In or Out? The Effects of Supranational Integration on National Politics and Political Polarization in Multi-party Systems

Nikitas Konstantinidis  Hande Mutlu-Eren
University of Cambridge  New York University
Konstantinos Matakos
King’s College London
August 29, 2016

Abstract

The European debt crisis ushered in a period of political instability, high cabinet turnover, and ideological polarization in European parliamentary democracies. The crisis has sapped the EU’s ‘output legitimacy’ without giving rise to any significant counterbalancing measures enhancing the Union’s so-called ‘input legitimacy’. As a result, the tightening policy constraints of EU membership coupled with the decline of public trust in the EU’s ability to deliver have reversed the trend towards higher levels of party-system polarization. In this paper, we first develop a formal version of the above argument and then proceed to test it against different aggregate measures of polarization in EU member states. Our findings suggest that, while moderate supranational policy constraints can initially speed up platform convergence and reduce polarization, deeper integration and extremely tight policy constraints can have the opposite effect. Under those circumstances, parties that care both about their vote share and their unity will no longer find it profitable to converge to the median,
since the loss from the disunity of the party will outweigh the gains from increased electoral influence.

Keywords: party-system polarization, multi-party systems, European integration, policy constraints

JEL classification:
“But for a man [...] not to be drawn too taut is no shame. Likewise, the man in command of a ship who draws the foot sheet taut and leaves no slack, capsizes and sails what is left with his decks upside down.”

Sophocles, Antigone

1 Introduction

The global financial crisis and the subsequent European debt crisis ushered in a period of political instability, high cabinet turnover, and ideological polarization. The rise of anti-systemic right-wing nationalism in the European ‘North’ and left-wing populism in the European ‘South’ (or even the Trump phenomenon in the US) can be viewed through the prism of the same underlying phenomenon of a populist backlash against the ‘straitjacket’ of globalization. In one European country after another, the spectre of ungovernability is forcing moderate pro-EU parties into uneasy coalitions. Even in a country that lies at the outer layers of European integration like the UK, the fundamental question of EU membership runs to the core of government and creates deep fissures in the cabinet, parties, and society at large in the context of the impending British referendum. Effectively, the common thread that connects all these recent domestic political developments in Europe and beyond is that of polarization in connection with increasingly constraining supranational integration and economic interdependence.

The broadening and deepening of the European integration process have reshaped the contours of domestic electoral competition by bundling up a number of policy areas and delegating them to the supranational level. This has several implications. First, national democracies and political systems have become increasingly interdependent through the emergence of potential negative spillovers vis-à-vis the European integration process. In other words, economic interdependence has brought about political interdependence, thus raising demand for further political integration. The global attention attracted by the Greek national elections of May and June 2012 (when Greece’s Eurozone membership was hanging on the balance) is a telling example of the heightened uncertainty and political salience of national elections for the regional (or even) global economy. Second, the tra-
ditional left-right dimension seems no longer relevant, since the traditional left and right extremes have seemed to converge in terms of their anti-European and anti-globalization rhetoric and join up against the moderate, pro-European center.\footnote{This kind of political dynamic has been in full display in the run-up to both the Greek referendum of June 2015 and the upcoming British referendum on EU membership, where euroskeptics and europhobes politicians from the extreme left and right of both countries have joined forces against the embattled moderate pro-EU camps.} The concomitant rise in populism (Pappas 2014), euroskepticism, and extremism in a number of European countries (Halikiopoulou et al. 2012) are all symptoms of an inherent process of polarization of national party systems. The latter will be the focus of this paper, which seeks to theorize and test the causal mechanism linking political polarization in parliamentary multi-party systems to the embedded process of economic and political integration in Europe.

The concept of political polarization has received much attention by political scientists of late. One may easily juxtapose this with the study of economic inequality in economics, insofar as it bears interesting positive and normative implications for the state of the political system. Higher polarization implies heightened policy uncertainty for citizens and market participants (or political country risk for potential investors). In light of the political instability affecting countries in the aftermath of the Eurozone debt crisis (as evidenced by the shocking results of the May 2014 European election), increased polarization along the pro/anti-EU axis and the potential rise of anti-European governments may generate negative spillover effects for the entire European project. This may spell the end of consensus politics at the supranational level, the rise in the political implementation costs of reforms and EU directives, and hence the political unraveling of the integration process. In parallel, the heightened incidence of negative referendums on EU treaties may have similar effects (Dür and Konstantinidis 2013), although when the incumbent government is pro-European these may be more easily mitigated through framing and minor treaty revisions.

The Eurozone debt crisis has highlighted the growing incongruence between the democratic functions of representative government (‘by the people’) and those of responsible government (‘for the people’). Policy prescriptions no longer seem compatible with the politics of representation and contestation. Executive dominance now seems to characterize the predominant current of Western politics in which emphasis has come to lie on
the so-called ‘output legitimacy’ of democratic institutions. In contemporary democracies responsible (or output-oriented) government takes prevalence over representative (or input-oriented) government (Mair 2009). Katz and Mair (1995) have described a perceived change in political party organization as part of a reciprocal development towards more output-oriented, technocratic politics.

‘Output legitimacy’ pertains to the notion that a democratically elected government should ‘care for the common good’ and can potentially be at tension with so-called ‘input legitimacy’, which amounts to a government’s need to ‘carry out the will of the people’ (Scharpf 1999). The two may be at odds when a government decides to pursue a policy that in its own assessment would produce the best functional outcome for society as a whole, even if the electorate would opt for a different policy direction, resulting in what Scharpf (1999) names a tension between efficiency and democracy.²

European integration has made leaps and bounds on the basis of so-called ‘output legitimacy’ as a successful technocratic project generating growth and stability across Europe. However, all this came to a halt in the face of economic stagnation, rising unemployment, and sustained welfare cuts. The crisis has sapped the EU’s ‘output legitimacy’ without giving rise to any significant counterbalancing measures enhancing the Union’s so-called ‘input legitimacy’. As a result, both technocracy and populism thrive in an institutional setting where the executive becomes increasingly insulated from parliamentary and societal pressures (Bickerton 2012). The lack of different policy options fuels the rise of populist parties that claim to still represent the ‘will of the people’ (Mudde 2004) but rarely deliver, as opposed to mainstream parties, which take responsibility but no longer seem to voice their voters concerns. By amplifying this dynamic, the emergency politics of the crisis era in short undermine democratic representation through national (or supranational) parliaments.

In this paper, we seek to derive the relationship between ‘input’ and ‘output’ legitimacy and party-system polarization in the context of national democracies. Although both concepts of legitimacy and accountability are a direct function of supranational processes and

²This trade-off is particularly evident in the current British referendum debate between ‘Brexit’ supporters’ insistence on democratic sovereignty outside of the EU and ‘Remain’ advocates’ emphasis on the sustained trade, prosperity, and economic stability afforded by EU membership.
centralization of policies, here we are interested in how they are projected onto the domestic level and mediated through national democratic institutions. In other words, we study the ‘second-image-reversed’ effects of systemic processes and supranational policies on national political outcomes. To that effect, we proxy for the idea of ‘input legitimacy’ by postulating (exogenously) imposed constraints that truncate the admissible policy space of domestic electoral competition. The less influence voters have on the formulation of those constraints through their national representatives, the lower the ‘input legitimacy’ of representative government. Especially for smaller countries with little political clout, the only indirect source of democratic ‘input legitimacy’ was their elected governments’ decision to abide by those externally imposed rules and international commitments that came with membership in the international union. Up until recently, that decision had rarely been put to the test by means of mechanisms of direct democracy such as referendums. Therefore, the recent trend towards the use of referendums on membership may be viewed as efforts to bolster the ‘input legitimacy’ of a multi-tiered democratic polity.

Especially during the early years of the euro and the booming 2000s, responsible governments seemed secure in the ‘output legitimacy’ of the multi-layered European structure as it was perceived as a bulwark of economic stability and sustained growth. Output-driven notions of democratic accountability and economic voting allowed moderate pro-EU parties to reap the valence benefits of EU membership and trust in EU institutions. Given that in an environment of low transparency and high complexity voters are not able to identify the locus of policy formation and thus to assess the relationship between treatment and effect, they often rely on non-economic proxies, cues, and public signals in order to form political preferences and beliefs (Ezrow 2010, Hellwig 2008, Hellwig 2014). By that account, in both the model and the empirical analysis that follows, we show how ideological convergence and declining party-system polarization came as the result of the popularity and the positive public perception of EU membership and how this trend has been reversed in the face of deteriorating economic conditions and declining trust in the EU’s ability to deliver.

Political polarization may manifest itself in various stages of the political process (electoral, government formation, legislative, media coverage, etc.). Thus, the concept lends

---

3See Indridason (2011).
itself to different levels of measurement: at the voter level (using public opinion data), at
the party level (using party manifesto data), at the electoral level (using vote share data),
or at the representation level (using seat allocation data). In this paper, we do not put
forward any argument about the causal electoral linkages in elite/mass interactions. In
other words, we are not trying to determine whether polarization in multi-party systems
is mostly driven by parties or voters. We just abstract away from these empirical regulari-

ties and correlations with respect to polarization in its various forms, focusing instead
on how external policy constraints (not constitutional ones) affect rhetorical party-level
polarization and platform divergence within the context of electoral competition.

The question of Europe and European integration has featured as a separate and or-
thogonal dimension in domestic electoral competition (along the pro/anti- integration
dimension). Yet, successive intergovernmental treaties at the EU level have increasingly
tied a number of issue areas together. In other words, transnational policy areas, such
as immigration, taxation, environmental regulation, etc., that used to be relatively inde-
pendent and orthogonal, are becoming increasingly intertwined and correlated as part of
the deeper commitments of EU membership. As a result, the degrees of freedom of elec-
toral competition have shrunk and the range of credible and internally consistent policy
platforms has become more constrained.

This process is even more pronounced in countries that belong to the European core
(namely EMU and Schengen Treaty), those subject to high levels of external condition-
ality, and those on the losing side of the globalization/integration debate (e.g., deficit
countries of the European ‘South’ with little influence over the formulation of EU poli-
cies). The external constraints on the traditional left-right policy dimension originate
from a range of sources and vary accordingly: they may stem from the formal or infor-
mal intergovernmental commitments of EU or Eurozone membership (e.g., the Growth
and Stability Pact or more recently the Fiscal Compact), the ‘hard’ process of compli-
ance, transposition, and implementation of EU rules and directives, the conditionality
agreements with the EU or the IMF, and finally the ‘soft’ imperatives of EU associa-
tion and globalization more generally.\footnote{Arguably, globalization increasingly exerts
pressure on governments to implement monetary discipline, deregulation, privatization,
and labor market flexibility in order to enhance competitiveness (Milner...}
else the degree of sovereign policy discretion) to vary with respect to country size, (formal or informal) influence in the EU political process, enforcement costs, the economic environment, and the perceived benefits of European membership (pro-European sentiment in public opinion). An implication of this transformation of the political space is that the traditional left-right dimension has become more blurred, as the traditional ‘left’ and ‘right’ extremes have seemed to converge in terms of their anti-European and anti-globalization rhetoric and join up against the moderate, pro-European center. Moreover, national elections function less effectively as a way of translating citizens’ preferences into governmental policy.

Unlike much of the existing formal literature that focuses on the effects of constitutional rules such as the electoral system on political polarization,\(^5\) we focus on extraneous policy rules and constraints that stem from a country’s set of international commitments. The concept of polarization plays out differently in the context of majoritarian and PR systems. Within the burgeoning polarization literature in US politics, political polarization becomes tantamount with ideological differentiation between the two main parties and the emphasis is on its institutional underpinnings (Hirano 2013, McCarty 2009) and its impact on electoral turnout, the menu of political choices, media coverage (Prior 2012), and the possibility of bipartisanship in Congress (McCarty et al. 2006). In that regard, it is easier to study polarization in systems with smaller number of effective parties and higher number of single-party governments, focusing on the effects of the disproportionality of the electoral system.

In the context of parliamentary PR systems, however, the concept of polarization is not as straightforward in terms of its measurement or implications. Polarization in PR systems may be the result of ideological differentiation and/or party fragmentation, whereby niche groups of extremist ideologues may splinter off from larger office-seeking parties. The aggregate picture becomes even more convoluted once we consider that convergence may take place among centrist parties and divergence between extremist and centrist ones. The combined effect of higher divergence in policy platforms and lower political influence

(in terms of vote share) among extremist parties has an ambiguous effect on aggregate levels of polarization. This might explain why existing measures of polarization proposed in the economics and political science literatures present shortcomings when applied to multi-party systems (Duclos et al. 2004, Esteban and Ray 1994, Stanig 2011). Until recently (e.g., Calvo and Hellwig 2011; Matakos et al. 2013), polarization in PR systems might not have featured high on the research agenda because coalition governments have for the most part been formed by moderate centrist parties and, therefore, the expected policy locus has not tended to shift as much.6

Another strand of the literature that we seek to address refers to the Europeanization of national party systems and the transformation of electoral competition as a consequence of regional integration (Bernhard 2004, Dorussen and Nanou 2006, Mair 2000). Nanou and Dorussen (2013), for example, argue that EU legislation limits the policies that parties in the Member States can pursue once in government, thus also constraining the policy positions that parties offer at national elections. Their main finding is that in policy domains where the involvement of the EU has increased, the distance between parties’ positions tends to decrease. We, on the other hand, take a broader view of polarization as a systemic variable and differentiate between convergence among moderate pro-European parties and divergence at the extremes of the ideological spectrum. By making use of aggregate measures of polarization weighted by party vote shares, we do obviate the need for an arbitrary distinction between mainstream and non-mainstream parties.

In what follows, we first develop a formal demonstration of the above argument by deriving the incentives of extremist parties to converge to (or diverge from) the position of the centrist party in a game of three-party electoral competition with proportional representation. We find that binding external constraints upon a unidimensional spectrum of policies will induce convergence up to the point where internal pressures form party activists and the party base will cause divergence back towards the party’s intrinsic ideology. We then proceed to test this non-linear effect against a panel dataset using different aggregate measures of polarization and integration. We do find rather strong evidence confirming our hypothesis of conditional convergence and divergence (polarization).

6A related argument put forward by Tsebelis (2002) is that a higher number of effective veto players in PR systems ensure more policy stability.
2 Theoretical framework

In this section we develop our theoretical argument regarding the relationship between the ‘input’ –captured by the extraneous policy constraints placed on the local polity by a supranational institution– and ‘output’ –proxied by the indirect popularity of union membership– legitimacy of an embedded national democratic system and political (party-system) polarization, as measured by the distance between parties’ political platforms. For purposes of expositional simplicity alone, and in order to illustrate the intuition behind our empirical implications, we will introduce some formal notation and develop a stylized model in order to fix ideas. Its purpose is simply to isolate the main forces that mediate the effect of supranational integration on party-system polarization and illustrate the trade-offs involved with greater clarity.

2.1 Political environment

Since our study applies mostly to multi-party parliamentary democracies (e.g., almost all EU countries), we consider that electoral competition takes place among three parties. We also implicitly assume that the electoral rule is some form of proportional representation, which is almost always the case in such multi-party parliamentary systems (see e.g., Iaryczower and Mattozzi 2013). Without wishing to complicate the analysis further at this stage, it follows from the latter that the government formation process will entail some sort of coalition building –it will become evident that our empirical predictions do not hinge on different assumptions regarding the government formation process. The electoral competition takes place along the unidimensional, left-right policy space $[0, 1]$. Voters are distributed over the unit interval according to a uniform distribution and vote sincerely for their most preferred party, the party whose proposed platform is closest to their ideal point. We assume that voters have symmetric single-peaked preferences over the policy space. That is, their utility symmetrically decreases as the proposed policy

---

7 Complete formal arguments are provided in the appendix.
8 The results and the main argument carry through even if we assume more than three parties. We only focus on the three-party case for analytical tractability and expositional simplicity since it is the simplest case of a non-degenerate (i.e., more than two) multi-party system.
moves further away (in any direction) from their most preferred policy point (bliss-point).

Parties occupy a position, denoted by $p_j$, in the left-right policy space and propose a platform $x_j$ with the primary goal of maximizing their vote share. We think of parties’ vote shares as a rough, but fair, approximation of the probability that they participate in a coalition government and, hence, of their ability to acquire office rents (e.g., the number of ministerial portfolios allocated to each coalition member usually depends on its vote share). Parties also seek to minimize the costs incurred the further away their proposed platform ($x_j$) from their initial policy position $p_j$ is. This cost, which is strictly increasing in the distance between the proposed platform and the party’s initial position, can reflect the communication or ‘transportation’ costs that a party must pay in order to re-brand itself. For example, this could reflect the increased cost that must be paid by party-members who have to sell the party’s new platform to the voters. $^9$ An alternative interpretation is that parties might pay a cost for displeasing some of their internal factions (e.g., Roemer 1998). $^10$ Without any loss in generality (see also Matakos et al. 2015), assume that initially parties are symmetrically positioned along the full range of the policy spectrum, i.e., the left party has an initial position $p_L = 0$, the centrist party has a position $p_C = 0.5$, and the right party has a position $p_R = 1$. Then a utility function for party $j \in \{L, C, R\}$ that captures this trade-off is depicted below:

$$U_j(x_j, p_j) = -c(|x_j - p_j|) + v_j(x_j, x_{-j}),$$

where $p_j$ is party $j$’s initial position, $x_j$ is its policy platform proposal, $c(\cdot)$ is a cost function that is increasing in the absolute distance $|x_j - p_j|$, and $v_j(x_j, x_{-j})$ is party $j$’s

$^9$In general, there is a clear analogy between these communication costs and the transportation costs that firms pay in the standard Hotelling (1952) model.

$^{10}$While the analysis of intra-party factions evades the scope of this present study, one can easily justify this assumption if one considers a party containing two factions: the ‘pure ideologues’—who only want to implement their ideal policy without any compromise—and the ‘pragmatists’ who only care about winning elections and thus maximizing the party’s vote shares (de Mesquita and Friedenberg 2011; Ghosh and Tripathi 2012). In such a case, if the ‘pragmatists’ faction is in control of party leadership, the proposed utility representation reflects the increasing costs that this faction has to pay as it moves away from the party’s initial position. Clearly, those costs are proportional to those deviations as the more the party drifts away from its initial position the higher the dissatisfaction of the ‘ideologues’ towards the more ‘pragmatic’ party leadership.

11
vote share as a function of all parties’ proposed platforms.\textsuperscript{11} The first part of the utility function captures the increase in the the ‘transportation’ costs when a party moves further away from its initial point, while the second part reflects the benefits of convergence –as a party moves towards the median, it increases its vote share. Given that the decisive voter is the median, extremist parties face the following trade-off: on the one hand, moving towards the center (and closer to the median voter) increases their vote share and, hence, their chances of being in government and sharing in the spoils of the office; on the other hand, this convergence towards the center increases their communication costs. Thus, in equilibrium, this trade-off will be reflected by some limited convergence towards the center of the policy space by extremist parties, as long as the ‘transportation’ costs are not very high.\textsuperscript{12}

2.2 Introducing policy constraints: The mechanism

In the absence of any other changes in the institutional set-up (such as changes in the electoral rule; e.g., see Matakos et al. 2015), the degree of policy convergence towards the center will depend on the size of spoils of office (centripetal force) and the ‘transportation’ costs (centrifugal force).\textsuperscript{13} But what happens when additional policy constraints are imposed exogenously as a result of a country’s participation in supranational integration schemes? Are the incentives, faced by extreme parties, to converge towards the center of the policy space amplified or mitigated? To answer this question, we need to shed more light on the mechanism at play. The supranational institution of which a country is a member can impose an exogenous constraint on the domain of admissible policies and, effectively, tie the hands of national governments. That is, the spectrum of admissible policies is truncated and, hence, the space of electoral competition is effectively constrained. For example, the supranational principal may require all national member-state governments to harmonize their tax systems or their welfare policies. For simplicity, we

\textsuperscript{11}For example, if $c(|x_j - p_j|) = (x_j - p_j)^2$, then we have a quadratic loss function.
\textsuperscript{12}In the appendix, we provide a complete characterization of the unique symmetric Nash equilibrium for different functional forms of the cost function. Nevertheless, as it will become evident, our argument does not depend on our choice of a specific functional form.
\textsuperscript{13}In the example above, the spoils of office were normalized to unity without any loss in generality and the cost function was assumed to be the Euclidean distance.
assume that the constrained set of admissible policies symmetrically excludes policies too much to the left or the right of the policy spectrum. That is, the policy platform that an extremist party would like to propose—absent any exogenously imposed constraint—might fall outside the permissible policy space. Then what will an extremist party do? Will it converge further towards the center in order to abide by the constraint? Or will it continue to advocate policies that fall outside the spectrum of admissible policies?

In order to examine this mechanism in more detail, we need to add some more structure to the problem. While maintaining the assumptions outlined above, we further assume that the exogenous policy constraint is measured by a parameter $\alpha \in (0, \frac{1}{2})$, such that the permissible space of domestic policies—if a country chooses to remain a member of the supranational institution—becomes the interval $\left[\frac{1}{2} - \alpha, \frac{1}{2} + \alpha\right]$. That is, $\alpha$ measures the slackness of the constraint: the larger its value, the lower the severity of the constraint and vice versa. Then, the utility function below summarizes the trade-off faced by parties:

$$U_j(x_j, p_j) = \begin{cases}  
-c(|x_j - p_j|) + v_j(x_j, x_{-j}) + b_2, & \text{if } x_j \in \left[\frac{1}{2} - \alpha, \frac{1}{2} + \alpha\right] \\
-c(|x_j - p_j|) + v_j(x_j, x_{-j}) + b_1, & \text{if } x_j \notin \left[\frac{1}{2} - \alpha, \frac{1}{2} + \alpha\right]
\end{cases}$$

where $b_1, b_2$ are popularity shocks that a party experiences if it proposes a platform that is consistent ($b_2$) or not ($b_1$) with the exogenous constraint and all other parameters are the same as defined above. That is, if ‘output legitimacy’ is high and the supranational institution is very popular, we have that $b_1 < 0 < b_2$, while the converse is true if the ‘output legitimacy’ is low and the supranational institution is not too popular ($b_2 < 0 < b_1$). Furthermore, without any loss in generality, one can define $\beta \equiv b_2 - b_1$ and normalize $b_1 = 0$. Then $\beta > 0$ implies high ‘output legitimacy’ (and a party gains in popularity if it proposes a platform that falls within the constraint), while $\beta < 0$ implies low ‘output legitimacy’ (the experiences a positive popularity shock if it proposes a platform outside of the constrained policy space). In other words, this formulation of preferences (which is a simplification of a standard valence model) implies that a party experiences

---

14In the limit case, where $\alpha = \frac{1}{2}$, the constraint is not binding and, hence, the second line becomes irrelevant. By analogy, if $\alpha = 0$, then the only permissible policy point is $\frac{1}{2}$, which implies that parties will always pay a cost, unless they all propose an identical platform.
a positive (negative) popularity shock if it proposes a platform that falls within (outside) the exogenously imposed policy constraint if and only if the supranational institution is sufficiently popular (high ‘output legitimacy’). The converse is true if the supranational institution suffers from low ‘output legitimacy’. Since it would have been trivial to focus on the case where \( \beta < 0 \) –there would never be any convergence–, henceforth, we assume that \( \beta > 0 \). Below, we present a simple example in order to expose the trade-off between the tightness of the policy constraint –a result of increasing integration– and extremism.\(^{15}\)

### 2.2.1 An Example

In order to fix ideas further and convey our point with greater clarity, start from a position of polarized divergence (i.e., \( x_L^* = p_L = 0 \) and \( x_R^* = p_R = 1 \)), in the absence of any exogenous constraint (Proposition 1 in the appendix), where there is no convergence to the center by the extreme parties \( L \) and \( R \), and thus, extremism—understood as the distance of the two most extreme platforms from the median position—is maximized.\(^{16}\) Now consider the case where a binding policy constraint (\( 0 < \alpha < \frac{1}{2} \)) is introduced. In such an environment, the previous platform proposals of the two extreme parties fall outside the constrained policy space \( \left[ \frac{1}{2} - \alpha, \frac{1}{2} + \alpha \right] \). This implies that they face the dilemma of whether to propose a policy platform on the boundary of the constrained policy space or to stick with their equilibrium platform proposal \( x_j^* \).\(^{17}\) Clearly, their decision to converge or not will depend on: a) the tightness of such a constraint (which we associate with the institutions’ ‘input legitimacy’), and b) the benefits (or costs) of compliance (which we

---

\(^{15}\) For the interested reader, we defer the presentation of detailed equilibrium characterization results to the appendix.

\(^{16}\) Formally speaking, this is indeed the unique symmetric Nash equilibrium of the game (when there are no policy constraints) if the cost function is the Euclidean distance (Proposition 1).

\(^{17}\) As we formally show in the appendix, all other platform choices that lie between \( x_R^* \) and \( \frac{1}{2} + \alpha \) are strictly dominated strategies for party \( R \) in expected utility terms. If it proposes a platform to the left of \( x_R^* \) that still falls outside the permissible policy space, it will suffer the increased ‘transportation’ cost without receiving the positive net popularity shock; if it proposes a platform strictly within the interior of the permissible policy space, then this only increases its transportation costs without further increasing the positive popularity shock that it experiences – once within the permissible domain a party receives no further benefits. This leaves only two options as candidates: the initial equilibrium platform \( x_j^* \) or the boundary of the constrained policy space.
associate with institution’s ‘output legitimacy’). Therefore, the two extreme parties, $L$ and $R$, face the following dilemma: should they converge towards the center and propose a platform that lies within the constrained policy space, or not?

The trade-off is the following: by proposing a policy platform that lies on the boundary of the constrained policy space they pay a ‘transportation cost’, which is increasing in the tightness of the constraint, but, at the same time, they can both: a) increase their vote shares by moving towards the median, and also b) enjoy a positive popularity shock as their proposed platforms appear to be congruent with the policies advocated by the supranational institution; by sticking to their initial platforms, which lie outside the constraint, they experience a (potentially) large but fixed negative popularity shock – by proposing a non-congruent platform– that can cost them some votes, but they do not have to suffer any additional ‘transportation’ costs. That is, the tightness of the policy constraint increases the cost of convergence, thereby creating centrifugal incentives, while the positive popularity shock creates centripetal incentives for extremist parties to capitulate to the will of the supranational institution and propose a congruent platform that satisfies it.

To see this trade-off more clearly, we have to compare the utility they receive from proposing a policy on the boundary of the constrained space with the utility of proposing the initial equilibrium platform. Due to the symmetry of the problem faced by the two extreme parties, we need only focus on the behavior of one of them, say the right-wing extremist party $R$. Party $R$ will chose among the two options depending on which one yields the higher utility. Recall that voters are uniformly distributed in $[0, 1]$ according to a uniform distribution function with c.d.f. $F(z) = z$. Formally, the indifference condition between these two options is summarized below:

$$U_R(\hat{x}_R = 1/2 + \alpha) = U_R(x^*_R = 1)$$

which, in turn, implies that

---

18 As we show in the appendix, while the equilibrium platform proposal of the centrist party will always fall within the constrained policy space (i.e. $x^*_C = 0.5$), the same is not necessarily true for the more extreme ones.
Then, solving for $\alpha$ yields the threshold value $\alpha^*(\beta) \equiv \frac{1}{2} - 2\beta$. That is, whenever $\alpha > \alpha^*(\beta)$ (the constraint is relatively loose), party $R$ will prefer to choose platform $\hat{x}_R$, that is, converge; otherwise, whenever the constraint is too tight, i.e., $\alpha < \alpha^*(\beta)$, then party $R$ is better off not converging towards the center and proposes platform $\hat{x}_R$ as in the case when no constraints were introduced. Clearly, the threshold value of $\alpha^*(\beta)$ is strictly decreasing in $\beta$; the higher the positive popularity shock (due to increased ‘output legitimacy’), the tighter the constraint can be without inducing the extremist party to diverge. That is, if the ‘output legitimacy’ of the supranational institution is high (higher value of the popularity shock $\beta$), then convergence towards the center can be sustained (and thus extremism need not increase by that much) even if the policy constraint becomes very tight –that is, even if integration intensifies. A threshold value of $\alpha^*$ close to zero implies that extremist parties will prefer to converge even if the constraint is very tight and the domain of permissible policies is extremely narrow around one-half.

We will return to the interplay between ‘output legitimacy’ and the degree of convergence in the next section when we discuss the policy implications of this relationship. Here, a brief comment regarding the (a)symmetry of the constraint is in order. Consider the case where the supranational institution has a policy bias. That is, it might allow more slack on the left (right) of the policy spectrum, such that the permissible policy space becomes the interval $[\frac{1}{2} - a_1, \frac{1}{2} + a_2]$ with $a_1 \neq a_2$ (without any loss in generality assume that the constraint is more slack on the left and, hence, $1/2 > a_1 > a_2$). Then, it is straightforward to argue that the main qualitative features of our result will go through. What will happen is that, in equilibrium, the two extremist parties will occupy asymmetric positions affecting the speed of convergence: the party that faces the more tight constraint will converge (and diverge when the constraint becomes too tight) faster. Yet, the overall effect on extremism (and polarization) will remain unchanged: once the constraint becomes (asymmetrically) too tight extremism will eventually increase.
Next, we focus on the main prediction of our theoretical framework, that is, the relationship between the degree of supranational integration (measured by the slackness, or tightness, of the constraints on domestic policies) and extremism for any given value of $\beta$. For the purposes of this comparative statics exercise, we define extremism as the distance of the two extremist platforms from the median party position (party C); that is, $v_L(|x_R^* - x_C^*|) + v_R(|x_L^* - x_C^*|)$. Moreover, following a vast body of literature (e.g., Andrews and Money 2009; Matakos et al. 2015), we will also define polarization as the distance between the two most extreme party-platforms. That is, polarization and extremism are positively and monotonically related. For illustrative purposes, we fix the value of $\beta = \frac{1}{8}$ (moderate ‘output legitimacy’) and we focus on the following three cases with low, intermediate, and high levels of integration implying low, medium, and high policy constraints such that $\alpha^L > \alpha^M > \alpha^H$ respectively. 19

**Case 1: Low level of integration ($\alpha^L = 0.4$)**

Consider first the case where the policy constraint placed by the supranational institution on policies at the national level is very slack (e.g., the early years of the EEC). In this case, the threshold value $\alpha^*(\beta = \frac{1}{8})$ that makes an extreme party indifferent between converging towards the center or remaining at its position is $\frac{1}{4}$, and, hence, we have $\alpha^L > \alpha^*$. This, in turn, implies that both extremist parties will have strong incentives to converge towards the center and propose a platform that lies on the boundaries of the constrained policy space. That is, party $R$ will propose $\hat{x}_R = 1/2 + \alpha = 0.9$ and (by symmetry) the extreme left party $L$ will propose $\hat{x}_L = 1/2 - \alpha = 0.1$. In this case, extremism shrinks from its maximal value of 0.5 (without any constraint) to 0.4. Hence, mild supranational integration decreases extremism: the extreme parties have strong incentives to converge towards the center as a result of the exogenously imposed constraint.

**Case 2: Intermediate level of integration ($\alpha^M = 0.3$)**

Consider now the case where supranational integration intensifies and the supranational institution places more tight constraints on national policies (e.g., the treaty of Maastricht in 1992), thus shrinking the permissible policy space even more. In this case, again $\alpha^M > \alpha^*(\beta = \frac{1}{8}) \equiv 1/4$, which implies that the two extreme parties will converge further

---

19 Recall that parameter $\alpha$ measures the slackness of the policy constraint; the lower its value the more tight the constraint is and vice-versa.
towards the center, proposing $\hat{x}_R = 1/2 + \alpha = 0.8$ and $\hat{x}_L = 1/2 - \alpha = 0.2$, respectively. As a result extremism further decreases (from 0.4 to 0.3) as a consequence of even closer supranational integration.

**Case 3: High level of integration ($\alpha^H = 0.2$)**

But what will happen if the process of supranational integration intensifies further? Will extreme parties continue to converge towards the center (and approach each other) thus reducing extremism, or not? In this case, the policy constraint is too tight ($\alpha^H < \alpha^*(\beta = 1/8) \equiv 1/4$). As a result, both extremist parties prefer to propose their initial positions, $x^*_R = 1$ and $x^*_L = 0$ respectively, instead of converging. Moreover, notice that the two parties will not return back to the positions they had when the constraint was moderate ($\alpha = 0.3$) or low ($\alpha = 0.4$) – doing so will only increase their ‘transportation’ costs without generating any positive popularity shock as their proposed policy would still lie outside the space of permissible policies. Hence, the two parties will go all the way back to the extremes and, therefore, extremism will be maximized. That is, closer supranational integration (beyond a certain point) not only did not result in more policy convergence and less extremism (and polarization), but, on the contrary, it can result in extreme levels of polarization and divergence that were only observed in the pre-integration era. The latter implies that the relationship between supranational integration and polarization is non-monotonic: while during the first stages closer integration leads to a decrease in extremism, ever closer integration, ceteris paribus, will result in higher levels of extremism (and polarization). We discuss the implications of this finding in the section that follows.

### 2.3 Discussion

As the example above illustrates, initially, if a slack constraint is introduced ($\alpha$ is very large) extreme parties have incentives to converge and propose a platform within the constrained space as the costs of doing so are relatively mild and can be offset by the positive popularity shock that they will experience (recall that $\beta > 0$).\(^{20}\) But as the constraint becomes tighter (and $\alpha$ decreases), as a result of intensified integration, the

\(^{20}\)By converging towards the median, the party will also experience an additional gain in votes, which only strengthens our argument.
'transportation’ costs of proposing a platform that lies within the constrained space—and closer to the median—cannot be offset by the positive popularity shock that the party will experience and the few additional votes that it will gain. The reason is that when the constraint is too tight, satisfying it requires an extreme level of convergence to the center of the policy space, which, in turn, increases the ‘transportation’ costs that a party has to incur so much that the party is better off proposing an extreme policy that caters to its base despite the incidence of a negative popularity shock. This dynamic gives rise to a non-monotonic relationship between the tightness of the exogenously imposed policy constraint (due to membership in a supranational institution) and the degree of platform convergence to the center—and, consequently, the degree of extremism.

Furthermore, notice that the driving force behind this result is independent of the specific parametrization that we chosen for the cost function and the popularity shock. As long as the positive popularity shock does not take extreme values, then there will always exist a threshold value of $\alpha^*(\beta) \in (0, \frac{1}{2})$ such that a) whenever $\alpha > \alpha^*(\beta)$, an increase in the intensity of supranational integration (policy constraints become tighter and $\alpha$ decreases) reduces extremism and speeds up convergence, and b) whenever $\alpha \leq \alpha^*(\beta)$, a further increase in the intensity of supranational integration results in more extremism and divergence.\(^{21}\) That is, extremism is minimized (and convergence is maximized) when $\alpha = \alpha^*(\beta)$.

In sum, our results shed light on an interesting trade-off between supranational integration, policy convergence, and extremism. Since office-seeking parties have strong incentives to converge towards the center, even in the absence of exogenously imposed policy constraints, introducing moderate and relatively slack policy constraints intensifies the process of policy convergence and reduces extremism (and polarization). This is the case where $\alpha$ is relatively large. Yet, as these constraints become tighter (and $\alpha$ becomes larger than the threshold value $\alpha^*(\beta)$), the situation can backfire; it will no longer be the case that parties will converge. Instead, we observe extreme divergence in equilibrium. That is, ever closer supranational integration may have some unintended consequences:

\(^{21}\)Moreover, note our results do not hinge in the slightest at the assumption that the popularity shock is fixed. One can assume that this is changing continuously—just like the vote share—as the parties move towards the center and still obtain the same results, as long as the value of this shock is relatively moderate and smooth.
instead of further decreasing extremism it, can lead to greater platform divergence and polarization. This non-monotonic relationship between closer integration and extremism, ceteris paribus, is a testable prediction that our framework offers. We set out to test it in the section that follows.

3 Empirical Analysis

3.1 Data and Empirical Model

In order to estimate the impact of EU membership on polarization, we collected data for 15 EU countries for the years between 1958 and 2013. We created a party-level dataset using the Manifesto Project Dataset (Volkens et al. 2013). This dataset contains information on the left-right positions of political parties in fifty-five democracies in their election manifestos since 1920. It relies on saliency theory, which estimates parties’ position by looking at both negative and positive emphases of issues. The ideological scores range from -100 to +100, where higher values indicate a more right-wing emphasis. In the subsequent analysis, we rescaled the parties’ position to the range of 1 to 10.\textsuperscript{22} Each observation in our dataset is a country-election-party. We estimate the following equation:

\[
Extremism_{i,e,j} = \beta_0 + \beta_1 Trend_{i,e,j} + \beta_2 EU\text{Membership}_{i,e,j} + \beta_3 Trend \times EU\text{Membership}_{i,e,j} + \beta_4 Controls + \lambda_i + \gamma_t + \varepsilon_{i,e,j}
\]

where \(i\) indicates countries, \(e\) indicates elections and \(j\) indicates parties.

Dependent Variable: Extremism

We conceptualize polarization as extremism of political parties. Accordingly, the more extreme the parties are, the more polarized the party system becomes.

\textsuperscript{22}To rescale the index, we added 100 to each party’s position and divided the sum by 20.
We construct two measures of extremism: *Extremism (Mean)*, operationalized as the absolute distance between a party’s policy position and the mean party position weighted by vote share, and *Extremism (Median)*, operationalized as the absolute distance between a party’s policy position and the median voter. The data on median voter position comes from Somer-Topcu and Williams (2014), which in turn extends Kim-Fording data (Kim and Fording 1998).

Our dataset that has Extremism (Median) as the main dependent variable includes 1284 cases, and the one that has Extremism (Mean) as the main dependent variable includes 1423 cases.

*Independent Variables: EU Membership and Trend*

*EU Membership* is a dummy variable that is coded 1 when a country becomes a member of the EU and for the years thereafter, and 0 otherwise. To account for the deepening of the EU, we estimate the equation using a *Trend* variable that starts with 1957, and increases by 1 unit in every subsequent year. We expect EU membership to impose more constraints as the European integration deepens. At the same time, we expect countries to face more constraints if they join the EU later than earlier. To capture these effects we include an interaction variable, *Trend* *EU Membership* (Hix, 2003). We expect the marginal impact of EU membership to decrease extremism for early years of the EU and increase it for later years, when the constraints imposed by the EU increased considerably.

We used several control variables to account for factors that are likely to affect parties’ policy position and extremism of parties. First, the *Effective Number of Parties* is the effective number of electoral parties as calculated based on Laakso and Taagepera formula. If a country has a high number of effective parties, party policy positions are expected to be more dispersed, hence increasing extremism. This variable also captures the effects of changes in the electoral system.

Second, increased global interdependence forces governments to follow certain macroeconomic policies, and therefore is expected to decrease extremism independent of EU

\[ \frac{1}{\sum v_i^2}, \text{ where } v_i \text{ is the percentage of the vote received by party } i. \]
membership. This effect is captured by the *Openness* variable, which is measured as the size of a country’s exports and imports relative to its gross domestic product. We coded this variable from Penn World Table Version 8.1. Third, to capture the idea that the impact of EU membership on party extremism might be felt more strongly in countries that have a higher voting weight in the council, we included in the regressions *Population* variable operationalized as the square root of a country’s population. Finally, we included country fixed effects $\lambda_{it}$ to account for country-specific effects as well as decade fixed effects, $\gamma_{t}$, to capture the decade-specific effects.

We estimated an ordinary least squares (OLS) model with robust standard errors clustered by an index of country-years. This option assumes that observations are not independent within cluster, although they may be independent between clusters. The findings are presented in the next section and in the appendix.

Table 1 presents the estimates from the regressions where we analyze the impact of EU membership on extremism.

[Insert Table 1]

The first pair of regressions (Columns 1 and 2) examine the effect of EU membership on extremism measured as the distance from the median voter position. The second pair of regressions (Columns 3 and 4) examine the effect of EU membership on extremism measured as the distance from the mean party position.

The first column shows the estimates of the model where the main dependent variable is extremism measured as the distance from the median voter. The coefficient on *EU Membership* is negative and significant at the 10% level, while the coefficient on *Trend* is negative but not significant. In the second column we run the regressions with additional controls, *EMU* and *Single Market*, however, none of these is found to be significant. The third and forth column show the estimates of the model where the main dependent variable is extremism measured as the distance from the mean party position. The coefficients as well as their significant levels are similar to the previous model.

We are particularly interested in the impact of EU membership on extremism. To do this, we compute the marginal effect of EU membership on extremism for different values of trend; Figure 1 shows the marginal effect of the EU membership on extremism variable.
measured as the distance from the median voter (regression on Column 1 in Table 1). The two-tailed 90% confidence intervals around the line indicate the areas under which this effect is statistically significant. Accordingly, as expected, the EU membership exerts a negative on extremism for the early years of the EU. However, as time passes, this effect becomes positive. In particular, the effect of the EU membership on extremism is significant and positive after 43 years. Furthermore, this effect continues to increase with time. The implication of this finding is that in the period after 1999 when the EU was fairly integrated and constraints posed on member states were high, political parties are pushed towards the extremes.

[Insert Figure 1]

Figure 2 replicates the same exercise for the regression where extremism is measured as the distance from the mean party position (regression on Column 3 in Table 1). Accordingly, the marginal effect of the EU membership on extremism is negative and significant for the first 22 years of the EU, and this effect becomes positive but not significant after the first 32 years of the EU. Put differently, the EU membership started pushing the political parties to the extremes after 1989. This finding lends further support to our hypothesis that for the early years of the EU, political parties moved towards the mean political party position, but they started to move towards the extremes as the European integration deepened.

[Insert Figure 2]

So far, we tested the main prediction of our theoretical model, that is, for low levels of European integration the EU membership decreases party extremism and for high levels of European integration the EU membership increases party extremism. This finding is valid for any level of output legitimacy of the EU ($\beta$). As far as the effect of this popularity shock on extremism is concerned, it is expected that as the output legitimacy rises, convergence towards the center can be sustained even for tighter policy constraints. Put differently, we expect that for low levels of EU popularity, Europeanization leads to more polarization and for high levels of popularity this effect slows down. Testing this claim requires data.
on polarization at the party system level rather than party level. To this end, we gathered a new dataset where our main dependent variable is polarization measured at the party system level. We use this dataset also as a robustness test to our previous analyses. In the following section, we first estimate the effect of the European membership on party system polarization in a model that includes a new measure of Europeanization and then estimate the effect of Europeanization on polarization for varying levels of EU popularity.

3.2 Alternative Test and Estimating the Effect of Output Legitimacy

We collected data for 15 EU countries for the years between 1957 and 1993. Each observation in our dataset is a country-election. The analysis covers 129 cases in total. The empirical model that examines the impact of EU membership on party system polarization has the following functional form:

\[
Polarization_{i,e} = \beta_0 + \beta_1 \text{Europeanization}_{i,e} + \beta_2 \text{EU Membership}_{i,e} \\
+ \beta_3 \text{Europeanization} \times \text{EU Membership}_{i,e} + \beta_4 \text{Controls} + \lambda_i + \gamma_t + \varepsilon_{i,e}
\]

where \(i\) indicates countries and \(e\) indicates elections.

**Dependent Variable: Polarization**

Our dependent variable *Polarization* is measured as the distance between the most extreme parties on the left-right dimension for every election year in each country.\(^{24}\) Increased values point to more polarization on the left-right dimension. For the policy positions of parties we used the most recent Manifesto Project Dataset (Volkens et al. 2013).

**Independent Variables: Europeanization and EU Member**

\(^{24}\)Other studies that use the distance between the two most extreme parties to measure polarization include Budge and McDonald (2006) and Andrews and Money (2009).
Our main explanatory variable is *Europeization*, which is derived from the Europeanization index initially computed by Schmitter (1996) and updated by Schakel et al. (2013). This Europeanization index provides information on the level of EU competency across 28 issue areas in four main categories: economic policy, social/industrial policy, legal/constitutional policy, and international relations/security policy. The index is constructed based on the treaty obligations that are undertaken so far. A Europeanization score ranging from 1 to 5, where 1 implies that all policy decisions are taken at the national level and 5 indicates that all policy decisions are taken at the EU level, is assigned to these policy areas for each milestone periods of EU: the initial phase (1950-56), the European Economic Community (1957-67), the European Community (1968-1991), and the EU after Maastricht (1992-2000).

We constructed two Europeanization scores for each period: *Europeization* is calculated by taking the average of all scores across four main policy categories, and *Europeization (Economic)* is calculated by taking the average of scores only in the economic policy area. This variable captures the deepening of the European integration.

We estimated an ordinary least squares (OLS) model with robust standard errors clustered by countries.

Table 2 presents the estimates from the regressions where we analyze the impact of Europeanization on party system polarization.

[Insert Table 2]

The first column shows the estimates of the model with control variables, as specified above. *Europeization* is positive but not significant. *EU Membership* exerts a negative but insignificant impact on party system polarization. *Europeization* × *EU Membership* is both positive and significant at the 10% level. We are particularly interested in the impact of Europeanization on party system polarization at the national level. We employ two strategies to test this: first, we compute the marginal effect of EU membership on polarization for different values of Europeanization; and second we estimate the marginal effect of Europeanization on polarization in member states in general.

We find that the marginal effect of EU membership on polarization is negative for low values of Europeanization, suggesting that the constraints imposed by the EU on
member states in the early years of European integration lead to convergence of parties, however, this effect is not significant. As the Europeanization variable increases from 1.15 to 1.53 (EU integration deepens), the marginal effect of EU membership on polarization becomes both positive (0.51) and significant (with a p-value of 0.03). Moreover, this effect continues to increase with Europeanization, and becomes 1.78 when Europeanization increases to 2.3 (with a p-value of 0.02), 2.55 (with a p-value of 0.03) when Europeanization increases to 2.76, and 2.86 (with a p-value of 0.04) when Europeanization increases to 2.95, implying that as European integration deepens and constraints on member states intensify, polarization increases in member states’ party system. 25

When we analyze the marginal impact of Europeanization for members states in general, we find that a unit increase in Europeanization variable increases polarization by 1.76 (with a p-value of 0.01). This effect is significant at the 5% level. The results also show that the Effective Number of Parties has no statistically significant impact on polarization. Trade Openness, on the other hand, has a negative and highly significant impact on party system polarization. This finding is consistent with our expectation.

The second column reports the estimates from the model where Europeanization (Economic) is used as the main independent variable. The results are similar to the previous results, where EU membership has a negative effect on polarization for early years of European integration and positive and significant effect as European integration deepens.

To get a meaningful interpretation of these results, we plotted the marginal effect of EU membership on party system polarization for a wide range of values of Europeanization (Economic) using the regression on Column 2 in Table 2. In Figure 3, the solid line indicates how this marginal effect changes with the level of Europeanization (Economic). Accordingly, EU membership increases polarization when Europeanization (Economic) measure is higher than 1.85. In particular, the effect of EU membership on polarization is -0.24 when Europeanization measure is 1.4. This effect becomes 0.55 (with a p-value of 0.03) when Europeanization increases from 1.4 to 2, 1.43 (with a p-value of 0.01) when Europeanization increases to 2.67, 2.30 (with a p-value of 0.02) when Europeanization

25 In an alternative specification, where we use a non-linear functional form with the square of Europeanization variable, we find that the coefficient on the Europeanization is negative, and the coefficient on its square is positive. This implies that polarization is decreasing with Europeanization, but increasing after a certain point.

26
increases to 3.33, and 2.76 (with a p-value of 0.02) when Europeanization increases to 3.67.

When we analyze the marginal impact of Europeanization on polarization in member states in general, we find that an increase in Europeanization measure increases polarization by 1.38 point (with p-value of 0.03). This finding lends further support to our hypotheses that polarization increases with higher levels of integration. In column 3 and 4, we estimated the regression with a lagged dependent variable. To a large extent, the coefficients and their statistical significance levels remain unchanged, though when Europeanization (Economic) variable is used, the coefficient on EU Membership turns out to be significant. As a further robustness check, we estimated the model using an alternative measure of polarization, which is calculated as the weighted standard deviation of parties’ policy position on the left-right dimension. The results are very similar to those reported in this paper.\(^{26}\)

3.2.1 Estimating the Effect of Output Legitimacy

In order to analyze the effect of Europeanization on polarization for different values of EU popularity we collected data on EU popularity from the Eurobarometer surveys, using the responses given to the question of whether membership of the EU is a good or a bad thing starting from 1973. We estimate the following equation:

\[
Polarization_{i,e} = \beta_0 + \beta_1 Europeanization_{i,e} + \beta_2 EUPopularity_{i,e} \\
+ \beta_3 Europeanization \times EUPopularity_{i,e} + \lambda_i + \gamma_t + \varepsilon_{i,e}
\]

where \(i\) indicates countries and \(e\) indicates elections.

The results are presented in Table 3. We find that the effect of Europeanization on polarization is positive for low values of EU popularity and continues to be positive but at

\(^{26}\)The results can be requested from the authors.
a decreasing rate for higher values of EU popularity. The interaction effect is significant across all levels of popularity.

[Insert Table 3]

Figure 4 plots the marginal effect of Europeanization on polarization for different values of EU popularity. In particular, the effect of Europeanization on polarization is 3.58 (with a p-value of 0.02) when EU popularity is at its lowest value, 0.3. This effect becomes 2.72 (with a p-value of 0.00) when EU popularity increases to 0.5, and 1.44 (with a p-value of 0.09) when EU popularity increases to 0.8. This finding suggests that popularity exerts a centripetal force, corroborating our hypothesis.

[Insert Figure 4]

4 Concluding Remarks

The Eurozone debt crisis marked the end of political stability and policy consensus in some European parliamentary democracies. Increasing economic interdependence and political integration in Europe has led to the erosion of well-entrenched national democratic institutions. With polarization legislative politics become increasingly adversarial, parties more ideologically cohesive, coalition government majorities thinner, and cabinet tenures more short-lived. In a big part of Europe, this has marked the end of the post-war social democratic consensus on the welfare state. The gradual erosion of national sovereignty and the deep-rooted frustration with the state’s inability to single-handedly tackle a range of festering problems (unemployment, fiscal imbalances, immigration, etc.) has left many people feeling effectively disenfranchised from the democratic process, without any real choice. The trite assertion that there are no dead-ends in democracy has started to ring hollow for many Europeans suffocating under the ‘straitjacket’ of Eurozone membership. This is a clear manifestation of Rodrik’s (2012) so-called ‘globalization trilemma’: financially integrated and economically interdependent polities will either have to relinquish their national sovereignty or their democratic institutions; something has to give! The effects of economic interdependence have long started to spill over to the political sphere.

28
rendering political integration and national democracy effectively incompatible. Ironically enough, we are witnessing the reversal of the EU’s oft-touted democracy-enhancing role in its expanded periphery and its degeneration into a subversive force against the legitimacy of national democratic structures (Hix 2003).

In this paper, we argue that one of the principal manifestations of this degenerative process is party-system polarization across the ideological spectrum within both the electoral and legislative arenas. We first theorize the relationship between tightening policy constraints stemming from EU membership and partisan polarization within the framework of a three-party game of electoral competition. We argue that while binding policy constraints and rapid integration initially lead to increasing moderation and speed up ideological convergence across the spectrum, it can be the case that extreme constraints and policy targeting can have the opposite effect and might backfire in terms of extreme polarization. This non-linear relationship is borne out by our empirical analysis. During the early stages of the European project, when the integration process centered around areas of ‘low politics’ (energy, trade, single market, etc.), we find a negative and significant effect on aggregate levels of political polarization. However, there is strong evidence that this process of convergence is reversed during the later stages of European integration that were marked by the broadening and deepening of the EU’s scope of policy competences (Maastricht Treaty, EMU, Lisbon Treaty, European Stability Mechanism). Our findings have wide-ranging normative implications in terms of the political feasibility of the policy centralization process, the sequencing of economic and political integration, and the design of conditionality agreements with indebted Member States.

The rise of anti-European extremist parties in Europe (e.g., the Front National in France, the Party for Freedom in the Netherlands, or the UK Independence Party) and recent developments in the politics of the indebted ‘South’ are giving us a glimpse into the future. In fact, Greece provides one of the more characteristic examples of the degenerative effects of extreme conditionality and policy centralization on political stability and ideological moderation. The Greek political system has been strained beyond its breaking point by the exigencies of its international commitments. As a result of increasing fiscal and supply-side conditionality and complex policy interdependence, the dimensionality of electoral competition has dwindled. The scope for unilateral political action has shrunk
to such an extent that political rhetoric has been greatly polarized around stark political
dilemmas between pro- and anti- Europeanism, pro- and anti- austerity, populism and
pragmatism, democracy and technocracy (Vasilopoulou et al. 2014). The electoral shift
towards extremist and radical parties (e.g., the extreme right-wing Golden Dawn party
witnessed a meteoric rise in its vote share from less than 1% in 2009 to 7% in 2012) pro-
vides clear evidence of the radicalization of domestic politics in the context of deepening
European integration.
Appendix

In this section we provide the formal results of the general version of our model together with the proofs of those results. We first start with the case of no exogenously imposed constraints (no supranational integration). Recall that, without any loss in generality we have assumed the policy space to be the $[0, 1]$ interval and that parties are symmetrically positioned such that $(p_L, p_C, p_R) = (0, 0.5, 1)$. Further assume that voters are distributed in the $[0, 1]$ interval according to some unimodal, symmetric about one-half, absolutely continuous and twice continuously differentiable function $F(\cdot)$. Proposition 1 characterizes the unique symmetric NE of the game (for the case where $-c(|x_j - p_j|)$ is the Euclidean distance).

**Proposition 1 (No constraints)** Let $x_L < x_C < x_R$ and also $f(\frac{1}{2}) < 2$. Then the unique symmetric Nash equilibrium of the game is the triplet $(x_L^*, x_C^*, x_R^*) = (0, 0.5, 1)$ and extremism (measured as the weighted average of the distance of the two extremist from the median) is maximized.

**Proof.** In any symmetric equilibrium it must be that $x_L^* = 1 - x_R^*$. Then we first examine the behavior of party $C$. We will show that strategy $x_C^* = 0.5$ is a strictly dominant strategy for this party. That is, we need to show that $x_C^* \in \arg \max U_C(x_C, p_C) = -|x_C - 0.5| + v_C(x_C, x_R^*, x_L^*)$. Notice that, given that voters have single-peaked preferences, we must have that

$$v_C(x_C, x_R^*, x_L^*) = F\left(\frac{x_C + x_R^*}{2}\right) - F\left(\frac{x_C + x_L^*}{2}\right) = F\left(\frac{x_C + x_R^*}{2}\right) - F\left(\frac{x_C + 1 - x_L^*}{2}\right),$$

where the last equality follows due to symmetry. Then to show that $x_C^* = 0.5$ is a strictly dominant strategy for C, it suffices to show that $v_C$ is maximized when $x_C^* = 0.5$. That is, it must be that $\frac{\partial v_C}{\partial x_C} |_{x_C^*=0.5} = 0$. We then compute:

$$\frac{\partial v_C}{\partial x_C} = f\left(\frac{x_C + x_R^*}{2}\right) - f\left(\frac{x_C + 1 - x_L^*}{2}\right) = 0 \text{ which implies } f\left(\frac{x_C + x_R^*}{2}\right) = f\left(\frac{x_C + 1 - x_L^*}{2}\right).$$

Then, because $F(\cdot)$ is unimodal and symmetric about one-half, this implies that it is monotonically increasing for $x < 1/2$ and monotonically decreasing for $x > 1/2$. Given
that $x_L < x_C < x_R$, symmetry and monotonicity of $F$ imply that the condition above is satisfied if and only if the following holds:

$$\frac{x_C + x_R^*}{2} = 1 - \frac{x_C + 1 - x_R^*}{2} \implies x_C^* = \frac{1}{2}.$$ 

Thus we have shown that $x_C^* = \frac{1}{2}$ is a strictly dominant strategy for party $C$. This completes the first part of the argument. Now by symmetry, we need only focus on the behavior of one of the two extreme parties. Suppose, without any loss in generality, that we focus on party $L$. Party $L$ chooses $x_L^* \in \arg \max U_L(x_L, p_L) = -|x_L - 0| + F\left(\frac{x_C + x_L}{2}\right)$. It is then easily checked that $\frac{\partial U_L}{\partial x_L} = -1 + \frac{1}{2} f \left(\frac{0.5 + x_L}{2}\right) < 0$ for every $x_L < x_C^* = 0.5$. This, in turn, implies that the utility function is strictly decreasing in $x_L$ and, hence, we have that $x_L^* = 0$ and $x_R^* = 1 - x_L^* = 1$. Thus, the unique symmetric equilibrium of the game is $(x_L^*, x_C^*, x_R^*) = (0, 0.5, 1)$. This completes the proof. ■

It is easily checked that the above result holds in the case which the distribution function is a uniform in $[0, 1]$.\footnote{Technically speaking, a separate proof is required because the uniform distribution does not have a twice continuously differentiable density. Yet, if one follows the steps of Proposition 1, the algebra is straightforward.} Proposition 2 studies the special case in which the distribution $F$ is a uniform on $[0, 1]$ and $-c(|x_j - p_j|) = -(x_j - p_j)^2$. The purpose is to show that while the extreme divergence (and polarization)\footnote{It is easily verified that, in equilibrium, extremism (measured as the weighted average of the distance of the two extremists from the median party position) and polarization (measured by the Most Distant Platform index) are positively and monotonically related: as extremism increases so does polarization.} equilibrium might depend on the assumption about the density of the distribution $F$ the qualitative features of the equilibrium in Proposition 1 are robust to different specifications of the cost function and the distribution function $F$.

**Proposition 2** Let $x_L < x_C < x_R$ and assume $F$ is a uniform distribution function on $[0, 1]$. Then the unique symmetric Nash equilibrium of the game is the triplet $(x_L^*, x_C^*, x_R^*) = \left(\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right)$.

**Proof.** The arguments for the behavior of party $C$ are identical to the ones in Proposition 1. Simply notice that when $F$ is uniform, we have $v_C(x_C, x_R, x_L^*) = F\left(\frac{x_C + x_R^*}{2}\right) =$
That is, its vote share does not vary with \( x_C \). As a result, \( \frac{\partial U_C}{\partial x_C} |_{x_C^*} = 0 \) implies that \(-2(x_C^* - p_C) = 0 \implies x_C^* = p_C = 0.5\). Then, it is easily checked that \( \frac{\partial U_L}{\partial x_L} |_{x_L^*, x_C^*} = 0 \) implies that \(-2(x_L^* - p_L) + \frac{1}{2} = 0 \implies x_L^* = \frac{1}{4}\).

Then, by symmetry \( x_R^* = 1 - x_L^* = \frac{3}{4} \). This completes the proof.

Next we prove our main result for the case of constraints. Recall that the constrained policy space is a subset \([\frac{1}{2} - \alpha, \frac{1}{2} + \alpha]\) \( \in [0, 1] \) where \( \alpha \in [0, \frac{1}{2}) \) and the utility function of party \( j \in \{L, C, R\} \) is given by the following expression:

\[
U_j(x_j, p_j) = \begin{cases} 
-c(|x_j - p_j|) + v_j(x_j, x_{-j}) + b_2, & \text{if } x_j \in \left[\frac{1}{2} - \alpha, \frac{1}{2} + \alpha\right] \\
-c(|x_j - p_j|) + v_j(x_j, x_{-j}) + b_1, & \text{if } x_j \notin \left[\frac{1}{2} - \alpha, \frac{1}{2} + \alpha\right]
\end{cases}
\]

Further assume, as in Proposition 1, that \(-c(|x_j - p_j|) = -|x_j - p_j|\) and that \( F \) is a uniform on \([0, 1]\). We also assume that \( 0 < \beta < \frac{1}{4} \), where \( \beta \equiv b_2 - b_1 \) with \( b_2 > 0 > b_1 \). Then we can state the following result that fully characterizes the unique symmetric Nash equilibrium of the game.

**Proposition 3 (Constraints)** Let \( x_L < x_C < x_R, \alpha \in [0, 1) \) and \( \beta \in (0, \frac{1}{4}) \). Then, there exists \( \alpha^*(\beta) \) such that: a) for every \( \alpha \geq \alpha^*(\beta) \) the unique symmetric Nash equilibrium of the game is \((x_L^*, x_C^*, x_R^*) = (\frac{1}{2} - \alpha, \frac{1}{2}, \frac{1}{2} + \alpha)\); b) for every \( \alpha < \alpha^*(\beta) \) the unique symmetric Nash equilibrium of the game is \((x_L^*, x_C^*, x_R^*) = (0, 0.5, 1)\).

**Proof.** The arguments for the behavior of party \( C \) are identical with those employed in the proof of Propositions 1 and 2. The reason is that the constraint is never binding for the centrist party (it is always a strictly dominant strategy for party \( C \) to choose \( x_C^* = 0.5 \). Then, due to symmetry, we need only examine the behavior of party \( L \). Recall that we have \( x_L < x_C^* = 0.5 \) the domain of \( x_L \) is the interval \([0, \frac{1}{2})\) and observe that: a)\(^{29}\)

---

\(^{29}\)This assumption is without any loss in generality. If the popularity shock is allowed to be larger than \(1/4\) then we will always have convergence. On the other hand, if it is negative we will always have maximum divergence. Therefore, the study of such cases is trivial.
for all $x_L \in [0, \frac{1}{2} - \alpha)$ we have \( \frac{\partial U_L}{\partial x_L} = -\text{sgn}(x_L - p_L) + \frac{1}{2}f\left(\frac{x_L + x_L^*}{2}\right) = -1 + \frac{1}{2} < 0 \) and, b) for all $x_L \in \left[\frac{1}{2} - \alpha, \frac{1}{2}\right)$ again, by an analogous computation, we have \( \frac{\partial U_L}{\partial x_L} < 0 \). That is, its utility is strictly (and monotonically) decreasing in $\left[0, \frac{1}{2} - \alpha\right)$ and in $\left[\frac{1}{2} - \alpha, \frac{1}{2}\right)$. This implies that there are only two candidates for maximum: $x_L^* = 0$ and $x_L^{**} = \frac{1}{2} - \alpha$ (i.e. the two corners).

Then, the indifference condition is formally summarized below:

$$U_L(p_L = 0; x_L^{**} = \frac{1}{2} - \alpha) = U_L(p_L = 0; x_L^* = 0),$$

which, in turn, implies that

$$-|\frac{1}{2} - \alpha - p_L| + F\left(\frac{x_L^* + x_L^{**}}{2}\right) + b_2 = -|0 - p_L| + F\left(\frac{x_L^* + x_L^*}{2}\right) - b_1 \implies$$

$$-\frac{1}{2} + \alpha + \left(\frac{1 - \alpha}{2}\right) + b_2 = \frac{0 + \frac{1}{2}}{2} - b_1 \implies \frac{\alpha}{2} + (b_2 - b_1) = \frac{1}{4} \implies \alpha^* = \frac{1}{2} - 2\beta$$

Define $\alpha^*(\beta) \equiv \frac{1}{2} - 2\beta$. Then for every $\alpha \geq \alpha^*(\beta)$ we have $U_L(x_L^{**}) > U_L(x_L^*)$ which implies that party $L$ chooses $x_L^{**} = \frac{1}{2} - \alpha$, while for every $\alpha < \alpha^*(\beta)$ we have $U_L(x_L^{**}) < U_L(x_L^*)$ which implies that party $L$ chooses $x_L^* = 0$. By a completely symmetric argument, one can show the same for party $R$. Hence, this completes the proof. ■

The next corollary summarizes how equilibrium polarization (measured as the distance between the equilibrium platforms of the two extreme parties, also known as the MDP index) varies with the tightness of the constraint – recall that parameter $\alpha$ captures the slackness of the constraint, implying that as $\alpha$ decreases the constraint becomestighter.

**Corollary 4** Let assumptions of Proposition 3 hold. Then extremism (and polarization measured by the MDP index) is decreasing in the tightness of the constraint (i.e. it is increasing in $\alpha$) for every $\alpha \in \left[\alpha^*(\beta), \frac{1}{2}\right)$; extremism (polarization) is maximal when $\alpha \in (\alpha^*(\beta), 0]$.

**Proof.** When $\alpha \geq \alpha^*(\beta)$ define extremism as

$$\frac{\nu_L(|x_L^* - x_L^{**}|) + \nu_R(|x_R^* - x_L^{**}|)}{\nu_L + \nu_R} = \frac{\alpha(v_L + v_R)}{\nu_L + \nu_R} = \alpha$$

and polarization as the distance $x_R^{**} - x_L^{**} = 2\alpha$. It is then straightforward to see that extremism (and polarization) are strictly increasing in $\alpha$ and, hence, strictly decreasing in the tightness of the constraint. It is also straightforward to see that for $\alpha \in \left[\alpha^*(\beta), \frac{1}{2}\right)$
the maximal value that extremism can take is strictly less than $\frac{1}{2}$ while for the MDP the maximal value is strictly less than 1. When $\alpha \in [0, \alpha^*(\beta))$ we have that $x^*_R = 1$ and $x^*_L = 0$ and, hence, extremism is equal to $\frac{1}{2}$ while the MDP index is equal to 1. This completes the argument. ■
Table 1: The Effect of EU Membership on Extremism

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU Membership</td>
<td>-0.27*</td>
<td>-0.28</td>
<td>-0.39***</td>
<td>-0.38**</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.16)</td>
<td>(0.14)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Trend</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>EU Membership*Trend</td>
<td>0.01*</td>
<td>0.01*</td>
<td>0.01**</td>
<td>0.01**</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Effective Number of Parties</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>-0.01***</td>
<td>-0.01***</td>
<td>-0.00**</td>
<td>-0.00**</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Population</td>
<td>-0.00</td>
<td>-0.00</td>
<td>-0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>EMU</td>
<td>0.14</td>
<td></td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td></td>
<td>(0.12)</td>
<td></td>
</tr>
<tr>
<td>Single-Market</td>
<td>-0.12</td>
<td>-0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td></td>
<td>(0.09)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.62**</td>
<td>1.73**</td>
<td>1.85***</td>
<td>1.93***</td>
</tr>
<tr>
<td></td>
<td>(0.67)</td>
<td>(0.68)</td>
<td>(0.59)</td>
<td>(0.60)</td>
</tr>
<tr>
<td>N</td>
<td>1284</td>
<td>1284</td>
<td>1423</td>
<td>1423</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.15</td>
<td>0.15</td>
<td>0.14</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Dependent variable *Extremism* is measured in column (1) and (2) as the absolute distance between a party’s policy position and the median voter, and in column (3) and (4) as the absolute distance between a party’s policy position and the mean party policy. Numbers in parentheses are standard errors clustered by election. All models include country fixed effects and decade fixed effects, which are omitted here. * significant at 10% ** significant at 5% ***significant at 1%.
Table 2: The Effect of EU Membership on Party System Polarization

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europeanization</td>
<td>1.11</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.91)</td>
<td>(0.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europeanization (Economic)</td>
<td>0.26</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.72)</td>
<td>(0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EU Membership</td>
<td>-2.01</td>
<td>-2.09</td>
<td>-2.13</td>
<td>-2.28*</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.11)</td>
<td>(0.12)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Europeanization*EU Membership</td>
<td>1.65*</td>
<td>1.32**</td>
<td>1.65*</td>
<td>1.35**</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Effective Number of Parties</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>(0.87)</td>
<td>(0.87)</td>
<td>(0.78)</td>
<td>(0.78)</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>-0.18**</td>
<td>-0.17**</td>
<td>-0.18**</td>
<td>-0.17**</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Population</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(0.47)</td>
<td>(0.46)</td>
<td>(0.46)</td>
<td>(0.43)</td>
</tr>
<tr>
<td>Lagged Polarization</td>
<td>0.09</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>7.75*</td>
<td>7.9*</td>
<td>7.88**</td>
<td>8.16*</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.05)</td>
<td>(0.08)</td>
</tr>
</tbody>
</table>

Dependent variable is Polarization as measured as the distance between the two most extreme parties. Numbers in parentheses are standard errors clustered by election. All models include country fixed effects and decade fixed effects, which are omitted here. * significant at 10% ** significant at 5% ***significant at 1%.
Table 3: The Effect of Europeanization on Party System Polarization for Varying Levels of EU Popularity

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europeanization</td>
<td>4.86*</td>
<td>(2.24)</td>
</tr>
<tr>
<td>Europeanization (Economic)</td>
<td>5.58*</td>
<td>(2.57)</td>
</tr>
<tr>
<td>EU Popularity</td>
<td>8.24</td>
<td>(6.16)</td>
</tr>
<tr>
<td>Europeanization*EU Popularity</td>
<td>-4.26</td>
<td>(3.36)</td>
</tr>
<tr>
<td>Effective Number of Parties</td>
<td>0.01</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>-0.20</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Population</td>
<td>0.00</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.61</td>
<td>(8.96)</td>
</tr>
</tbody>
</table>

Dependent variable is Polarization as measured as the distance between the two most extreme parties. Numbers in parentheses are standard errors clustered by election. All models include country fixed effects and decade fixed effects, which are omitted here. * significant at 10% ** significant at 5% ***significant at 1%.
Figure 1: Marginal Effect of EU Membership on Extremism from the Median Voter
Figure 2: Marginal Effect of EU Membership on Extremism from the Mean Party Position
Figure 3: Marginal Effect of EU Membership on Polarization for Varying Levels of Europeanization
Figure 4: Marginal Effect of EU Membership on Polarization for Varying Levels of EU Popularity
References


[27] Hotelling, H. (1952)


