

# The Rise of Identity Politics\*

Tim Besley  
LSE and CIFAR

Torsten Persson  
IIES, Stockholm and CIFAR

March 29, 2019

## Abstract

We develop a dynamic model of multi-dimensional politics to explore how political cleavages, policies, and social identities interact over time. The model yields insights into the profound political changes we have witnessed around the world, as economic shocks and trend-wise breakdowns of traditional social hierarchies have reinvigorated nationalist sentiments. We show how such sentiments can shape immigration policy, and how this may shape the outlook for the next generation. Shifting outlooks are reflected in political preferences implied by social-identity choices, which are modeled as a process of cultural evolution. Expected policy thus feeds back to political preferences, rooted in endogenous social identities. Once we allow for endogenous political organization – formation of social movements or new political parties – the model can also encompass mechanisms of hysteresis, such that temporary shifts in nationalism can have permanent effects.

---

\*We have benefitted from comments by participants in the 2018 Munich Lectures and a public lecture in Münster, and from financial support from the ERC.

# 1 Introduction

A striking feature of recent political events across a range of countries is a new form of identity politics. Although these experiences certainly differ, a common feature is a weakening of traditional political cleavages rooted in economic differences in favor of cleavages in a second dimension where resurgent nationalism or nativism plays an important part. The exact reasons for this development are not well understood, but two drivers are frequently invoked (e.g., Gidron and Hall 2018). Increased economic polarization has been a short-run consequence of the recent financial crisis and a long-term consequence of manufacturing jobs being swept away by globalization. And heightened concerns about the loss of social status have gradually emerged among traditionally dominant groups, who feel threatened by immigration and gender equality. The election of Donald J. Trump in the US, the Brexit vote in the UK, and the increasing support for radical-right populist parties and politicians in many countries all across the world, are all commonly attributed to such shocks and trends.

Traditional theoretical approaches to electoral politics are poorly equipped for thinking through these events. Standard approaches generally rely on static models built around a single stable political dimension, usually an economic cleavage. Moreover, existing political actors – like political parties – and the political preferences among voters are generally treated as given. However, the minimal ingredients needed to think about recent events is a multi-dimensional model that encompasses identity politics and allows for evolving policy preferences, as well as a changing party structure. Ideally, we need a framework that explains under which specific circumstances, and at which specific times, these changes occur, even though the underlying grievances can be long-standing.

In this paper, we develop a dynamic model of electoral competition and identity formation. This model generates a two-way feedback between political culture and policy, fuelled by endogenous organizational change. Our approach sees social identification as capable of altering policy preferences. In this way, we break with the long tradition in political economics of taking policy preferences as exogenous and given. To encapsulate some of the ongoing discussion, we instead allow preferences to form endogenously as a key part of identity politics.

**Main results** The key equilibrium outcome from our model is a set of policies that reflect the political conflicts between traditional political elites, who control established parties formed around economic interests, and an evolving share of citizens who unite on a non-economic dimension of politics. To breathe life in this model, we apply it to the political battles over immigration, where cosmopolitan political elites hold opposing preferences to a set of citizens with nationalist sentiments. But the general approach is malleable and applies to a range of cleavages, as long as a second policy issue – independent of economic circumstances – can evolve and become more or less salient over time. Importantly, that evolution can reflect social forces and affective reasoning rather than rational self-interest. The formation of groups or social movements around nationalist, anti-immigration issues is one way to enhance their political salience. A new party, which focuses on a nationalist cause, can also boost the leverage over immigration policy by nationalists. The strength of nationalist feelings can thus promote change of political organization, in the form of new groups forming or new parties entering.

Our model delivers a number of insights. In the baseline version of the model, immigration policies reflect two key factors: (i) the polarization on economic issues between traditional elites and their political supporters, and (ii) the salience of nationalism. In each period, citizens identify as nationalists or cosmopolitans, depending on the advantage or disadvantage, in terms of immigration policy, of being a nationalist rather than a cosmopolitan type. This leads to an expanding or contracting share of nationalists over time. Permanent shifts in polarization and salience can change this dynamic process. When nationalist groups or social movements are allowed to form, we can study the conditions for multiple steady states such that even temporary shocks can have permanent effects. When parties are allowed to enter, we can explore how immigration policies differ across polities that do and do not have a nationalist party. As may be expected, party entry may occur more readily under proportional electoral rules than under plurality rule. Party entry too can lead to legacy effects from surges in nationalism, which may leave a permanent imprint on policy.

**Related research** The different components of our modeling are related to well-established lines of work in several social sciences beyond economics, including anthropology, political science, social psychology, and sociology.

As for individual-level behavior, the sociological research on identity theory (see, e.g., Stryker and Burke 2000) is mostly about the roles that accompany belonging to a certain group, focusing on intragroup structures and behavioral outcomes. Related work in social psychology includes so-called social-identity theory (Tajfel 1974, Tajfel and Turner 1979) and its extension self-categorization theory (Turner 1985). This work deals more with in-group members and their relation to out-groups and focuses on intergroup structures and cognitive outcomes. Both of these literatures are based on experiments rather than on formalized modeling. One of the striking findings is that it does not seem to take much for experimental subjects – when assigned to membership of a particular group – to adopt group-specific preferences and behaviors that benefit in-group members at the expense of out-groups. However, there is relatively little work on the drivers of this identification process and few direct applications to politics. Our approach to *individual* social identities tied to particular groups – and the adoption of policy preferences associated with those groups – in the baseline model presented in Section 2 of this paper is related to the findings in these two literatures.

Our modeling also borrows ideas from research closer to economics. Akerlof and Kranton (2000) introduced identity to economists, while Shayo (2009) proposed a formalized social-identification model. The latter applied this individual-level model to macro-level redistribution policy – people who identify with their nation, rather than their class, demand less redistribution – and showed that the predictions are broadly consistent with individual as well as cross-country survey data. Very recent research by economists indeed set out to explain the rise of populism and nationalist policies in terms of social identification. This work is complementary to ours and focuses on the link between identity and beliefs (Gennaioli and Tabellini 2018) and the mapping from identity to protectionism (Grossman and Helpman 2018). But like Shayo (2009) these papers rely on static models, while our modeling throughout the paper is dynamic.

By now, there is a large literature on populism, especially radical right-wing populism, and its origins. The contributions in Rydgren (2018) provide a set of comprehensive and up-to-date reviews of different strands in this literature, which is mostly due to political scientists and sociologists. Our approach follows some of the key ideas in Kitschelt and McGann (1997), who essentially use a two-dimensional political space where one dimension is the traditional left vs. right politics based on economic issues, and the other is labeled as liberalism vs. authoritarianism. They argue that the preferences

in this second dimension have been shifting and that it is important to study the strategic responses by incumbent parties, but their theoretical setup is quite loosely described. Along the same way, Wagner and Meyer (2017) emphasize how incumbent parties have accommodated the greater salience of the liberalism-authoritarianism dimension of politics by moving “to the right” on these issues. Our baseline model in Section 2 tries to make precise the circumstances under which existing parties adapt their proposed policies to such shifting preferences.

Of course, an important goal of many papers in the existing populism literature is to study the emergence and drivers of new nationalistic parties (Kitschelt 2018). An important idea is that economic and political modernization changed politics such that the ideology of existing parties converged (Kitschelt 1995, Kitschelt and McGann 1997, Carter 2005). Such platform convergence created room for new parties to gain electoral support via anti-globalization, anti-immigration, anti-austerity, and anti-establishment platforms. This is the argument e.g., of Hobolt and Tilley (2017). Our extended framework in Section 4 explores such party entries, with the added twist that the driving anti-liberal sentiment shift is endogenously related to earlier political responses.

Another relevant branch of political sociology is the research on social movements, especially their mobilization for political action, which goes back to key contributions like Tilly (1978), Skocpol (1979), and McAdam (1982). Walder (2009) provides a useful overview of the main ideas in this research. In line with the approach taken here, later strands of this literature emphasize that social movements foster the formation of *collective* social identities, which are stronger than individually held social identities. While the notion of social movements has been applied to explain progressive mobilizations like early revolutionary movements in Europe, and the contemporary civil-rights movement in the US, it has not applied extensively to the recent surge in nationalism. However, the recent survey articles by Caiani and della Porta (2018) and by Gattinara and Pirro (2018) suggest that such an application is natural and productive.

The extension of our baseline model that allows for the formation of groups around social identities in Section 3 of the paper can be readily interpreted as allowing social movements that foster collective identities to form. In terms of the classification used by Caiani and della Porta (2018), the two extensions of our baseline model that allow new identity-based groups, as well as new parties to enter, and therefore emphasize endogenous political

organization at the meso level, rather than the individual or macro level.

More generally, the modeling in this paper builds on our earlier work on the interplay between culture and the design of institutions, which itself draws heavily on ideas from the literature on evolutionary anthropology (Boyd and Richerson 1985, Cavalli-Sforza and Feldmann 1981). There, the design of democratic institutions interact with democratic values (Besley and Persson 2019), and the design of organizations interact with organizational culture (Besley and Persson 2018).

**Attitudinal survey data** Multi-country individual surveys, such as the World Values Survey (WVS), provide a useful window on nationalist sentiments. A commonly used question allows respondents in the survey to declare how proud they are of their nation on a five point scale. Not surprisingly, the responses vary with nationality and with individual characteristics. For example one finds a robust negative correlation between nationalism, expressed in this way and education. There also seem to be cohort or generational effects with nationalism stronger, controlling for age, among those born before World War II. It is natural to think about such attitudes as the product of socialization and life experience. While WVS data do not give us a direct insight into the causes of such dynamics, they do suggest that it will be fruitful to look into time-varying drivers of nationalism.

A key aspect of our paper is that nationalism manifests itself in policy preferences, specifically around attitudes to immigration. Although the WVS only asked the question in three of its waves, respondents were polled about their favored immigration policy among the alternatives: “Let anyone come”, “As long as jobs available”, “Strict limits”, and “Prohibit people from coming”. If we code this variable from 1 to 4 (i.e., higher values for more closedness), the responses are strongly positively correlated with professing pride in one’s nation. Likewise, being proud of one’s nation and being anti-immigrant are correlated with classifying oneself as belonging to the political right on a self-declared scale. Anti-immigrant and right-wing attitudes show similar correlations with education and birth cohort as pride in one’s nation, consistent with the idea of generational cultural dynamics.

**Structure of the paper** Section 2 of the paper formulates our baseline model, where citizens/voters have conflicting interests in two dimensions. One is their exogenous income: they can be rich or poor. The other dimen-

sion is their endogenous identity: they can be nationalists or cosmopolitans. Electoral competition determines one policy in each dimension: redistributive taxes (and transfers) favoring poor or rich, and immigration policy favoring nationalists or cosmopolitans. In each period, two existing political parties strategically propose these policy instruments, and these parties may or may not appeal to the preferences of existing nationalist voters. Over time, the share of nationalist voters may change, as new generations decide to individually identify as nationalists or cosmopolitans. However, the immigration policy, and the share of nationalists, in the long run, reflect only the exogenous parameters of the model.

Section 3 introduces a form of endogenous organization in the model, namely a social group formed by nationalist citizens at a fixed (sunk) cost. This group does not exercise its influence via campaign contributions, as in standard models of lobbying (e.g., Baron 1994, Grossman and Helpman 1996). Rather, the group strengthens the nationalist identification of its members, which makes them internalize policy outcomes more strongly. In terms of the literature on social movements, being part of a social group boosts the members' collective identity. This, in turn, makes it more attractive for parties to cater for their vote. The equilibrium formation of a group is associated with a discontinuous jump in equilibrium immigration policy. This model can have multiple equilibria (but unique dynamics), where long-run outcomes depend on initial conditions. Moreover, it has a kind of hysteresis: group formation creates an interplay between the share of nationalists and stricter immigration policy, which mutually reinforce each other.

Section 4 considers another endogenous organization, by allowing a new party that represents the nationalist voters to enter into the electoral race. Like in the case of endogenous nationalist groups, this brings about discontinuous policy shifts, the possibility of multiple equilibria, and a kind of hysteresis. We compare the policy and political outcomes with endogenous party entry under two alternative electoral systems: plurality rule and proportional representation. Our analysis shows that nationalistic parties may well form sooner – meaning for a lower share of nationalist voters – under proportional representation than under plurality rule.

Section 5 discusses a number of possible extensions and variations of our models, while Section 6 concludes. All proofs of lemmas and propositions are relegated to the Appendix

## 2 Baseline Model

We develop a baseline model of two-dimensional politics, which deliberately abstracts away from a range of realistic features. This core model also imposes a range of strong (simplifying) symmetry assumptions. Two parties are organized and propose policies across two dimensions. The first is the traditional left-right dimension, where policy redistributes (disposable) income, and people belong to given exogenous groups. In the second dimension, policy preferences are instead endogenous to people’s identities. To fix ideas, we think of the second policy as the regulation of immigration, mediating between the opposing interests of those who identify as “nationalists” and “cosmopolitans”. However, other interpretations are possible, including the regulation of environmental pollution, or of labor-market condition, striking a balance between the conflicting interests of those who identify as “environmentalists” vs. “polluters”, or as “feminists” vs. “traditionalists”, respectively. A key issue is how far the policies adopted on the second dimension meet the preferences of the new, non-traditional identity group.

### 2.1 Equilibrium with Exogenous Nationalism

In this subsection, we introduce the basic assumptions of the model and study its equilibrium in a single time period, where the share of nationalists in the population is given. The next subsection will turn to the dynamics when the share of nationalists adapts over time, and to the long-run equilibria of this share and of policy.

**Population sub-groups** A continuum of citizens is split into two economic groups  $J \in \{1, 2\}$ , where  $J = 1$  are the “poor” and  $J = 2$  are the “rich”. All economic characteristics are fixed over time. Incomes are denoted by  $y^J$  and the two groups have equal shares of the population.

In addition, at date  $t$  each citizen can individually adopt one of two social identities: “nationalist”,  $N$ , or “cosmopolitan”,  $C$ . The fraction of nationalist citizens at date  $s$  is called  $\mu_s$  and this fraction is identical in each income group (this will be an equilibrium feature of our dynamic model). Some of the population are “irreducibly” nationalist and cosmopolitan, with relatively small and constant shares  $\underline{\mu}$  and  $1 - \bar{\mu}$ , respectively, such that  $\underline{\mu} < \frac{1}{3}$  and  $\frac{2}{3} < \bar{\mu}$ . This means that  $\mu_t$  is bounded between  $\underline{\mu}$  and  $\bar{\mu}$  and that  $\bar{\mu} - \underline{\mu}$  is the maximal cultural leverage for nationalism.



**Policies and preferences** Governments pursue policies in two dimensions: a redistributive policy and an immigration policy.

The redistributive policy is an income tax rate,  $t \in [0, 1]$ , which pays for welfare-enhancing government spending or lump-sum transfers. Preferences over this policy dimension are  $\bar{U}^J(t)$ , which is a group-specific indirect-utility function, which could be derived from proper micro-foundations. Assume that each economic group has a well-defined interior optimum  $t^J$ , the group's bliss point. Because  $y^1 < y^2$ , we must have  $t^1 > t^2$ . Moreover, let  $\bar{U}^J(t) = U(t - t^J)$ , be a loss function, which is *symmetric* in distance from the bliss point. Sometimes, we will simplify further by making this distance function *linear* – e.g.,  $U(t - t^J) = -|t - t^J|$ . In either case, the distance of bliss points  $t^1 - t^2$  grows with income inequality  $y^2 - y^1$ , which can undergo (exogenous and unexpected) shifts in the analysis to follow.

The immigration policy is denoted by  $x \in [0, 1]$ . Specifically, we think about  $x$  as regulating the openness to migration. We assume that nationalists prefer  $x = 0$  while cosmopolitans prefer  $x = 1$ . These preferences are captured by decreasing, convex payoff functions,  $W(1 - x)$  for  $C$  and  $\theta W(x)$  for  $N$ . Thus,  $\theta$  indexes the (relative) “salience” of this dimension for nationalists, which could represent mere sentiments and beliefs about migration. Note that  $W$  is the same function although it takes different arguments for groups  $C$  and  $N$ . We will also allow (exogenous and unexpected) shifts in  $\theta$ .

Putting these pieces together, cosmopolitans from group  $J$  have preferences

$$V^{C,J}(t, x) = U(t - t^J) + W(1 - x),$$

while nationalists from group  $J$  have preferences

$$V^{N,J}(t, x) = U(t - t^J) + \theta W(x).$$

To simplify notation in the subsequent analysis, we define polarization on the redistributive policy as

$$z = U(0) - U(t^1 - t^2). \tag{1}$$

Parameter  $z$  thus measures the utility gain – for a rich or a poor voter – from the group's own preferred redistributive policy over the preferred policy of the out-group. Symmetry conveniently implies that polarization on the left-to-right dimension is the same for both rich and poor.

We assume throughout that

$$U(0) - U(t^1 - t^2) > W(0) - W(1). \quad (2)$$

This condition will imply (see below) that cosmopolitans always prefer to vote based on their income.

**Parties** Two traditional parties are organized along economic lines. The parties of the poor and the rich are labeled  $J = 1, 2$ . Each party is controlled by an existing political elite, who does not care for the nationalist voters as such. In the baseline case, we take the party landscape as exogenously given. (This is generalized in Section 4, where a party representing the nationalists can enter into the picture.) The parties are thus run by cosmopolitan, economic-class, citizen-candidates.

In a given period, each party offers a platform  $\{t_J, x_J\}$  to maximize the expected utility (minimize the losses) of its underlying economic group. The objective function of party  $J$  is thus  $E[U(t - t^J) + W(1 - x)]$ , where the expectation reflects uncertainty about the electoral outcome. The only credible redistributive policies are  $t_J = t^J$ . For simplicity, we assume that parties can *commit* to a nationalistic policy  $x_J$ . To offer this credibly, some nationalists may have to be recruited as candidates, at a risk a takeover of the party by the nationalists. However, we abstract from that possibility in the baseline model (see the possible extensions in the Section 5).

**Electoral strategies** As party objectives coincide with voter utilities, all the poor (rich) cosmopolitans always vote for party 1 (2), provided that (2) holds. While the cosmopolitans are loyal voters, the nationalists become swing voters: they vote for the party that offers them the highest utility. But these utilities are subject to random shocks, as in the standard probabilistic-voting models (e.g., Lindbeck and Weibull 1987, Persson and Tabellini 2000).

Formally, let  $v_J^K$  be swing-voter utility offered by party  $J$  to group  $K$  nationalists:

$$v_J^K(t_J, x_J) = U(t^J - t^K) + \theta W(x_J). \quad (3)$$

For example, a nationalist from group 1 (the poor) votes for party 1 if  $v_1^1 + \omega + \eta \geq v_2^1$ , where the *idiosyncratic* shock  $\omega$  in favor of party 1 has symmetric c.d.f.  $H(\omega)$ , with unimodal p.d.f.  $h(\omega)$ , and the *aggregate* shock  $\eta$  in favor of party 1 has log-concave symmetric c.d.f.  $G(\eta)$  with p.d.f.  $g(\eta)$ .

As the nationalist shares among poor and rich are equal, the parties offer total swing-voter utilities

$$v(x_1) = \frac{1}{2}(v_1^1 + v_1^2) = \frac{1}{2}[U(0) + U(t^1 - t^2)] + \theta W(x_1) \quad (4)$$

$$v(x_2) = \frac{1}{2}(v_2^1 + v_2^2) = \frac{1}{2}[U(0) + U(t^2 - t^1)] + \theta W(x_2) \quad (5)$$

we write these utilities as  $v(x_j)$ , because the first term in square brackets is exogenous and equal across parties by  $U(t^1 - t^2) = U(t^2 - t^1)$ . Therefore, the only way for parties to offer more or less swing-voter utility is to vary their migration policy. The maximum swing-voter utility on offer to the nationalists is thus  $\frac{1}{2}[U(0) + U(t^1 - t^2)] + \theta W(0)$ , while the minimum is  $\frac{1}{2}[U(0) + U(t^1 - t^2)] + \theta(W(1))$ .

**Winning probabilities** The poor-party vote share among nationalists is  $n_1 = H(\eta + v(x_1) - v(x_2))$ . It wins the election if its overall vote share is larger than that of the rich party – i.e., if

$$\frac{1}{2}(1 - \mu) + \mu n_1 \geq \frac{1}{2}(1 - \mu) + \mu(1 - n_1).$$

The condition is thus  $n_1 \geq \frac{1}{2}$ , or

$$\eta + v(x_1) - v(x_2) \geq H^{-1}\left(\frac{1}{2}\right) = 0, \quad (6)$$

where the equality takes advantage of our assumption that c.d.f.  $H(\omega)$  is symmetric.

As the aggregate shock  $\eta$  is drawn from c.d.f.  $G(\eta)$ , the win probability for poor party 1 is

$$P(x_1, x_2) = G(v(x_1) - v(x_2)) = G(\theta(W(x_1) - W(x_2))). \quad (7)$$

This probability is independent of  $\mu$ , the population fraction of nationalists. This is because both parties symmetrically lose loyal cosmopolitan voters when the share of nationalist swing voters rises. Symmetry of  $G$  implies that the win probability of rich party 2 is:

$$1 - P(x_1, x_2) = G(\theta(W(x_2) - W(x_1))).$$

**Gains from winning** To study the political equilibrium, we write the surplus from winning the election for party 1 as

$$Z^1(x_1, x_2) = z + W(1 - x_1) - W(1 - x_2).$$

Similarly, the surplus for party 2 is

$$Z^2(x_1, x_2) = z + W(1 - x_2) - W(1 - x_1).$$

We can thus write the poor party's objective function as  $P(x_1, x_2) Z^1(x_1, x_2)$  plus a constant.

Condition (2) implies that both parties want to win the election for all  $x_J \in [0, 1]$ . Note that  $W(1 - x)$  is increasing in  $x$  – i.e., a policy closer to the cosmopolitan loyal voters' bliss point  $x = 1$ . But the probability of winning is decreasing in  $x$  – i.e., a policy closer to the nationalist swing voters' bliss point  $x = 0$ . This means that parties face a standard trade-off: pushing policy closer to the party-members' preference diminishes the chance of winning the election.

**Political equilibrium** We now look for a Nash equilibrium in the choices of  $\{x_1, x_2\}$ , with

$$x_1 \in \arg \max_{x \in [0, 1]} \{[Z^1(x_1, x_2)]G[\theta(W(x_1) - W(x_2))]\} \quad (8)$$

$$x_2 \in \arg \max_{x \in [0, 1]} \{[Z^2(x_1, x_2)]G[\theta(W(x_2) - W(x_1))]\}. \quad (9)$$

The symmetry of economic payoffs  $U(\cdot)$ , and c.d.f.  $G$ , imply that parties face symmetric problems.

A political equilibrium is defined by a winning probability for the poor party (party 1) and two immigration policies

$$\{\widehat{P}(\theta, z), \widehat{x}_1(\theta, z), \widehat{x}_2(\theta, z)\}.$$

We have (all Lemmas and Propositions are proved in the Appendix)

**Lemma 1** *A Nash equilibrium exists and is unique.*

Lemma 1 relies on the fact that the electoral-strategy game is (log) supermodular. Thus the immigration policies are strategic complements: one party pursuing a more nationalistic policy raises the motive of the other party to do the same. The strategic complementarity reflects two effects. A tougher migration policy by one party: (i) increases polarization, which induces the other party to compete more intensely for office; (ii) appeals more to swing voters and hence reduces the other party's probability of winning, lowering the cost of setting a policy against the interests of its loyal voters. Supermodularity also makes it straightforward to handle corner solutions where  $x = 0$  or  $x = 1$ .

**Characterization** The political equilibrium has a very simple form. Given the model symmetry, it is perhaps not surprising that the two parties choose the same migration policies  $\hat{x}_1(\theta, z) = \hat{x}_2(\theta, z) = \hat{x}(\theta, z)$  and have the same chance of winning the election  $\hat{P}_1(\theta, z) = \frac{1}{2}$ . To develop some intuition for the next result, consider the first-order condition for party 1 (that defines its reaction function)

$$\begin{aligned} & -W_{1-x}(1-x_1)G(\theta(W(x_1) - W(x_2))) \\ & + \theta[z + W(1-x_1) - W(1-x_2)]W_x(x_1)g(\theta(W(x_1) - W(x_2))) \lesseqgtr 0. \end{aligned} \quad (10)$$

The first term is the cost of conceding to the nationalists while the second term is the benefit of a larger win probability.

As a preliminary step, define function  $h(m)$  for  $m \in [\underline{m}, \bar{m}]$  from

$$\frac{W_{1-x}(1-h(m))}{W_x(h(m))} = m \frac{g(0)}{G(0)},$$

where  $\bar{m} = \frac{W_{1-x}(1)/g(0)}{W_x(0)/G(0)}$  and  $\underline{m} = \frac{W_{1-x}(0)/g(0)}{W_x(1)/G(0)}$ . It is straightforward to see that  $h(\cdot)$  is a decreasing function. This function together with (10) allows us to write the Nash-equilibrium strategies of the parties in a convenient way:

**Proposition 1** *Optimal electoral strategies  $\hat{x}(\theta, z)$  are the same for both parties and characterized by:*

$$\hat{x}(\theta, z) = \begin{cases} 0 & \theta z \geq \bar{m} \\ h(\theta z) & \theta z \in (\underline{m}, \bar{m}) \\ 1 & \theta z \leq \underline{m}. \end{cases}$$

Proposition 1 says that there are three ranges for  $\theta z$ , the product of nationalistic salience and economic polarization. If this product is high, both parties give the nationalists a strict immigration policy with minimum immigration:  $x = 0$ . Here, the nationalist vote is very sensitive to concessions and/or winning the elections matters a lot because of high economic polarization across groups. If  $\theta z$  is low, the opposite is true and political elites ignore the nationalists for electoral purposes, setting  $x = 1$ . The intermediate case has an interior solution with  $x \in (0, 1)$ , which optimally trades off the winning probability against policy concessions to the nationalists.

Thresholds  $\underline{m}$  and  $\bar{m}$  depend on two things. They reflect the density of swing-voter shifts around the symmetric equilibrium point – i.e., how many swing voters shift in response to a stricter immigration policy. They also reflect the extent to which voters value extreme positions.<sup>1</sup>

**Implications** A convenient feature of the model is that its equilibrium is fully determined by  $\theta z$ , the product of the nationalistic salience and economic polarization. This gives us the following comparative-statics result

**Corollary** *Parties set a stricter immigration policy –  $x$  closer to 0 – when the nationalistic salience of immigration policy  $\theta$  is higher, and when income inequality and thus economic polarization  $z$  is greater.*

The corollary spells out a non-trivial insight about the rise of identity politics. Neither economic polarization between parties nor salience of nationalism can by itself drive the pursuit of nationalist policies. Polarization alone is not enough, because to run a immigration-restrictive platform parties have to attract sufficiently many nationalist voters from the party that represents their class interest. Salience alone is not enough, because parties have to care sufficiently about the redistributive gains from winning to court the nationalist electorate.

If we think about nationalist identification as the root of populism, we should expect the strength of a populist surge to reflect two complementary shifts in the political climate. One is an increasing salience of immigration among nationalists – a higher  $\theta$  in the model – perhaps due to media portrayal of immigration threats, or an actual immigration wave. Another is an increasing polarization of traditional parties – a higher  $z$  through a higher

---

<sup>1</sup>Note, for example that if  $W(y) = \frac{1}{2}(y)^2$  then  $W_y(0)/W_y(1) = W_y(1)/W_y(0)$ .

$y^2 - y^1$  in the model – via greater income cleavages perhaps due to globalization or new technologies.

Already this simple model predict different responses to nationalism across polities. In our baseline model, what matters is not the size of the nationalist group but the underlying elements of salience and polarization. At least for some time – even in the run up of the 2015 European refugee crisis – traditional political elites in some countries, like Germany and Sweden, appear to have resisted the temptation to court nationalists via clamp-downs on immigration. Perhaps this can be (partly) attributed to weaker polarization across traditional parties. However, Donald Trump made immigration a central plank of his US Presidential campaign, as did the leave supporters in the UK Brexit campaign. Perhaps this can be (partly) attributed to prolonged periods of increasing income cleavages across rich and poor and higher party polarization in these countries. But other differences across these countries may also have mattered. In particular, the analysis in Section 4 zooms in on the role of different electoral systems.

## 2.2 Endogenous Nationalism

We now endogenize the share of nationalists through a simple evolutionary model of social identity, focusing on the link between these dynamics and policy choices. Thus, we will ignore other, autonomous influences that could shift the share of nationalists  $\mu$ , or the salience of nationalism  $\theta$ .

**Timing** We consider a sequential-generation model, with an infinite sequence of periods indexed by  $s$ . In each  $s$ , the timing of events is as follows:

1. The polity enters period  $s$  with a share  $\mu_s$  of nationalists in the current generation.
2. Parties offer platforms  $\{t_s^1, x_{1,s}\}, \{t_s^2, x_{2,s}\}$ .
3. Individual and aggregate shocks  $\omega$  and  $\eta$  are realized.
4. An election is held where party 1 wins with probability  $\hat{P}(\theta, z)$ .
5. Payoffs are realized.
6. The next generation of citizens decide to identify as nationalists or cosmopolitans. This determines  $\mu_{s+1}$ .

As the previous subsection characterized the outcomes within a certain period, we now focus on the dynamics.

**Equilibrium fitness** The payoff difference between nationalist and cosmopolitan identity plays a key role in the analysis and represents the ex ante (i.e., before the realization of  $\omega$  and  $\eta$  at stage 3) “fitness” of being a nationalist rather than a cosmopolitan. The symmetric objective functions and the symmetric equilibrium policies,  $\hat{x}_1(\theta, z) = \hat{x}_2(\theta, z) = \hat{x}(\theta, z)$ , make it straightforward to show that this fitness is given by:

$$\Delta(\theta, z) = \theta W(\hat{x}(\theta, z)) - W(1 - \hat{x}(\theta, z)). \quad (11)$$

Function  $\Delta(\theta, z)$  is more likely to take a positive value, the higher is  $z\theta$  – e.g., if  $\theta = 1$ , then  $\Delta(\theta, z) \gtrless 0$ , as  $\hat{x}(\theta, z) \gtrless \frac{1}{2}$ . Given the comparative statics in Proposition 1, this is more likely the higher is nationalistic salience  $\theta$  and economic polarization  $z$ . Thus,  $\Delta(\theta, z)$  is increasing in each of its two arguments. Note also that  $\Delta(\theta, z)$  is time independent.

**Evolution over time** Cultural transmission of identities is a deterministic map from  $\mu_s$  to  $\mu_{s+1}$ . We specify a very general evolutionary process, but are agnostic about the specific mechanism behind it – as Besley and Persson (2019) show, one can microfound such a process by either forward-looking socialization or backward-looking imitation. In particular, the dynamics follow a revision protocol (Sandholm 2010):

$$\mu_{s+1} - \mu_s = (1 - \mu_s) \zeta^{C,N} - \mu_s \zeta^{N,C} \text{ for } \mu \in [\underline{\mu}, 1 - \bar{\mu}] \quad (12)$$

$$\zeta^{C,N} > 0 \iff \Delta(\theta, z) > 0 \text{ and } \zeta^{N,C} > 0 \iff \Delta(\theta, z) < 0.$$

This model is quite simple and makes no *a priori* assumptions about the strength of the link between the fitness of nationalism and the share of nationalists, and hence about the speed at which these dynamics occur.

Given this dynamic process, we obtain the following dynamic path for nationalism:

**Proposition 2** *For all  $\mu \in [\underline{\mu}, 1 - \bar{\mu}]$ , there are two cases*

1. *If  $\Delta(\theta, z) > 0$  the polity converges to maximal nationalism  $\bar{\mu}$  from any starting point  $\mu \in [\underline{\mu}, 1 - \bar{\mu}]$*



2. If  $\Delta(\theta, z) < 0$  the polity converges to minimal nationalism  $\underline{\mu}$  from any starting point  $\mu \in [\underline{\mu}, 1 - \bar{\mu}]$ .

This is a useful benchmark for the results to follow in the next two sections, as the only drivers of the dynamics are parameters  $\theta$  and  $z$ . A polity with a combination of weak nationalist salience and weak economic polarization will not create any movements towards nationalism and the reverse is true when  $\theta$  and  $z$  are high. Note that nationalism can disappear even without  $\hat{x}(\theta, z) = 0$ . All we need is that  $\hat{x}(\theta, z)$  is sufficiently low – i.e., the political equilibrium does not give sufficient recognition to nationalism.

The steady-state level of nationalism can change following permanent shocks to  $\theta$  and  $z$  which change  $x$ . Specifically suppose that we begin with  $(\theta, z)$  which changes to  $(\theta', z')$  such that

$$\hat{x}(\theta', z') > \frac{1}{2} > \hat{x}(\theta, z).$$

This can raise the proportion of nationalists up to  $\bar{\mu}$ . However, in this baseline model, such endogenous shifts in nationalism do not feed back onto immigration policy. The following two sections show how allowing a response of political institutions can substantially enrich the feedbacks in the model.

### 3 Nationalist Groups

Our baseline model supposes that identification with nationalists (or cosmopolitans) is tacit and individualistic: it only alters individual voting behavior. We now suppose that nationalists can get together and form an organized group at a fixed and sunk cost. We have in mind organizations such as the US Tea Party, the German Pegida demonstrators, or the early stages of the Italian Five-Star Movement. As we mentioned in the introduction, a long tradition in political sociology studies social movements which increase the salience of certain social issues. Such groups/movements may become influential via different channels of political action. One channel would be the traditional role of campaign contributions to influence election outcomes or policies, as emphasized in the political-economics literature (see Baron 1994 or Grossman and Helpman 1996).

Here, we invoke a different mechanism in tune with the research on social identity. Specifically, we suppose that being part of a formal group fosters a

collective identity among nationalists, which increases the salience of issues they care about. This view of political organization is consonant with the literature on mobilizing social movements, with group formation acting as a “meso” phenomenon in between individual behavior and macro structures (e.g., the electoral rule).<sup>2</sup> To capture this idea formally, we assume that members of an organized group internalize the payoffs of other nationalists as an expression of solidarity. This enhances their sensitivity to nationalist policies. As group formation induces political parties to cater more forcefully to group members, a discrete jump in immigration policy follows the organization of a nationalist group.

This framework links the policy outcome  $x$  and the share of nationalists  $\mu$ . For some parameters, this link produces multiple steady states. Moreover, group formation can lead to two-way feedbacks between stricter immigration policy and the share of nationalists. Most significantly, it now becomes possible for temporary shocks to either  $\theta$  or  $z$  to have permanent effects – i.e., the model entails hysteresis.

We develop the analysis in two steps. First, we study a given time period with a predetermined share of nationalists (Subsection 3.1). Second, we allow the share of nationalists to evolve over time (Subsection 3.2).

### 3.1 Group Formation and Policy

To model group formation, we first suppose that the share of nationalists  $\mu$  is fixed (and equally split between the rich and the poor). In addition to voting, we allow nationalists to form a group at a per-period total cost of  $F$ , which has to be shared among the group. In our symmetric framework, either all or no nationalists will want to join the organization

The per-capita cost of forming an interest group is thus  $F/\mu$ . We focus on the case where  $F$  is small but positive. This guarantees that if a group does not form, then it is not for the trivial reason that collective action is costly.<sup>3</sup> In effect, it rules out a group forming that would not have any

---

<sup>2</sup>Political organization can also increase communication among like-minded people leading to the creation of “echo-chambers” as discussed, for example, in Levy and Razin (2019).

<sup>3</sup>In practice, the costs of collective action could be substantial, especially in countries where government puts barriers in the way of civil society. Hence, our results can be thought of as an upper bound on the prospects for group formation. In practice, we would also expect the costs to differ substantially across people. If this were the case, then only a share of nationalists would join the group. Even then, the existence of an organized

political influence.<sup>4</sup>

**Utility of group members** We postulate that when an organized group is formed, then this increases the salience of nationalist identity by enhancing in-group attachments. To capture the sense of solidarity that a group can create, we assume that – once a group is formed – its members internalize the utility of the other in-group nationalists (but not those of the “out-group” cosmopolitans).<sup>5</sup> This may be the consequence of formal group activities – such as newsletters, advertising, or joint actions – that make members more aware of being part of nationalist collective. Forming a group may also intensify word-of-mouth or social-media communication.

Formally, we write the utility of a nationalist-group member from income class  $K$  of the policy offered by party  $J$  as

$$v_J^K(t_J, x_J) = U(t^J - t^K) + \theta W(x_J) + \int_i \xi(i) v_J^L(t_J, x_J) di, \quad (13)$$

where  $i$  runs across all individuals, whether they come from the same income class  $L = K$ , or not  $L \neq K$ . To capture identification with the nationalistic in-group, the weights are:

$$\xi(i) = \begin{cases} \xi & \text{if } i \in N \\ 0 & \text{if } i \notin N, \end{cases}$$

where the parameter  $\xi$  reflects the strength of cohesion among the nationalists when organized as a group.<sup>6</sup>

---

group could create a spillover to non-group-members if the latter feel that their views gain legitimacy from the group’s existence.

<sup>4</sup>Some of the theories of social movements reviewed in Walder (2009) do not require a group to be instrumentally powerful. A social movement could be created purely by affective behavior in response to a sense of grievance whether or not it actually affects policy.

<sup>5</sup>This is similar in spirit to the assumption about group-member preferences in Besley and Persson (2018).

<sup>6</sup>Introducing out-group hostility – i.e.  $\xi(i) < 0$  for  $i \notin N$  – would only serve to enhance the effect that we are studying. In this symmetric model, all  $\mu$  nationalists would either want to form a group, or not. It would also be possible to generalize the model in other ways. For example, we could allow either the costs of joining the group  $F$ , or the group cohesion parameter  $\xi$ , to vary across individuals. Then, the distribution of these idiosyncratic parameters would decide which individuals would join the group, and the group could have only a subset of nationalists join the group so that the organized element of nationalists would be smaller than  $\mu$ .

**Modified swing-voter utility** Evaluating (13) for each income class  $K = 1, 2$  and summing across across these classes gives

$$v(x_J) = (1 + \mu\xi)\left\{\frac{1}{2}[U(0) + U(t^1 - t^2)] + \theta W(x_J)\right\}.$$

The previous utility expressions in (4) and (5) are thus multiplied by  $1 + \mu\xi$ , an expression which is increasing in  $\mu$  and  $\xi$ . Repeating the steps outlined in Subsection 2.1, and exploiting the model symmetry, the poor party's probability of winning, in the presence of a nationalist group, is now

$$P(x_1, x_2) = G[v(x_1) - v(x_2)] = G[\Theta(\mu)(W(x_1) - W(x_2))]$$

where  $\Theta(\mu) = (1 + \xi\mu)\theta$  is the enhanced salience of nationalism once nationalists operate as an organized group. Clearly, the function  $\Theta(\mu)$  is increasing in  $\theta$  and  $\xi$  for all  $\mu$ .

**Timing** The within-period timing, incorporating the possibility of group formation, is as follows:

1. There is a given share of nationalists  $\mu$ , salience  $\theta$  and economic polarization  $z$ .
2. Nationalists choose whether to form an interest group, at per-capita cost  $F/\mu$ .
3. Parties offer platforms  $\{t^1, x_1\}$ ,  $\{t^2, x_2\}$ .
4. Individual and aggregate shocks  $\omega$  and  $\eta$  are realized.
5. An election is held where party 1 wins with probability  $\hat{P}(\theta, z)$  or  $\hat{P}(\Theta(\mu), z)$
6. Payoffs are realized.

If the nationalist group does not form, we get the same outcome as in the baseline model. However, if the group does form, we get an outcome that is similar to Proposition 1, but with salience increased to  $\Theta(\mu) = (1 + \xi\mu)\theta > \theta$ .

**Policy with group formation** For the static political equilibrium, the analysis in Section 2.1 applies unchanged, but with  $\theta$  replaced by  $\Theta(\mu)$ . Moreover, since  $\Theta(\mu) > \theta$ , it follows from Proposition 1, that formation of an organized nationalist group increases nationalist influence over immigration policy. Moreover, once the group has organized, and  $x \in (0, 1)$ , a higher share of nationalists  $\mu$  now implies a stricter immigration policy since  $\Theta(\mu)$  is increasing. There is still a threshold requirement for  $z$  and  $\theta$ . Specifically, if  $\Theta(\mu)z = (1 + \xi\mu)\theta z < \underline{m}$ , a nationalist group will not have any effect on immigration policy. So we still need that  $\theta$  and  $z$  are sufficiently high for nationalists to influence policy.

**Group formation** A nationalist group will organize if the policy gains from doing so exceed the cost of forming. Since we focus on  $F \rightarrow 0$ , we need only check whether the gain from organizing a group is strictly positive. For this, we need to be in the range where having an active nationalist has an impact on immigration policy.

Hence, we have:

**Proposition 3** *As  $F \rightarrow 0$ , a sufficient condition for a nationalist group to organize is that*

$$\mu > \left[ \frac{\underline{m}}{\theta z} - 1 \right] \frac{1}{\xi}.$$

This says that a group forms when  $\mu$  passes a threshold value which depends on  $\theta z$ . If  $\theta z > \underline{m}$ , this will occur for all  $\mu \in [\underline{\mu}, \bar{\mu}]$ . Then the sole effect of the group is to magnify the policy concession made to nationalists by the incumbent parties. A more interesting case is where  $(1 + \xi\mu)\theta z > \underline{m} > \theta z$ . Then, the group has to organize for the political parties to offer a stricter immigration policy, i.e. one that moves in the direction of nationalist preferences, i.e.,  $x < 1$ . Now, group formation is associated with a policy change when  $\mu$  is large enough.

A period in which a nationalist group organizes will see a discontinuous jump in immigration policy as the salience of the issue that the nationalists care about most jumps from  $\theta$  to  $\Theta(\mu) = (1 + \xi\mu)\theta$ , i.e. a nationalist group enhances the sense of nationalist identity. Since group formation is costly, the nationalist group is only formed when doing so changes the equilibrium immigration policy. But this does not change the insight of the baseline model that nationalists are more likely to obtain policy concessions whenever

economic polarization is large,  $z$  is high, and immigration policy is salient,  $\theta$  is high.

### 3.2 Coevolving Nationalist Identities and Groups

Our baseline model had no link between cultural evolution of nationalist identities and strategic policy design. However, in this version the value of  $\mu$  affects whether a nationalist group organizes. This dependence, in turn, adds a new dimension to the analysis, namely a coevolution of institutions and culture. Moreover, this coevolution creates the possibility of multiple steady states.

**Timing** Incorporating institutional dynamics in response to  $\mu$ , modifies the timing as follows:

1. The polity arrives to period  $s$  with a given share of nationalists  $\mu_s$  equally split among the rich and the poor, and with given values of salience  $\theta$  and economic polarization  $z$ .
2. Nationalists choose whether to organize a group at per-capita cost  $F/\mu_s$ . (They can also abandon a pre-existing group.)
3. Parties offer platforms  $\{t_s^1, x_{1,s}\}, \{t_s^2, x_{2,s}\}$ .
4. Individual and aggregate shocks  $\omega$  and  $\eta$  are realized.
5. An election is held where party 1 wins with probability  $P(\theta, z)$  or  $\hat{P}(\Theta(\mu), z)$
6. Payoffs are realized.
7. The next generation of citizens decide to identify as nationalists or cosmopolitans. This determines  $\mu_{s+1}$ .

**Fitness** We continue to assume that the dynamics of nationalist identities follow the revision protocol in Sandholm (2010). Without an existing group, the relative fitness of being a nationalist is still given by (11). When nationalist groups have formed (and are expected to be maintained), fitness is modified to:

$$\Delta(\Theta(\mu), z) = \theta W(\hat{x}(\Theta(\mu), z)) - W(1 - \hat{x}(\Theta(\mu), z)). \quad (14)$$

Moreover, it should be clear that  $\Delta(\Theta(\mu), z) \geq \Delta(\theta, z)$  – i.e., the sense of solidarity created by an organized nationalist group provides a fitness advantage to nationalists. However, the signs of the expressions in (11) and (14) remain ambiguous *a priori*.

**Dynamics of groups and values** As in Section 2.1, the polity will always converge to a steady state. To make the problem non-trivial, assume that

$$\bar{\mu} > \left[ \frac{m}{\theta z} - 1 \right] \frac{1}{\xi} > \underline{\mu}. \quad (15)$$

The expression between inequalities is, of course, the critical value of  $\mu$  at which it is optimal to form an interest group, according to Proposition 3. We thus assume that this critical value lies within the feasible range for  $\mu$ . The possible long-run outcomes are then given by:

**Proposition 4** *If (15) holds, the model with endogenous groups has three possible steady states:*

1. *If  $\Delta(\Theta(\bar{\mu}), z) < 0$ , then the polity monotonically approaches the unique steady-state  $\mu = \underline{\mu}$  from all  $\mu_0 \in [\underline{\mu}, \bar{\mu}]$ . No nationalistic group will organize. If such a group already exists, it will eventually be disbanded.*
2. *If  $\Delta(\theta, z) > 0$ , the polity monotonically approaches the unique steady-state  $\mu = \bar{\mu}$  for all  $\mu_0 \in [\underline{\mu}, \bar{\mu}]$ . A nationalist group will organize along the equilibrium path.*
3. *If  $\Delta(\Theta(\bar{\mu}), z) > 0 > \Delta(\theta, z)$ , then there exists a critical value  $\hat{\mu} \in [\underline{\mu}, \bar{\mu}]$ , such that the polity converges to the maximal-nationalism steady state  $\bar{\mu}$ , if and only if  $\mu_0 \geq \hat{\mu}$ . A nationalistic group will organize along the equilibrium path only if  $\mu_0 \geq \hat{\mu}$ . If  $\mu_0 < \hat{\mu}$ , the polity approaches the minimal-nationalism steady state  $\underline{\mu}$  without a nationalist group organizing.*

In the first clause of Proposition 4, policy does not move far enough if a nationalist group forms, to create a fitness advantage to nationalist citizens. In the second clause the opposite is true and nationalism grows monotonically over time, which eventually creates an incentive for a nationalist group to organize which results in a discrete policy jump. Once the group has

organized, the complementarity between salience and group size sets in so that the share of nationalists and immigration policy start to reinforce each other.

The third clause is arguably the most interesting since the steady state outcome depends on the starting point. Here, too, the prospective complementarity between the share of nationalists and immigration policy plays a key role. If a polity begins with a high enough share of nationalists  $\mu$ , this guarantees that  $\mu$  will eventually cross the threshold for organizing a nationalist group (if it has not reached that point already). This reinforces nationalism by shifting to a tighter immigration policy and starts the two-way feedback process between  $\mu$  and  $x$ . However, if the polity begins with sufficiently weak pro-nationalist momentum, any nationalist group is eventually disbanded, the polity loses its nationalist orientation, and immigration policy becomes more open.

Once a polity has crossed the threshold  $\hat{\mu}$ , no forces restrain nationalism. A *necessary* condition for this nationalist outcome is that a group organizes. This is because  $\Delta(\Theta(\underline{\mu}), z) < 0$  implies  $\Delta(\theta, z) < 0$ . That means that nationalism could not take off, even if a nationalist group were to (counterfactually) exist.

**Temporary shocks can have permanent effects** One implication of the third case in Proposition 4 is that a temporary change in the key parameters  $(\theta, z)$  can have permanent consequences. To see this, suppose there are two “states”, with values  $(\theta', z') \gg (\theta, z)$  such that

$$\Delta(\theta[1 + \xi\mu_0], z) < 0 < \Delta(\theta'[1 + \xi\mu_0], z').$$

Moreover, suppose that the initial level of nationalism,  $\mu_0$ , satisfies

$$\left[ \frac{m}{\theta'z'} - 1 \right] \frac{1}{\xi} < \mu_0 < \left[ \frac{m}{\theta z} - 1 \right] \frac{1}{\xi}$$

Under these circumstances, a switch to  $(\theta', z')$  from  $(\theta, z)$  at time 0 leads to a nationalist group being organized and hence an increase in  $\mu$ . If a switch back from  $(\theta', z')$  to  $(\theta, z)$  occurs at date  $s$  when there is a fraction  $\mu_s$  of nationalists such that

$$\Delta(\theta[1 + \xi\mu_s], z) > 0$$

and

$$\left[ \frac{m}{\theta z} - 1 \right] \frac{1}{\xi} < \mu_s,$$



then nationalism will continue to grow even when  $\theta$  and/or  $z$  have fallen back from  $(\theta', z')$  to  $(\theta, z)$ .

The logic explored here can be brought to bear on the attempts to combat the extreme nationalism that had formed in some countries in the 1930s. In post-war Japan and Germany, external forces imposed policies equivalent to  $x = 1$ . This helped to quell nationalist sentiments in young generations that were forming their identities afresh. On top of that, nationalist groups were outlawed. The model could rationalize such legal constraints, in a situation when one wants to reduce nationalist fitness, but the political system cannot commit to  $x = 1$ .

The model could also illustrate the role that a supra-national body – like the European Union – could play during events like the recent refugee crisis. By trying to maintain a policy with  $x = 1$ , such a body might prevent nationalist sentiments from translating into policy in individual countries.

## 4 Nationalist Parties

So far, we have only allowed political organization to change via the creation of nationalist groups that strengthen solidarity among nationalists, even if these groups do not try to directly influence policy. We now study the participation of nationalist parties in elections. These are more common in countries with proportional elections compared to those with majoritarian electoral systems.<sup>7</sup> To shed light on this, we propose an extension of our baseline model that allows us to contrast these two electoral systems. This model shows that entry may indeed span a wider range of nationalist shares under proportional elections. The electoral system matters because it affects how three-party systems work – i.e., the political equilibrium after the entry of a nationalist party – including not only the electoral stage but also government formation after an election outcome.

We first introduce the basic assumptions that allow us to study party entry in Subsection 4.1. Given these basics, Subsections 4.2 and 4.3 establish conditions for party entry under plurality rule and proportional representation, respectively. Then, Subsection 4.4 considers the dynamics and long-run outcomes in these models with endogenous party entry.

---

<sup>7</sup>A similar logic would apply to explain why Green parties thrive in proportional election systems.

## 4.1 Basics

Suppose that nationalist citizens can form a party, at a per-capita cost of  $B/\mu$ . We will again focus on the case of small costs of collective action where  $B \rightarrow 0$ . In this section, we assume that the economic preferences are given by the linear distance function:

$$U(t - t^J) = -|t - t^J|. \quad (16)$$

This rules out party entry being driven by a desire to change the redistributive policy so that we can focus on the immigration policy.

We assume that rich and poor nationalists maximize their joint utility when they form a party. This implies that the tax rate they agree upon will lie in the interval  $[t^2, t^1]$ . We denote their optimal, i.e. unconstrained, compromise in this redistributive dimension by  $\tilde{t}$ .

**Entry** Let  $e \in \{0, 1\}$  denote the nationalists' decision to enter (or leave) the electoral race in a given period. Our baseline model in Section 2 characterized the case with  $e = 0$ . Associated with party entry  $e = 1$  will be a probability distribution over a finite set of equilibrium policies denoted by  $\mathcal{T}$ . In general, we denote the probability of each possible policy outcome  $(t, x) \in \mathcal{T}(\mu, \theta, z)$  by  $p(t, x, \mu, \theta, z)$ . The specific cases we study below will generate different such probability distributions. But for now it is useful to proceed in this more abstract way.

Given the symmetry of the model, there are equally many poor and rich individuals in the nationalist segment  $\mu$  of the population. If a nationalist party enters and competes in the election, the joint expected payoff to nationalists can thus be written as

$$N(\mu) = \sum_{(t,x) \in \mathcal{T}(\mu,\theta,z)} \left( \frac{U(t - t^1) + U(t - t^2)}{2} + \theta W(x) \right) p(t, x, \mu, \theta, z). \quad (17)$$

Below, we will give two specific examples of  $N(\mu)$  under different electoral systems.

**Timing** The within-period timing is similar to that in the model with formation of nationalist groups.

1. There is a given share of nationalists  $\mu$ , salience  $\theta$  and economic polarization  $z$ .

2. Nationalists choose whether to form a party at per capita cost  $B/\mu$ .
3. If  $e = 0$ , the two existing parties offer platforms  $\{t^1, x_1\}$ ,  $\{t^2, x_2\}$ , individual and aggregate shocks  $\omega$  and  $\eta$  are realized and the election is held where party 1 wins with probability  $\hat{P}(\theta, z)$ .
4. If  $e = 1$ , the election involves three parties offering policies, with  $\{p(x, t, \mu, \theta, z)\}_{(t,x) \in \mathcal{T}(\mu, \theta, z)}$  representing the (electoral-rule dependent) probability distribution over these policy outcomes. These probabilities may also involve a government-formation phase.
5. Payoffs are realized.

**General conditions for party entry** A nationalist party enters if and only if

$$N(\mu) - \frac{B}{\mu} > \frac{U(t^1 - t^2) + U(0)}{2} + \theta W(\hat{x}(\theta, z)).$$

where the left-hand side is the net expected payoff with entry ( $e = 1$ ), while the right-hand side is the expected payoff without entry ( $e = 0$ ).<sup>8</sup>

We will work with models where the following condition holds:

**Condition 1** *The payoff after entry is (weakly) increasing in  $\mu$ , with  $N(\mu) = U(\tilde{t}) + \theta W(0)$  at some  $\mu < \bar{\mu}$ .*

The first part of Condition 1 says that the expected payoff to a nationalist party is increasing in the share of nationalists. The second part says that, if a nationalist party has entered, then at some point, before the nationalist share reaches its maximum  $\bar{\mu}$ , it gets its preferred policy – i.e.,  $t = \tilde{t}$  and  $x = 0$  – with certainty – e.g., when the share reaches above one half (recall that  $\bar{\mu} > \frac{2}{3}$ ). We will verify that Condition 1 holds in the specific models below.

We can now state a useful provisional result which will be applied below:

**Lemma 2** *If Condition 1 holds, then for small enough  $B$ , there exists  $\hat{\mu} \leq \bar{\mu}$  such that  $e = 1$  for  $\mu > \hat{\mu}$ , whenever  $\hat{x}(\theta, z) > 0$ .*

---

<sup>8</sup>We assume that no entry occurs in the case of indifference.

As in the case of a nationalist group, entry is motivated by the desire to shift to a tougher immigration policy. Moreover, because it is costly, entry will only occur if it induces a strictly tougher policy  $x < \hat{x}(\theta, z)$ . In the two following subsections, we characterize  $\hat{\mu}$  and party entry – first under plurality rule, then under proportional representation.

## 4.2 Plurality Rule

**The electoral system** We consider the simplest case of plurality rule where each voter sincerely casts her ballot for one and only one of the participating parties. All voters belong to a single electoral district which comprises the whole polity. Moreover, the election is based on winner-takes-all with a very simple government formation phase where the party with the largest vote share wins all the seats, forms a government, and acquires an unchecked right to choose policy.

Whenever a nationalist party has entered, we assume that the stochastic shocks to nationalist voter preferences in our baseline model become irrelevant. This means that the nationalists are no longer swing voters, but instead – like the rich and the poor cosmopolitans before – vote sincerely for the party that represents their own interests. If we were to maintain some neutral non-loyal voters in the model, this would likely strengthen the results (on entry) derived below.

**Equilibrium policies** The sincere voting by all voter groups means that each party pursues policies equal to the preferred outcomes of these voters: even though commitments in immigration policy are possible, there is no value of policy compromise in a citizen-candidate world. Parties 1 and 2 thus offer policy outcomes  $\{t^1, 1\}$  and  $\{t^2, 1\}$  respectively.

That said, because the poor and rich nationalists have different views about redistribution, they have to reach a compromise within the nationalist party. We suppose that if they win, they pursue policies  $\{\tilde{t}, 0\}$ .

In our simple model of plurality rule, the largest party wins the election (otherwise there is a randomization among parties with equal numbers of voters). This implies that the policy probabilities defined above become

$$p(\tilde{t}, 0, \mu, \theta, z) = \begin{cases} 1 & \text{if } \mu > 1/3 \\ 0 & \text{if } \mu \leq 1/3. \end{cases}$$

and

$$p(t^J, 1, \mu, \theta, z) = \begin{cases} 0 & \text{if } \mu > 1/3 \\ \frac{1}{2} & \text{if } \mu \leq 1/3 \end{cases} \quad J = 1, 2.$$

By the symmetry of the model, the poor and rich parties each carry half of the cosmopolitan vote  $(1 - \mu)$ . This means that the nationalist party wins for sure when the share of nationalist voters  $\mu$  is higher than a third. We can thus write the post-entry payoff in (17) as

$$N(\mu) = \begin{cases} U(\tilde{t}) + \theta W(0) & \text{if } \mu > 1/3 \\ \frac{U(t^2 - t^1) + U(0)}{2} + \theta W(1) & \text{if } \mu \leq 1/3, \end{cases}$$

which clearly satisfies the two parts of Condition 1.

**Equilibrium entry** Drawing on Lemma 2, we can now state

**Proposition 5** *Under plurality rule, when  $B \rightarrow 0$  a nationalist party is active if and only if  $\mu > 1/3$  and  $\hat{x}(\theta, z) > 0$ .*

In terms of Lemma 2, the critical limit for entry is  $\hat{\mu} = 1/3$ . Entry can only occur when the nationalist party is strong enough to collect a plurality of the votes for the legislature. But entry only takes place when it triggers a tougher immigration policy. Bearing in mind the comparative statics in Proposition 1, this means that entry of a nationalist party is more likely when nationalism is *not* very salient and economic polarization is *not* very strong. Under the opposite circumstances, the nationalists already get the immigration policy they want as the other parties are eager to compete for their votes, so it is not necessary to bear the entry costs. The entry of a nationalistic party thus becomes a substitute for policy concessions by existing parties when the social conditions are different.

### 4.3 Proportional Representation

Next, we consider party entry under proportional representation.

**The electoral rule** Let us assume that the election is still fought in a single polity-wide district. But now, each party obtains a seat share in the legislature that (closely) corresponds to its vote share. Moreover, the setting of policy must be backed by at least 50 percent of the seats in the legislature.

We are thus implicitly considering a parliamentary form of government. As our model has full party discipline, this can only be achieved by a single party if this party commands more than half the seats, which requires half the votes. Otherwise, setting of policy requires a coalition of two parties. We can think about the determination of the coalition and its policy as a process of government formation under positive parliamentarism.

As in Subsection 4.2, we assume that the electoral shocks to the nationalist voters in the two-party baseline model are no longer relevant when these voters are represented by their own nationalist party. In the same way as under plurality rule, all voters thus cast a single ballot sincerely in favor of their preferred party.

**Equilibrium policies** If  $\mu > 1/2$ , the nationalist party will control policy in the legislature without having to form a coalition with some other party. Hence, the policy outcome will be  $\{\tilde{t}, 0\}$ . Under these circumstances, the logic of entry is the same as under plurality rule, albeit with a different threshold value of  $\mu$ . The most interesting question is therefore whether entry can occur when  $\mu < 1/2$  and, in particular, when  $\mu < 1/3$  – i.e., whether entry is feasible under proportional representation but not under plurality rule.

When  $\mu < 1/2$ , parties 1 and 2 can always choose to form a coalition that will defeat the nationalists in the legislature. We suppose such a coalition would maximize its surplus and hence set policy at  $\{\bar{t}, 1\}$ . To keep a level playing field, we assume that  $\bar{t} = (t^1 + t^2) / 2$  – i.e., the compromise tax rate is equidistant between the ideal points of the cosmopolitan parties.<sup>9</sup>

Can there still be entry of a nationalistic party in this case? To study this, we must specify the policy choices made by a coalition between the nationalist party and one of the cosmopolitan parties  $J$ . Intuitively, a nationalist party has something more to offer in the redistributive dimension than the other cosmopolitan party since latter represents only one class of voters, while a nationalist party represents both rich and poor voters. The best possible outcome for the nationalists in a coalition with party  $J$  is therefore given by

---

<sup>9</sup>If this did not hold, it would be easier for a nationalist party to form a coalition as it would choose to coalesce with the party that was forced to make the largest compromise on redistribution.

$$\widehat{N}^J = \max_{(t,x)} \left\{ \frac{[U(t-t^1) + U(t-t^2)]}{2} + \theta W(x) \right\} \quad (18)$$

subject to  $U(t-t^J) + W(1-x) \geq U(\bar{t}) + W(0)$ .

This is the party's maximal payoff, subject to the participation constraint of the cosmopolitan coalition partner  $J$ , i.e. that it gets at least as much as it would by forming a coalition with the other cosmopolitan party.

**Lemma 3** *The solution to (18), has  $t = t^J$  and*

$$x = \widehat{x}^J(z) = \begin{cases} 0 & \text{if } \frac{z}{2} \geq W(0) - W(1) \\ 1 - W^{-1}(W(0) - \frac{z}{2}) & \text{otherwise.} \end{cases}$$

Intuitively, one possibility for mutually profitable exchange is that the nationalist party offers the cosmopolitan coalition partner its preferred outcome in the redistributive dimension in exchange for the nationalists' preferred outcome in the immigration dimension. (Recall that the nationalist party is indifferent between all  $t \in [t^2, t^1]$ .) This outcome is sufficient for the cosmopolitan party if it makes them better off compared to making a redistributive compromise with the other incumbent party. This condition is more likely to hold when economic polarization, represented by  $z$ , is large. If this condition does not hold, then the nationalist party will have to make a concession on immigration policy to satisfy the participation constraint of the cosmopolitan party, i.e. pick  $x^J > 0$ .

The symmetry of the model makes a coalition with the poor and the rich party equally attractive, for the nationalists. We can therefore think about  $\widehat{N}^J$  coming about by having the two cosmopolitan parties competing (a la Bertrand) with each other for including the nationalists in a coalition. Because of the symmetry, the nationalists are equally well off in either coalition, so it is natural to assume that their actual choice of coalition partner (should a coalition occur) is random.

Since  $\widehat{N}^J$  is an upper bound on the nationalists' coalition payoff, entry will never be optimal if it is not optimal with  $(t^J, \widehat{x}^J(z))$ . To show that entry is possible, we focus on the case where this is the outcome. That is to say, we postulate equilibrium policy probabilities

$$p(\tilde{t}, 0, \mu, \theta, z) = \begin{cases} 1 & \text{if } \mu > 1/2 \\ 0 & \text{if } \mu \leq 1/2. \end{cases}$$

and

$$p(t^J, \hat{x}^J(z), \mu, \theta, z) = \begin{cases} 0 & \text{if } \mu > 1/2 \\ 1/2 & \text{if } \mu \leq 1/2. \end{cases} \quad J = 1, 2.$$

Putting these pieces together, we have

$$N(\mu) = \begin{cases} -\frac{z}{2} + \theta W(0) & \text{if } \mu > 1/2 \\ -\frac{z}{2} + \theta W(\hat{x}^J(z)) & \text{if } \mu \in [\underline{\mu}, 1/2]. \end{cases}$$

**Equilibrium entry** We can now state the following result:

**Proposition 6** *Under proportional representation, as  $B \rightarrow 0$ , a nationalist party enters for all*

1.  $\mu > 1/2$ , unless  $\hat{x}(\theta, z) = 0$
2.  $\mu \in [\underline{\mu}, 1/2]$ , if  $\hat{x}(\theta, z) > \hat{x}^J(z)$ .

In terms of Lemma 2, the critical limit here is either  $\hat{\mu} = 1/2$  or  $\hat{\mu} = \underline{\mu}$ . The proposition says that entry is possible, but not guaranteed, when  $\underline{\mu} < \mu < 1/3$ . If a nationalist party forms, it enters into a coalition either with the rich or the poor. This coalition is based on trading policy favors, whereby the cosmopolitan party achieves a better deal in the redistributive dimension (compared to the non-nationalist coalition) in exchange for a tougher policy in the nationalist dimension. However, these policy motives may not be sufficiently strong compared to the outcome without entry for the nationalists to form their own party. However, when  $\hat{\mu} = \underline{\mu}$ , we see entry of a nationalist party at a lower population share of nationalists when the electoral system is based on proportional representation rather than plurality rule.

As under plurality rule, we expect a nationalist party to enter when the immigration policy in the two-party equilibrium is pretty lax, which in turn reflect relatively low salience  $\theta$  and/or relatively low economic polarization  $z$ . At the same time, low polarization limits the gains from a nationalist-cosmopolitan coalition (recall the discussion of Lemma 3). In the same way as under plurality rule, the entry of a nationalist party is associated with a discrete change towards a tougher immigration policy.



## 4.4 Coevolving Parties and Nationalist Identities

As a final step, we may consider the joint evolution of nationalist identities and the party system. To do so, we use a timing structure analogous to that in Subsection 3.2 where we studied the dynamics with endogenous interest groups.

Following a similar approach as in that section, it suffices to specify the relative fitness of a nationalist relative to a cosmopolitan. In the current context, this depends on whether or not the nationalists have formed a party, i.e.,

$$\Delta(\mu, \theta, z) = \begin{cases} \sum_{(x,t) \in T(\mu, \theta, z)} [\theta W(x) - W(1-x)] p(t, x, \mu, \theta, z) & \text{if } \mu \geq \hat{\mu} \\ \theta W(\hat{x}(\theta, z)) - W(1 - \hat{x}(\theta, z)) & \text{otherwise.} \end{cases}$$

This is a piece-wise linear function with an upward jump at  $\hat{\mu}$  – i.e., the point of entry of the nationalist party. The jump in fitness reflects the fact that – under both electoral systems – entry occurs only when it triggers a tougher immigration policy, i.e. a lower value of  $x$ . Using this insight in the same way as in Section 3, we obtain

**Proposition 7** *The model with endogenous party entry has three possible steady states:*

1. *If  $\Delta(\hat{\mu}, \theta, z) < 0$  the unique steady-state has  $\mu = \underline{\mu}$  for all  $\mu_0 \in [\underline{\mu}, \bar{\mu}]$  and no nationalist party forms. If such a party exists initially, it exits along the equilibrium path.*
2. *If  $\Delta(\underline{\mu}, \theta, z) > 0$ , then the unique steady-state has  $\mu = \bar{\mu}$  for all  $\mu_0 \in [\underline{\mu}, \bar{\mu}]$  and a party forms along the equilibrium path (at  $\mu = \hat{\mu}$ ).*
3. *If  $\Delta(\hat{\mu}, \theta, z) > 0 > \Delta(\underline{\mu}, \theta, z)$ , then the polity converges to  $\bar{\mu}$  with a nationalist party if  $\mu_0 \geq \hat{\mu}$  and to  $\underline{\mu}$  without a nationalist party if  $\mu_0 < \hat{\mu}$ . A nationalist party may either enter or exit along the equilibrium path.*

The result in Proposition 7 with endogenous party formation is very similar to our result in Proposition 4 with endogenous group formation. A growing share of nationalists can change political institutions – here, by creating a nationalist party – and this will feed back onto the level of nationalism in the population. Drawing on Propositions 5 and 6, however, entry only happens

when the pre-entry equilibrium does not recognize the forces of nationalism and instead exhibits a lax immigration policy, a high value of  $x$ .

If a nationalist party enters, this is associated with a tougher immigration policy, which provides a boost to the nationalist cause. This can create hysteresis, as in case 3 of Proposition 7. As already mentioned, the impact of economic polarization  $z$  is now somewhat subtle. Specifically, a high value of  $z$  produces a tougher immigration policy (a low value of  $x$ ) in the pre-entry equilibrium, but also magnifies the post-entry equilibrium gain to a nationalist party that forms a coalition with one of the class-based cosmopolitan parties.

The key insight from Proposition 7 is that endogenous parties can create policy hysteresis. Once a nationalist party has been created, perhaps in response to a temporary shock to  $z$  or  $\theta$ , nationalism can maintain a toe-hold even after the shock subsides as the party continues to exert influence on policy outcomes. This illustrates vividly why it is important to consider endogenous institutions to get a complete understanding of the issues surrounding the sustainability of nationalist culture.

## 5 Extensions and Future Directions

The model that we have developed here is extremely simple. But the findings can be generalized and enriched to look at wider range of issues.

**Relaxing economic symmetry** We have put a lot of symmetry restrictions on the model so that we could focus on a new set of issues. But the core structure can be expanded to consider the possibility of nationalists belonging disproportionately to different income groups, a less stylized shape of income inequality, and differential salience of nationalism across income groups. We can also allow for asymmetric polarization – i.e. the rich or poor party fears more or less having the other party in power. Moreover, there is no need to have equal numbers across these groups. With all these extensions, the model structure that leads up Proposition 1 remains unchanged, and the game between parties remains log supermodular. We can thus easily uncover versions of the comparative static results in the corollary to Proposition 1, namely that a combination of increased polarization and salience of nationalism drive nationalist policy.

**Relaxing political symmetry** Once the preferences and sizes of the underlying groups are allowed to differ this also destroys the symmetry of the political model. On the other hand, this allows us to investigate new questions, such as which of the existing parties would be most anxious to cater to the nationalist voters.

The simplifying assumptions on the political side of the model can be relaxed as well. To take but one example, we could introduce more than one electoral district and an asymmetric distribution of economic groups across these districts. This would make the comparison between plurality rule and proportional representation richer and more subtle than in Section 4 (see e.g., Persson and Tabellini 1999).

**Citizen candidates** We have supposed that parties can commit to a policy in the  $x$  dimension. Arguably, this may not be credible unless the party takes on citizens candidates from among the nationalists. Relaxing the commitment assumption, the only credible policy may otherwise be  $x = 1$  which – as we have seen in Section 4 – would make entry of nationalist parties more likely. Extending the framework in this direction goes beyond issues of policy credibility. There is now a risk of a takeover of a party by nationalists: once they become a large enough fraction of the party, they can challenge a party leadership made up by cosmopolitans. If such takeovers were realized, this would introduce a mechanism for hysteresis already in the baseline model with only two parties.

**Multiple issues** As long as political preferences remain separable, it is possible to introduce additional new dimensions beyond the single nationalist issue in the model. Each of these new issues would give rise to the same basic tension between political elites and those who feel strongly about a new dimension of politics. The kinds of dynamics that the paper studies could then apply to a range of dimensions, where identity politics could arise.

**Institutions and commitment** A conventional feature of the framework is that all issues are determined in political equilibrium by a single governmental authority. It would be interesting to investigate the consequences of more than one level of policymaking the commitments entailed in alternative institutional arrangements. For example, there is effective commitment to  $x = 1$  within the European Union. In this case, policy in the second di-

dimension is no longer determined in a national political equilibrium. Within the logic of our model, this reduces the motives for citizens to become nationalists. Of course, the model perspective is too simple, in that a country can pursue many policies beyond formal openness that all have a bearing on the welcoming of people. But the general point remains that auxiliary institutions, beyond policies decided in national elections, help shape social identification and cultural dynamics.

**Untangling causality** More generally, our framework is useful for thinking about pathways of causality in relation to nationalist dynamics of the kind we have recently witnessed. The public discussion makes much of things that – in our framework – are only proximate causes of nationalism. One example is a preoccupation with the share of nationalists  $\mu$  without trying to get behind why  $\mu$  is varying. Another example is that the creation of nationalist parties is discussed as if this is a cause of policy change. Our model has a minimalist approach: the root causes are captured by parameters  $\theta$  and  $z$ . Parameter  $\theta$  represents autonomous changes in nationalist sentiments which could be rooted in new technologies, such as social media, or exogenous shocks, such as the Syrian refugee crisis. Parameter  $z$  represents a range of economic factors that have increased economic polarization, such as fallouts of globalization, and the recession and policies driven by the financial crisis.

**Critical junctures and timing** A related point is that our framework suggests sharply non-linear responses at “critical junctures”. These can induce organizational change, even for small changes in  $\mu$  caused by underlying shocks to  $\theta$  and  $z$ . This model may help us think about puzzles regarding the timing of events. Observers frequently point out that pressures on manufacturing and other industries due to migration and trade liberalization predate many of the recent political developments. However, this is exactly what we should expect with threshold effects of the kind suggested by our model. Tensions could build gradually, but suddenly erupt into organized political activity.

## 6 Concluding Comments

This paper has developed a model framework with three key elements: (i) multi-dimensional politics with one non-economic dimension based on po-

litical identity, (ii) dynamics of social identity, a type of cultural change, and (iii) party entry or group formation, depending on the evolving culture. We are therefore able to capture some elements of the disaffection between incumbent elites, who control major parties, and an evolving share of the electorate, who are less cosmopolitan in their outlook. This rhymes with a large literature in political science which has argued that a major shift has been underway for some time shifting from traditional left-right spectrum towards greater importance of an authoritarian-liberalism dimension.

Our model delivers a number of insights. These paint a multi-faceted picture of when and how we should expect a second policy dimension – like strict immigration control based on nationalism – to become salient in national politics. When economic polarization is strong, we might expect existing parties to pander to nationalists, under the condition that immigration policy is salient enough in their preferences. When these conditions are not present, a shift in immigration policy may require the formation of a nationalist social movement or a new nationalist party, which in turn hinge on the evolution of the share of citizens who identify as nationalists rather than cosmopolitans. Taken together, the findings suggest we are more likely to see such policy shifts pushed by existing parties in countries with plurality rule, but by new nationalist parties in countries with proportional representation.

From a substantive point of view, the paper stresses the need to study political dynamics in a systematic way. Existing political commentators do a poor job at identifying endogenous and exogenous aspects in the observed process of political change. Our model assumes that the fundamental shifters are exogenous shocks to two economic and political conditions: the economic drivers of political polarization, and the salience of the non-traditional political dimension. Of course, in a bigger picture, these shifters themselves are direct or indirect products of policy. For example, globalization due to liberal policies in the domains of trade or technology may drive the disgruntlement among economically displaced groups, and social change due to liberal policies in the domains of gender or immigration may drive disaffection among those who see their status decline.

From a methodological point of view, the paper stresses how political dynamics can lead to new steady states. The possibility of hysteresis due to endogenous political organization seems particularly important. In earlier generations, it took a world war and a long time of conscious dismantling to get rid of the institutions that had evolved to support nationalism in the 1930s. Once these were stamped out, we saw an unprecedented march of a

liberal world order. Unless existing elites understand the kinds of dynamics that underpin the rise of identity politics, they may see history repeat itself.

## References

- [1] Akerlof, George and Rachel Kranton (2000). “Economics and Identity,” *Quarterly Journal of Economics* 115(3), 715–753
- [2] Akerlof, George and Rachel Kranton (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*, Princeton, NJ: Princeton University Press.
- [3] Algan, Yann, Sergei Guriev, Elias Papaioannou, and Evgenia Passari (2017). “The European Trust Crisis and the Rise of Populism”, *Brookings Papers on Economic Activity* 48(2), 309-400.
- [4] Autor, David. H., David Dorn, Gordon Hanson, and Kaveh Majlesi (2016). “Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure”, NBER Working Paper 22637, National Bureau of Economic Research.
- [5] Baron, David (1994). “Electoral Competition with Informed and Uninformed Voters”, *American Political Science Review* 88(1), 33-47.
- [6] Barth, Fredrik (1998). *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*, Waveland Press.
- [7] Becker, Sascha. O., Thiemo Fetzer, and Dennis Novy (2017). “Who Voted for Brexit? A Comprehensive District-Level Analysis”, *Economic Policy* 32(92), 601–650.
- [8] Besley, Timothy and Torsten Persson (2018). “Organizational Dynamics: Culture, Design, and Performance,” unpublished manuscript.
- [9] Besley, Timothy and Torsten Persson (2019). “Democratic Values and Institutions,” forthcoming in *American Economic Review: Insights*.

- [10] Bisin, Alberto and Thierry Verdier (2011). “The Economics of Cultural Transmission and Socialization”, Chapter 9 in Benhabib, Jess (ed.), *Handbook of Social Economics*, Volume 1A, Elsevier.
- [11] Boyd, Robert and Peter Richerson (1985). *Culture and the Evolutionary Process*, Chicago, IL: University of Chicago Press.
- [12] Caiani, Manuela and Donatella della Porta (2018). “The Radical Right as Social Movement Organizations,” in Rydgren, Jens (ed.), *The Oxford Handbook of the Radical Right*, Oxford, UK: Oxford University Press.
- [13] Caplin, Andrew and Barry Nalebuff (1991). “Aggregation and Imperfect Competition: On the Existence of Equilibrium”, *Econometrica* 59(1), 25-59.
- [14] Cavalli-Sforza, Luigi L. and Marcus Feldman (1981). *Cultural Transmission and Evolution*, Princeton, NJ: Princeton University Press.
- [15] Gattinara, Pietro C. and Andrea L.P. Pirro (2018). “The Far Right as a Social Movement,” *European Societies* <https://doi.org/10.1080/14616696.2018.1494301>.
- [16] Gennaioli, Nicola and Guido Tabellini (2018). “Identity, Beliefs and Political Conflict,” unpublished manuscript.
- [17] Gidron, Noam and Peter A. Hall (2018). “Populism as a Problem of Social Integration,” available at <https://scholar.harvard.edu/files/hall/files/gidronhallmay2018.pdf>
- [18] Grossman, Gene and Elhanan Helpman (1996). “Electoral Competition and Special-Interest Politics”, *Review of Economic Studies* 63(2), 265-286.
- [19] Grossman, Gene and Elhanan Helpman (2018). “Identity Politics and Trade Policy,” unpublished manuscript.
- [20] Hobolt, Sara 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

- [21] Inglehart, Ronald and Pippa Norris (2017). “Trump and the Populist Authoritarian Parties: The Silent Revolution in Reverse”, *Perspectives on Politics* 15(2), 443-454.
- [22] Kitschelt, Herbert (1993). “Class Structure and Social Democratic Party Strategy,” *British Journal of Political Science* 23(3), 299-337.
- [23] Kitschelt, Herbert and Anthony McGann (1997). *The Radical Right in Western Europe: A Comparative Analysis*, Ann Arbor, MI: University of Michigan Press.
- [24] Kitschelt, Herbert (2018). “Party Systems and Radical Right-wing Parties”, in Rydgren, Jens (ed.), *The Oxford Handbook of the Radical Right*, Oxford, UK: Oxford University Press.
- [25] Levy, Gilat and Ronny Razin, [2019], “Echo Chambers and Their Effects on Economic and Political Outcomes,” *Annual Review of Economics* forthcoming.
- [26] Lindbeck, Assar and Jörgen Weibull (1987). “Balanced Budget Redistribution as the Outcome of Political Competition,” *Public Choice* 52(3), 273-297.
- [27] McAdam, Doug (1982). *Political Process and the Development of Black Insurgency, 1930–1970*, Chicago, IL: University of Chicago Press
- [28] Mutz, Diana C. (2018). “Status Threat, Not Economic Hardship, Explains the 2016 Presidential Vote,” *PNAS* 115(19), E4330-E4339.
- [29] Persson, Torsten and Guido Tabellini (1999). “The Size and Scope of Government: Comparative Politics with Rational Politicians”, 1998 Marshall Lecture, *European Economic Review* 43(4-6), 699-735.
- [30] Persson, Torsten and Guido Tabellini (2000). *Political Economics: Explaining Economic Policy*, Cambridge, MA, MIT Press.
- [31] Pettigrew, Thomas F. (2017). “Social Psychological Perspectives on Trump Supporters,” *Journal of Social and Political Psychology* 5(1).
- [32] Reny, Tyler T., Loren Collingwood, and Ali Valenzuela (2018). “Vote Switching in the 2016 Election: How Racial and Immigration Attitudes,



- Not Economics, Explain Shifts in White Voting,” forthcoming in *Public Opinion Quarterly*.
- [33] Rueda, David (2005). “Insider-outsider Politics in Industrialized Democracies: The Challenge to Social Democratic Parties”, *American Political Science Review* 99(1), 61-74.
- [34] Rydgren, Jens (2007). “The Sociology of the Radical Right”, *Annual Review of Sociology* 33(1), 241-262.
- [35] Rydgren, Jens (ed.) (2018). *The Oxford Handbook of the Radical Right*, Oxford, UK: Oxford University Press.
- [36] Sandholm, William H. (2010). *Population Games and Evolutionary Dynamics*, Cambridge, MA: MIT Press.
- [37] Shayo, Moses (2009). “A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution”, *American Political Science Review* 103(2), 147-174.
- [38] Skocpol, Theda R. (1979). *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*, New York, NY: Cambridge University Press
- [39] Stryker, Sheldon and Peter J. Burke (2000). “The Past, Present, and Future of an Identity Theory,” *Social Psychology Quarterly: Special Millenium Issue on the State of Sociological Social Psychology* 63(4), 284-297.
- [40] Tilly, Charles (1978). *From Mobilization to Revolution*, Reading, MA: Addison-Wesley
- [41] Tjafel, Henri (1974). “Social Identity and Intergroup Behaviour,” *Social Science Information* 13(2), 65-93.
- [42] Tjafel, Henri and John C. Turner (1979). “An Integrative Theory of Intergroup Conflict,” 33-47 in W.G. Austin and S. Worcel (eds.) *The Social Psychology of Intergroup Relations*, Monterey CA: Brooks/Cole.
- [43] Turner, John C. (1985). “Social Categorization and the Self-concept: A Social Cognitive Theory of Group Behaviour,” Lawler, Edward J. (ed.), *Advances in Group Processes*, volume 2, Greenwich, CT: JAI Press.

- [44] Vives, Xavier (2005). “Games with Strategic Complementarities: New Applications in Industrial Organization,” *International Journal of Industrial Organization* 23(7-8), 625-637.
- [45] Vliegthart, Rens and Hajo G. Boomgaarden (2007). “Explaining the Rise of Anti-Immigrant Parties: The Role of News Media Content in the Netherlands,” *Tijdschrift voor Communicatiewetenschap* 26(2), 404-417.
- [46] Wagner, Markus and Thomas M. Meyer (2017). “The Radical Right as Niche Parties? The Ideological Landscape of Party Systems in Western Europe, 1980-2014,” *Political Studies* 65(1), 84-107.
- [47] Walder, Andrew (2009). “Political Sociology and Social Movements,” *Annual Review of Sociology* 35, 393–412.

# Appendix

**Proof of Lemma 1** This electoral game is log supermodular

$$L(x_1, x_2) = \log[z + W(1 - x_1) - W(1 - x_2)] + \log G(\theta(W(x_1) - W(x_2)))$$

with  $\frac{\partial^2 L(x_1, x_2)}{\partial x_1 \partial x_2} =$

$$-\frac{\partial \left[ \frac{g(\theta(W(x_1) - W(x_2)))}{G(\theta(W(x_1) - W(x_2)))} \right]}{\partial W(x)} \theta \frac{\partial W(x)}{\partial x_1} \frac{\partial W(x)}{\partial x_2} + \frac{1}{[z + W(1 - x_1) - W(1 - x_2)]^2} \frac{\partial W(1 - x)}{\partial x_1} \frac{\partial W(1 - x)}{\partial x_2}$$

$> 0$ . The first term in  $\frac{\partial^2 L(x_1, x_2)}{\partial x_1 \partial x_2}$  is positive, because distribution  $G$  is log concave and  $W(x)$  is decreasing in  $x$ , while the second term is positive because  $W(1 - x)$  is increasing in  $x$ . We can then appeal to the fixed-point theorem of Topkis (Vives 2005, Caplin and Nalebuff 1991).

**Proof of Proposition 1** To prove the Proposition, it suffices to note that at an interior solution (10) can be arranged so that

$$\frac{W_{1-x}(1 - h(\theta z))}{W_x(h(\theta z))} = \theta z \frac{g(0)}{G(0)}.$$

Moreover, since  $h(\cdot)$  is increasing, there exists  $\bar{m} = \frac{W_{1-x}(1) / g(0)}{W_x(0) / G(0)}$  and  $\underline{m} = \frac{W_{1-x}(0) / g(0)}{W_x(1) / G(0)}$ . The result follows by noting that we obtain a corner solution if  $\theta z > \bar{m}$  and  $\theta z < \underline{m}$ .

**Proof of Proposition 2** To see why Proposition 2 is true, note that if  $\Delta(\theta, z) < 0$ , then (12) implies  $\mu_{s+1} < \mu_s$  for all  $\mu \in [\underline{\mu}, 1 - \bar{\mu}]$ . And if  $\Delta(\theta, z) > 0$ , then for all  $\mu \in [\underline{\mu}, \bar{\mu}]$   $\mu_{s+1} > \mu_s$ .

**Proof of Proposition 3** To prove Proposition 3, observe that for a nationalist, the payoff without a group is

$$\frac{U(0) + U(t^1 - t^2)}{2} + \theta W(\hat{x}(\theta, z)),$$

while the net payoff with such a group is

$$\frac{U(0) + U(t^1 - t^2)}{2} + \theta W(\hat{x}(\Theta(\mu), z)) - \frac{F}{\mu}.$$

As  $F$  goes to zero, the net payoff to group formation is

$$\theta [W(\hat{x}(\Theta(\mu), z)) - W(\hat{x}(\theta, z))] > 0,$$

as long as the following condition is also fulfilled:

$$\theta z (1 + \xi \mu) > \underline{m}.$$

This latter condition is the inequality in the proposition.

**Proof of Proposition 4** The first and second parts are straightforward. If  $\Delta(\Theta(\bar{\mu}), z) < 0$  then  $\Delta(\theta, z) < 0$ . Hence, whether or not an interest group forms, (12) implies  $\mu_{s+1} < \mu_s$  for all  $\mu$ . If  $\Delta(\theta, z) > 0$ , then for all  $\mu \in [\underline{\mu}, \bar{\mu}]$   $\mu_{s+1} > \mu_s$  using (12). To prove the third part, define

$$\hat{\Delta}(\theta, z, \mu) = \begin{cases} \Delta(\Theta(\mu), z) & \text{for } \mu \geq \left[\frac{m}{\theta z} - 1\right] \frac{1}{\xi} \\ \Delta(\theta, z) & \text{otherwise.} \end{cases}$$

This function is increasing and hence can switch sign at most once. From (12)  $\mu_{s+1} > \mu_s$ , if and only if  $\hat{\Delta}(\theta, z, \mu) > 0$ . As  $\Delta(\Theta(\bar{\mu}), z) > 0 > \Delta(\theta, z)$  by (15), there must exist  $\mu_H \in \left[\left[\frac{m}{\theta z} - 1\right] \frac{1}{\xi}, \bar{\mu}\right]$  such that  $\Delta(\Theta(\tilde{\mu}), z) > 0$  and by (12) we have  $\mu_{s+1} > \mu_s$ . When  $\Delta(\theta, z) < 0$ , (15) implies that there exists  $\mu_L$  such that  $\mu_{s+1} < \mu_s$  for all  $\mu_L \in \left(\underline{\mu}, \left[\frac{m}{\theta z} - 1\right] \frac{1}{\xi}\right]$ . Hence, there must exist a unique value  $\hat{\mu} \in [\underline{\mu}, \bar{\mu}]$  at which  $\hat{\Delta}(\theta, z, \mu)$  switches from strictly negative to strictly positive so that  $\mu_{s+1} > \mu_s$  for  $\mu > \hat{\mu}$  and  $\mu_{s+1} < \mu_s$  for  $\mu < \hat{\mu}$ .

**Proof of Lemma 2** To see why Lemma 2 applies, note that

$$U(\tilde{t}) + \theta W(0) - \frac{B}{\mu} > \frac{U(t^1 - t^2) + U(0)}{2} + \theta W(\hat{x}(\theta, z)),$$

holds for small enough  $B$  since  $U(\tilde{t}) = \frac{U(t^1 - t^2) + U(0)}{2}$  with (16) and  $W(0) > W(\hat{x}(\theta, z))$  if  $\hat{x}(\theta, z) > 0$ .

**Proof of Proposition 5** With  $\mu > 1/3$ , then for small enough  $B$ , the condition

$$\theta [W(0) - W(\hat{x}(\theta, z))] > \frac{B}{\mu}$$

holds for small enough  $B$ , as long as  $\hat{x}(\theta, z) > 0$ . Moreover, if  $\mu \leq 1/3$ , entry is always deterred because

$$\theta [W(1) - W(\hat{x}(\theta, z))] < \frac{B}{\mu}$$

for all  $\hat{x}(\theta, z) \in [0, 1]$ .

**Proof of Lemma 3** First, note that given (16),  $\frac{[U(t-t^1)+U(t-t^2)]}{2}$  is independent of  $t$  for  $t \in [t^2, t^1]$ . To fix ideas, consider  $J = 1$  (the argument for  $J = 2$  is analogous). This means that  $t = t^1$  and  $\hat{x}^1(z) = 0$  solves (18) if

$$U(0) + W(1) \geq U(\bar{t}) + W(0). \quad (19)$$

Because  $U(0) = 0$  and

$$U(\bar{t} - t^1) = \frac{t^1 + t^2}{2} - t^1 = -\frac{t^2 - t^1}{2} = -\frac{z}{2}.$$

we need that

$$\frac{z}{2} \geq W(0) - W(1).$$

If (19) does not hold, the solution for  $\hat{x}^1(z)$  is implicitly defined by

$$W(1 - x) = W(0) - \frac{z}{2}.$$

This expression yields the formula for  $\hat{x}^J(z)$ .

**Proof of Proposition 6** As the proof of part 1 parallels the proof of Proposition 5, we focus on part 2. Entry will occur if:

$$\hat{N}^J - \frac{B}{\mu} > \frac{U(t^1 - t^2) + U(0)}{2} + \theta W(\hat{x}(\theta, z)) = -\frac{z}{2} + \theta W(\hat{x}(\theta, z)).$$

This requires

$$\theta [W(\hat{x}^J(z)) - W(\hat{x}(\theta, z))] > \frac{B}{\mu},$$

which holds as  $B$  goes to zero whenever  $\hat{x}^J(z) < \hat{x}(\theta, z)$ .

**Proof of Proposition 7** The proof is very similar to Proposition 4. The first and second parts are straightforward. If  $\Delta(\hat{\mu}, \theta, z) < 0$ , then no nationalist party forms and (12) implies  $\mu_{s+1} < \mu_s$  for all  $\mu$ . If  $\Delta(\underline{\mu}, \theta, z) > 0$ , then for all  $\mu \in [\underline{\mu}, \bar{\mu}]$   $\mu_{s+1} > \mu_s$  using (12). To prove the third part, recall that

$$\Delta(\mu, \theta, z) = \begin{cases} \sum_{(x,t) \in T(\mu, \theta, z)} [\theta W(x) - W(1-x)] p(t, x, \mu, \theta, z) & \text{if } \mu \geq \hat{\mu} \\ \theta W(\hat{x}(\theta, z)) - W(1 - \hat{x}(\theta, z)) & \text{otherwise.} \end{cases}$$

It follows that  $\Delta(\mu, \theta, z) = \Delta(\underline{\mu}, \theta, z) < 0$  for all  $\mu < \hat{\mu}$ . Hence from (12)  $\mu_{s+1} > \mu_s$  if and only if  $\mu \geq \hat{\mu}$ . When  $\hat{\Delta}(\theta, z, \mu) > 0$ ,  $\mu_{s+1} < \mu_s$  for  $\mu < \hat{\mu}$ .