that Republican survey respondents in 2008 expressed more negative feelings toward their own party than in the preceding surveys, which reflected President George W. Bush's declining popularity. However, we are skeptical that this represented declining affective polarization among Republicans: the surge in out-party animosity related to the Tea Party movement following 2008 suggests that Republican partisans were indeed affectively polarized in 2008

lower legislative chamber. This is a measure of how Party *A*'s partisans evaluate out-parties, weighted by their sizes. This *weighted out-party dislike* measure will be higher – denoting greater out-party dislike – to the extent that party *A*'s partisans dislike out-parties, particularly larger out-parties. (Recall that we reverse partisans' thermometer ratings of parties so that 10 denotes maximum dislike, while zero denotes maximum liking.)

To compute the overall intensity of out-party dislike across all the partisan constituencies in the system, we take the average of each partisan constituency's weighted out-party dislike score, weighted by each party's vote share:[8]

$$\text{Country Level Affective Polarization} = \sum_{i=A}^{K}(\text{Affective Polarization Party}_i * \text{Vote Share Party}_i)$$

(2)

Figure 1 displays affective polarization scores based on our out-party dislike measure for the 20 western publics in our study, computed over the 75 CSES national election surveys in our data set. The dots represent the mean affective polarization scores for each country averaged across the available surveys, while the bars represent the range between the minimum and maximum computed values in each country (recall that we have at least two surveys per country). The United States scores near the median value on affective polarization, while the Spanish and Greek publics appear the most affectively polarized. The mean value for Spain, 7.34 on the 0-10 scale where 10 denotes maximum dislike and 5 denotes neutral feelings, indicates that Spanish partisans intensely dislike out-parties (on average), while the score for Greece, 7.21, denotes almost equally intense affective

_____

[8] Normalized by the total vote share across the parties included in the election survey (for example, if the parties that are included in the survey captured 97% of the vote, we divide each party's vote share by .97 to make scores comparable across countries and across time).

polarization. By contrast the northern European and Scandinavian publics display far less intense affective polarization: most have scores below 6.0 with the Netherlands scoring the lowest at 5.03, i.e., Dutch partisans display neutral feelings towards out-parties, on average. We note that the three least affectively polarized countries – The Netherlands, Finland, and Norway – feature consensual institutions with proportional voting systems, whereas several of the most affectively polarized countries – including France, Canada, and Greece – have less proportional systems and govern-ments that at times lack broad popular support. This pattern provides preliminary support for the majoritarian institutions hypothesis. Moreover, several of the most affectively polarized countries including Greece, Spain, and Portugal, were among those hardest hit by the global economic reces-sion and all suffered enormous spikes in unemployment, which provides preliminary support for the unemployment hypothesis. Below we show that these patterns persist in multivariate analyses that additionally control for income inequality and for elite ideological polarization.

While the mean affective polarization score for the United States, 6.69, implies considerable partisan dislike towards out-parties, it only slightly exceeds the mean and median values across the 20 countries in our study: the US mean score is 8th highest out of 20 countries, and is slightly above the overall country mean of 6.25. The US public's comparative standing is surprising (but in line with findings from Lauka et al. 2018, Reiljan forthcoming, and Wagner 2017), given scholarly at-tention to growing affective polarization in the US. To ensure that this conclusion was not an arti-fact of comparing the US two-party system to the multiparty systems in the other democracies in our study – some of which feature radical populist parties that are intensely disliked by their oppo-nents – we also made comparisons across countries based on a two-party measure of affective po-larization, that was restricted to the inter-party ratings provided by the partisans of the largest main-stream center-left and center-right party in each election. (For instance, for the UK we restricted our

analyses to Labour partisans' thermometer ratings of the Conservative Party, and Conservative partisans' ratings of Labour.) The two measures of affective polarization are significantly correlated (r=0.49, p <.01). When using the two-party measure of affective polarization, the United States actually appears slightly <u>less</u> affectively polarized than our twenty-country average. We conclude that while affective polarization in the US – defined as out-party dislike – has sharply intensified over time (Iyengar et al. 2012), it is moderate in comparative perspective.

Figure 1: Affective Polarization Scores by Country
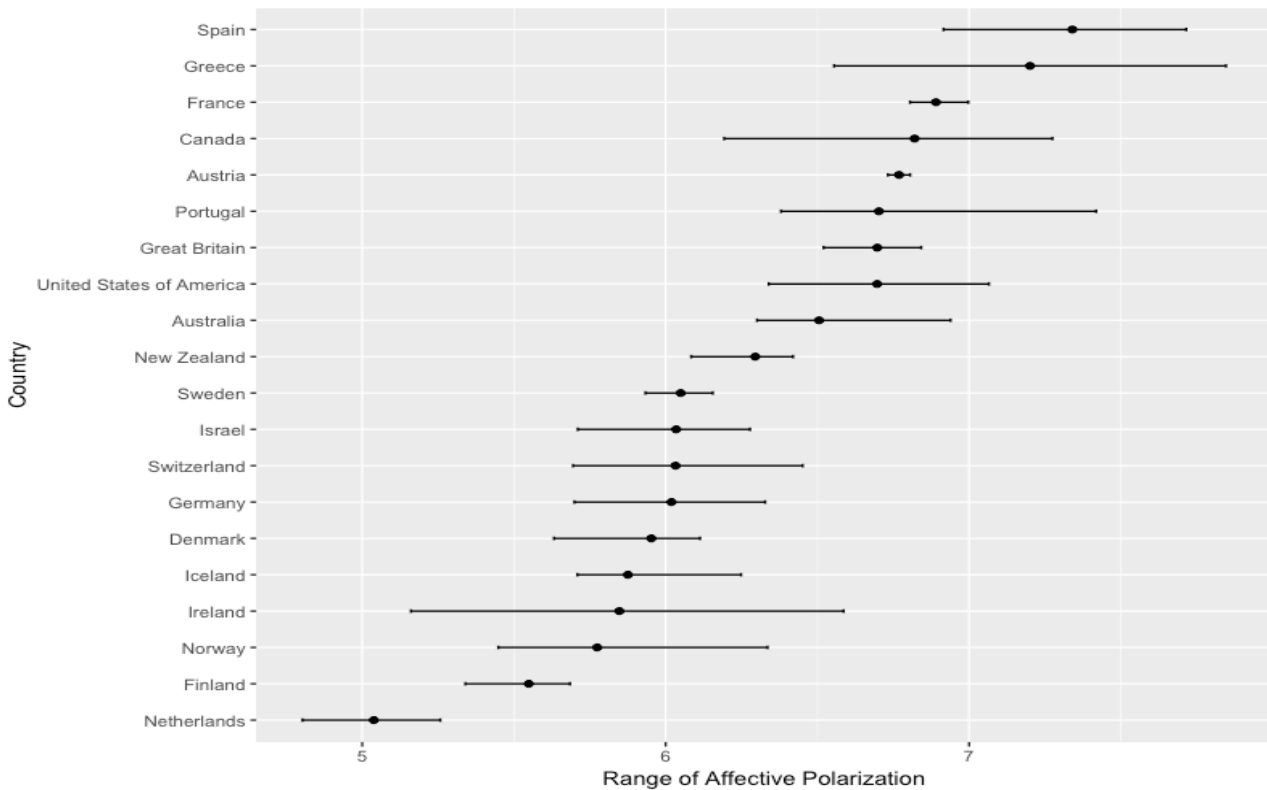
Notes. Figure 1 displays out-party dislike scores for the countries in our study, on a 0-10 scale where 10 denotes maximum out-party dislike, i.e., more intense affective polarization. The dots represent the mean score for each country averaged across the available election surveys from the country, while the bars represent the range between the minimum and maximum values in each country. Table S1 in the Appendix lists the national election surveys included in our analyses.

Independent variables: elite ideological polarization, economic conditions, and institutions

Our measure of elite ideological polarization in the year *t* of a country's current election survey, which we label [*elite ideological polarization* (*t*)], was developed by Dalton (2008) to capture the divergence of parties' positions on Left-Right ideology, which is the only dimension for which we have respondents' party placements across the CSES surveys. We measure this variable using citizens' <u>perceptions</u> of party positions, because previous cross-national research finds that citizens respond to the party positions they perceive, rather than to objective party position measures such as those based on content analyses of party manifestos (Adams et al. 2011). (Below we report robustness checks using this manifesto-based party position measure). In the CSES, each respondent is asked to place all relevant parties on a 0-10 Left-Right scale, where higher numbers denote more right-wing positions. Using these party placements, Dalton defines party system polarization – our [*elite ideological polarization* (*t*)] measure – as the weighted average of the squared distance of each party's mean perceived position from its country mean, transformed so that the maximum score of 10 denotes complete elite polarization – as when parties are perceived as evenly split between the two most extreme Left-Right positions – and 0 when all parties are perceived at identical positions. The supplementary information memo provides the details of this variable's construction, and also presents the CSES party placement survey question on which it is based. Of the cases we study, the country year with the highest perceived elite ideological polarization is Sweden in 1998 (5.19), while the least polarized country year is Canada in 1997 (1.83).

With respect to economic conditions, we measure income inequality using Gini coefficient values from the Standardized World Income Inequality Database (SWIID), where a Gini coefficient of 0% denotes perfect equality (i.e., all income values are the same) while 100% denotes maximum inequality (Solt 2016). The SWIID has been extensively used in studies of European political econ-

omy (e.g., Mosimann and Pontusson 2017). We label the income inequality value in the year t of a country's current election survey [*income inequality* (t)]; the highest country-year value in our data set is for the US in 2012 (37.7%); the lowest is for Iceland in 1999 (22.2%). Unemployment data are from the World Bank, with the country's unemployment rate in year t denoted [*unemployment* (t)]. During the 1996-2016 period of our study, which encompasses the global financial crisis, the highest unemployment rate was for Greece in 2012 (24.4%) and the lowest was for Iceland in 2002 (2.2%).

Our political institutions measure is (logged) average district magnitude (DM), which is an indicator of the degree to which representation is more majoritarian (lower district magnitude, as in single-member plurality voting systems) or more proportional, (higher district magnitude). (Below we report robustness checks using an alternative measure devised by Lijphart (2010) that controls for additional aspects of power-sharing including executive-legislative relations, federalism, and interest-group pluralism.) We use the (logged) average district magnitude of the first tier, which we label [*logged DM* (t)], based on the dataset compiled by Bormann and Golder (2013). The single-member district-based countries in our study (the US, UK, Canada, Australia, and France) all score at zero on the [*logged DM* (t)] variable, while the countries with the highest values (most proportional systems) include Germany, the Netherlands and Norway which all score slightly above five.

Table 1 reports descriptive statistics for our variables, including both within- and between-country standard deviations. These show that unemployment levels vary substantially both between and within countries. The within-country variation reflects, to some degree, the global economic crisis during the period of our study, when unemployment soared in many western democracies. Elite ideological polarization also varies substantially both between and within countries. This within-country variation reflects the many factors that can drive changes in our elite polarization

measure including new parties entering the system (such as the radical left- and right-wing parties that have entered many western party systems in recent decades) along with vote shifts and/or position shifts between elections by existing parties. These within-country variations in unemployment and elite polarization suggest that we have leverage to parse out these variables' predictive power to explain mass affective polarization in statistical analyses that control for country fixed effects, i.e., analyses driven entirely by variations within countries.

Table 1. Descriptive Statistics

	N	Mean	Standard Deviation	Min	Max	Within country SD	Between country SD
Elite ideological polarization	75	3.45	0.80	1.83	5.19	0.45	0.68
Income inequality	76	29.63	4.21	22.20	37.70	0.81	4.15
Unemployment level	76	6.95	3.79	2.20	24.40	2.26	3.31
Logged (DM)	76	2.45	1.89	0	6	0.08	1.85
Out-party dislike	76	6.25	0.63	4.80	7.85	0.29	0.57

By contrast income inequality varies substantially between countries but very little within countries, which reflects the sticky, slow-moving, character of this variable. The (logged) district magnitude variable also varies sharply between countries but is essentially static within countries, since most western democracies did not meaningfully change their electoral systems during the period of our study. This lack of within-country variation implies that we cannot parse out the effects of income inequality and (especially) electoral systems on affective polarization based on within-countries analyses that control for country fixed effects; instead we must rely on comparisons between countries. However, as we discuss below, such between-country comparisons may be problematic due to unmeasured differences between countries, as well as questions about whether sur-

vey respondents' thermometer ratings of parties – which we use to construct our affective polarization measure – have similar substantive interpretations cross-nationally.

These considerations prompt us, in the next section, to conduct statistical analyses both with and without controls for country fixed effects. The country fixed-effects models are arguably most reliable for estimating the effects of unemployment and elite ideological polarization, which vary substantially within countries. The analyses without country fixed effects allow us to estimate cross-national associations between electoral laws, income inequality, and affective polarization, while controlling for unemployment and for elite ideological polarization. We discuss the (cautious) inferences to be drawn from such cross-national comparisons below. On this basis we proceed.

Table 2 reports correlations between the values of the independent variables in our data set. Consistent with Iversen and Soskice (2015), income inequality is associated with *less* intense elite ideological polarization ($r = -0.38$, $p < .01$), while the negative relationship between district magnitude and income inequality ($r = -0.32$, $p < .01$) supports Lijphart's (1999) arguments about the beneficial effects of proportional, power-sharing systems. And as expected, unemployment is associated with income inequality ($r = +0.32$, $p < .01$). These significant correlations underline the importance of estimating multivariate models that jointly control for these various predictive factors.

Table 2. Correlations between the Independent Variable Levels

	Elite ideological polarization (*t*)	Income inequality (*t*)	Unempl-oyment (*t*)	Logged District Magnitude
Elite ideological polarization (*t*)	--	-0.38**	0.08	0.24*
Income inequality (*t*)		--	0.32**	-0.32**
Unemployment (*t*)			--	-0.12

*p < .05 ; ** p < .01

Hypothesis Tests

We estimate the relationship between affective polarization (defined as out-party dislike) and the independent variables defined above. First, we estimate a <u>basic model</u> to evaluate our hypotheses pertaining to elite ideological polarization (H1) and economic conditions (H2a, H2b):

$$\textit{Out-party dislike } (t) \; = \; b_1 + b_2[\textit{elite ideological polarization } (t)]$$
$$+ \; b_3[\textit{income inequality } (t)] \; + \; b_4[\textit{unemployment } (t)] \qquad . \qquad (3)$$

Note that the basic model does not include the [*logged DM* (*t*)] variable, because as discussed above some scholars argue that countries that feature less inter-group hostility and competition tend to choose consensual institutions (e.g., Boix 1999). However given that most of the countries in our data set chose their institutions long before the time period of our study (1996-2015), which lessens concerns that they are endogenous to contemporary levels of out-party dislike, we also estimate the following <u>fully-specified model</u> which includes our institutional measure:

$$\textit{Out-party dislike } (t) \; = \; b_1 \; + \; b_2[\textit{elite ideological polarization } (t)] \; + \; b_3[\textit{income inequality } (t)]$$
$$+ \; b_4[\textit{unemployment } (t)] \; + \; b_5[\textit{logged (DM)}] \qquad . \qquad (4)$$

The elite ideological polarization hypothesis (H1) implies that the coefficient on the [*elite ideological polarization* (*t*)] variable will be <u>positive</u> in both models, denoting that elite ideological polarization intensifies out-party dislike. (Recall that higher values of the [*Out-party dislike* (*t*)] variable denote more intense out-party dislike.) The income inequality hypothesis (H2a) and the unemployment hypothesis (H2b) imply <u>positive</u> coefficients on the [*income inequality* (*t*)] and [*unemployment* (*t*)] variables, denoting that out-party dislike intensifies in response to higher levels of income inequality and unemployment. Finally, the majoritarian institutions hypothesis (H3) implies

that the coefficient on the [*logged DM*] variable will be <u>negative</u>, denoting that out-party dislike is more intense in countries that feature majoritarian electoral systems with low district magnitudes.

We initially estimate the parameters of the basic and fully specified models (equations 3-4 above) without controlling for country fixed-effects. As discussed above this is the best strategy for reliably estimating the relationship between affective polarization (our dependent variable) and levels of income inequality and district magnitude, which display little to no variation within countries over the period of our study. However ignoring country fixed effects raises two concerns about our analyses: namely, a specification issue pertaining to possible omitted variable bias, and a measurement issue concerning the cross-national comparability of our affective polarization measure. To foreshadow our findings, we report below that – controlling for unemployment, income inequality, and elite polarization – the countries with lower district magnitudes (notably the US, UK, Australia, Canada, and France) display more intense levels of affective polarization than do the countries with higher district magnitudes (such as The Netherlands and the Scandinavian countries), based on survey respondents' thermometer ratings of parties. Yet this comparison is invalid if respondents differ cross-nationally in their interpretations of the thermometer scales, or if social desirability pressures push their party ratings in different directions. Moreover, while we control for unemployment, income inequality, and elite ideological polarization in addition to district magnitude, the countries in our study differ in many other ways we cannot systematically measure, raising concerns about omitted variable bias.

The above considerations prompt us to re-estimate our basic model while controlling for country fixed effects. (For these analyses we do not control for district magnitude since – as reported in Table 1 above – this variable displays essentially zero within-country variation.) This specification gives us purchase to parse out the effects of unemployment and elite ideology although we

will have little ability to reliably estimate the effects of income inequality, which moves only slowly within countries over the time period of our study:

Out-party dislike (t) = b_1 + b_2[*elite ideological polarization* (t)]

$$+ \ b_3[\textit{income inequality } (t)] \ + \ b_4[\textit{unemployment } (t)]$$

$$+ [\text{country-specific intercepts}] \qquad . \qquad\qquad (5)$$

Results

Table 3 reports parameter estimates for the basic model (column 1) and the political institutions model (column 2), without country fixed effects, while column 3 reports estimates for the basic model with country fixed effects. We estimate the models using OLS with robust standard errors clustered by country. We report the reduced models in the supplementary information memo.

   With respect to the elite ideological polarization hypothesis (H1), our estimate on the [*elite ideological polarization* (t)] variable is statistically insignificant across all three models, which does not support H1. Moreover while the estimates on elite polarization are positive for each model, as H1 implies, these estimates suggest that – even if elite polarization does indeed intensify mass affective polarization – this effect is modest: the estimate on the [*elite ideological polarization* (t)] variable for the fully-specified model (col. 3 in Table 3) is +0.113, which implies that a change from one standard deviation below the mean value in our data set (2.65) to one standard deviation above the mean (4.25) intensifies partisans' predicted out-party dislike by only about 0.2 units on the 0-10 thermometer scale, all else equal. And, again, this predicted effect is not statistically significant. Below we report robustness checks using an alternative, manifesto-based measure of elite ideological polarization, which again fails to detect significant effects on affective polarization.

Our estimates provide much stronger evidence when we shift to economic factors. The estimate on the [*unemployment* (*t*)] variable is statistically significant in all of our models (*p* < .01), i.e., there is strong support for the unemployment effects hypothesis (H2b) that unemployment intensifies affective polarization, based on comparisons both between and within countries. The magnitudes of these estimates are consistent across our three models (ranging from +0.65 to +0.70), and they imply that a change from one standard deviation below to one standard deviation above the mean unemployment rate in our data set (from 3.2% to 10.7%) will intensify partisans' out-party dislike by about 0.5 units on the 0-10 thermometer rating scale, all else equal. Moreover, the estimates imply that the unemployment spikes in some democracies during the recent global economic crisis greatly exacerbated mass affective polarization in these countries. In Greece, for instance, the unemployment increase between the 2009 and 2012 national elections (from 9.6% unemployment in 2009 to 24.4% in 2012) is predicted to intensify out-party dislike by about one point on the 0-10 thermometer scale – which is nearly the difference between the average levels of affective polarization for Greece versus Sweden during the period of our study, as displayed in Figure 1 above.[9]

There is also evidence, based on comparisons <u>between</u> countries but not <u>within</u> countries, to support the income inequality hypothesis (H2a) that income inequality intensifies affective polarization: the coefficient on the [*income inequality* (*t*)] variable is significant (*p* < .01) in the models without country fixed effects (columns 1-2 in Table 3), and these estimates (+0.69 for the basic

---

[9] In fact, the actual level of affective polarization in Greece increased by 0.9 units on the 0-10 thermometer scale between 2009 and 2012 (from 6.55 to 7.85), based on analyses of the Greek CSES surveys from these years – almost exactly the increase predicted from our parameter estimates. We note that we find substantively similar results when excluding from the sample Portugal, Spain and Greece, which suffered especially high unemployment spikes after 2008 (see below).

model, +0.55 for the fully-specified model) imply substantial effects: a change from one standard deviation below to one standard deviation above the mean income inequality level in our data set (from a Gini coefficient of 25.4% to one of 33.8%) intensifies partisans' predicted out-party dislike by 0.5 to 0.6 units on the 0-10 thermometer rating scale (all else equal). However, as expected, we do not detect statistically significant effects in the country fixed-effects model (column 3), where the lack of in-country variation on income inequality makes precise estimation difficult.

Finally, our positive estimate on the [*logged (DM)*] variable for the political institutions model (column 2, $p < 0.01$), supports the majoritarian institutions hypothesis (H3) that affective polarization is more intense in countries featuring majoritarian institutions, controlling for economic conditions and for elite ideological polarization. A change from one standard deviation below to one standard deviation above the mean [*logged (DM)*] value in our data set increases partisans' predicted out-party thermometer ratings by about 0.5 units on the 0-10 scale, all else equal. As discussed above we do not infer a causal relationship, given research suggesting that institutional design may respond to pre-existing social and political cleavages; and moreover we cannot evaluate this relationship while controlling for country fixed effects. Nevertheless, this association is consistent with Lijphart's (2010) argument that majoritarian institutions promote more contentious politics than the "kinder, gentler" politics in consensual systems. In particular, given Lijphart's empirical finding that majoritarian institutions are associated with greater income inequality – a relationship we confirm in our data set (see Table 2 above) – it is striking that district magnitude displays a direct association with affective polarization in our analyses that control for inequality (and also for unemployment and elite polarization).

Table 3. Analyses of Affective Polarization

INDEPENDENT VARS	Basic Model (1)	Fully-specified Model (2)	Country Fixed Effects Model (3)
Party system ideological polarization	0.119 (0.108)	0.177 (0.109)	0.113 (0.095)
Income inequality	0.069** (0.024)	0.055** (0.019)	-0.058 (0.030)
Unemployment level	0.070** (0.013)	0.065** (0.014)	0.068** (0.015)
Logged (DM) (*t*)		-0.146** (.043)	
Intercept	3.30 (0.91)	3.91 (0.82)	7.68 (0.99)
N	75	75	75
Adjusted R^2	0.48	0.65	0.81

** $p \leq .01$; * $p \leq .05$, two-tailed tests.

<u>Notes</u>. The dependent variable in these analyses is [*Out-party dislike* (*t*)], defined as the mean respondent evaluation of out-parties in the current election survey (with the scale reversed so that higher numbers denote greater out-party dislike), computed over the set of election surveys listed in Table S1 in the Appendix. The independent variables are defined in the text. The top number in each cell is the unstandardized coefficient, the number in parentheses below is the standard error. The OLS regression models were estimated with standard errors clustered on countries.

Figure 2 illustrates, based on the full model without fixed effects (column 2 in Table 3), the predicted relationships between affective polarization (the vertical axis) and perceived elite ideological polarization (Figure 2A); income inequality (Figure 2B); unemployment (Figure 2C); and district magnitude (Figure 2D), over the range of values for these independent variables in our data set.[10] (The dotted lines denote 95% confidence intervals.) We see that the predicted effects of elite

---

[10] For these computations the values of all other independent variables were set to the mean values.

ideological polarization (Figure 2A) are insignificant, and are weaker than the predicted effects of

income inequality and unemployment (Figures 2B-2C) as well as the relationship between district

magnitude and affective polarization (Figure 2D). (Recall, however, that the predicted inequality

effects become insignificant for the country fixed effects model.)

Figure 2. Predicted Levels of Affective Polarization
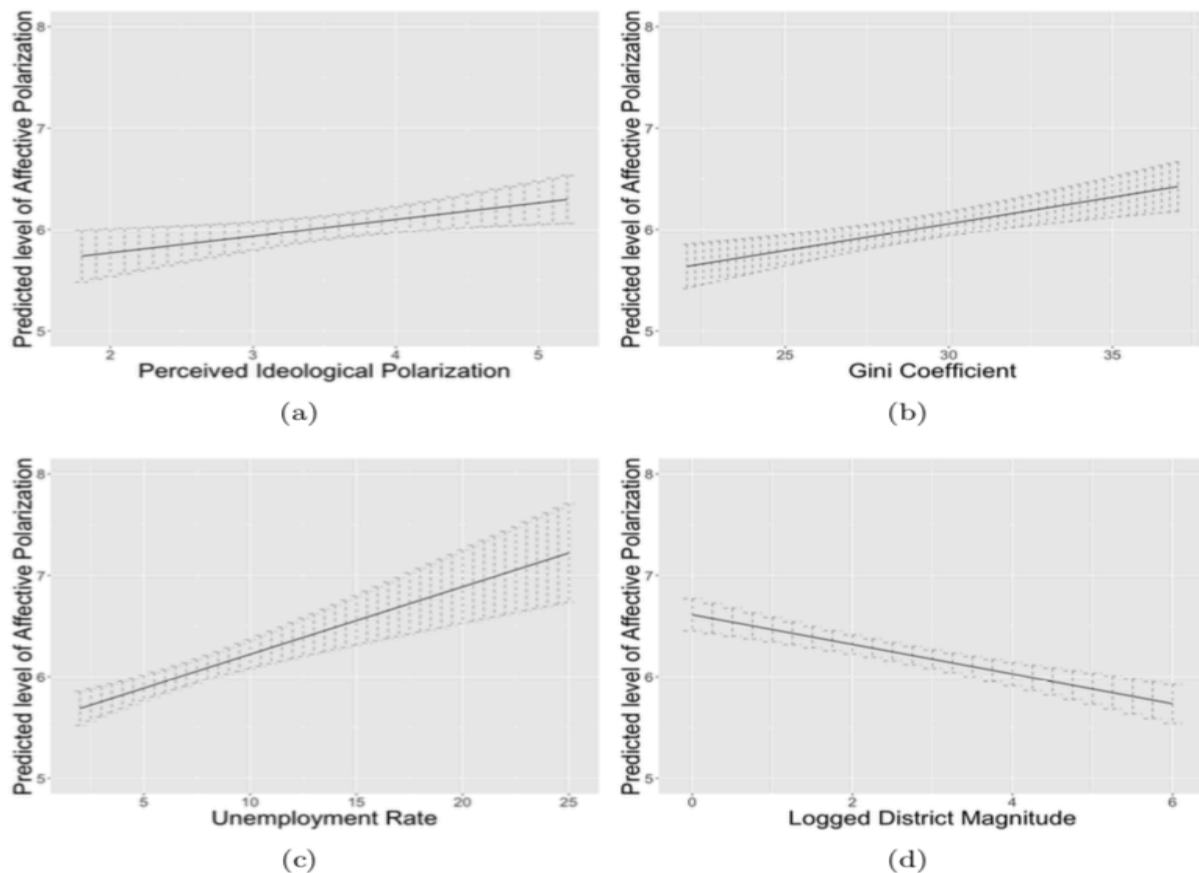
Notes. The figure displays the predicted level of affective polarization (the [*out-party dislike* (*t*)] variable), computed for the coefficient estimates for the fully-specified model reported in column 2 of Table 3. The dashed lines denote 95% confidence intervals.

Robustness checks

(*Note to reviewers*: *We present all of the analyses described below in the supplementary infor-*

mation memo.) We performed several robustness checks to evaluate the effects of using alternative

variable measures; of excluding problematic cases from our analyses; and of controlling for foreign-born population. With respect to variable measures, we first re-estimated our models using an alternative elite ideological polarization measure based on the Comparative Manifesto Project codings of the Left-Right tones of parties' election manifestos (Budge et al. 2001). Next, we re-estimated our political institutions model using an alternative measure of majoritarian/consensual institutions, Lijphart's (2010) executive-parties dimension, which considers factors beyond district magnitude including executive-legislative relations, federalism, and interest-group pluralism. Finally, because some scholars conceptualize affective polarization as the difference between partisans' in-party and out-party thermometer ratings (e.g., Iyengar et al. 2012), we re-estimated our models using this alternative dependent variable measure. All of these analyses continued to substantiate our substantive conclusions.

We note that the analyses using a dependent variable measured as the difference between in-party like and out-party dislike provided stronger support for the elite ideological polarization hypotheses, at least in the analyses without country fixed effects. (However we again estimated insignificant effects in the country fixed-effects model.) It may be that elite ideological polarization increases out-party dislike but also prompts more positive feelings toward in-parties, which is then captured in the statistically significant relationship between elite level ideological polarization and the difference-based partisan dislike measure. This calls for further attention to the different drivers of dislike toward out-parties and positive feelings toward in-parties.

Next, we estimated models controlling for the national proportion of foreign-born population, which might be related to social cleavages pertaining to immigration/multiculturalism, and also to elite polarization (McCarty et al. 2006). We find no evidence that the share of foreign born population is predictive of affective polarization.

Finally, we explored the effects of omitting countries from our analyses for which causal processes plausibly differed, in ways that are not captured by our independent variables. Portugal, Spain, and Greece were hit especially hard by the recent economic crisis, and moreover these countries have experienced non-democratic governments much more recently than the other countries in our data set, which might mediate the effects of our independent variables on affective polarization. We re-estimated our models while omitting the data from these countries, and our analyses continued to support our substantive conclusions.

Discussion and Conclusions

This paper is a first step towards integrating the United States into comparative analyses of affective polarization across western publics. While inter-party hostility, distrust, and contempt are disturbing features of many contemporary polities, nearly all extant research addresses the single American case. We believe that embedding the US in a comparative context is promising, first because affective polarization in other western publics is intrinsically important, second because the US case in isolation is overdetermined: Many of the proposed explanations for intensifying US affective polarization – including rising income inequality, elite-level ideological polarization, and the rise of partisan media – have occurred roughly concurrently, raising difficulties in parsing out these causal factors in a US-centered study. By conducting cross-national analyses, we can compare affective polarization levels across countries that display different combinations of these explanatory variables. By analyzing trends within countries over time, we can analyze how affective polarization evolves as the levels of these proposed independent variables change.

We have applied our comparative strategy to analyze the ideological, economic, and institutional roots of affective polarization across 20 western democracies between 1996 and 2015, a period for which the Comparative Study of Electoral Systems includes common survey items measur-

ing citizens' partisanship, their party thermometer ratings, and their perceived party Left-Right positions. In comparisons both between countries and within countries over time, we find that higher unemployment is strongly associated with intensified mass-level affective polarization. These effects, moreover, appear substantively significant: our estimates imply that the unemployment shocks that countries such as Greece, Spain, Ireland, and Portugal suffered during the global recession greatly intensified affective polarization in these countries. By contrast we do not detect especially strong effects of elite ideological polarization on mass affective polarization, when either comparing levels between countries or comparing level changes within countries over time. This suggests that partisans' dislike and contempt for out-parties is not primarily driven by Left-Right policy debates. Finally, we estimate strong cross-national associations between affective polarization, income inequality, and majoritarian institutions that feature disproportional, single-member district voting systems. However, we cannot triangulate these findings via longitudinal analyses within countries, since income inequality evolved very slowly (and electoral systems hardly at all) within the countries in our data set, across the 1996-2015 time period of our study.

Our findings raise several causal inference questions. We detect a strong relationship between unemployment and affective polarization, in both in comparisons between countries and within countries over time – and hence we conclude that this relationship is likely causal. However while we also document strong links between affective polarization, income inequality, and majoritarian institutions in comparisons between countries, we cannot reliably estimate these relationships via comparisons within countries: institutions, and to a lesser degree also inequality, are too sticky. Our between-country comparisons support Lijphart's (2010) arguments about the beneficial effects of consensual (i.e., non-majoritarian) institutions, as well as Hitlin and Harkness's (2017) arguments about the corrosive emotional effects of income inequality. Yet these cross-national compari-

29

sons may be problematic, because the countries in our study differ in ways we cannot measure (or only measure imperfectly): a partial list includes differences in partisan media, in economic regimes, in the structure of issue dimensions beyond overarching Left-Right ideology, and in countries' democratic histories that influence current patterns of mass partisanship. Moreover, our cross-national comparisons are only valid provided that survey respondents' thermometer ratings of parties are cross-nationally comparable. While we address some issues pertaining to omitted variable bias in our robustness checks,[11] we are under no illusion that we have fully controlled for between-country differences. Yet if we discount cross-national comparisons, it may be impossible to precisely estimate how income inequality and institutions influence affective polarization; there is simply too little within-country variation on these variables. This is a subject for future research.

Our analyses also speak to debates about the role of elite ideological polarization in driving mass-level affective polarization. Given the strong theoretical reasons to expect this causal relationship we do not reject this hypothesis, particularly since our coefficient estimates are in the expected direction, albeit statistically insignificant (when using a measure of out-party dislike). It seems plausible that if we could expand our study to additional countries or a longer time period, we could more precisely estimate these elite-driven effects. At the same time, our non-significant findings extend across comparisons both between countries and within countries. We believe this is well worth knowing. In pushing forward this research agenda, we plan to additionally analyze elite polarization on cultural, identity-related issues that may cross-cut the overall Left-Right dimension. Previous work shows that identity-related disagreements generate more emotionally heated re-

---

[11] These include analyses that omit the countries with post-WWII experiences with authoritarian regimes, and analyses that control for foreign-born populations. These are reported in the supplementary information memo.

sponses than disputes on economic issues (Hetherington et al. 2016; Sides et al. 2018), and these emotional responses may be translated into affective evaluations of out-parties.

Our analyses of economic conditions open avenues for future research on the political economy of mass-level polarization, and the rise of radical populist parties (e.g., Iversen and Soskice 2015; McCarty et al. 2006). Our findings that unemployment and (in cross-national comparisons) income inequality strongly predict partisan dislike are relevant for political developments in the aftermath of the Great Recession. The post-recession decline in unemployment implies less intense affective polarization, as western democracies recover from the financial crisis. Yet if the economic recovery is uneven (Saez 2012), then the long-term negative effects of growing income inequality may counteract the beneficial short-term effects of employment gains. Our findings also suggest a link between economic downturns, affective polarization, and support for radical populist politics. There is an ongoing debate about the mechanisms connecting economic hardship and the rise of populist parties and candidates (e.g., Gidron and Hall 2017; Mutz 2018). Our results propose an indirect mechanism through which economic deprivation may feed into support for these political challengers: since economic downturns that feature rising unemployment intensify out-party dislike, they may provide fertile grounds for political actors to mobilize support based on populist appeals that adopt an 'us-versus-them' tone.

We close with two observations about the United States, the country which has prompted virtually all extent research on affective polarization. On the one hand, our finding that the US is not unduly polarized in cross-national perspective may comfort observers who are dismayed by the levels of inter-party hostility, scorn, and distrust in our contemporary politics: many other western polities appear equally (or even more intensely) polarized. On the other hand, our finding that unemployment increases tend to intensify mass-level affective polarization – the strongest, most con-

sistent finding in our study – is arguably ominous for the future of American politics. American unemployment levels are currently at a nearly 50-year low – yet affective polarization in the US appears to be at a postwar high (although it remains moderate in comparative perspective). If American unemployment levels revert towards their historic norms, let along levels associated with another recession, this development may further amplify inter-party hostility between Democrats and Republicans. While the future of American civic life will turn on many unforeseeable developments, this consideration highlights the challenges we confront as we struggle to build a more civil society.

Bibliography

Abramowitz, Alan and Steven Webster. (2017). "The Ideological Foundations of Affective Polarization in the US Electorate." *American Politics Research*, 45(4):621–647.

Adams, James, Lawrence Ezrow and Zeynep Somer-Topcu. (2011). 'Is Anybody Listening? Evidence That Voters Do Not Respond to European Parties' Policy Statements During Elections.' *American Journal of Political Science*, *55*(2), 370–382.

Anderson, Christopher J., and Christine A. Guillory. 1997. "Political Institutions and Satisfaction with Democracy: A Cross-National Analysis of Consensus and Majoritarian Systems." *American Political Science Review* 91(1): 66-82.

Boix, Carles. (1999). "Setting the Rules of the Game: The Choice of Electoral Systems in Advanced Democracies." *American Political Science Review,* 93(3), 608-624

Bonica, Adam. (2014). "Mapping the Ideological Marketplace." *American Journal of Political Science*, 58(2), 367-386.

Bormann, Nils-Christian, and Matt Golder. (2013). "Democratic Electoral Systems around the World, 1946–2011." *Electoral Studies,* 32(2), 360-369

Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Eric Tannenbaum, and Judith Bara. 2001. *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford: Oxford University Press.

Carlin, Ryan E. and Gregory J. Love. (2018). 'Political Competition, Partisanship and Interpersonal Trust in Electoral Democracies.' *British Journal of Political Science*, 48(01), 115–139.

Clarke, Harold, Dutt Nitish, and Allan Kornberg. 1993. "The Political Economy of Attitudes Toward Polity and Society in West European Democracies." *Journal of Politics* 55(4): 998-1021.

Colomer, Josep M. 2018. "Party System Effects on Electoral Rules." In *The Oxford Handbook of Electoral Systems*. Erik Herron, Robert Pekkanen, and Matthew Shugart, editors. Oxford University Press.

Dalton, Russell J. (2008). 'The Quantity and the Quality of Party Systems: Party System Polarization, Its Measurement, and Its Consequences.' *Comparative Political Studies*, 41(7):899–920.

Dow, Jay K. (2011). "Party System Extremism in Majoritarian and Proportional Electoral Systems." *British Journal of Political Science* 41(2): 341-61.

Eichengreen, Barry. (2018). *The Populist Temptation*. Oxford: Oxford University Press.

Ezrow, Lawrence. 2008. "Parties' Policy Programmes and the Dog that Didn't Bark: No Evidence that Proportional Systems Promote Extreme Party Positioning. *British Journal of Political Science* *38*(3):479 – 497.

Fenzi, Michele.2018. "Income Inequality and Party (De)Polarization." *Western European Politics*.

Fiorina, Morris P., and Samuel J. Abrams. (2008). 'Political polarization in the American public.' *Annual Review of Political Science*. 11: 563-588.

Fiske, Susan T., Amy J.C. Cuddy, and Peter Glick. (2007). 'Universal Dimensions of Social Cognition: Warmth and Competence.' *Trends in Cognitive Sciences* 11:77–83.

Funke, Manuel, Moritz Schularick and Christoph Trebesch. (2016). 'Going to Extremes: Politics after Financial Crises, 1870–2014.' *European Economic Review*, 88:227–260.

Gidron, Noam and Peter A. Hall. (2017). 'The politics of social status: economic and cultural roots of the populist right.' *British Journal of Sociology* 68:57-84.

Gilley, Bruce. (2006). "The Meaning and Measure of Democratic Legitimacy." *European Journal of Political Research* 45(3): 499-525.

Hetherington, Marc J. (2009). 'Review Article: Putting Polarization in Perspective.' *British Journal of Political Science*, 39(02):413–36.

Hetherington, Marc J., and Thomas J. Rudolph. (2015) *Why Washington Won't Work: Polarization, Political Trust, and the Governing Crisis*. Chicago: University of Chicago Press.

Hetherington, Marc J., Meri T. Long, and Thomas J. Rudolph. (2016). 'Revisiting the Myth: New Evidence of a Polarized Electorate'. *Public Opinion Quarterly*, 80: 321-350.

Hitlin, Steven and Sarah K. Harkness. (2017). *Unequal Foundations: Inequality, Morality, and Emotions Across Cultures*. Oxford: Oxford University Press.

Huddy, Leonie, Alexa Bankert and Caitlin Davies. (2018). 'Expressive Versus Instrumental Partisanship in Multiparty European Systems.' *Political Psychology*, 39(3), 173–199.

Iversen, Torben and David Soskice. (2006). "Electoral Institutions and the Politics of Coalitions: Why some Democracies Redistribute more than Others." *American Political Science Review,* 100(2), 165-181.

Iversen, Torben and Soskice, David. (2015). 'Information, Inequality, and Mass Polarization: Ideology in Advanced Democracies.' *Comparative Political Studies*, 48(13):1781–1813.

Iyengar, Shanto, Gaurav Sood and Yphtach Lelkes. (2012). 'Affect, Not Ideology.' *Public Opinion Quarterly*, 76(3):405–431.

Iyengar, Shanto and Sean Westwood. (2014). 'Fear and Loathing across Party Lines: New Evidence on Group Polarization.' *American Journal of Political Science* 59(3):690–707.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhorta and Sean Westwood. (2019). 'The Origins and Consequences of Affective Polarization in the United States.' *Annual Review of Political Science*. 22.

Kinder, Donald R. and Nathan P. Kalmoe. *Neither Liberal nor Conservative: Ideological Innocence in the American Public*. Chicago: University of Chicago Press, 2017.

Lau, Richard R., David J. Andersen, Tessa M. Ditonto, Mona S. Kleinberg, and David P. Redlawsk. (2017). 'Effect of Media Environment Diversity and Advertising Tone on Information Search, Selective Exposure, and Affective Polarization.' *Political Behavior* 39(1): 231-255.

Lauka, Alban, Jennifer McCoy, and Rengin B. Firat. (2018). 'Mass Partisan Polarization: Measuring a Relational Concept.' *American Behavioral Scientist* 62(1): 107-126.

Lelkes, Yphtach. (2016). 'Mass Polarization: Manifestations and Measurements.' *Public Opinion Quarterly*, 80(S1), 392–410.

Lelkes, Yphtach. (2018). 'Affective Polarization and Ideological Sorting: A Reciprocal, Albeit Weak, Relationship.' *The Forum*, 16(1), 67–79.

Lelkes, Yphtach and Sean Westwood. (2017). 'The Limits of Partisan Prejudice.' *Journal of Politics*, 79(2), 485–501.

Levendusky, Matthew. (2013). 'Partisan Media Exposure and Attitudes Toward the Opposition.' *Political Communication*, 30(4), 565–581.

Levitsky, Steven, and Daniel Ziblatt. (2018). *How Democracies Die*. New York: Crown.

Lieberman, Robert C. Suzanne Mettler, Thomas B. Pepinsky, Kenneth M. Roberts and Richard Valelly. (2018). 'The Trump Presidency and American Democracy: A Historical and Comparative Analysis.' *Perspectives on Politics* (2018): 1-10.

Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven: Yale University Press.

Lijphart, Arend. 2010. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries, 2nd edition*. New Haven: Yale University Press.

López, Edward J. and Carlos D. Ramírez. (2004). 'Party Polarization and the Business Cycle in the United States.' *Public Choice* 121(3-4): 413-430.

Lupu, Noam. (2015). "Party Polarization and Mass Partisanship: A Comparative Perspective." *Political Behavior*, 37(2)3: 331-356.

Mason, Lilliana. (2015). ""I disrespectfully agree": The differential effects of partisan sorting on social and issue polarization." *American Journal of Political Science* 59(1): 128-145

Mason, Lilliana. (2018). *Uncivil Agreement: How Politics Became Our Identity*. Chicago: Chicago University Press.

McCarty, Nolan, Keith T. Poole and Howard Rosenthal. (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge: MIT Press.

McConnell, Christopher, Yotam Margalit, Neil Malhotra and Matthew Levendusky. (2018). 'The Economic Consequences of Partisanship in a Polarized Era.' *American Journal of Political Science*, 62(1):5–18.

McCoy, Jennifer, and Murat Somer. (2019). 'Toward a Theory of Pernicious Polarization and How It Harms Democracies: Comparative Evidence and Possible Remedies.' *The Annals of the American Academy of Political and Social Science*, 681(1): 234-271.

Mickey, Robert, Steven Levitisky and Lucan Ahmad Way. (2017). 'Is America Still Safe for Democracy: Why the United States Is in Danger of Backsliding,' *Foreign Affairs* 96.

Mosimann, Nadja and Jonas Pontusson. (2017). 'Solidaristic Unionism and Support for Redistribution in Contemporary Europe.' *World Politics*, 69(03), 448–492.

Mounk, Yascha. (2018). *The People Vs. Democracy: Why Our Freedom is in Danger and How to Save it*. Cambridge: Harvard University Press.

Mutz, Diana. (2018). 'Status Threat, Not Economic Hardship, Explains the 2016 Presidential Vote.' *Proceedings of the National Academy of Sciences* 115 (19), E4330-E4339

North, Douglass. (1990). *Institutions, Institutional Change and Economic Performance.* Cambridge: Cambridge University Press

Pop-Eleches, Grigore and Joshua A. Tucker. (2011). 'Communism's Shadow: Postcommunist Legacies, Values, and Behavior.' *Comparative Politics*, 43(4), 379–408.

Reiljan, Andres. (forthcoming). 'Fear and Loathing across Party Lines" (also) in Europe: Affective Polarisation in European Party System*s*.' *European Journal of Political Research*.

Rogowski, Jon C. and Joseph L. Sutherland. (2015). 'How Ideology Fuels Affective Polarization.' *Political Behavior*, 38(2), 485–508.

Saez, Emmanuel. 2012. 'Striking It Richer: The Evolution of Top Incomes in the United States.' Working paper.

Sides, John, Michael Tesler, and Lynn Vavreck. (2018). *Identity Crisis: The 2016 Presidential Campaign and the Battle for the Meaning of America*. Princeton: Princeton University Press.

Solt, Frederick. (2016). 'The Standardized World Income Inequality Database,' at http://fsolt.org/swiid/,

Somer, Murat and Jennifer McCoy. (2018). 'Déjà vu? Polarization and Endangered Democracies in the 21st Century.' *American Behavioral Scientist*, 62(1), 3–15.

Sunstein, Cass R. (2015). 'Partyism.' *University of Chicago Legal Forum* 1:1-27.

Wagner, Markus. 2017. "Affective Polarization in Multiparty Systems." Typescript.

Westwood, Sean J., Shanto Iyengar, Stefaan Walgrave, Rafael Leonisio, Luis Miller and Oliver Strijbis. (2017). 'The Tie that Divides: Cross-national Evidence of the Primacy of Partyism.' *European Journal of Political Research* 70(2):542–22.

Zakharova, Maria and Paul V. Warwick. (2014). "The Sources of Valence Judgments: The Role of Policy Distance and the Structure of the Left–Right Spectrum." *Comparative Political Studies*, 47(14), 2000-2025.

How Ideology, Economics and Institutions Shape Affective Polarization in Democratic Polities: Online Supplementary Information Memo

This Supplementary Information memo describes the measurement, data and robustness checks described in our manuscript. The memo is structured as follows:

1. Section S1 lists the countries and elections included in the analyses of the Comparative Study of Electoral Systems [CSES] Surveys.
2. Section S2 provides additional information on our measure of elite-level ideological polarization.
3. Section S3 presents reduced models for our main empirical analyses (Table 3 in the body of the text).

We then turn to a series of robustness checks:

4. Section S4 presents the results of our analyses using a measure of elite ideological polarization based on the Comparative Manifesto Project codebook of parties' positions (instead of survey respondents' perceived party positions, the measure used in the body of the text).
5. Section S5 presents the results of our analyses using an alternative measure of political institutions (Lijphart's Executive Dimension Score instead of logged district magnitude).
6. Section S6 presents the results of our analyses using a difference-based measure of affective polarization as the dependent variable (instead of a measure based on out-party dislike).
7. Section S7 presents the results of our analyses controlling for the share of foreign born population.
8. Section S8 presents the results of our analyses after excluding Spain, Portugal and Greece from the data-set.

Section S1. **Countries and Elections Included in the Analyses of the Comparative Study of Electoral Systems Surveys**

Table S1 lists the countries and elections included in our analyses. As explained in the body of the text, we subset the Comparative Study of Electoral Systems and include only Western democracies.

Table S1: Countries and Parties Included in the
Analyses of National Election Study Surveys

Country		Elections included
Australia		1996, 2004, 2007, 2013
Austria		2008, 2013
Canada		1997, 2004, 2008, 2011
Denmark		1998, 2001, 2007
Finland		2003, 2007, 2011, 2015
France		2002, 2007, 2012
Germany		1998, 2002, 2005, 2009, 2013
Great Britain		1997, 2005, 2015
Greece		2009, 2012
Iceland		1999, 2003, 2007, 2009, 2013
Ireland		2002, 2007, 2011
Israel		1996, 2003, 2006, 2013
Netherlands		1998, 2002, 2006, 2010
New Zealand		1996, 2002, 2008, 2011, 2014
Norway		1997, 2001, 2005, 2009, 2013
Portugal		2002, 2005, 2009, 2015
Spain		1996, 2000, 2004, 2008
Sweden		1998, 2002, 2006, 2014
Switzerland		1999, 2003, 2007, 2011
United States		1996, 2004, 2008, 2012

Section S2. Measuring Party System Ideological Polarization

Dalton (2008) defines party system polarization as the weighted average of the squared ideological distance of each party from its country mean, transformed so that the polarization score equals a maximum of 10 which denotes complete party system polarization – as when parties are perceived as evenly split between the two most extreme Left-Right positions – and 0 when all parties are perceived at identical Left-Right positions:

$$\text{Party System Ideological Polarization} = \sqrt{\sum_{i=1}^{n} (\text{Party Vote Share}_i) * [(\text{Party Left Right Score}_i - \text{Country Average Left Right Score})/5)^2]}$$

The question wording for the CSES common content module is as follows: "In political matters people talk of "the left" and "the right." What is your position? Please use a scale from 0 to 10, where '0' means "left" and "10" means "right." Which number best describes your position? And about where would you place [INSERT PARTY NAME] on this scale?"

Section S3. Reduced Models

Table 3 in the body of the text reports the results of multivariate regression analyses. Table S2 presents the results of the reduced models: that is, regressing separately our dependent variable (out-party dislike) on each independent variable (elite ideological polarization, income inequality, unemployment, logged district magnitude). The results are in line with the analyses presented in Table 3 in the body of the text: income inequality, unemployment and electoral institutions are statistically significant predictors of affective polarization, while elite ideological polarization is not.

Table S2. Reduced univariate models

INDEPENDENT VARS	(1)	(2)	(3)	(4)
Elite ideological polarization	0.00 (0.093)			
Income inequality		0.081** (0.015)		
Unemployment level			0.097** (0.016)	
Logged (DM) (t)				0.180** (0.032)
Intercept	6.216** (0.329)	3.843** (0.435)	5.576** (0.123)	6.692** (0.100)
N	75	76	76	76
Adjusted R^2	-0.014	0.287	0.334	0.285

** $p \leq .01$; * $p \leq .05$, two-tailed tests.

Notes. The dependent variable in these analyses is [*Out-party dislike* (*t*)], defined as the mean respondent evaluation of out-parties in the current election survey (with the scale reversed so that higher numbers denote greater out-party dislike), computed over the set of election surveys listed in Table S1 above. The independent variables are defined in the text. The top number in each cell is the unstandardized coefficient, the number in parentheses below is the standard error. The OLS regression models were estimated with standard errors clustered on countries.

Section S4. Robustness checks: CMP-based measure of ideological polarization

We now turn to a series robustness checks described in the body of the text. First, we use an alternative measure to capture elite ideological polarization. In the body of the text, we use a measure based on survey respondents' perceived party positions. Table S3 below reports analyses based on an alternative party position measure, in which we rely on party Left-Right positions as coded in the Comparative Manifesto Project. We rely on the left-right (RILE) dimension of the Comparative Manifesto Project. Our variable of CMP-based party system ideological polarization ranges from 0.663 (Ireland in 2007) to 7.125 (Greece in 2012). Attentive readers will note that the N drops to 73 when using this variable; this is because the CMP does not include scores for the following cases: Finland (2015), Norway (2013) and Portugal (2015). The results are in line with those presented in Table 3 in the text. Again, the coefficient for elite ideological polarization as measured by the CMP data is positive but fails to reach statistical significance, while the estimates on the remaining variables continue to support the same conclusions we reported in the paper.

Table S3. Robustness checks: CMP-based ideological polarization

INDEPENDENT VARS	Basic Model (1)	Fully-specified Model (2)	Fixed Effects Model (3)
CMP-based elite ideological polarization	0.020 (0.040)	0.030 (0.033)	0.036 (0.042)
Income inequality	0.056** (0.014)	0.037** (0.012)	-0.050 (0.040)
Unemployment level	0.074** (0.015)	0.073** (0.013)	0.060** (0.015)
Logged (DM) (t)		-0.137** (0.025)	
Intercept	4.026** (0.422)	4.888** (0.389)	7.615** (1.295)
N	73	73	73
Adjusted R^2	0.438	0.603	0.800

** $p \leq .01$; * $p \leq .05$, two-tailed tests.

Notes. The dependent variable in these analyses is [*Out-party dislike (t)*], defined as the mean respondent evaluation of out-parties in the current election survey (with the scale reversed so that higher numbers denote greater out-party dislike), computed over the set of election surveys listed in Section S1 in the Supplementary Information Memo. The independent variables are defined in the text. The top number in each cell is the unstandardized coefficient, the number in parentheses below is the standard error. The OLS regression models were estimated with standard errors clustered on countries.

Section S5. Robustness checks: Lijphart's measure of institutions

Next, we turn to an alternative measure of political institutions. In the body of the text, we use the (logged) district magnitude in order to capture the proportionality of electoral institutions. Our theoretical expectations for the relationship between electoral institutions and affective polarization draw heavily on Lijphart's work, which encompasses a broader range of institutional factors. In Table S4, instead of (logged) district magnitude we use Lijphart's (2010) executive-parties dimension, which considers factors beyond district magnitude including executive-legislative relations, federalism, and interest-group pluralism. Lijphart's Executive Dimension Score ranges from -1.48 in the UK to 1.67 in Switzerland. The parameter estimates are similar to those reported in the body of the text: more power-sharing (less majoritarian) institutions are associated with less intense affective polarization, while income inequality and unemployment are associated with greater affective polarization. The one interesting difference is that we now estimate statistically significant effects of elite ideological polarization.

Table S4. Robustness checks: Lijphart's measure of institutions

INDEPENDENT VARS	(1)
Elite ideological polarization	0.155**
	(0.060)
Income inequality	0.034*
	(0.013)
Unemployment level	0.057**
	(0.012)
Lijphart's Executive Dimension Score	-0.347**
	(0.055)
Intercept	4.417**
	(0.490)
N	75
Adjusted R^2	0.663

** $p \leq .01$; * $p \leq .05$, two-tailed tests.

Notes. The dependent variable in these analyses is [*Out-party dislike* (*t*)], defined as the mean respondent evaluation of out-parties in the current election survey (with the scale reversed so that higher numbers denote greater out-party dislike), computed over the set of election surveys listed in Section S1 in the Supplementary Information Memo. The independent variables are defined in the text. The top number in each cell is the unstandardized coefficient, the number in parentheses below is the standard error. The models were estimated with standard errors clustered on countries.

Section S6. Robustness checks: Difference-based affective polarization

Given that many scholars conceptualize affective polarization as the difference between partisans' in-party and their out-party thermometer ratings, we re-estimated our models using this differenced dependent variable. These analyses, as shown in Table S5 below, continued to support our conclusions, although they identified statistically significant effects of elite ideological polarization based on citizens' party perceptions (see discussion of this finding in the body of the text).

Table S5. Robustness checks: Difference-based affective polarization

INDEPENDENT VARS	Basic Model (1)	Fully-specified Model (2)	Fixed Effects Model (3)
Elite ideological polarization	0.179** (0.072)	0.208** (0.070)	0.162 (0.083)
Income inequality	0.070** (0.015)	0.063** (0.014)	-0.109* (0.046)
Unemployment level	0.045** (0.015)	0.042** (0.014)	0.061** (0.016)
Logged (DM) (t)		-0.073** (0.028)	
Intercept	1.368* (0.553)	1.676** (0.545)	7.419** (1.511)
N	75	75	75
Adjusted *R²*	0.404	0.404	0.695

** $p \leq .01$; * $p \leq .05$, two-tailed tests.

Notes. The dependent variable in these analyses is [*In-party versus out-party dislike differential* (*t*)], defined as the difference between the mean partisan respondent's thermometer rating of their in-party and their mean ratings of out-parties in the current election survey (with the scale reversed so that higher numbers denote greater dislike), computed over the set of election surveys listed in Section S1 in the Supplementary Information Memo. The independent variables are defined in the text. The top number in each cell is the unstandardized coefficient, the number in parentheses below is the standard error. The OLS regression models were estimated with standard errors clustered on countries.

Section S7. Robustness checks: Foreign-born population

Previous work has shown a relationship between immigration and elite ideological polarization in the US (McCarty et al. 2006). To the best of our knowledge, no study has examined whether immigration is also associated with mass-level affective polarization. In order to account for this possibility, we run the same model as in Table 3 in the body of the text but this time also with a measure of the share of foreign born population, based on OECD data. The results in Table S6 below find no statistically significant between foreign born population and affective polarization. In the OECD foreign born data, the following cases were missing: Finland (2015), Sweden (2014), Greece (2009), Portugal (2015), New Zealand (2014), Great Britain (2015), Iceland (1999), Israel (1006), Spain (1996), leaving us with an N of 66. The results for the other variable are substantively similar as to those in our previous analyses. As shown in model 2, when controlling for the level of foreign population, our measure of elite ideological polarization becomes statistically significant.

Table S6. Robustness checks: Foreign-born population

INDEPENDENT VARS	Fixed Effects Model (1)	District Magnitude Model (2)
Foreign born population	0.024	0.005
	(0.030)	(0.009)
Elite ideological polarization	-0.017	(0.152)*
	(0.124)	(0.121)
Income inequality	-0.026	0.045**
	(0.058)	(0.018)
Unemployment level	0.064**	0.075**
	(0.025)	(0.017)
Logged (DM) (*t*)		-0.146**
		(0.044)
Country Fixed Effects	Yes	No
Intercept	6.391	4.178
	(1.791)	(0.916)
N	66	66
Adjusted R^2	0.781	0.608

** $p \leq .01$; * $p \leq .05$, two-tailed tests.

Notes. The dependent variable in these analyses is [*Out-party dislike* (*t*)], defined as the mean respondent evaluation of out-parties in the current election survey (with the scale reversed so that higher numbers denote greater out-party dislike), computed over the set of election surveys listed in Section S1 in the Supplementary Information Memo. The independent variables are defined in the text. The top number in each cell is the unstandardized coefficient, the number in parentheses below is the standard error. The OLS regression models were estimated with standard errors clustered on countries.

Section 8. Robustness checks: Excluding Portugal, Spain, and Greece

As explained in the body of the text, some readers may be concerned that our results are shaped by three potential outliers: Portugal, Spain, and Greece. These three countries democratized significantly later than other countries in our sample, and were also among the hardest hit by the global financial crisis. We there for estimate the same regression analyses reported in the body of the text but this time excluding these three countries from our sample. The results, as shown in Table S7 below, continued to support our conclusions

Table S7. Robustness checks: Excluding Portugal, Spain, and Greece

INDEPENDENT VARS	Basic Model (1)	Fully-specified Model (2)	Fixed Effects Model (3)
Elite ideological polarization	0.046 (0.134)	0.116 (0.120)	0.008 (0.086)
Income inequality	0.062** (0.030)	0.049** (0.021)	-0.041 (0.027)
Unemployment level	0.057* (0.032)	0.057* (0.030)	0.073*** (0.022)
Logged (DM) (t)		-0.142** (0.044)	
Intercept	3.813** (1.287)	4.317** (0.957)	7.385** (0.927)
N	65	65	65
Adjusted R^2	0.260	0.507	0.756

** $p \leq .01$; * $p \leq .05$, two-tailed tests.

Notes. The dependent variable in these analyses is [*Out-party dislike* (*t*)], defined as the mean respondent evaluation of out-parties in the current election survey (with the scale reversed so that higher numbers denote greater out-party dislike), computed over the set of election surveys listed in Section S1 in the Supplementary Information Memo. The independent variables are defined in the text. The top number in each cell is the unstandardized coefficient, the number in parentheses below is the standard error. The OLS regression models were estimated with standard errors clustered on countries.