

“Take Back Control?” The Effects of Supranational
Integration on Political Extremism and Party-System
Polarization

Nikitas Konstantinidis*

Hande Mutlu-Eren[†]

IE University

New York University

Konstantinos Matakos[‡]

King’s College London

This draft: September 18, 2017

*Address: School of Politics and International Relations, IE University, C/ Pedro de Valdivia 21, Madrid 28006, Spain; e-mail: nikitas.konstantinidis@ie.edu

[†]Address: Wilf Family Department of Politics, New York University, 19 West 4th St. 2nd Floor New York, NY 10012 USA; e-mail: hande.mutlu@nyu.edu

[‡]Corresponding author’s address: Department of Political Economy, King’s College London, Strand S.2.43, London WC2R 2LS, United Kingdom; e-mail: konstantinos.matakos@kcl.ac.uk

Abstract

In this paper we study the relationship between economic and political integration and party system polarization, focusing on the key trade-off between the “output legitimacy” of economic efficiency and welfare-enhancing openness and the inevitable loss of democratic responsiveness and representation caused by externally imposed constraints on the domestic sources of “input legitimacy”. The formal version of our argument highlights the role of (ideological) “transportation costs” emanating from the discrepancy between core ideological positions and an externally circumscribed set of feasible policy options. We test our key prediction on the curvilinear effect of integration on political polarization in multi-party systems 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, the further tightening of policy constraints – coupled with heightened volatility in economic outcomes – may reverse the trend towards higher levels of party-system polarization.

Keywords: party-system polarization, extremism, output legitimacy, input legitimacy, European integration, globalization, policy constraints

Advanced liberal democracies have entered a period of political turbulence, where fundamental questions pertaining to globalization and supranational integration run to the core of domestic government and politics, creating in the process deep fissures in cabinets, parties, and society at large. The contemporaneous rise of *right*-wing nativism in the European “North” and *left*-wing populism in the European “South” (or even “Brexit” in the UK and Trump’s victory in the US) can be viewed through the same prism of a populist backlash against the “straitjacket” of economic and political globalization (Guiso et al. 2017; Rodrik 2017) and the inevitable conflict about political sovereignty between the rich “North” and the poor “South” (Tyson 2017). Effectively, the common thread that connects all these electoral results is that of increasing party-system polarization and political extremism in connection with increasingly constraining supranational integration and economic interdependence.

Political scientists have shifted their attention to the study of political polarization and extremism because, even if policy outcomes remain the same, higher polarization implies heightened political uncertainty for citizens (or political country risk for foreign investors) but also better defined political platform choice. With polarization, legislative politics become increasingly adversarial, (coalition) governments more unstable, and policy outputs more potentially volatile. Domestic political instability may have multiplicative effects in the context of an international union – such as the European Union (EU) – as increased polarization along the pro/anti- integration axis may generate negative spillovers for the entire supranational project: policy deadlock, the rise in the political implementation costs of policy reforms, the undermining of public support, even the political unraveling of the integration process itself (Hobolt and de Vries 2016).

Political polarization may manifest itself at various stages of the domestic political process: in elections (Cox 1990), the government formation process (Indridason 2011), legislative bargaining (McCarty et al. 2006), and political communication (Prior 2012).¹ Within the

¹Thus, the concept lends 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

burgeoning polarization literature in US politics (Layman et al. 2006), political polarization becomes tantamount with ideological differentiation between the two main congressional parties. In the context of parliamentary multi-party systems, however, the concept of polarization is not as straightforward in terms of measurement as it needs to account for both ideological differentiation and party fragmentation, whereby niche groups of extremist ideologues may splinter off from larger office-seeking parties (Ezrow 2010).² In both contexts, much of the existing literature has studied the effects of constitutional rules and institutions, such as the (dis)proportionality of the electoral system (Calvo and Hellwig 2011; Curini and Hino 2012; Dow 2011; Ezrow 2008; Iaryczower and Mattozzi 2013; Matakos et al. 2016), gerrymandering (McCarty 2009), and primary elections (Hirano et al. 2013), on party-system compactness, polarization, and extremism.

We focus instead on how *extraneous* policy rules and constraints – stemming from a country’s level of economic integration and international commitments – affect rhetorical party-level polarization and platform divergence within the context of electoral competition. Effectively, this paper derives the domestic political limits to globalization by highlighting the trade-off between democratic “inputs” and “outputs”. In other words, we show that there is point beyond which the “output legitimacy” of the globalization-bound pursuit of economic efficiency and welfare-enhancing openness will not be sufficient to counterbalance the inevitable loss of democratic responsiveness caused by externally imposed constraints on the domestic sources of “input legitimacy”. We formally derive the implicit trade-off between input and output legitimacy by relying on the mechanism of (ideological) “transportation costs” and also conduct empirical tests of the non-monotonic effect of integration on party-

 (using vote share data), or at the legislative representation level (using seat allocation data).

²The aggregate measurement of polarization in multi-party systems need also consider that convergence may take place among centrist parties and divergence between smaller extremist parties and centrist ones. This might explain the shortcomings of existing measures of polarization in multi-party systems (Dalton 2008; Esteban and Ray 1994; Stanig 2011).

system polarization. All in all, we focus on the behavior of parties and the characteristics of party systems as intervening variables that capture the underlying relationship between the input and output legitimacy of national democracies in a globalizing environment.

In doing so, we also seek to contribute to the literature on the convergence of national party systems and the transformation of electoral competition as a consequence of regional integration (e.g., Dorussen and Nanou 2006; Hix 2003; Mair 2000). Nanou and Dorussen (2013), for example, find that, because EU legislation limits the set of policies that parties can pursue once in government, the distance between parties' positions has tended to decrease in policy domains where EU involvement has increased. We actually expand on their results by taking a systemic view of party system polarization (and extremism) – thus obviating the need for an arbitrary distinction between mainstream and non-mainstream parties –, and by deriving the conditions for both conditional convergence and divergence of party platforms.

In what follows, we first theorize about the trade-off faced by advanced liberal democracies between input and output legitimacy in the context of the “political trilemma of globalization” (Rodrik 2011). We then provide a formal demonstration of the above argument by deriving the incentives of extreme 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 show that binding policy constraints will induce platform convergence up to the point where internal pressures for representation and input legitimization by party activists and the party base will cause divergence back towards the party's intrinsic ideology. Finally, we proceed to test this non-monotonic effect against a panel dataset of different aggregate measures of political polarization and extremism. We do find strong evidence confirming our hypothesized *non-monotonic* effect of integration on platform polarization (extremism).

Inputs, Outputs, and the Globalization Trilemma

In the current era, politics is no longer “business as usual” and that applies not just to the *supply* side of electoral competition and policy formation, but also the *demand* side of

partisanship and electoral behavior (Guiso et al. 2017). Recent electoral contests in Western Europe, Latin America, and the US showcase the political limits of globalization (Acemoglu and Yared 2010). Avowedly, the key factor that has introduced “noise” and unexplained volatility into the post-war paradigm of liberal democracy and output legitimacy has been the corrosive effect of globalization on national sovereignty and democratic politics.

Output legitimacy effectively pertains to the notion that a democratically elected government should “care for the common good” and can potentially come at the expense of 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 a tension between efficiency and democracy (Eichengreen and Leblang 2008).³ Policy responses to the recent financial crises highlighted the growing incongruence between the democratic functions of “representative” government (“*by* the people”) and those of “responsible” government (“*for* the people”). 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 has taken prevalence over *representative* (or *input-oriented*) government (Mair 2009), thus fueling the rise of populist parties that claim to still represent the “will of the people” but rarely deliver, as opposed to mainstream parties that take responsibility but no longer seem to voice their voters’ concerns (Caramani 2017).

So, in terms of framing our theoretical argument, we take Dani Rodrik’s “globalization trilemma” (2011) as our point of departure in order to demonstrate how the inescapable constraints of economic globalization and political integration shape the input-output trade-off

³This trade-off was particularly evident in the recent British referendum debate between the Brexiters’ insistence on democratic sovereignty outside of the EU and the Remainers’ emphasis on the sustained prosperity and economic stability afforded by EU membership.

and thereby affect the aggregate level of party-system polarization. Rodrik (2011) argues that there are inherent trade-offs in how we choose to organize our political decision-making structures in the face of economic globalization. The flattening logic of globalized competition for arbitrage opportunities, footloose capital, and minimal transaction costs renders economic integration, national sovereignty, and democracy simultaneously incompatible. Financially integrated and economically interdependent polities will either have to relinquish their national sovereignty or their democratic institutions; something has to give.

Figure 1 applies the incontrovertible logic of this trilemma to the domestic political arena by interpreting the nature of politics associated with the pursuit of Rodrik's three aforementioned goals of sociopolitical organization, namely, economic integration (i.e., *efficiency*), democratic politics (i.e., *equity*), and sovereign national statehood (i.e., *identity*). The horizontal axis of political contestation captures the traditional left-right spectrum of electoral competition where ideology, identity, and partisanship matter in the pursuit of lofty goals ranging from left-wing egalitarianism to right-wing nationalism. This dimension effectively constitutes the main source of input legitimacy of sovereign democratic systems. On one hand, equity politics emphasizes issues such as economic inequality, social justice, and democratic representation, while identity politics is mostly about the reification of the nation-state as the ultimate symbol of political and social self-identification and belonging. The clash between the various ideologies associated between these two types of politics left an indelible mark in the violent political history of the 20th century and gave rise to the traditional political cleavages still associated with established democratic party systems.

On the other hand, the vertical axis of efficiency politics captures the materialist orthogonal dimension of performance- or outcome- based electoral politics. Avowedly, the salience of this dimension tends to rise in direct proportion with economic globalization and the deepening of supranational coordination and integration. The output legitimacy that emanated from a long period of prosperity and growth allowed national politicians to steadily craft a well-entrenched consensus over the desirable level of openness to global markets and inte-

gration within supranational governance structures, while at the same time leaving enough room for traditional left-right political vacillation within moderate levels of polarization.

As a country moves further up the triangle presented in Figure 1, the menu of feasible ideological inputs of equity and identity gets increasingly constrained by the exigencies of exposure to global markets for goods, services, capital, and labor. In other words, the pursuit of efficiency by means of economic integration and rule harmonization trumps popular demands for democratic institutions of regulation, stabilization, and social protection (part of the traditional agenda of the socialist *left*) and undermines ethnic homogeneity and national self-determination (part of the traditional agenda of the nationalist *right*). In addition, countries that join international unions like the EU or the European Monetary Union (EMU) commit to formal constraints over their domestic policy space limiting what parties can promise in elections and deliver in government and, thus, reducing their “room to manoeuvre” (Hellwig 2014). For as long as this process remains embedded within a broader liberal democratic consensus, whereby the efficiency gains of openness are broadly shared among voters, then the enhanced output legitimacy of efficiency politics induces parties to advocate a feasible menu of policies within the constrained input space circumscribed by the given level of integration, albeit at the expense of unfettered democratic responsiveness.

As we proceed to demonstrate both theoretically (through the formal mechanism of transportation costs) and empirically, however, there is a threshold point beyond which the overall level of output legitimacy of any given degree of supranational integration is not sufficient to compensate for the loss in input legitimation and responsiveness, causing a democratic backlash in the form of increased platform polarization. The implicit trade-off between inputs and outputs gets reversed, giving rise to atavistic tendencies for *left*- and *right*-wing parties to revert to their ideological roots and core agendas of equity and identity respectively.⁴ This is due to the heightened pressure for popular representation by an increasingly

⁴This “truel” between equity, efficiency, and identity has materialized in different forms during the turbulent post-Crisis years of the 2010s, where left and right extremes have

disenfranchised part of their core electorate and the unwieldy nature of policy compromises necessary to maintain a country's liberal orientation.

This “breaking point” may be brought about by stochastic shocks to the aggregate welfare benefits of integration (e.g., in the context of capital account liberalization and financial crises), excessive constraints on the core democratic functions and policy choices of a country, or both. Countries further integrated within global markets (mostly due to the small relative size of their domestic markets) and systems of governance (often due to the political spillovers generated by the unfettered flows of capital, people, and ideas) will be subject to more narrow constraints on their domestic sources of input legitimacy. However, higher levels of integration also imply higher exposure to exogenous macroeconomic shocks and risks (Kim 2007), i.e., more noise in their political system's overall level of output legitimacy. Therefore, countries further up the triangle in Figure 1 are more likely to move beyond the aforementioned breaking point and experience heightened political volatility and democratic instability in the form of increased party-system polarization.⁵

Moreover, we surmise that the level of ideological polarization on the input dimension will affect the nature of electoral competition. At lower levels of integration (see Figure 1), where a national democratic polity is subject to much softer and wider supranational constraints, seemed to converge in terms of their anti-globalization rhetoric and join up against the moderate liberal center. Some glaring recent examples of this triadic nature of national democratic politics in a globalized environment include Sanders/Clinton/Trump in the US, Labor Brexiters/Remainers/UKIP and Conservative Brexiters in the UK, the Euroskeptic No camp/Yes camp/Nationalist No camp in the 2015 Greek referendum, and most recently Mélenchon/Macron/Le Pen in France.

⁵Note that in this paper we do not seek to explain why countries integrate further *per se*. We surmise that higher levels of economic and political integration come about as a result of pro-active policies of market liberalization, openness to trade, and the extraneous pressures of globalization (e.g., capital flows, immigration flows, environmental externalities, etc.).

we expect the traditional left-right dimension of electoral politics to be more salient in terms of explaining electoral outcomes. On the other hand, as a political system comes under more strain as a result of the stifling trade-offs of globalization, i.e., farther up the triangle in Figure 1, the ensuing backlash towards electoral platforms of *equity* and *identity* politics – now effectively incompatible with the country’s level of integration – will render the pro- anti-globalization dimension of electoral competition much more salient. Since the rise in platform polarization may run counter to the policies imposed by the “golden straitjacket” of economic and political globalization, then the necessity of those constraints and the desirability of a country’s overall liberal orientation will be put into question, thereby further politicizing the pro- anti- globalization axis (Kriesi 2016). In other words, the juxtaposition between the “vector” of *output*-oriented pro-globalization *efficiency* politics and the “resultant” of the “vectors” of *input*-oriented anti-globalization *equity* and *identity* politics will become a lot more pronounced in countries that are subject to stronger supranational constraints. And, even if policy outcomes and existing integration trends may not be severely affected by the rise in platform polarization, they become increasingly tenuous and fragile.

A Formal Model of Integration and Polarization

In this section we develop our formal argument regarding the relationship between the scope of supranational policy constraints imposed on a country and its party-system polarization and extremism. In order to fix ideas further and illustrate the intuition behind our empirical implications, we will introduce some formal notation and develop a stylized model. Its purpose is simply to isolate the main forces that mediate the effect of supranational integration on party-system polarization (and extremism) and uncover the trade-offs involved. Complete formal arguments and equilibrium characterization results are provided in the Appendix.

Political environment

Since our empirics draw primarily from European multi-party parliamentary democracies, we propose a model of electoral competition among three parties, the simplest form of a non-degenerate (i.e., more than two) multi-party system (although our key findings and results carry through to systems with more than three parties). Parties compete along two dimensions: the traditional left-right policy space and a second (binary) dimension that captures whether a party’s platform is congruent with the supranational institution’s policies.⁶ Voters have symmetric single-peaked preferences, with ideal points distributed over the unit interval $[0, 1]$ of the left-right dimension according to a *uniform* distribution (we relax this assumption in the Appendix), and vote sincerely for the party whose proposed platform is closest to their ideal point.

Parties occupy a core ideological position, denoted by p_j , in the left-right spectrum and propose a platform x_j with the primary goal of maximizing 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.⁷ Parties also seek to minimize the costs incurred by proposing a platform (x_j) away from their core policy position p_j . These are essentially communication or “transportation” costs that a party must pay in order to re-brand itself and sell its new policy program to its voters or, alternatively in order to appease its more militant ideological factions; they are, thus, strictly increasing in the distance between the proposed platform x_j and the party’s initial position p_j .⁸ Without

⁶This formulation is analytically very similar to standard spatial models of electoral competition with valence (e.g., Aragonès and Palfrey 2002) used in the formal literature on polarization and extremism (e.g., Aragonès and Xefteris 2017).

⁷For example, according to the proportionality norm (Gamson’s Law), cabinet portfolios are distributed among government parties in strict proportion to the number of seats that each party contributes to the government’s legislative majority.

⁸In general, there is a clear analogy between these communication costs and the trans-

any loss in generality, assume that initially parties are symmetrically positioned along the full range of the policy spectrum, i.e., $(p_L, p_C, p_R) = (0, 0.5, 1)$, where L, C, R denote the left-wing, centrist, and right-wing parties respectively.

A utility function for party $j \in \{L, C, R\}$ capturing this trade-off between *office-seeking* (vote share) and *policy-seeking* (transportation costs) incentives is depicted as follows:

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

where $c(\cdot)$ is a cost function that is monotonically increasing in the absolute distance $|x_j - p_j|$ and $v_j(x_j, x_{-j})$ denotes party j 's vote share as a function of all parties' proposed platforms.⁹ The first part of the utility function captures the transportation costs incurred when a party moves further away from its core ideology, while the second part reflects the office-seeking (vote share) benefits of convergence towards the median. Extreme parties face the following trade-off: moving towards the position of the centrist party (and also that of the median voter) increases not only their vote share and, hence, their expected office rents, but also their transportation (communication) costs. That is, in the absence of any other changes in the institutional set-up (e.g., changes in the electoral rule), the degree of policy convergence towards the center will depend on the relative size of the spoils of office (*centripetal* force) and the transportation costs (*centrifugal* force).

Thus, in the equilibrium without policy constraints (x_j^*), this trade-off will be reflected by some limited convergence by extreme parties towards the center of the policy space so long as transportation costs are not too high. In the Appendix we provide a complete characterization of the unique symmetric Nash equilibrium for different functional forms of $c(\cdot)$. We also show that for a Euclidean distance cost function $c(\cdot)$ the unique symmetric Nash equilibrium of the game without policy constraints is such that the two extreme parties L and R have *no incentive to converge* towards the center, and, hence, extremism – i.e., the distance of the two extreme platforms from the median position – is maximized (see Proposition 1).

portation costs that firms pay in the standard Hotelling model.

⁹For example, if $c(|x_j - p_j|) = (x_j - p_j)^2$, then we have a quadratic loss function.

Introducing the mechanism of supranational policy constraints

But what happens when additional policy constraints are imposed as a result of a country's participation in supranational integration schemes and overall level of economic globalization? Are the incentives faced by extreme parties to converge towards the center of the policy spectrum 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 domestic policies (e.g., harmonization of taxation and welfare spending) and, effectively, "tie the hands" of national governments. That is, the spectrum of admissible policies is truncated and the dimensionality of electoral competition increases to account for the compatibility of proposed platforms with a country's international commitments. For simplicity, we assume that the constrained set of admissible policies symmetrically excludes policies too much to the left or the right of the policy spectrum. How will then an extremist party behave if those restrictions start to bind and the policy platform that it would propose in an unconstrained equilibrium (x_j^*) falls outside the permissible policy space? Will it further converge towards the center in order to abide by the constraints or will it continue to advocate policies in the knowledge that they are not compatible the country's level of economic and political integration?

In order to examine this mechanism in more detail, we need to add some more structure to the model. 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 admissible space of domestic policies – if a country chooses to remain a member of the supranational institution – becomes $[\frac{1}{2} - \alpha, \frac{1}{2} + \alpha]$. That is, α measures the *slackness* of the constraint: the lower its value, the tighter the constraint, and *vice versa*.¹⁰

Then, the expanded utility function below summarizes the trade-off faced by parties:

¹⁰Note that $\alpha = \frac{1}{2}$ denotes the unconstrained case and $\alpha = 0$ denotes the limit case where parties will always pay a cost unless they propose the unique permissible platform of $\frac{1}{2}$.

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

where b_1 and 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 constraints and all other parameters are defined as above. If the output legitimacy associated with the supranational institution is high, we have that $b_1 < 0 < b_2$, while the converse is true if the output legitimacy of the supranational institution is low ($b_2 < 0 < b_1$). 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 (a party experiences a positive popularity shock if it advocates breaking ties with the supranational institution by proposing to implement a non-admissible policy platform). In other words, this formulation of preferences (which is a simplification of a standard valence model) implies that a party experiences a positive (negative) popularity shock if it proposes a platform that falls within (outside) the exogenously imposed policy constraint if and only if the output legitimacy associated with the supranational institution is sufficiently high (low). Since the case of $\beta < 0$ yields the trivial result of no convergence, henceforth, we assume that $\beta > 0$.

To preclude non-trivial results, we also proceed on the assumption that the policy constraint ($0 < \alpha < \frac{1}{2}$) introduced is *binding* in the sense that the unconstrained equilibrium policy platforms of the two extreme parties fall outside the constrained policy space, i.e., $x_j^* \notin [\frac{1}{2} - \alpha, \frac{1}{2} + \alpha]$.¹¹ This implies that they face the dilemma of whether to propose a policy platform on the *boundary* of the constrained policy space or stick with their unconstrained equilibrium platform proposal x_j^* .¹² Clearly, the extreme parties' (L and R) decision to con-

¹¹As we show in the Appendix, the unconstrained equilibrium platform proposal of the centrist party will always fall within the constrained policy space since $x_C^* = 0.5$.

¹²Note that all other platform choices are strictly dominated strategies for both extreme

verge or not will depend on (a) the tightness of such constraints on their domestic sources of input legitimacy and (b) the valence benefits of compliance, which are a function of the output legitimacy of supranational integration. On one hand, the tightness of the policy constraint creates *centrifugal* incentives by raising the transportation cost of convergence, while the positive popularity shock creates *centripetal* incentives by raising the valence benefits of advocating a moderate platform that is compatible with the country's international commitments. In other words, any shift away from the optimized unconstrained trade-off between office-seeking (vote shares) and policy-seeking objectives (communication costs) will only occur *if and only if* the orthogonal valence benefits of advocating integration-friendly policies (*gain* in *output* legitimacy) outweigh the additional (ideological) transportation costs imposed by *binding* supranational policy constraints (*loss* in *input* legitimacy). In what follows, we present the equilibrium that emerges as a result of this trade-off for the polar case of maximum unconstrained polarization.

Equilibrium analysis of the game with policy constraints: An example

To demonstrate this trade-off more clearly, we consider the case of a Euclidean distance cost function, i.e., $c(\cdot) = |x_j - p_j|$, which yields a unique symmetric Nash equilibrium of *maximum polarization* and *no convergence* in the game without policy constraints, i.e., $x_L^* = p_L = 0$ and $x_R^* = p_R = 1$. Due to the symmetry of the problem faced by the two extreme parties, we proceed to focus on the behavior of one of them, say the right-wing party R , and compare its utility from proposing a policy on the boundary of the constrained space (\hat{x}_R) with its utility of proposing the unconstrained equilibrium platform (x_R^*). Recall that voters are distributed

parties in expected utility terms. Take for example party R : if it proposes a platform x_R , such that $\frac{1}{2} + \alpha < x_R < x_R^*$, it will incur higher transportation costs without receiving the positive net popularity shock; if it proposes a platform strictly within the interior of the permissible policy space, i.e., $\frac{1}{2} < x_R < \frac{1}{2} + \alpha$, then it only incurs higher transportation costs without further increasing the positive popularity shock that it experiences.

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 as follows:

$$\begin{aligned} \underbrace{U_R(\hat{x}_R = 1/2 + \alpha)}_{\text{Convergence}} &= \underbrace{U_R(x_R^* = 1)}_{\text{No convergence}} \iff \\ -|\frac{1}{2} + \alpha - 1| + \left[1 - F\left(\frac{\hat{x}_R + x_C^*}{2}\right)\right] + b_2 &= -|1 - 1| + \left[1 - F\left(\frac{x_R^* + x_C^*}{2}\right)\right] + b_1 \iff \\ -\frac{1}{2} + \alpha + \left[1 - \frac{(\frac{1}{2} + \alpha) + \frac{1}{2}}{2}\right] + b_2 &= \left[1 - \frac{1 + \frac{1}{2}}{2}\right] + b_1 \iff \frac{\alpha}{2} + (b_2 - b_1) = \frac{1}{4} \end{aligned}$$

Then, solving for α yields the threshold value $\alpha^*(\beta) = \frac{1}{2} - 2\beta$. So, whenever the constraint is relatively loose, i.e., $\alpha > \alpha^*(\beta)$, party R will choose to converge to platform $\hat{x}_R = \frac{1}{2} + \alpha$; otherwise, whenever the constraint is too tight, i.e., $\alpha < \alpha^*(\beta)$, then party R will be better off not converging and sticking with its unconstrained platform $x_R^* = 1$. Therefore, political polarization (extremism) is minimized when $\alpha = \alpha^*(\beta)$. Clearly, the threshold value of $\alpha^*(\beta)$ is *strictly decreasing* in β : the larger the positive popularity shock associated with the supranational institution is high, the more slack the input-output trade-off allowing conditional platform convergence to be sustained even as integration deepens further and policy constraints become tighter.¹³ If, as argued earlier, we take β to be stochastic, then we should expect countries subject to narrower constraints of integration α (due to small size, explicit policy conditionality, or lack of political clout in the policy-formation process) and exposed to the higher external risk and volatility of macroeconomic shocks or business cycles to be more likely to go past the breaking point, i.e., $\alpha < \alpha^*(\beta)$.

Before returning to the interplay between input output legitimacy and discussing the policy implications of this relationship, we offer a brief comment regarding the (a)symmetry of the supranational policy constraints. Consider the case of supranational institution with a policy bias (e.g., the widely alleged “neoliberal” bias of the EU). That is, it might allow more slack on the right (left) of the policy spectrum such that the permissible policy space becomes the interval $[\frac{1}{2} - a_1, \frac{1}{2} + a_2]$ with $a_1 < (>)a_2$, $a_1 < \frac{1}{2}$, and $a_2 < \frac{1}{2}$. Then, it is straightforward

¹³In the limit, for a threshold value of α^* close to zero, conditional convergence can be sustained even as the domain of permissible policies becomes extremely narrow around $\frac{1}{2}$.

to argue that the main qualitative features of our results still hold. What will happen is that, in equilibrium, the two extreme parties will occupy asymmetric positions at a different speed of convergence: the party that faces the tighter constraint will converge (and diverge when the constraint becomes too tight) faster. Yet, the overall effect on political extremism (polarization) will remain unchanged; once the constraint becomes (asymmetrically) too tight, extremism will eventually increase.

Next, we provide a numerical example to illustrate the relationship between the depth of integration (captured by the tightness of supranational constraints on domestic policies) and political extremism (polarization) for *any* given value of β . For the purposes of this comparative statics exercise, we define extremism as the weighted average of the distance of the two extreme platforms from the median party position, i.e., $\frac{v_L \times (|x_L^{**} - x_C^*|) + v_R \times (|x_R^{**} - x_C^*|)}{v_L + v_R}$. Moreover, we use the following measures of polarization: (a) the distance between the two most extreme platforms (MDP) (Andrews and Money 2009; Matakos et al. 2016) and (b) the Dalton index (2008).¹⁴ Political polarization and extremism are, therefore, positively and monotonically related. For illustrative purposes, we fix the value of $\tilde{\beta} = \frac{1}{8}$ (moderate output legitimacy), so that $\alpha^*(\tilde{\beta}) = \frac{1}{4}$, and focus on the following three cases of low (L), intermediate (M), and high (H) levels of integration yielding low, medium, and high policy constraints $\alpha^L > \alpha^M > \alpha^H$ respectively.

Case 1 (Low level of integration $\alpha^L = 0.4$): When the constraint on domestic policies imposed by the supranational institution is very slack, we have $\alpha^L > \alpha^*(\tilde{\beta})$. This, in turn, implies that both extreme parties will have strong incentives to converge towards platforms that lies on the boundaries of the constrained policy space. That is, R will propose $\hat{x}_R = 1/2 + \alpha = 0.9$ and (by symmetry) L will propose $\hat{x}_L = 1/2 - \alpha = 0.1$. In this case, (i) political extremism shrinks from its maximal value of 0.5 in the unconstrained equilibrium

¹⁴The formula proposed by Dalton (2008) to operationalize party-system polarization is as follows: $P = \sqrt{\sum_{j=1}^N v_j (\frac{x_j - \bar{x}}{5})^2}$, where v is party j 's vote share, x is party j 's proposed platform, \bar{x} is the weighted mean of all parties' position, and N is the number of parties.

to 0.4 and (ii) polarization decreases (MDP decreases from its maximal value of 1 to 0.8, while the Dalton index goes from 0.71 down to 0.62). Hence, at low levels of integration both extremism and polarization *decrease* as a result of the exogenously imposed constraint.

Case 2 (Intermediate level of integration $\alpha^M = 0.3$): Even as integration deepens further and supranational constraints become tighter, we may still have $\alpha^M > \alpha^*(\tilde{\beta})$, which implies that the two extreme parties will converge further towards policy platforms $\hat{x}_R = 1/2 + \alpha = 0.8$ and $\hat{x}_L = 1/2 - \alpha = 0.2$ respectively. As a result, we get a further decrease in extremism (from 0.4 to 0.3) and polarization (MDP from 0.8 to 0.6 and Dalton from 0.62 to 0.5) at intermediate levels of supranational integration.

Case 3 (High level of integration $\alpha^H = 0.2$): However, if integration deepens even further past a certain breaking point, then policy constraints become too tight, i.e., $\alpha^H < \alpha^*(\tilde{\beta})$ and both extreme parties prefer to diverge back to their original unconstrained platforms of $x_R^* = 1$ and $x_L^* = 0$ respectively.¹⁵ Therefore, political extremism and party-system polarization (in terms of both the MDP and Dalton indices) will jump back up to their maximal value. In other words, there is a threshold point α^* beyond which closer supranational integration does not induce further policy convergence but instead triggers maximal levels of extremism (polarization). We now proceed to discuss the implications of this effectively *non-monotonic* relationship between supranational integration and political extremism (polarization).

Discussion of the theoretical results

In sum, our results shed light on how the process of supranational integration gives rise to an interesting trade-off between the *output* legitimacy of efficiency politics and responsible government and the *input* legitimacy of democratic responsiveness to the core left and right ideological agendas of equity and identity respectively. Low levels of integration with

¹⁵Notice that the two parties will not revert to the positions they had when the constraint was moderate ($\alpha = 0.3$) or low ($\alpha = 0.4$) as doing so would only increase their transportation costs without generating any positive popularity shock.

relatively slack policy constraints (i.e., α relatively large) intensify the process of policy convergence and reduce extremism (polarization) since the extreme parties' gain in output legitimacy (valence) is high enough to compensate for their loss in input legitimacy. Yet, as these constraints become tighter and α exceeds the threshold value $\alpha^*(\beta)$, the situation can backfire leading to extreme levels of platform divergence (and polarization). Simply put, the corresponding loss in input legitimacy associated with extreme levels of policy convergence is too high relative to the gain in output legitimacy so that extreme parties are better off proposing a policy platform that caters to their core constituency and reflects their ideological roots. This dynamic gives rise to the following *non-monotonic* relationship: while at first relatively permissive supranational constraints lead to a decrease in extremism (polarization), ever closer integration and extremely narrow constraints will all else equal result in higher levels of political extremism and party-system polarization. This is the key prediction of our analysis, which we set out to test in the following section.

Although so far we have been focusing on the convergence and divergence of party electoral platforms, we still expect government policy (ω) to remain at the pro-integration center of gravity of the left-right spectrum – whether that be the position of the median party, parliamentary mean, or the expected dominant party j selected to form a government (and implement its proposed platform x_j) with probability proportional to its vote share v_j (Merrill and Adams 2007). Therefore, all the equilibria we have presented above are identical with respect to the locus of expected policy outcomes ($\omega^* = \frac{1}{2}$) due to their symmetric structure. Nevertheless, from the point of view of moderate voters, market participants, and even supranational principals – and for any degree of risk aversion –, there is a clear welfare ranking among mean-preserving Nash equilibria with respect to other moments of the distribution of policy outcomes such as policy variance. For any generic specification of the government and policy formation process in parliamentary multi-party systems, platform convergence by extreme parties implies less policy variance (see Corollary 2 in the Appendix), and *vice versa*, in light of the symmetric nature of the equilibria and the simple specification of a uniform

voter distribution. Hence, platform convergence towards equilibria of low to medium levels of integration and decreasing party-system polarization implies decreasing policy variance, increasing government stability, but also less responsiveness to the policy preferences of extreme voters. On the other hand, for even deeper levels of integration and narrower policy constraints, the jump in equilibrium levels of political extremism and party-system polarization comes with higher policy variance, increasing government instability, but also higher policy congruence with extreme voter preferences.¹⁶

Integration and Polarization in European Party Systems

In this section, we seek to test our hypothesis on the non-monotonic relationship between supranational integration – insofar as it comes with extraneous constraints on the set of admissible domestic policies – and party-system polarization in the context of advanced liberal democracies. No other region has integrated politically further than Europe – while at the same time exhibiting substantial levels of political and economic diversity –, which renders it a most suitable and empirically rich setting for a cross-sectional and intertemporal study of the relationship between economic and political integration and political polarization and extremism in parliamentary multi-party systems.

The European integration process made leaps and bounds on the basis of so-called output legitimacy as a successful technocratic project generating growth and stability across Europe. Especially during the early years of the single market and the euro (late 1980s to early 2000s), 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.

¹⁶These implications have been nowhere more so evident than in Greece, a country that has changed seven governments and teetered on the brink of exit from the Eurozone and painful currency reform several times since the beginning of the crisis in 2010.

As of late, however, this technocratic process of “integration by stealth” (Majone 2005) has come under intense scrutiny as a result of stagnating labor productivity, sustained austerity, rising unemployment and debts levels, Brexit, and growing levels of third-country migration. These “multiple crises” have sapped the EU’s output legitimacy without having so far triggered any significant counterbalancing measures enhancing the Union’s so-called input legitimacy. Vivien Schmidt (2012) captures this nicely by pointing towards a precarious political imbalance: whereas output legitimacy (or “governing *for* the people”) has largely come to reside at the European level, which constitutes the main locus of “policy without politics”, input legitimization (or “governing *by* the people”) still takes place primarily at the domestic level, where national democracies engage in “politics without policy”.

The broadening and deepening of the integration process – through successive intergovernmental EU treaties – has reshaped the contours of domestic electoral competition by tying a number of policy areas together as part of the deeper commitments of EU membership, shrinking the degrees of freedom of electoral competition have shrunk, and constraining the range of admissible and internally consistent policy platforms.¹⁷ In what follows, we proxy for these domestic policy constraints through the overall depth of integration at the European level and a country’s level of openness and membership of E(M)U. We also expect the tightness of these constraints (or else the degree of sovereign policy discretion) to vary with respect to country size, (formal or informal) influence in the EU policy-formation process, participation in the European core policy areas (namely EMU), and overall macroeconomic conditions (e.g., balance of payments and debt burden). If our theory holds water, then

¹⁷Such constraints on the traditional left-right dimension may stem from (i) formal or informal intergovernmental commitments of EU or EMU membership (e.g., the fiscal coordination mechanisms of the European Semester and the Fiscal Compact), (ii) the hard process of compliance, transposition, and implementation of EU rules and directives, (iii) (as of late) bailout conditionality agreements signed with the EU or the IMF, and (iv) the softer imperatives of EU association, financial liberalization, and globalization more generally.

European political parties should converge towards the center (decreasing party-system polarization) during the early to middle stages of European integration (e.g., Treaty of Rome, Single European Act) and they should diverge towards the extremes (increasing party system polarization) during its later stages (e.g., Maastricht Treaty, Lisbon Treaty, Fiscal Compact).

Data and empirical model

In order to estimate the impact of EU membership on polarization, we collected data for 15 EU countries (Austria, Belgium, Denmark, Finland, France, Germany, Great Britain, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain and Sweden) for the years between 1958 and 2013. We created a party-level dataset using the Manifesto Project Dataset (Volkens et al. 2013).¹⁸ 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 analysis to follow, we rescale parties' position to a range between 1 and 10.¹⁹ Each observation in our dataset is a country-election-party. We estimate the following equation:

$$\begin{aligned} Extremism(Mean)_{i,e,j} = & \beta_0 + \beta_1 Trend_{i,e,j} + \beta_2 EUMembership_{i,e,j} + \beta_3 Trend * EUMembership_{i,e,j} + \\ & \beta_4 Controls + \lambda_i + \gamma_t + \varepsilon_{i,e,j}, \end{aligned}$$

where i indicates countries, e indicates elections, and j denotes parties.

We conceptualize polarization as extremism of political parties. The more extreme parties are, the more polarized the national party system becomes. Our dependent variable *Extremism (Mean)* is operationalized as the absolute distance between a party's policy position and the mean party position weighted by vote share. Our dataset includes 1420 cases.

EU Membership is a dummy variable that is coded 1 when a country becomes a member

¹⁸Data for Greece, Portugal and Spain start in 1974, 1975, and 1977 respectively.

¹⁹To rescale the index, we added 100 to each party's position and divided the sum by 20.

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 one 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 during the early years of the EU and increase it during the later years, when the policy constraints imposed by the EU became tighter.

We use 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 calculated based on Laakso and Taagepera (1979) formula.²⁰ If a country has a high number of effective parties, party policy positions are expected to be more dispersed, thereby 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 is, therefore, expected to decrease extremism independent of EU membership.²¹ This effect is captured by the *Trade Openness* variable, which is measured as the size of a country's exports and imports relative to its gross domestic product. This variable was coded 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 lower voting weight in the European Council, we regress on the square root of a country's population (*Population* variable) as a proxy for a country's political clout. We also include control variables for the position of the *Median Voter* – coded from Kim and Fording (1998) –, the *Disproportionality* of the electoral system – coded using the Gallagher index –, and

²⁰It is as follows: $1/\sum v_i^2$, where v_i is the percentage of the vote received by party i .

²¹Arguably, globalization increasingly exerts pressure on governments to implement monetary discipline, deregulation, privatization, and labor market flexibility in order to enhance competitiveness and productivity (Mosley 2005).

Party Vote Share – coded from the Manifesto Project Database. Finally, we include country fixed effects (λ_i) to account for country-specific effects as well as decade fixed effects (γ_t) to capture the decade-specific effects. Table 1 in the Appendix includes summary statistics of the data used in the subsequent empirical analyses.

We estimate an ordinary least squares (OLS) model with robust standard errors clustered by countries. This option assumes that observations are not independent within cluster although they may be independent between clusters. Table 1 presents the estimates from the regressions testing the impact of EU membership on political extremism. The first column shows the estimates of the model where the main dependent variable is extremism measured as the distance from the mean party position and standard errors are clustered by country. The coefficient on *EU Membership* is negative and significant at the 5% level, while the coefficient on *Trend* is negative but not significant. The interaction variable is both positive and significant at the 1% level. The second column shows the results of the estimation after controlling for the *GDP Growth*. Column (3) includes additional controls, namely *EMU* and *Single Market*, which are coded as dummy variables taking the value of 1 only in years when EMU and the Single Market respectively were in effect. Of these two variables, only *Single Market* is found to exert a negative and significant effect on party extremism. Columns (4)-(6) replicate the earlier regressions, but cluster standard errors by election. Estimated coefficients as well as their significant levels are similar to the previous models. Column (7) shows the estimates of the model presented in column (2) using bootstrapped standard errors. The results again remain the same.

We are particularly interested in assessing the impact of EU membership on political extremism. To do this, we calculate the marginal effect of EU membership on extremism over time, starting from 1957. Figure 2a shows the marginal effect of EU membership on extremism measured as the party distance from the mean party position (regression on Column 1 in Table 1). The two-tailed 95% confidence intervals around the line indicate the areas under which this effect is statistically significant. Accordingly, as expected, the

EU membership exerts a negative effect on extremism during the early years of the EU. In particular, the effect of the EU membership on extremism is negative and statistically significant for the first 18 years of the EU. As time passes, the effect becomes positive. It becomes statistically significant after the first 43 years of the EU and continues to increase with time.

Put differently, the EU membership helped political parties move towards the center until 1975 and started pushing them to the extremes after 2000. This finding lends support to our hypothesis that during the early years of the EU political parties moved towards the mean political party position, but then started to move towards the extremes as European integration deepened further.²²

So far, we have tested the main prediction of our theoretical model, that is, for low levels of European integration EU membership decreases party extremism and for high levels of European integration EU membership increases party extremism. Polarization, however, can also be conceptualized as a party system phenomenon. To this end, we put together a new dataset as a robustness test against our previous analysis, where our main dependent variable is platform polarization measured at the level of the party system. We also include in our analysis a new measure of Europeanization.

An alternative test

We collected data for the same 15 EU countries (Austria, Belgium, Denmark, Finland, France, Germany, Great Britain, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain and Sweden) for the years between 1957 and 1992. Each observation is now a country-election. The analysis covers 126 cases in total. The empirical model that examines the impact of EU membership on party-system polarization has the following functional form:

²²As a robustness check, we estimate the model with the dependent variable *Extremism* (*Median*) operationalized as the absolute distance between a party's position and the median party position. The results (Table 2 of the Appendix) are similar to those reported here.

$$\begin{aligned}
Polarization_{i,e} = & \beta_0 + \beta_1 Europeanization_{i,e} + \beta_2 EUMembership_{i,e} \\
& + \beta_3 Europeanization * EUMembership_{i,e} + \beta_4 Controls + \lambda_i + \gamma_t + \varepsilon_{i,e},
\end{aligned}$$

where i indicates countries and e denotes elections.

We employ two measures for our dependent variable *Polarization*. Our first measure amounts the distance between the most extreme parties on the left-right dimension for every election year in each country.²³ For our second measure, we use the widely accepted Dalton index of polarization (2008) that conceptualizes polarization as the degree of ideological differentiation among political parties. Since the Dalton index measures deviations from the party-system mean and weights them by party vote share, it is clearly sensitive to the distribution of voters: if that is very dense around the mean, a party's movement towards the extremes might have a minimal contribution to overall polarization as it will lose many votes to centrist parties whose small deviations will now be overweighted. To account for this, those models using the Dalton index also include an additional control variable (*Median Voter*) that measures the position of the median voter. In terms of measuring parties' policy positions we used the most recent Manifesto Project Dataset (Volkens et al. 2013). Higher values of our dependent variable denote higher levels of party-system polarization on the left-right dimension.

Our main explanatory variable is *Europeanization*, which is derived from the Europeanization index computed by Schmitter (1996). 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 scope of treaty obligations undertaken up that point. 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 those policy areas for each milestone period of the EU: the initial phase

²³Other studies that use the distance between the two most extreme parties to measure polarization include Budge and McDonald (2006) and Andrews and Money (2009).

(1950-56), the European Economic Community (1957-67), and the European Community (1968-1991).

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

We estimate 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. The first column shows the estimates of the model with control variables as specified above. *Europeanization* is positive but not significant. *EU Membership* exerts a negative and significant effect (at the 10% level) on party-system polarization. *Europeanization*EU Membership* is both positive and significant at the 5% 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 general.

We find that the marginal effect of EU membership on polarization is negative for low values of Europeanization, suggesting that the constraints imposed on member states during 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 (0.58) and significant (with a p -value of 0.02). Moreover, this effect continues to increase with Europeanization, implying that as European integration deepens and constraints on member states intensify, polarization increases in member states' party systems.²⁴

²⁴In an alternative specification, where we use a non-linear functional form including the square of the *Europeanization* variable, we find that the coefficient on the *Europeanization*

When we analyze the marginal impact of Europeanization for members states in general, we find that a unit increase in the *Europeanization* variable increases polarization by 2.11 (with a p -value of 0.002). 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 significant impact on party-system polarization. This finding is consistent with our theoretical expectations.

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: EU membership has a negative effect on polarization during the early years of European integration and a positive and significant effect as European integration deepens.

To get a meaningful interpretation of these results, we plot the marginal effect of EU membership on party-system polarization for a wide range of *Europeanization (Economic)* values using the regression results reported in Column 2 of Table 2. In Figure 2b, the solid line indicates how this marginal effect changes with the level of economic Europeanization. Accordingly, EU membership increases polarization when the *Europeanization (Economic)* measure is higher than 1.88. In particular, the effect of EU membership on polarization is -0.29 when the Europeanization measure is 1.4. This effect becomes 0.6 (with a p -value of 0.02) when Europeanization increases from 1.4 to 2, 1.59 (with a p -value of 0.005) when Europeanization increases to 2.67, 2.57 (with a p -value of 0.01) when Europeanization increases to 3.33, and 3.08 (with a p -value of 0.01) 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 the *Europeanization* measure increases polarization by 1.74 point (with a p -value of 0.01). This finding lends further support to our hypothesis that polarization increases at higher levels of integration. In columns 3 and 4, we estimate the regression with a lagged dependent variable. To a large extent, the coefficients is negative and the coefficient on its square is positive. This finding also implies that polarization is decreasing with Europeanization but increasing after a certain point.

and their statistical significance levels remain unchanged. Columns 5 and 6 show the estimates of the models using the Dalton index of polarization as the dependent variable. The results are consistent with the models using the alternative measure of polarization based on the distance between the most extreme parties.²⁵ As a further robustness check, we estimate the model using an alternative measure of polarization, that is, the weighted standard deviation of parties' position on the left-right dimension. The results (available upon request) are very similar to those reported in this paper.

Concluding Remarks

Throughout the post-WWII era all the way into the early 2000s the inherent tensions of modern democratic politics aptly captured by Rodrik's "globalization trilemma" (2011) – namely, the balance between efficiency, equity, and identity – appeared to be bottled up within a broader consensus on the merits of economic globalization, political integration, and technological advancement. In recent times, however, many advanced liberal democratic systems seem to have experienced a breaking point beyond which many of these trade-offs become politicized, juxtaposing the previously unchecked forces of globalization, liberalization, and integration against the counterforces of nativism, populism, and illiberalism and polarizing public discourse around stark political dilemmas between populism and pragmatism, democracy and technocracy (Caramani 2017). Within the current fluid environment of political turbulence and disequilibrium, such tensions have translated into a widening spatio-temporal gap between capital markets and democratic politics and an increasing incompatibility between the functions of *responsible* and *representative* government.

This "paradigm shift" is underway for several reasons. First and foremost, the recessionary effects of the Global Financial and Eurozone Debt Crises on the real economy have called

²⁵When we plot the marginal effect of EU membership on party-system polarization for a wide range of values of *Europeanization (Economic)* using the regression results reported in Column 6 of Table 2, we get a replication of Figure 2b.

into question the design of economic and financial globalization on *efficiency* grounds. Second, the persistence of austerity policies, rising debt burdens, stagnating labor incomes, bank bailouts, and the augmenting incidence of taxation on unskilled labor have given prominence to issues of income inequality, fairness, and social justice that form part and parcel of democratic *equity* politics. Finally, the surge in migration flows – coupled with the escalation of religiously-motivated terrorist incidents – have rekindled the rise of *identity* politics against the multiculturalism of most advanced liberal democracies (Rodrik 2017).

In this paper we postulate that one of the principal manifestations of this degenerative process is party-system polarization in the electoral arena. We model the relationship between integration and polarization by deriving the Nash equilibria of a three-party game of electoral competition with (ideological) transportation costs as a function of the slackness of supranational policy constraints. We show that, while rapid integration in the form of tightening policy constraints may initially speed up ideological convergence across the spectrum, there is a threshold level of policy constraints beyond which they can have the opposite effect and backfire in terms of extreme platform polarization. This non-monotonic relationship is borne out by our empirical analysis of party extremism and party-system polarization in European multi-party systems. During the early phases of the European project, when the integration process centered around areas of “low politics” (e.g., energy, trade, common market regulation), we find a negative and statistically significant effect of integration on aggregate levels of extremism and polarization. However, we also find strong evidence of a reversal of the process of ideological convergence during the later stages of European integration marked by the broadening and deepening of the EU’s scope of policy competences.

Our work helps better understand the potential repercussions of unfettered crisis-ridden globalization coupled with political volatility and polarization. As long as the locus of democratic inputs remains at the national level, then basic questions of job (in)security, social protection, income inequality, and national identity will keep roaring back to the forefront

of politics in advanced liberal democracies. We should, therefore, no longer be surprised by the political success of populist candidates advocating trade protectionism, immigration restrictions, and retrenchment from global policy coordination agreements on issues such as financial regulation, tax policy, defense, and climate change. Moreover, we should expect national elections to function less effectively as a way of translating citizens' preferences into government policy as a result of the growing incongruence between input-oriented democratic representation and efficiency-oriented responsible government.

Our findings also have wide-ranging normative implications in terms of the political feasibility of the policy centralization process and the sequencing of economic and political integration. The EU's "multiple crises" (namely, the euro crisis, the migration crisis, Brexit, terrorism, and the rise of populism) have highlighted the complex nature of economic interdependence and cross-border policy spillovers. On one hand, the "emergency politics" of crisis management, predominantly driven by considerations of output legitimacy and policy coordination, have led to a power shift towards unelected and unaccountable executive bodies at the European level. On the other hand, there is a backlash of growing popular demand for input legitimization in the form of national "beggar-thy-neighbor" solutions, often voiced (and in some cases put into effect) by populist and nationalist politicians. In fact, the recent trend towards the use of EU referendums may be viewed as an effort to bolster the input legitimacy of representative government by means of mechanisms of direct democracy.

Europe, therefore, finds itself at a critical juncture whereby any further pursuit of efficiency-oriented policy centralization will have to be cushioned by institutional structures of democratic responsiveness and representation. However, this conundrum cannot be overcome as long as these mechanisms of input legitimization reside at the level of the nation-state. Therefore, the only way of pushing beyond the boundaries of the current input-output trade-off seems to be the reinstating of democracy at the European level and the infusion of supranational institutions with direct mechanisms of democratic legitimization.

References

- Acemoglu, Daron, and Pierre Yared. 2010. "Political Limits to Globalization." National Bureau of Economic Research Working Paper No. 15694.
- Andrews, Josephine T., and Jeannette Money. 2009. "The Spatial Structure of Party Competition: Party Dispersion within a Finite Policy Space." *British Journal of Political Science* 39 (4): 805-824.
- Aragonès, Enriqueta, and Thomas R. Palfrey. 2002. "Mixed Strategy Equilibrium in a Downsian Model with a Favored Candidate." *Journal of Economic Theory* 103 (1): 131-161.
- Aragonès, Enriqueta, and Dimitrios Xefteris. 2017. "Imperfectly Informed Voters and Strategic Extremism." *International Economic Review* 58 (2): 439-471.
- Budge, Ian, and Michael D. McDonald. 2006. "Choices Parties Define: Policy Alternatives in Representative Elections, 17 Countries 1945-1998." *Party Politics* 12 (4): 451-466.
- Calvo, Ernesto, and Timothy Hellwig. 2011. "Centripetal and Centrifugal Incentives under Different Electoral Systems." *American Journal of Political Science* 55 (1): 27-41.
- Caramani, Daniele. 2017. "Will vs. Reason: The Populist and Technocratic Forms of Political Representation and Their Critique to Party Government." *American Political Science Review* 111 (1): 54-67.
- Cox, Gary W. 1990. "Centripetal and Centrifugal Incentives in Electoral Systems." *American Journal of Political Science* 34 (4): 903-35.
- Curini, Luigi, and Airo Hino. 2012. "Missing Links in Party-system Polarization: How Institutions and Voters Matter." *The Journal of Politics* 74 (2): 460-473.
- 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.
- Dorussen, Han, and Kyriaki Nanou. 2006. "European Integration, Intergovernmental Bargaining, and Convergence of Party Programmes." *European Union Politics* 7 (2): 235-256.
- Dow, Jay K. 2011. "Party System Extremism in Majoritarian and Proportional Electoral Systems." *British Journal of Political Science* 41 (2): 341-361.
- Eichengreen, Barry, and David Leblang. 2008. "Democracy and Globalization." *Economics & Politics* 20 (3): 289-334.
- Esteban, Joan, and Debraj Ray. 1994. "On the Measurement of Polarization." *Econometrica* 62 (4): 819-851.
- 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.

- Ezrow, Lawrence. 2010. *Linking Citizens and Parties: How Electoral Systems Matter for Political Representation*. Oxford, UK: Oxford University Press.
- Guiso, Luigi, Helios Herrera, and Massimo Morelli. 2017. "Demand and Supply of Populism." Centre for Economic Policy Research Discussion Paper 11871.
- Hellwig, Timothy. 2014. *Globalization and Mass Politics: Retaining the Room to Maneuver*. Cambridge, UK: Cambridge University Press.
- Hirano, Shigeo, James M. Snyder, Jr., Stephen Ansolabehere, and John Mark Hansen. 2013. "Primary Elections and Partisan Polarization in the U.S. Congress." *Quarterly Journal of Political Science* 5 (2): 169-191.
- Hix, Simon. 2003. "The End of Democracy in Europe? How the European Union (As Currently Designed) Restricts Political Competition." Unpublished manuscript.
- Hobolt, Sara B., and Catherine E. de Vries. 2016. "Public Support for European Integration." *Annual Review of Political Science* 19: 413-432.
- Hobolt, Sara B., and James Tilley. 2016. "Fleeing the Centre: The Rise of Challenger Parties in the Aftermath of the Euro Crisis." *West European Politics* 39 (5): 971-991.
- Iaryczower, Matias, and Andrea Mattozzi. 2013. "On the Nature of Competition in Alternative Electoral Systems." *The Journal of Politics* 75 (3): 743-756.
- Indridason, Indridi H. 2011. "Coalition Formation and Polarisation." *European Journal of Political Research* 50 (5): 689-718.
- Kim, Heemin, and Richard C. Fording. 1998. "Voter Ideology in Western Democracies, 1946-1989." *European Journal of Political Research* 33 (1): 73-97.
- Kim, So Young. 2007. "Openness, External Risk, and Volatility: Implications for the Compensation Hypothesis." *International Organization* 61 (1): 181-216.
- Kriesi, Hanspeter. 2016. "The Politicization of European Integration" *Journal of Common Market Studies* 54 (S1): 32-47.
- Laakso, Markku, and Rein Taagepera. 1979. "'Effective' Number of Parties: A Measure with Application to West Europe." *Comparative Political Studies* 12 (1): 3-27.
- Layman, Geoffrey C., Thomas M. Carsey, and Juliana Menasce Horowitz. 2006. "Party Polarization in American Politics: Characteristics, Causes, and Consequences." *Annual Review of Political Science* 9: 83-110.
- Mair, Peter. 2000. "The Limited Impact of Europe on National Party Systems." *West European Politics* 23 (4): 27-51.
- Mair, Peter. 2009. "Representative versus Responsible Government." Max Planck Institute for the Study of Societies Working Paper 8.

- Majone, Giandomenico. 2005. *Dilemmas of European Integration: The Ambiguities and Pitfalls of Integration by Stealth*. Oxford, UK: Oxford University Press.
- Matakos, Konstantinos, Orestis Troumpounis, and Dimitrios Xefteris. 2016. "Electoral Rule Disproportionality and Platform Polarization." *American Journal of Political Science* 60 (4): 1026–1043.
- McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: MIT Press.
- McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2009. "Does Gerrymandering Cause Polarization?" *American Journal of Political Science* 53 (3): 666-680.
- Merrill III, Samuel, and James Adams. 2007. "The Effects of Alternative Power-sharing Arrangements: Do Moderating Institutions Moderate Party Strategies and Government Policy Outputs?" *Public Choice* 131 (3-4): 413-434.
- Mosley, Layna. 2005. "Globalisation and the State: Still Room to Move?" *New Political Economy* 10 (3): 355–362.
- Nanou, Kyriaki, and Han Dorussen. 2013. "European Integration and Electoral Democracy: How the European Union Constrains Party Competition in the Member States." *European Journal of Political Research* 52 (1): 71-93.
- Prior, Markus. 2013. "Media and Political Polarization." *Annual Review of Political Science* 16: 101-127.
- Rodrik, Dani. 2011. *The Globalization Paradox: Why Global Markets, States, and Democracy Can't Coexist*. Oxford, UK: Oxford University Press.
- Rodrik, Dani. 2017. "Populism and the Economics of Globalization." Centre for Economic Policy Research Discussion Paper 12119.
- Scharpf, Fritz W. 1999. *Governing in Europe: Effective and Democratic?* Oxford, UK: Oxford University Press.
- Schmitter, Philippe C. 1996. "Imagining the Future of the Euro-polity with the Help of New Concepts." In Gary Marks, Fritz W. Scharpf, Philippe C. Schmitter, and Wolfgang Streeck, eds. *Governance in the European Union*. London, UK: Sage, 121-150.
- Stanig, Piero. 2011. "Measuring Political Polarization in Comparative Perspective." Presented at the Annual Meeting of the American Political Science Association, Seattle.
- Schmidt, Vivien A. 2013. "Democracy and Legitimacy in the European Union Revisited: Input, Output, and 'Throughput'." *Political Studies* 61 (1): 2-22.
- Tyson, Scott A. 2017. "The Strategic Foundations of Political Sovereignty." Mimeo.
- Volken, Andrea, Pola Lehmann, Nicolas Merz, Sven Regel, and Annika Werner with Onawa P. Lacewell, and Henrike Schultze. 2013. *The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR)*. Version 2013b. Berlin: WZB Berlin Social Science Center.

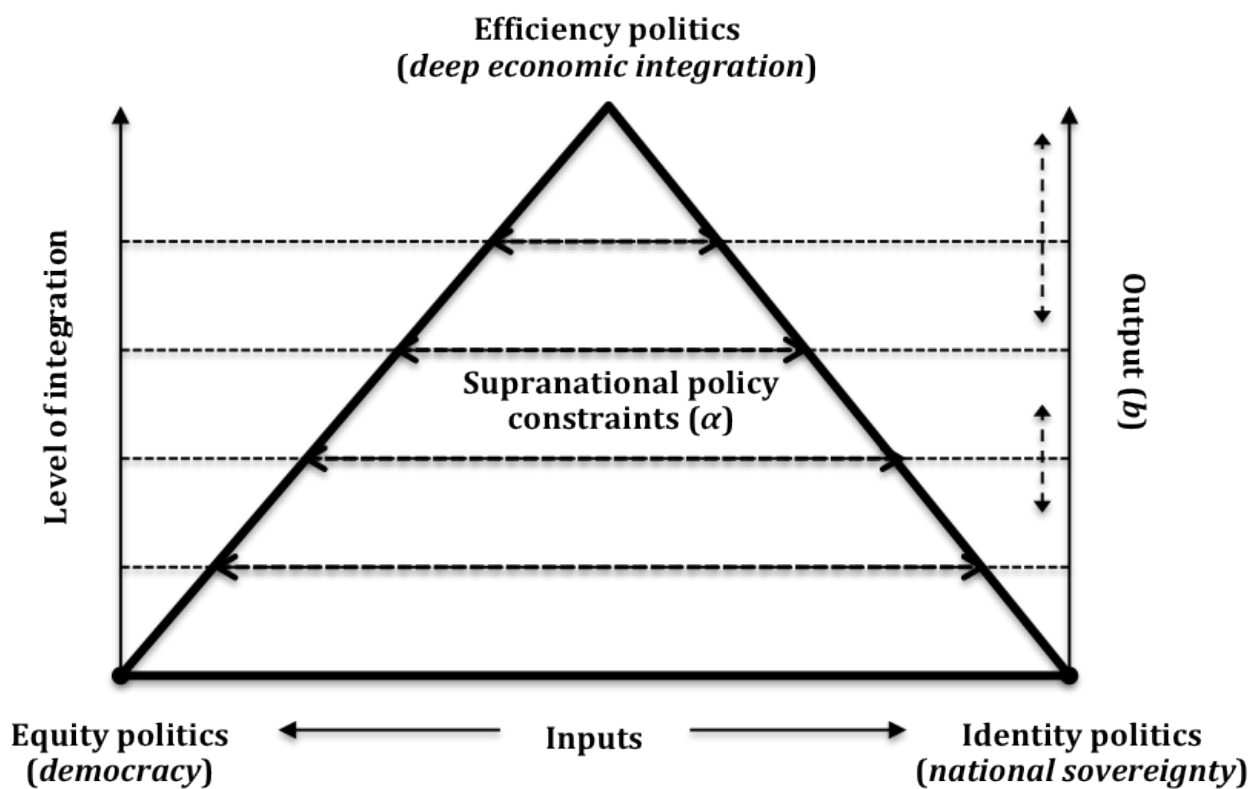
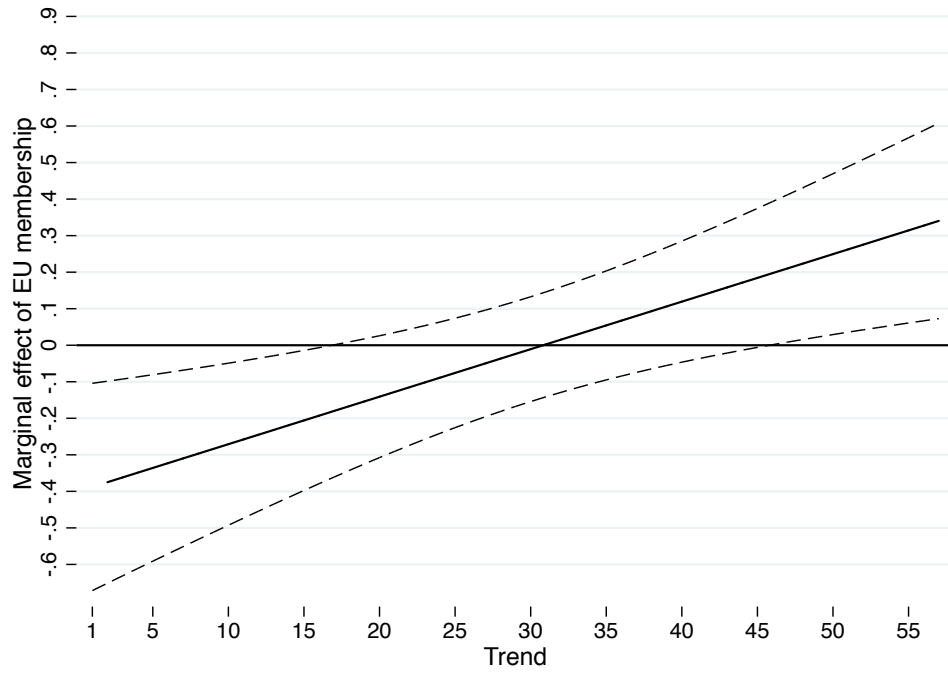
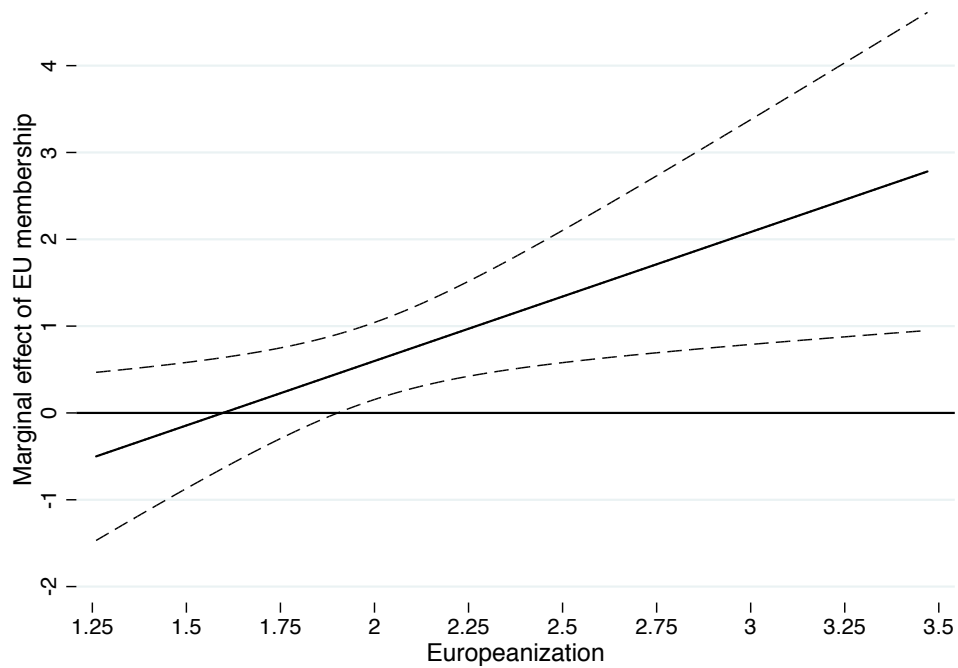


Figure 1: The political trilemma of globalization



(a) Marginal Effect of EU Membership on Extremism from the Mean Party Position



(b) Marginal Effect of EU Membership on Party-System Polarization

Figure 2

Table 1: The Effect of EU Membership on Extremism (Mean)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
EU Membership	-0.40** (0.15)	-0.39** (0.17)	-0.39** (0.17)	-0.40*** (0.15)	-0.39** (0.16)	-0.39** (0.16)	-0.39*** (0.13)
Trend	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)
EU Membership*Trend	0.01*** (0.00)	0.01** (0.00)	0.01** (0.01)	0.01** (0.01)	0.01** (0.01)	0.01** (0.01)	0.01*** (0.01)
Effective Number of Parties	0.01 (0.02)	0.02 (0.02)	0.02 (0.02)	0.01 (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)
Trade Openness	-0.00* (0.00)	-0.01** (0.00)	-0.01** (0.00)	-0.00** (0.00)	-0.01** (0.00)	-0.01** (0.00)	-0.01*** (0.00)
Population	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)
Median Voter	0.02 (0.02)	0.02 (0.02)	0.01 (0.02)	0.02 (0.02)	0.02 (0.02)	0.01 (0.02)	0.02 (0.02)
Party Vote Share	-0.64*** (0.19)	-0.60*** (0.20)	-0.60** (0.20)	-0.64*** (0.12)	-0.60*** (0.13)	-0.60*** (0.13)	-0.60*** (0.11)
GDP Growth	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
EMU			0.10 (0.11)			0.10 (0.11)	
Single-Market			-0.16** (0.07)			-0.16* (0.10)	
Constant	2.05*** (0.59)	2.15*** (0.64)	2.29*** (0.66)	2.05*** (0.60)	2.15*** (0.69)	2.29*** (0.70)	2.15*** (0.70)
N	1420	1313	1313	1420	1313	1313	1313
R ²	0.16	0.17	0.17	0.16	0.17	0.17	0.17

Dependent variable *Extremism* is measured as the absolute distance between a party's policy position and the mean party position. Numbers in parentheses are standard errors, which are clustered by country in models (1), (2) and (3) and by election in models (4), (5) and (6). Model (7) uses bootstrapped standard errors. All models include country fixed effects, decade fixed effects, and the *disproportionality* variable which are not reported here. * significant at 10% ** significant at 5% ***significant at 1%.

Table 2: The Effect of EU Membership on Party-System Polarization

	(1)	(2)	(3)	(4)	(5)	(6)
Europeanization	0.14 (0.97)		0.21 (0.96)		-0.03 (0.09)	
Europeanization (Economic)		0.25 (0.66)		0.3 (0.65)		-0.00 (0.95)
EU Membership	-2.43* (1.28)	-2.37* (1.23)	-2.59* (1.32)	-2.62* (1.26)	-0.19 (0.13)	-0.2 (0.11)
Europeanization*EU Member	1.96** (0.84)	1.48** (0.61)	1.99** (0.86)	1.55** (0.63)	0.16* (0.09)	0.13** (0.06)
Effective Number of Parties	-0.05 (0.09)	-0.04 (0.1)	-0.06 (0.09)	-0.06 (0.09)	-0.01 (0.01)	-0.01 (0.01)
Trade Openness	-0.18** (0.07)	-0.17** (0.07)	-0.18** (0.08)	-0.17** (0.08)	-0.01 (0.01)	-0.01 (0.01)
Population	-0.04 (0.04)	-0.05 (0.05)	-0.04 (0.04)	-0.05 (0.05)	-0.01** (0.00)	-0.01** (0.00)
Lagged Polarization			0.1 (0.07)	0.1 (0.07)		
Median Voter					0.01 (0.01)	0.01 (0.01)
Constant	9.02** (4.01)	9.36* (4.63)	9.23** (3.98)	9.75* (4.77)	1.03** (0.35)	1.09** (0.39)
<i>N</i>	126	126	122	122	126	126
<i>R</i> ²	0.63	0.63	0.64	0.63	0.44	0.44

Dependent variable is *Polarization* is measured in columns (1) through (4) as the distance between the two most extreme parties, and in columns (5) and (6) as the Dalton index of polarization. Numbers in parentheses are standard errors clustered by country. All models include country fixed effects and decade fixed effects, which are not reported here. Models in columns (5) and (6) also include the *disproportionality* variable, which is not reported here. * significant at 10% ** significant at 5% ***significant at 1%.

Online Appendix (not for publication)

for manuscript

*“Take Back Control?” The Effects of Supranational
Integration on Political Extremism and
Party-System Polarization*

Nikitas Konstantinidis

Hande Mutlu-Eren

IE University

New York University

Konstantinos Matakos

King’s College London

September 18, 2017

A1. Proofs of formal results

In this section we provide the formal results of the general version of our model together with the proofs of those results and some additional empirical analyses. 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 Nash equilibrium 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 need to show that strategy $x_C^* = 0.5$ is a strictly dominant strategy for this party, i.e., $x_C^* \in \arg \max U_C(x_C, p_C) = -|x_C - 0.5| + v_C(x_C, x_R^*, x_L^*)$. Notice that, given that voters have single-peaked preferences, we must have that

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

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:

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

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

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 $[0, 1]$.¹ 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)² 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*

¹Technically 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.

²It 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.

$[0, 1]$. Then the unique symmetric Nash equilibrium of the game is the triplet $(x_L^*, x_C^*, x_R^*) = (\frac{1}{4}, \frac{1}{2}, \frac{3}{4})$.

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 $v_C(x_C, x_R^*, x_L^*) = F\left(\frac{x_C + x_R^*}{2}\right) - F\left(\frac{x_C + 1 - x_R^*}{2}\right) = x_R^* - \frac{1}{2}$. 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 $-2(x_L^* - p_L) + \frac{1}{2} = 0 \iff x_L^* = \frac{1}{4}$. 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 $[\frac{1}{2} - \alpha, \frac{1}{2} + \alpha] \subset [0, 1]$ where $\alpha \in [0, \frac{1}{2})$ and the utility function of party $j \in \{L, C, R\}$ is given by the following expression:

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

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 $b_2 > 0 > b_1$.³ Then we can state the following result that fully characterizes the unique symmetric Nash equilibrium of the game.

Proposition 3 (Constraints) *Let $x_L < x_C < x_R$, $\alpha \in [0, \frac{1}{2})$ 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 $(x_L^*, x_C^*, x_R^*) = (0, 0.5, 1)$.*

³This 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.

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) for all $x_L \in [0, \frac{1}{2} - \alpha)$ we have $\frac{\partial U_L}{\partial x_L} = -\text{sgn}(x_L - p_L) + \frac{1}{2}f\left(\frac{x_L + x_C^*}{2}\right) = -1 + \frac{1}{2} < 0$ and (b) for all $x_L \in [\frac{1}{2} - \alpha, \frac{1}{2})$ 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 $[0, \frac{1}{2} - \alpha)$ and in $[\frac{1}{2} - \alpha, \frac{1}{2})$. 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:

$$\begin{aligned} U_L(p_L = 0; x_L^{**} = \frac{1}{2} - \alpha) &= U_L(p_L = 0; x_L^* = 0) \iff \\ -\left|\frac{1}{2} - \alpha - p_L\right| + F\left(\frac{x_L^{**} + x_C^*}{2}\right) + b_2 &= -|0 - p_L| + F\left(\frac{x_L^* + x_C^*}{2}\right) - b_1 \iff \\ -\frac{1}{2} + \alpha + \frac{(\frac{1}{2} - \alpha) + \frac{1}{2}}{2} + b_2 &= \frac{0 + \frac{1}{2}}{2} - b_1 \implies \frac{\alpha}{2} + (b_2 - b_1) = \frac{1}{4} \implies \alpha^* = \frac{1}{2} - 2\beta \end{aligned}$$

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 α captures the slackness of the constraint, implying that as α *decreases* the constraint becomes *tighter*.

Corollary 1 *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 α) for every $\alpha \in [\alpha^*(\beta), \frac{1}{2})$; extremism (polarization) is maximal when $\alpha \in (\alpha^*(\beta), 0]$.*

Proof. When $\alpha \geq \alpha^*(\beta)$ define extremism as $\frac{v_L(|x_L^{**}-x_C^*|)+v_R(|x_R^{**}-x_C^*|)}{v_L+v_R} = \frac{a(v_L+v_R)}{v_L+v_R} = a$ 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 α and, hence, strictly decreasing in the tightness of the constraint. It is also straightforward to see that for $\alpha \in [\alpha^*(\beta), \frac{1}{2})$ 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.

■

Finally, we perform a normative comparison of the above mean-preserving Nash equilibria with respect to higher moments of the government policy output distribution, such as policy variance. As long as the expected policy outcome ω of a generic (coalition) government and policy-formation process in parliamentary multi-party system, where a *dominant* party j with electoral platform x_j is selected with probability v_j (Merrill and Adams 2007), lies at the center of the policy spectrum, i.e., $E[\omega] = \frac{1}{2} = x_C^{**}$ and parties are able to fully commit to the implementation of their electoral programs once in office, then the following corollary shows that in our model increasing policy convergence by the two extremes necessarily implies decreasing policy variance.

Corollary 2 *For a dominant-party (coalition) government and policy-formation process, perfect commitment technology, and a uniform voter distribution on the $[0, 1]$ interval, symmetric mean-preserving Nash equilibria $\mathbf{x}^{**} = (x_L^{**}, x_C^{**}, x_R^{**})$, where $x_C^{**} = \frac{1}{2}$ and $x_L^{**} = 1 - x_R^{**}$, with increasing convergence of extreme party platforms x_L^{**} and x_R^{**} will exhibit lower levels of policy variance $V(\omega)$.*

Proof. A simple derivation of the policy variance of any mean-preserving symmetric Nash equilibrium in this game of three-party electoral competition with proportional representation and a uniform distribution of voters will suffice to demonstrate this result:

$$\begin{aligned}
V(\omega) &= E \left[(x_j^{**} - E(x_j^{**}))^2 \right] = v_L(\mathbf{x}^{**}) \times (x_L^{**} - \frac{1}{2})^2 + v_R(\mathbf{x}^{**}) \times (x_R^{**} - \frac{1}{2})^2 = \\
&2v_L(\mathbf{x}^{**}) \times (x_L^{**} - \frac{1}{2})^2 = 2 \left(\frac{x_L^{**} + \frac{1}{2}}{2} \right) (x_L^{**} - \frac{1}{2})^2 = (x_L^{**} + \frac{1}{2}) (x_L^{**} - \frac{1}{2})^2.
\end{aligned}$$

The, taking the first derivative, we get $\frac{\partial V(\omega)}{\partial x_L^{**}} = 3x_L^{**2} - x_L^{**} - \frac{1}{4} < 0$, for any $x_L^{**} \in [0, \frac{1}{2})$. In other words, as extreme parties converge towards the center in equilibrium, policy variance is monotonically decreasing. This completes the proof. ■

A2. Additional results

Table 1: Summary Descriptive Statistics of the Data Used in the Empirical Analyses

	Mean	Std. error	Min	Max	N
Extremism (Mean)	0.80	0.64	0	4.7	1473
Extremism (Median)	0.73	0.72	0	4.7	1473
EU Membership	0.76	0.43	0	1	1473
Trend	30.82	15.15	1	57	1473
EU Membership*Trend	26.45	19.37	0	57	1473
Effective Number of Parties	4.81	1.76	2.26	10.35	1473
Trade Openness	75.35	41.84	22.89	317.22	1420
Population	129	72.64	17.62	287.16	1433
Median Voter	4.37	1.24	1.21	8.05	1473
Disproportionality	4.63	4.33	0.41	25.25	1473
Party Vote Share	0.14	0.13	0	0.54	1473
GDP Growth	2.19	2.60	-12.10	9.39	1313
EMU	0.41	0.49	0	1	1473
Single-Market	0.52	0.50	0	1	1473

Table 2: The Effect of EU Membership on Extremism (Median)

	(1)	(2)	(3)	(4)
EU Membership	-0.28*	-0.26	-0.23	-0.26*
	(0.13)	(0.15)	(0.15)	(0.14)
Trend	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
EU Membership*Trend	0.01**	0.01*	0.01	0.01*
	(0.00)	(0.00)	(0.00)	(0.01)
Effective Number of Parties	0.04**	0.04**	0.04**	0.04**
	(0.01)	(0.01)	(0.02)	(0.02)
Trade Openness	-0.00	-0.01*	-0.00**	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
Population	-0.00	-0.00	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Median Voter	0.01	0.01	0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Party Vote Share	0.33*	0.38*	0.36	0.36**
	(0.18)	(0.20)	(0.20)	(0.14)
GDP Growth		0.00	0.00	0.00
		(0.01)	(0.01)	(0.01)
EMU			0.17*	
			(0.08)	
Single-Market			-0.09	
			(0.08)	
Constant	1.26**	1.24***	1.31**	1.24*
	(0.52)	(0.55)	(0.57)	(0.63)
N	1420	1313	1313	1313
R^2	0.11	0.12	0.12	0.12

Dependent variable *Extremism* is measured as the absolute distance between a party's policy position and the median party's policy. Numbers in parentheses are standard errors, which are clustered by country in models (1), (2) and (3) and by election in model (4). All models include country fixed effects, decade fixed effects, and the *disproportionality* variable which are not reported here. * significant at 10% ** significant at 5% ***significant at 1%.