Evolving Parties

Giovanna M. Invernizzi* Federica Izzo†

August 2, 2024

Abstract

Multiparty systems often experience significant changes due to parties' splits and mergers. While sometimes splitting provides a clear electoral benefit, other times factions split without an immediate return. We present a dynamic model to analyze the evolution of political parties, accounting for both scenarios. In our model, two factions decide over time whether to split or stay together based on incentives to cultivate their individual brands. Intuitively, our results show that if splits are too damaging to factional brands factions stay together. Conversely, a united party can not be sustained if splitting significantly enhances a faction's brand. Surprisingly, if the impact on brands is neither too negative nor too positive, the equilibrium exhibits cycles: factions split today and re-merge tomorrow. Notably, these cycles can even be initiated by factions expecting electoral damage from splitting. Additionally, our analysis shows that, contrary to prior intuition, majoritarian electoral systems and increased internal power-sharing can encourage party splintering.

*Bocconi University

†Vanderbilt University

1. Introduction

Political parties are not monolithic entities. Rather, they are composed of different factions, each with their own structure and organization, and holding specific preferences that might differ from those of the broader party. This factionalism is an important driver of the fluidity of multiparty systems. Conflicts between different factions in the same party can lead to splintering, contributing to the system's fragmentation and instability; conversely, different groups may choose to merge together within the same party, aiding in the consolidation of the system. Such splits and mergers are quite frequent in multiparty systems. In Europe, mergers have occurred on average every third electoral period since World War II, and splintering is even more common (Ibenskas, 2016).

In some cases, factional splits are easy to understand, as the newly formed splinter party goes on to be very successful, attract supporters from the original party and even cause a shift in the balance of power. Consider for example the history of the UK Labour party. In the 1980s, a group of centrist Labour Party MPs broke away and formed the Social Democratic Party, citing concerns over the party's leftward shift. The split ultimately led to the formation of the Liberal Democrats, which grew to be an important political force in the UK.

Other cases are more puzzling, as the splinter party fails to gain significant support or influence. For example, in 2017, Pierluigi Bersani, the former leader of the Italian *Partito Democratico* (PD), departed from the party. Alongside a group of like-minded politicians, he founded a new political party called *Articolo 1*, positioning it as a left-wing alternative to the more centrist PD. This split had significant repercussions for both parties involved. While the PD weakened its position in Italian politics, *Articolo 1* struggled to gain momentum and failed to make a significant impact in subsequent elections. One may be tempted to interpret these developments as the consequence of strategic mistakes or miscalculations by the faction's leaders. Interestingly, however, *Articolo 1*'s poor electoral performance was widely anticipated, and this did not deter the split from occurring.[1]

[1] In a public interview, Bersani explained the necessity of establishing a separate political entity, despite the grim electoral prospects: https://www.ilfattoquotidiano.it/2017/02/28/scissione-bersani-perche-abbiamo-detto-addio-al-pd-con-renzi-stavamo-andando-a-sbattere-contro-un-muro/3420638/.

This paper presents a theory of party evolution that focuses on factions' incentives to stay together or apart. We characterize the circumstances under which parties are stable and the conditions under which, instead, we should observe factional splintering. Our theory can account for both intuitive cases of party splits where the splinter party is successful, and for puzzling cases whereby the split is initiated by a faction that does not expect to gain electorally. Furthermore, we identify cases in which splinter factions choose to re-merge within the same party. Existing theories of party system evolution usually emphasize the impact of exogenous changes in institutions or voter demands. We provide a new perspective by emphasizing the supply-side effects generated by factions' dynamic incentives, which emerge even absent such exogenous changes.

Our theory starts from the observation that different factions within the same party have distinct identities, characterized by their unique ideologies or policy positions, as well as their separate resources, organizational capacities, and bases of support in the electorate (Clarke, 2020). We refer to these distinguishing characteristics as a faction's *brand*. While they operate under the umbrella of the broader party, establishing a strong brand allows the different factions to "construct a network of political support independent of party influence. Party sub-branding is thus a crucial element in the factional politics of resource capture." (Clarke, 2020, p. 456). As such, the strength of a faction's brand determines its relative power within the party, as well as its expected success if the faction decides to split and run alone. Against this backdrop, a primary consideration for a faction is the need to preserve and cultivate its own brand.

In particular, deciding to split from the main party can either strengthen or weaken a faction's brand, as the examples above illustrate. Splitting can allow a faction to freely establish its separate ideological identity, potentially enhancing its unique brand by increasing the ideological clarity of its manifesto (Lo, Proksch and Slapin, 2016). Additionally, a faction can build its brand by capitalizing on increased attention and the opportunity to more directly allocate its resources toward improving its standing with the electorate. On the other hand, the splinter faction may struggle to consolidate its image and establish a viable identity and organization separate from the original party. Voters may also penalize the split itself or punish the new party for an ideological position less in line with their own (Duell et al., 2023). Our model allows us to analyze how such strategic considerations influence the stability or instability of political parties.

Formally, we study the repeated interaction between two factions that belong to the same ideological camp. In each period in which factions are together, they can unilaterally decide to split, forming a separate party. In each period in which factions are split, they can merge again if they both agree to do so. Each faction is characterized by a distinct brand; as described above, the evolution of these brands is influenced by the decision to split (or merge).

Whether they split or stay together, the factions' brands affect their payoff. When factions split, their brand determines their electoral strength. When they are together, their brand contributes to the strength of the party *and* influences the internal division of the spoils (exogenously in the baseline model, and endogenously in an extension where the division is determined by bargaining between the factions). Further, net of their brands, when factions are in the same party they enjoy an efficiency gain, i.e., the party's strength is more than the sum of the individual factions'. For instance, we could think of this gain as a consequence of institutional factors such as the electoral system's disproportionality.

Our analysis uncovers a rich set of equilibria. Intuitively, if splits are too damaging to factional brands, in equilibrium we have a stable party (with no splits). In contrast, if one of the factions anticipates that a split will significantly enhance its brand, we observe a split at the beginning of the game and the party system remains fragmented over time.

Suppose instead that the impact of a split on the factions' brands is not too detrimental nor too beneficial. Here, we find that the equilibrium must exhibit cycles. The two factions (which belong to the same party at the beginning of the game) split today only to re-merge tomorrow. Importantly, these cycles are not caused by shifts in environmental conditions that alter the factions' statically optimal strategy. Rather, cycles are driven solely by the faction's dynamic incentives. Furthermore, we find that in our model a cycling equilibrium can emerge even (yet not only) if the splinter faction anticipates its brand will be *damaged* by the split.

To see why, suppose that splitting damages both factions' brands. Then, when the factions reunite in the second period, the party is weaker than if it had remained unified. However, one faction may still choose to split if it anticipates that doing so will cause more harm to its opponent. In this case, the splinter faction is prepared to incur a cost in the present to improve its relative position within the party in the future, even if this harms the party as a whole. That is, the splinter faction pays a cost today to get a bigger share of a smaller pie tomorrow. Notice that this damaging split may only be sustained in anticipation of a future re-merger (i.e., as part of a cycle), which is incentivized by the efficiency premium factions enjoy when running together.

This result provides a potential rationale for the split of *Articolo 1* from Italy's Democratic Party (*PD*). As anticipated, this split primarily benefited the opposing right-wing camp. Bersani, acknowledging this, often referred to the split as part of a long-term strategy: "It's legitimate to think that we are barking at the moon, there are no tangible results yet. We're doing this more for future memory than for the concrete present."[2] In line with our theory, the split was viewed as temporary: the faction had found itself with little to no leverage in the party, and the split was part of a strategy aimed at regaining bargaining power rather than permanently breaking ranks. Indeed, Bersani himself often suggested that *Articolo 1* would be open to rejoining the *PD* in the future under the right conditions.[3] These conditions realized in 2023, when *Articolo 1* was able to merge back into a weakened *PD* from a position of *relative* strength.

This type of cycle reveals a profound inefficiency that political parties may fall victim to when internal factions compete over the distribution of spoils. In equilibrium, both factions incur costs in the short run, and even after re-merging, the party remains weaker than it would have been without the split. In the baseline model, this damaging cycle arises because the factions' relative brands exogenously determine the internal division of spoils.

A natural question is whether endogenizing this division by allowing factions to bargain in each period can eliminate the inefficiency. We consider this possibility in an extension and show that, under certain conditions, bargaining fails and the damaging cycle persists. Specifically, our results highlight a dynamic resource curse, whereby factions are unable to prevent damaging splits today when they anticipate an abundance of resources tomorrow. Anticipating high future gains, the splinter faction prefers to break away from the party even if the opposing faction is willing to offer the entire pie today.

[2] Party national assembly, November 16, 2019 (minute 9): https://www.youtube.com/watch?v=T5rrWSuyH6M. In another interview, the former party secretary acknowledges that a splinter party associated with more progressive platforms could backfire in the short term: https://video.repubblica.it/dossier/referendum-costituzionale/riforme-bersani-vi-spiego-che-significa-non-vedere-la-mucca-in-corridoio/257880/258158.

[3] For example, in this interview Bersani considers re-entering the *PD* as a potentially interesting perspective: https://www.repubblica.it/politica/2022/04/24/news/articolo1.

One might think that the trade-off of splitting versus staying together in the first period applies only to the faction that expects to improve its relative power within the party, and that the faction standing to lose bargaining power always prefers to remain in the party. Perhaps surprisingly, we instead show that splits could also be initiated by the faction that expects to weaken its relative position, as part of an equilibrium cycle. This happens when the cycle damages the splinter faction's brand, but benefits the other one and the entire camp.

In this type of cycle, the splinter faction is willing to temporarily distance itself to allow the other to consolidate its brand, so that the ideological camp can become stronger in the future. This is done in anticipation of reuniting tomorrow to enjoy the gains, which will be a smaller portion of a larger pie. Notice that this type of cycle is sustained by a collusion dynamic, whereby the splinter faction is willing to 'take one for the team'. In other words, both factions dynamically benefit from the cycle. Thus, allowing factions to bargain over the internal division of resources can never eliminate this cycle, as neither faction has incentives to do so.

The logic behind this cycle provides a potential framework for interpreting the evolution of the Japanese Democratic Party (DP), the Party of Hope (PH) and the Constitutional Democratic Party of Japan (CDP). Faced with challenging electoral prospects, the DP declared they would not contest the 2017 election and would instead merge with the newly-formed PH. Interestingly, the leadership of the PH refused to admit the more progressive members of the DP, effectively inducing a split that led to the formation of the CDP.[4] The PH's refusal to join forces may seem puzzling, in light of the fact that the CDP immediately emerged as a much stronger electoral presence (winning almost twice as many seats in the House of Representatives). According to our framework, the split was part of a long-term strategy, aimed at allowing the CDP to establish its electoral progressive brand, with the goal of joining forces again in the future once the gains would be consolidated. Indeed, the CDP ultimately became the second-largest party, and the primary center-left party. In line with a 'smaller fish in bigger pond' logic, the PH[5] eventually merged back into the CDP in 2020, becoming a less influential faction within a larger party.[6]

[4]https://www.japantimes.co.jp/opinion/2017/11/13/commentary/japan-commentary/koikes-new-political-party-lost-hope/.

[5]We note that the PH renamed itself as the Democratic Party for the People in 2018.

[6]Beyond the two cases discussed above, there are several other examples of parties around the world that have split and later re-merged. For instance, the Norwegian Liberal People's Party and the Liberal Party (1972-1988),

Finally, we build on these results to analyze the effect of institutions on party evolution. First, we look at the impact of electoral institutions. The classic Duvergerian intuition suggests that increasing the disproportionality of the system should discourage splits, by increasing the efficiency premium the factions gain by staying together (Duverger, 1951). This intuition remains valid in our framework if factions are myopic, but may fail when we consider their dynamic incentives. To see why, consider a cycling equilibrium sustained by the first, 'bigger fish in smaller pond' dynamic. If a faction can initiate a split today expecting to re-merge tomorrow from a stronger bargaining position within the party, increasing the value of the pie will only strengthen its incentives to split. Scholars have often pointed out that Duvergerian patterns are not always observed in the data (see, e.g., Cox, 1997; Singer, 2013; Diwakar, 2007). Existing explanations argue that the predictions may fail when voters, or parties, do not behave strategically or are unable to solve coordination problems (Cox, 1997; Ordeshook and Shvetsova, 1994). In contrast, our theory highlights that it is precisely due to factions' strategic incentives that Duverger's intuition may fail in a dynamic world.

Our analysis of the effect of parties' internal institutions reveals similarly counterintuitive findings. If a cycle is sustained by a 'smaller fish in bigger pond' dynamic, making the party organization less egalitarian (or increasing ideological divisions in the party) reduces splits, because in this case the splinter faction's cost of damaging its relative power is amplified. In contrast, decreasing intra-party power sharing encourages a split by the faction expecting to become a bigger fish in a smaller pond. These findings are in line with the mixed results from the empirical literature (e.g., Key, 1949; Burden, 2004), again emphasizing how considering factions' dynamic incentives may be crucial in fully understanding observed patterns of party evolution.

We conclude by emphasizing that our proposed mechanism is one among many factors explaining party evolution. Existing research has emphasized the role of institutional factors such as electoral systems (Golder, 2006*a*,*b*; Blais and Indridason, 2007) and changes in voter preferences (Rokkan and Lipset, 1967; Pedersen, 1979; Taagepera and Grofman, 2003; Invernizzi, 2023). By highlighting the importance of factional brand cultivation, we complement this literature and offer new lens through which understand party system dynamics.

---

the Dutch Catholic National Party and the Catholic People Party (1948-1955) and the Progressive Party in the US and the Republican Party (1912-1920).

2. Related Literature

Our theory is based on the premise that parties are internally divided into competing factions. The formal literature has increasingly acknowledged the importance of factions to understand political parties' nomination processes (Caillaud and Tirole, 2002; Crutzen, Castanheira and Sahuguet, 2010; Hirano, Snyder Jr and Ting, 2009), intra-party power sharing (Invernizzi, 2022; Invernizzi and Prato, 2024), and competition, both over resources (Persico, Pueblita and Silverman, 2011) and ideology (Izzo, 2023). We share with this literature the focus on within-party actors, political factions. We show how considering factional incentives to develop their brand leads to unexpected predictions on party evolution.

The literature on American and comparative politics has put forward a few alternative hypotheses for why parties emerge and change. One approach focuses on the demand side, highlighting voters' heterogeneous preferences as the key explanation for party emergence. According to this primordialist account (Rokkan and Lipset, 1967), parties originate as a consequence of social cleavages, and the more numerous the cleavages, the higher the number of parties.

An opposite "top-down" approach is the one taken by Downs (1957) and subsequently revisited by Aldrich (1995), according to whom parties are set in motion by career-concerned politicians who need an institutional machinery to support them in elections and once in office. In this tradition, Snyder and Ting (2002) study how the party leadership uses control of the party platform to more effectively signal the candidates' preferences to voters. Levy (2004) analyzes party formation in the presence of a multidimensional policy space, where policy-motivated politicians can form coalitions (parties) to credibly commit to a broader set of policies (the Pareto set of the coalition). We also model party formation, and dissolution, as a top-down process, but focus on a different mechanism and uncover novel results.

Related models have typically focused on *party entry* as a determinant of party system evolution. For instance, Buisseret and Van Weelden (2020) study how an outsider candidate decides to enter the electoral contest (either via primaries or via a third-party), while Kselman, Powell and Tucker (2016) focus on party entry in Proportional Representation systems. Closest to our model, Forand and Maheshri (2015) consider how party systems evolve in a dynamic setting under different electoral systems. In their model, dynamic considerations arise due two key ingredients: exogenous stochastic changes to voters' ideological preferences, and frictions in the electoral process for newly formed parties (i.e., an electoral penalty and higher resource demands). We abstract away from these frictions and complement this paper by focusing on factions' dynamic incentives to cultivate their brand, which leads to new theoretical findings such as cycles of splits and mergers.

Finally, our paper connects to research on party switching, where candidates or legislators change party affiliation. This literature focuses on the incentives of individual politicians, emphasizing the immediate electoral, office, and policy benefits and costs associated with party switching (e.g., Desposato, 2006; Mershon and Shvetsova, 2013*a*,*b*). Our theory differs in two respects. First, we highlight the importance of considering actors' dynamic incentives, showing that we may observe party splintering even when it is costly in the short-run. Second, we think about coordinated groups of party members (i.e., factions), rather than individual politicians, as the key actors. As such, our model can inform recent empirical work that analyzes collective switches from legislatures into new parliamentary groups.[7]

3. A model of Party Evolution

We study the interaction between two factions, a and b, belonging to the same ideological camp. For ease of exposition, we refer to the ideological camp as the left-wing one. The factions interact over two periods ($t = 1, 2$).[8] At the beginning of the game, the two factions are together in the same party, so the party and camp coincide. In each period in which factions are together, they can unilaterally decide to split and form a separate party. In each period in which factions are split, they can merge again if they both agree to do so. For clarity of exposition, we begin by assuming that only faction a can initiate a split. Below, we analyze an extension where both factions are allowed to initiate a split.

Factions' Brands. In each period t, each faction ($i = a, b$) is characterized by a brand, $\mathcal{B}_t^i$. Factions have the same initial 'stock' of brand, which we normalize to 1: $\mathcal{B}_0^a = \mathcal{B}_0^b = 1$. The evolution of these brands then depends on factions' decision to split or merge in a given period.

[7]See the Party Instability in Parliaments (INSTAPARTY) Project: https://instapartyproject.com.

[8]Our main insights are robust to considering an arbitrary number of periods. We return to this point after the presentation of the results in Section 5.1.

In particular, if factions are together at t, their individual brands remain fixed. That is, faction i's brand (for $i = a, b$) is:

$$\mathcal{B}_t^i = \mathcal{B}_{t-1}^i.$$

Instead, in every period t in which factions are split, i's brand evolves by a factor $\sigma^i > 0$:

$$\mathcal{B}_t^i = \mathcal{B}_{t-1}^i \sigma^i.$$

If $\sigma^i > 1$, then a split helps faction i to build its own brand. If instead $\sigma^i < 1$, then a split damages i's individual brand.

Payoffs. Faction i's brand in period t determines the faction's political strength, $\mathcal{S}_t^i(\mathcal{B}_t^i)$. This will, in turn, influence its payoff not only when it decides to run independently but also when it operates within the larger party structure. For simplicity, we set $\mathcal{S}_t^i(\mathcal{B}_t^i) = \mathcal{B}_t^i$.[9]

Formally, let u_t^a be faction a's payoff in period t (the utility u_t^b for faction b is defined symmetrically). Then:

$$u_t^a = \begin{cases} \left(\mathcal{B}_t^a + \mathcal{B}_t^b + \alpha_m r^t\right) \left[\frac{1}{2} + \phi(\mathcal{B}_t^a - \mathcal{B}_t^b)\right] & \text{if in the same party in } t \\ \\ \mathcal{B}_t^a + \alpha_s r^t & \text{if running alone in } t. \end{cases} \tag{1}$$

When factions remain within the same political party, their payoff is commonly linked to the overall strength of the party, $\left(\mathcal{B}_t^a + \mathcal{B}_t^b + \alpha_m r^t\right)$, which can be understood as a function of two elements. First, the strength of the two composing factions, $\mathcal{B}_t^a$ and $\mathcal{B}_t^b$; second, an exogenous component that captures the overall strength of the ideological camp in a given time period, $r^t > 0$. When r is greater than 1 the camp is gaining support over time, whereas when r is less than 1 the camp is losing support. We think of r^t as capturing loyalist voters, who ideologically identify with the left-wing camp and are already committed to voting for a left-wing party in period t. Other voters, instead, may support the party because of its specific platforms, resources, or the attractiveness of its candidates. These elements are captured by the

[9]We introduce the notation $\mathcal{S}_t^i(\mathcal{B}_t^i)$ to simplify the mapping of this reduced-form model to a microfounded version, which we discuss briefly in the next section and formally analyze in Appendix D.

brands, $\mathcal{B}_t^a + \mathcal{B}_t^b$. Finally, as we describe in more detail below, the parameter $\alpha_m > 0$ captures an efficiency premium reaped by factions when staying merged.

Factions' brands also determine the internal division of the spoils within the party, with the parameter $\phi \geq 0$ indicating the payoff elasticity to factional strength. A higher ϕ may thus capture a less egalitarian party organization where resource distribution heavily depends on factions' relative power. It may also represent more significant ideological divisions, increasing the cost for weaker factions to maintain their disadvantaged position. Finally, a higher ϕ may indicate more intense rent-seeking motivations, which incentivize factions to improve their bargaining power within the party even absent ideological divisions. We note that ϕ is appropriately bounded to ensure that a faction's share of rents is between 0 and 1. In particular, we assume that $\sigma^a, \sigma^b \in [0, \bar{\sigma}]$, and $\phi < \frac{1}{2\bar{\sigma}^2}$.

Suppose instead that factions are split in period t. When a faction runs alone, its strength is determined by its own brand and the strength of the camp as a whole. Obviously, when running alone each faction gets to keep the entirety of the pie it gains. Thus, i's payoff if split in period t is $\mathcal{B}_t^i + \alpha_s r^t$. We will make the assumption that $\alpha_m > 2\alpha_s$ in this context, to represent the efficiency premium that results from the two factions joining forces. In other words, the overall strength of the party is more than the sum of the individual components. In this setting, the difference $\alpha_m - 2\alpha_s$ could represent institutional factors, such as the degree of disproportionality in the electoral system, creating economies of scale that incentivize factions to stay together. In a highly disproportional electoral system, for example, the advantage of belonging to a larger party is significant, as the party's combined vote share translates into an even larger number of seats. To reduce notation and simplify the exposition of the results, we will set α_s to zero. Thus, a higher efficiency premium is captured by a larger value of α_m.

Timing. To sum up, the timing of the game is as follows:

1. Period 1

   1.1 Faction a decides whether to split and run alone, or remain in the party

   1.2 Factions receive their first-period electoral payoffs

2. Period 2

2.1 Split/merge decisions

- If split in period 1, a decides whether to remain split or ask b to merge. If a wants to merge and b agrees, factions re-merge. Otherwise, factions remain split
- If no split in period 1, a decides whether to split and run alone or remain in the party

2.2 Factions receive their second-period electoral payoffs

The history of the game at the beginning of period t (h_t) is then the list of split/merge decisions up to period t. A strategy for faction i determines i's action after every possible history. We focus on subgame-perfect Nash equilibria of the game.

Discussion of the Assumptions

In our model, σ^i represents in reduced form how splits affect factions' political success. For ease of presentation, in the baseline model we black-box the electoral process, and the impact of each faction's brands on their payoff. We present a potential microfoundation in Appendix D, where we introduce a probabilistic voting model where each voter chooses whether to support one of the left-wing factions, abstain, or vote for a right-wing party. Each voter's payoff from supporting faction i is a function of i's brand. A split exogenously impacts the factions' brands and thus, endogenously, their support in the electorate. Each faction then obtains a share of the spoils proportional to the number of votes it brings to the party.

We show that the micro-founded model can be fully mapped onto the reduced-form setup. Furthermore, for appropriately chosen parameter values, we can sustain all four possible cases: a split may increase both factions' political strength, damage both, or benefit one while hurting the other.[10] The impact of a split on the factions' *relative* brands determines how voters evaluate the two left-wing factions against each other. Here, improving one faction's standing must damage the other. However, the *absolute* improvement of a faction's brand always increases each voter's propensity to support the faction, relative to abstaining or voting for a party from the

[10]This is in addition to changes in the overall ideological leaning in the electorate, captured in our model by the parameter r^t.

opposed ideological camp. As a consequence, a split may improve (or hurt) both factions' overall performance at the same time.

Figure 1 complements this discussion by illustrating that examples of all four possible scenarios are observed in real-world political parties. First, both factions may be damaged by a split, $\sigma^a < 1$ and $\sigma^b < 1$. This case is illustrated by the evolution of the already mentioned Italian Democratic Party, where both the parent party and the splinter *Articolo 1* recorded disappointing electoral performances following the split. Second, both factions may benefit, $\sigma^a > 1$ and $\sigma^b > 1$. As an example of this case, consider the exit of the notorious 'Gang of Four' UK Labour Party moderates, who left Labour to found the the Social Democratic Party (SDP). While the split initially caused upheaval within the Labour Party, both Labour and the SDP ultimately benefited from going their separate ways, as it allowed each party to appeal to different segments of the electorate.

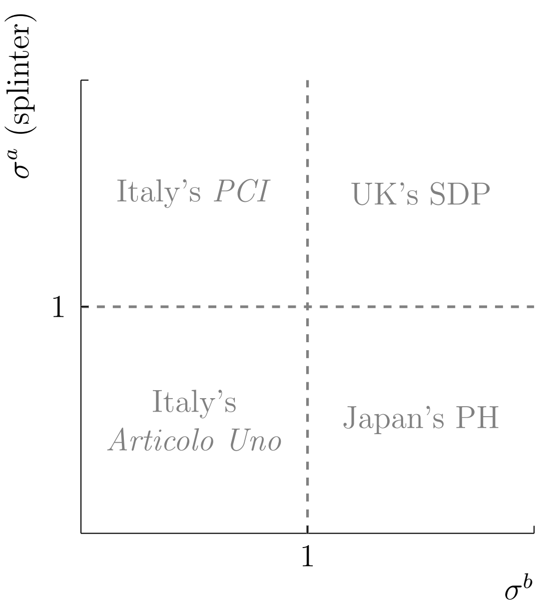

Figure 1 – Typology of different splits, as a function of the value of σ^a and σ^b, each factions' per-period shift in brand.

Third, the splinter may benefit, $\sigma^a > 1$, but hurt the parent party, $\sigma^b < 1$. An example of this is the split within the Italian Socialist Party (PSI) in the early 1920s, which led to the formation of the Communist Party of Italy (PCI) and the subsequent decline of the PSI. While the PCI benefited from the support of radical socialists and emerged as a significant political force in interwar Italy, the PSI experienced a constant decline in influence and electoral support. Finally, the splinter may suffer, $\sigma^a < 1$, but the opposing faction may benefit, $\sigma^b > 1$. The trajectory of the Japanese Party of Hope (PH) and the Constitutional Democratic Party of Japan (CDP) illustrates this case. As previously described, both descended from the Democratic Party, and the split was effectively initiated by the PH's refusal to admit the more progressive members, who went on to form the CDP. While the conservative PH lost electoral support following the split, establishing a more progressive brand benefited the Constitutional Democratic Party, which became one of the most influential parties in Japan.

In the baseline model we assume that σ^i does not vary across periods. That is, the effect of a split on the factions' brands compounds linearly over time. This assumption is useful to simplify notation and streamline the proofs, but is not necessary for the logic of our results. In Appendix C, we relax this assumption and show that our baseline findings remain robust.

We also emphasize that the assumption that factions' brands do not evolve while they are in the same party is solely for notational simplicity. Our theory's key ingredient is that the brands' evolution is *different* depending on whether factions are together or apart. Then, we can interpret the parameters σ^a and σ^b as representing the *net* effect of a split. We briefly return to this point in Section 7.3.

Finally, let us highlight that in our setup factions can perfectly foresee the consequences of a split for their relative brands (i.e., σ^a and σ^b are known). Of course, in real life these splits inevitably involve some uncertainty. We abstract away from this uncertainty in order to more clearly illustrate the mechanism behind the results, and show that dynamic incentives may generate splits in equilibrium even if factions anticipate that this will be costly in the short run (i.e., the split is statically damaging).[11] As long as the uncertainty factions may face is not too big (e.g., assuming that players' priors are sufficiently precise) our qualitative conclusions remain valid. In concluding the paper, we then briefly discuss how a large amount of uncertainty (e.g., considering priors with a high variance) may enrich our dynamics.

[11]Factions' expectations about the consequences of the split may be hard to observe, but a good proxy could come from surveys of the likeability of different factional leaders, which are often conducted even absent a split.

4. Analysis

We assume that both factions belong to the same party at the outset of the game. The aim of the analysis is then to examine how this party evolves over time. Specifically, we seek to identify the conditions under which the following scenarios occur:

1. A stable, non-fragmented party, where the factions remain merged in both periods;
2. An unstable party, where the factions remain merged in the first period but split in the second;
3. A stable fragmentation, where the factions split in the first period and do not re-merge;
4. Cycles of fragmentation, where the factions split in the first period and subsequently re-merge in the second.

Before fully characterizing the equilibrium, let's examine the strategic incentives that factions encounter in this setting. To do so, it is useful to start from a *static* benchmark. Suppose that factions only consider their current period payoff. Then, faction a would prefer to split in period t if and only if

$$\mathcal{B}^a_{t-1}\,\sigma^a > \left(\mathcal{B}^a_{t-1} + \mathcal{B}^b_{t-1} + \alpha_m r^t\right)\left[\frac{1}{2} + \phi(\mathcal{B}^a_{t-1} - \mathcal{B}^b_{t-1})\right]. \tag{2}$$

Faction a's incentives in this static benchmark are quite straightforward. If σ^a is large enough, the faction will gain a lot from running alone and will therefore choose to split from the party. On the other hand, if σ^a is small, the faction prefers to remain within the party. It is important to note that, when $\mathcal{B}^a_{t-1} = \mathcal{B}^b_{t-1}$ (as is the case in our setting in period 1), condition 2 can never be satisfied if $\sigma^a < 1$. Such damaging splits hurt the faction in the current period, hence would not emerge in equilibrium with myopic factions.

The comparative statics in this benchmark case are also intuitive. A larger α_m signifies a higher efficiency premium from the factions staying together, making it easier to maintain the equilibrium where the party stays united. Conversely, intensifying internal divisions or making the party organization less egalitarian (i.e., increasing ϕ) always incentivizes the weaker faction to split.

In our analysis below, we will see that while some of these intuitive results hold for forward-looking factions, others need to be qualified. In particular, we will show that damaging splits can emerge in equilibrium: dynamic incentives may push a faction to split even when this behavior would never be statically optimal (i.e., when $\sigma^a < 1$). Furthermore, the comparative statics results are richer, and under some conditions go in the opposite direction of what described above.

5. Equilibria

We now go back to the assumption that factions are forward-looking.

For clarity of exposition and to reduce the number of cases under consideration, but without much loss of generality for our qualitative results, we will assume that $r \geq 1$, so that the ideological camp is (weakly) gaining support over time. It is easy to demonstrate that this implies we can never have an equilibrium where factions remain merged in the first period but split in the second. Under $r \geq 1$, we then have:

Proposition 1. *There exist unique $\underline{\sigma}^a \leq \overline{\sigma}^a$ such that, in equilibrium:*

- *The factions remain merged in both periods if $\sigma^a < \underline{\sigma}^a$;*
- *The factions split in the first period and remain split if $\sigma^a > \overline{\sigma}^a$;*
- *Finally, if $\sigma^a \in [\underline{\sigma}^a, \overline{\sigma}^a]$, the factions split in the first period and re-merge in the second.*

The first and second bullet-points are intuitive, and mirror the static benchmark. When σ^a is sufficiently large, faction a can run alone and avoid the need to share the pie while also cultivating its brand. Thus, the party splits in the first period and remains split. In contrast, when σ^a is too low a split is too damaging (or not sufficiently beneficial to compensate for the loss of the efficiency premium $\alpha_m r^t$). Factions prefer to stay in the same party enjoying the efficiency gains from being together, and the party remains united for both periods.

More interesting, we see that for intermediate values of σ^a, a cyclic equilibrium can emerge, in which the factions begin the game united, then split in the first period, only to re-merge in the second. Importantly, this cycle does not emerge because factions' statically optimal strategy changes over time.

This is evident from the fact that, as depicted in Figure 2 below, for some values of r this cyclic equilibrium is sustainable when $\sigma^a < 1$, i.e., when the split always hurts the faction's static payoff. Put differently, the emergence of the cycle equilibrium is not driven by the faction's short-term gains or static considerations. Rather, it reflects a dynamic strategic behavior where faction a expects that splitting in the first period will pave the way for a future, more advantageous re-merger in the second period. As a result, faction a may find it beneficial to split, even if this entails a static loss in the short run. Of course, for a to prefer a merger in the second period, σ^a cannot be too large. At the same time, σ^a cannot be too small, to ensure that the static cost of the split in the first period is not too high.

Further, notice that for the cycle to be sustained, faction b must be able to credibly commit to merging in the second period. Corollary 1 follows straightforwardly:

Corollary 1. *There exists a unique* $\widetilde{\sigma}^b$ *s.t.* $\underline{\sigma}^a < \overline{\sigma}^a$ *only if* $\sigma^b < \widetilde{\sigma}^b$*.*

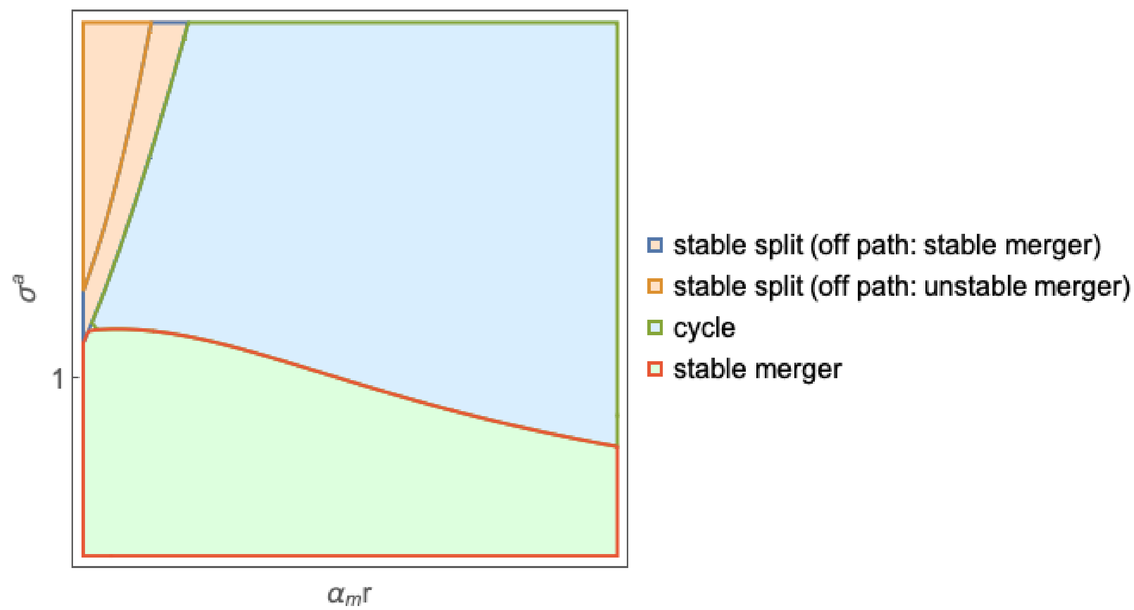

Figure 2 – Equilibria illustration for $\sigma^a \in [0.5, 2]$, $\alpha_m r \in [0.2, 2.2]$. The orange region correspond to the stable split equilibrium, the blue region to the cycling equilibrium, and the green region to the stable merger equilibrium. The other parameters are set to $\sigma^b = 0.5$ and $\phi = 0.13$.

5.1. The Logic of Cycles and Damaging Splits

Having established our equilibrium characterization, we now delve deeper into the logic underlying the emergence of cycles of mergers and splits. To provide a clearer understanding of the dynamics that give rise to these cycles we will specifically examine the scenario where $\sigma^a < 1$, which results in a *statically damaging* split in the first period. As we will see, there are two possible types of dynamics that may drive this cycle, depending on whether $\sigma^a > \sigma^b$ or $\sigma^a < \sigma^b$.

1. $\sigma^a > \sigma^b$: bigger fish in smaller pond. First, suppose that a split damages *both* the splinter faction and the party as a whole: i.e., it depletes both factions' brands ($\sigma^a, \sigma^b < 1$). Here, when factions reunite in the second period, the party will be weaker than it would have been without the split, and faction a's own brand will be less valuable. Nonetheless, it can still be advantageous for faction a to instigate the split if it inflicts an even greater damage on the opposing faction b (i.e., $\sigma^b < \sigma^a$). In this case, even if a's *absolute* brand will have weakened, its position *relative* to the opposing faction b will have strengthened.

In other words, even if the party becomes weaker as a result of the cycle, faction a chooses to split today to enhance its standing in the party in the future. This decision is made precisely because faction a expects to merge back and reap the benefits of a weakened opposing faction: grabbing a bigger share of a smaller pie.

2. $\sigma^b > \sigma^a$: smaller fish in bigger pond. Next, suppose $\sigma^a < 1 < \sigma^b$, meaning that the split is damaging to the splinter faction's brand but improves the opposing faction's. In this case, the cost of the split is very high for faction a. Not only it imposes an immediate cost in the first period, but it also puts the faction in a weaker bargaining position within the party following the re-merge in the second period.

Although the cycling equilibrium may seem counterproductive, there exist circumstances that sustain it. Specifically, if $\sigma^a < 1 < \sigma^b$ and $\sigma^a + \sigma^b > 2$ the cycle damages the splinter but allows the other faction to consolidate its brand and ultimately helps the camp overall. Thus, when the party re-merges in the second period it is stronger than it would have been absent the split. In this scenario, faction b benefits from the cycling equilibrium since it strengthens both the camp and its relative power. Meanwhile, the splinter faction a is prepared to bear the cost of splitting today and weakening its position because it expects that doing so will strengthen the party in the future. Again, a opts to split today precisely because it anticipates re-merging tomorrow, this time gaining a smaller share of a larger pie. We emphasize that this result shows that cycling equilibria with damaging splits may arise even when the relationship between factional brands has zero-sum features, in the sense that if the split damages a faction's brand it must instead help the other's.[12]

We conclude this section with a comment on the length of the factions' time-horizon. For ease of illustration, in this paper we assume that factions only interact over two periods. One may worry that this assumption is not innocuous in our setting, and in particular that end-period effects may be essential in sustaining damaging cycles in equilibrium. Absent such effects, one may worry, the factions' 'commitment' to re-merge in the future after a split today may not be credible. Reassuringly, this is not the case. To see this, suppose that our factions interact over an infinite horizon. If $r > 1$, the efficiency premium for joining forces becomes increasingly valuable over time. This implies that, in the long run, the factions must re-merge after a split, as the efficiency premium provides a strong enough incentive.[13] Then, it is intuitive to see that conditions analogous to those described above would continue to sustain our damaging cycles, whether under a bigger-fish logic, or under a collusion dynamic.

6. Institutions and Splits

We now leverage our theoretical results from the previous sections to study how institutional features regulating competition across and within parties influence party stability and fragmentation.

[12] Note that in this baseline model we assume that only faction a can split. Straightforwardly, this implies that a cycle can never emerge when $\sigma^a < \sigma^b < 1$: in fact, in this case a split would damage the camp as a whole, hurt factions in the short run *and* damage the splinter faction's standing within the party. However, this is not the case if both factions can initiate cycles, as the same bigger-fish logic described above may induce faction b to initiate a cycle when $\sigma^a < \sigma^b < 1$. We analyze this version of the model in Section 7.1.

[13] If $r < 1$, or if r can fluctuate over time, we can obviously also sustain equilibria in which the factions remain split over time.

6.1. Electoral Institutions

First, we study the impact of electoral institutions. In our setting, α_m quantifies the efficiency premium the factions gain when running in the same party. As such, this parameter can be interpreted as representing the disproportionality of the electoral system. We find that, surprisingly, increasing α_m can incentivize splits and decrease the likelihood of a stable-merger equilibrium.

Proposition 2. *Suppose that $\sigma^a > \sigma^b$ and r is sufficiently large. Then, the stable-merger equilibrium is harder to sustain as α_m increases, and the cycle equilibrium is easier to maintain.*

As discussed above, under $\sigma^a > \sigma^b$, faction a has incentives to initiate a cycle and pay a cost today to grab a larger piece of the pie tomorrow. The larger α_m, the larger the pie, the stronger the incentives underlying this dynamics. Thus, increasing α_m will sometimes increase the parameter region sustaining a cycling equilibrium, instead eroding the stable-merger region.

This result suggests a simple yet neglected relationship between electoral institutions and intra-party incentives. Constitutional design scholars typically focus on the *static* incentives that institutions produce at the party level. A powerful intuition in this literature is that increasing the stakes of winning the election should reduce fragmentation by encouraging factions not to break ranks. In particular, Duverger's law states that we should expect a less fragmented party system under more majoritarian electoral rules (Duverger, 1951). While this intuition is upheld in our model if factions merely consider their static payoffs, Proposition 2 highlights that dynamic considerations may generate the opposite results, whereby α_m may induce a split in the first period.

Importantly, the empirical evidence on the Duvergerian proposition is nuanced. Some studies found consistent results in cross-sectional analyses (Cox, 1997; Lijphart, 1994). Others, however, identify cases where the proposition fails, such as India (Diwakar, 2007) or Canada (Gaines, 1999). Furthermore, Singer (2013) provides mixed evidence on Duverger's law from single-member district election outcomes in fifty-three countries.

Existing scholarship explains this mixed evidence by suggesting that Duverger forces may be *dampened* when voters or parties fail to act strategically (Cox, 1997), or because of societal cleavages that interact with electoral institutions (Ordeshook and Shvetsova, 1994). In contrast, Proposition 2 emphasizes that, even if a Duvergerian logic is statically upheld, the effect of disproportionality on the effective number of parties may go in the *opposite* direction once we consider factions' dynamic incentives. Thus, the Duvergerian intuition may need to be qualified, due to previously overlooked strategic considerations.

6.2. Intra-Party Institutions

Following a similar logic, the next result shows that making the party organization more egalitarian (or reducing ideological divisions in the party) will only encourage a split initiated by the faction expecting to become a smaller fish in a bigger pond.

Proposition 3. *Suppose that $\sigma^b < \widetilde{\sigma}^b$. Then,*

(i) For $\sigma^a > \sigma^b$, the stable-merger equilibrium is harder to sustain as ϕ increases, and the cycle equilibrium is easier to maintain.

(ii) For $\sigma^a < \sigma^b$, the stable-merger equilibrium is easier to sustain as ϕ increases, and the cycle equilibrium is harder to maintain.

A larger ϕ captures a less egalitarian internal organization, with more resources going to the strongest faction, and/or a party platform that reflects less the ideological preferences of smaller factions. As such, naïve intuition would suggest that increasing ϕ would consistently render a stable merger equilibrium more challenging to sustain. This is because, in a less egalitarian organization, the weaker faction's participation constraints are harder to satisfy, potentially leading to factional splits. This intuition is validated in our model in the case of $\sigma^a > \sigma^b$, but not when $\sigma^a < \sigma^b$.

Recall that, under $\sigma^a < \sigma^b$, if a cycle emerges it is sustained by a 'smaller fish in bigger pond' dynamic. The splinter faction is willing to pay an immediate cost and damage its standing within the party, in order to strengthen the party's position. This dynamic is more profitable for the splinter faction as ϕ decreases: when the party is more cohesive or has a more egalitarian structure, a's cost of damaging its relative bargaining power (i.e., reducing $\mathcal{B}_t^a - \mathcal{B}_t^b$) is reduced. Thus, as ϕ increases this type of cycle equilibrium is harder to sustain, while the stable merger is easier to maintain.

In contrast, when $\sigma^a > \sigma^b$ the incentives underlying the cycle are for a to strengthen its relative standing within the party. As ϕ increases, these incentives become stronger. The cycling equilibrium region expands, eroding both the stable split and stable merger regions.

Our result, which shows that enhancing internal power sharing within political parties may either prevent or provoke splits, aligns with the heterogeneity observed in different empirical settings. For instance, scholars specializing in Southern U.S. politics have argued that Democrats endorsed primaries to maintain their one-party dominance by averting factional defections (Key, 1949). Conversely, other researchers have posited that introducing primaries might actually exacerbate internal conflicts (Burden, 2004).

Existing explanations for both the negative and positive impacts of intraparty power sharing typically rely on factions' static ideological motives. Introducing primaries is thought to help build consensus among party members and legitimize the chosen candidate in the eyes of those not selected (Carey and Polga-Hecimovich, 2006).[14] However, defections might also be encouraged if the primary electorate selects a nominee who appeals to a smaller, more extreme group of party activists (Brady, Han and Pope, 2007). Proposition 3 complements these theories by underscoring the importance of taking into account factions' dynamic incentives to fully understand the consequences of party internal institutions, and showing that both sets of results may be rationalized within the same theoretical framework.

7. Robustness and Extensions

In this section we discuss the results' robustness to relaxing some of our assumptions from the baseline model.

7.1. Both Factions Can Split

The baseline model assumes for clarity of exposition that only faction a can initiate a split. We now examine the robustness of our results to the possibility of both factions initiating a split. Recall that, in the original model, faction a can unilaterally decide to split, but both factions must agree for a re-merger to occur. Consequently, it is clear that the parameter values that generate a

[14]For additional evidence indicating that primaries unify parties, see Hortala-Vallve and Mueller (2015) and Ascencio (2023).

stable split or a cycle in the original model also support these equilibria in the expanded version. However one may wonder whether the stable merger equilibrium can be sustained in this enriched setup, or whether it is always the case that, if one faction does not want to initiate a split, then the other does.

In the Appendix, we analyze this modified version of the model and demonstrate that, although the stable-merger region shrinks, there exist parameter values sustaining a stable merger equilibrium.

Proposition 4. *There exist unique* $\underline{\sigma}^a$ *and* $\underline{\sigma}^b$ *s.t. a stable merger equilibrium exists iff* $\sigma^a < \underline{\sigma}^a$ *and* $\sigma^b < \underline{\sigma}^b$.

Intuitively, for the equilibrium to persist when b can also split, it is crucial that both σ^a and σ^b are small enough such that neither faction has an incentive to initiate a split and break away from the party.

7.2. Bargaining Among Factions

Thus far, we have assumed that the distribution of spoils within the political party is determined by its internal organization and institutions (represented by the parameter ϕ), which remain constant over time. However, this raises the question of whether allowing factions to bargain over the internal division of the spoils could prevent the costs associated with a cycle that involves a damaging split in the first period.

Intuitively, there are circumstances where allowing for bargaining does not alter factions' strategic problem. To see why, consider a cycle sustained by a 'smaller fish in a bigger pond' dynamic. In this case, cycles are welfare-improving for both factions, who effectively collude to split in the first period and re-merge in the second. Because neither faction has any reason to avoid a first-period split, allowing for bargaining has no impact on the results. However, this is not necessarily true in the case of a 'bigger fish in a smaller pond' cycle, where one faction initiates a split that harms both its opponent and the party as a whole. In this case, the faction whose brand is most damaged by the first-period split has a clear incentive to avoid the efficiency losses associated with a cycle.

To verify whether a 'bigger fish in a smaller pond' cycle is robust to factional bargaining, we explore an extension of our baseline model where the division of the pie within the party is determined endogenously in each period. More precisely, the factions start the game together as in the baseline model. In each period, one faction is selected to make a proposal. If this proposal is accepted by the other faction, the party remains united/re-merges and the proposed division is implemented. If not, the factions split/remain split. The likelihood of a faction being selected as the proposer in each period is a function of the relative strength of its brand at the beginning of that period. We note that the equilibrium of this game is equivalent to the one that would emerge if the internal division of the pie were determined by Nash Bargaining, under appropriately chosen weight parameters.

In this section, it is useful to generalize the assumption on how the strength of the ideological camp evolves over time. In particular, we assume that, in $t = 2$, the strength of the camp is given by $\alpha_m r^\gamma$. As in the baseline, in $t = 1$ the strength of the camp is given by $\alpha_m r$. The parameter $\gamma > 0$ then captures the speed at which the ideological camp strengthens, or weakens, over time (in the baseline, $\gamma = 2$). We maintain the assumption that $r > 1$, so that a higher γ implies that the camp gets stronger in the second period.

Then, we have:

Proposition 5. *There exists a $\underline{\gamma}$ such that, if $\gamma > \underline{\gamma}$, a cycle emerges in equilibrium.*

It is intuitive that, for a sufficiently large γ, factions will be incentivized to merge (or remain merged) in the second period, in order to enjoy the efficiency premium. Therefore, their second-period payoff depends on the outcome of the internal bargaining process which will determine the division of the pie. Intuitively, the more a faction consolidates its brand in the first period, the stronger its bargaining position in the second, and the higher its equilibrium share of the pie. Consider then the factions' first-period incentives. Suppose that $1 > \sigma^a > \sigma^b$, with a symmetric logic applying to $\sigma^a < \sigma^b < 1$. A split in the first-period imposes a large cost on faction b, both today and tomorrow. Thus, we can always find parameter values such that faction b is willing to strike a bargain in the first period: offer part (or even all) of its first-period share to faction a in order to avoid a split. However, we find that faction a is often unwilling to take the deal. Even if b were to offer the entire pie in the first period, this may not be enough to dissuade a from initiating a damaging cycle.

In particular, when γ is sufficiently high, bargaining in the first period fails and cycles become inevitable.[15] Faction a anticipates that improving its relative brand will improve its bargaining position in the second period, and thus its equilibrium share of the pie. If γ is very high, the second-period pie is far more valuable than the first period's. As a consequence, faction a faces dynamic incentives very similar to those highlighted in the baseline model, and it is willing to pay a cost today to get a larger share of the pie tomorrow.

The party thus falls victim to a dynamic resource course: The resources available today are insufficient to compensate the splinter faction for the higher gains it expects in the future. This highlights a fundamental issue with the allocation of resources within the party: while future party resources are high, factions cannot make binding agreements today on how to split the spoils tomorrow. In fact, in the context of a cycle, any promise made by faction b to offer a larger share of the pie (than it is statically optimal) to faction a in the second period is not credible. This is because faction a is willing to re-merge even without this offer, which means that faction b has no incentives to follow through on the promise. This commitment problem leads to an inefficiency, as the party's overall resources are depleted over time.[16]

7.3. Other Considerations

We now consider a few additional potential extensions of the model and discuss how they might affect our qualitative insights.

First, we could allow σ^a and σ^b to be a function of r^t, the ideological leaning of the electorate. Intuitively, one may expect the success of the two factions after a split to be related to the ideological strength of the camp as whole. Importantly, this need not fundamentally change our qualitative insights. The model focuses on the net effect of the split, taking into account the ideological leaning of the electorate and all other relevant factors. As long as these factors are captured by the parameter σ, the precise functional form of σ is less important.

[15] Another possible reason for bargaining failure, which we do not consider here, is information asymmetry. If factions have private information about the possible consequences of a split, the familiar logic from the conflict bargaining literature may lead to an inefficient outcome (see e.g., Fearon 1995).

[16] The result is also reminiscent of Powell (2004), who shows how large, rapid changes in the bargainers' relative power cause inefficiency, even with complete information.

Second, one might think that the effect on a faction's brand generated by a split is not long-lasting if factions re-merge in the same party again. Suppose for example, that the splinter faction presents itself as appealing to a more extreme and 'ideologically pure' portion of the electorate. This allows the faction to cultivate a certain identity, or brand. It seems plausible, then, that some of this identity may be lost should the faction re-merge again with its moderate counterpart. To capture this intuition, we could allow re-merging to dampen some of the effect of splits on the factions' brands. This could be done by introducing a parameter that captures the degree to which a re-merged faction retains the same brand accumulated during the split. Again, this extension would not change our qualitative insights as long as the effect of re-merging is not so strong as to completely erase the gains or losses incurred from the split.

Finally, as briefly discussed above, we could allow brands to evolve somewhat during merger periods. That is, factions could still develop their own ideological identity, or organizational capacity, thus differentiating from each other while in the same party. This is especially true in parties that allow for internal competition, for example through primaries. Similarly to the previous case, we could introduce a parameter that captures the extent to which factions' brands evolve when factions are within the party. As long as the evolution is muted compared to after a split, our main results would still hold.

8. Conclusion

Most party systems frequently witness significant political changes, with splits and mergers of political parties taking center stage. This has led to a growing interest among scholars and political observers in understanding the complex dynamics of party politics and factionalism. This paper develops a theory to explain why factions belonging to the same party might choose to split, and when instead we should expect party unity.

Our model produces some intuitive results. On the one hand, when a faction benefits a lot from splitting (e.g., because the new faction leadership gains visibility or voters perceive the new party as ideologically pure), in equilibrium party unity is not sustainable, and different factions diverge on their own separate paths. On the other hand, when the benefit of running together is very high (perhaps because voters tend to punish a divided camp), factions do not split in equilibrium, thus providing a possible explanation for party system stability.

Other results generated by our model are more surprising. We show that factions may split from their party even when, by doing so, they anticipate damaging both themselves and their ideological camp. These damaging splits can only be sustained in equilibrium as part of a 'cycle', whereby factions split today only to re-merge tomorrow. One dynamic that might sustain such a cycle is when a split damages both factions, but the splinter faction anticipates that by temporarily distancing itself it will improve its relative standing within the party, thus becoming a bigger fish in a smaller pond. A second, perhaps less intuitive dynamic occurs when a split harms the splinter faction but benefits the remaining one. The splinter faction pays a cost today to strengthen the party tomorrow, thus becoming a smaller fish in a bigger pond.

These different dynamics generate rich comparative statics, whereby forces that would intuitively push towards party stability in a static setting may actually incentivize splits in equilibrium in our model. For instance, we find that more majoritarian electoral systems may induce factions to split today as part of a cycle equilibrium. In a similar fashion, enhancing power-sharing within the party (for example, through the adoption of primaries) does not necessarily foster party unity. In light of mixed empirical evidence on the effect of institutions on party stability, our findings propose an explanation which focuses on factions' dynamic incentives.

Our model predominantly concentrates on factions within parties, but its insights can be extrapolated to pre-electoral coalitions. Each party within a coalition possesses a distinct brand, which may evolve differently depending on its coalition status. Analogous to our model, the decision to exit a coalition can either fortify or diminish a party's brand. Our model's focus on dynamic incentives faced by coalition partners then offers valuable insights for understanding the frequent formation and dissolution of alliances in multi-party systems.

While our results would persist in a world with limited uncertainty about the consequences of splits, a natural question to ask is how facing substantive uncertainty would change factions' incentives to split. We speculate that the equilibria we uncover persist, but a richer set of incentives may emerge in a high-uncertain world. For example, in such a setting splits might happen for experimentation: while factions have a perception of what could happen if they split when they belong to the same party, it is only when the split actually occurs that these perceptions become unequivocal signals and can thus impact their bargaining power. We believe that analyzing factions' incentives to experiment in this more complex setup is perhaps the most promising avenue for future theoretical research building on our model.

References

Aldrich, John H. 1995. *Why parties?: The Origin and Transformation of Political Parties in America*. University of Chicago Press.

Ascencio, Sergio J. 2023. "Retaining Political Talent: A Candidate-Centered Theory of Primary Adoption." *American Journal of Political Science* .

Blais, André and Indridi H Indridason. 2007. "Making Candidates Count: the Logic of Electoral Alliances in Two-Round Legislative Elections." *The Journal of Politics* 69(1):193–205.

Brady, David W, Hahrie Han and Jeremy C Pope. 2007. "Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?" *Legislative Studies Quarterly* 32(1):79–105.

Buisseret, Peter and Richard Van Weelden. 2020. "Crashing the Party? Elites, Outsiders, and Elections." *American Journal of Political Science* 64(2):356–370.

Burden, Barry C. 2004. "Candidate Positioning in US Congressional Elections." *British Journal of Political Science* 34(2):211–227.

Caillaud, Bernard and Jean Tirole. 2002. "Parties as Political Intermediaries." *The Quarterly Journal of Economics* 117(4):1453–1489.

Carey, John M and John Polga-Hecimovich. 2006. "Primary Elections and Candidate Strength in Latin America." *The Journal of Politics* 68(3):530–543.

Clarke, Andrew J. 2020. "Party sub-brands and American party factions." *American Journal of Political Science* 64(3):452–470.

Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge University Press.

Crutzen, Benoit, Micael Castanheira and Nicolas Sahuguet. 2010. "Party Organization and Electoral Competition." *The Journal of Law, Economics, & Organization* 26(2):212–242.

Desposato, Scott W. 2006. "Parties for rent? Ambition, Ideology, and Party Switching in Brazil's Chamber of Deputies." *American Journal of Political Science* 50(1):62–80.

Diwakar, Rekha. 2007. "Duverger's law and the size of the Indian party system." *Party Politics* 13(5):539–561.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York; Harper and Row.

Duell, Dominik, Lea Kaftan, Sven-Oliver Proksch, Jonathan Slapin and Christopher Wratil. 2023. "Communicating the Rift: Voter Perceptions of Intraparty Dissent in Parliaments." *The Journal of Politics* 85(1):76–91.

Duverger, Maurice. 1951. *Les Partis Politiques.* Armand Colin.

Fearon, James D. 1995. "Rationalist explanations for war." *International organization* 49(3):379–414.

Forand, Jean Guillaume and Vikram Maheshri. 2015. "A dynamic Duverger's Law." *Public Choice* 165:285–306.

Gaines, Brian J. 1999. "Duverger's Law and the Meaning of Canadian Exceptionalism." *Comparative Political Studies* 32(7):835–861.

Golder, Sona Nadenichek. 2006*a*. *The Logic of Pre-Electoral Coalition Formation.* Ohio State University Press.

Golder, Sona Nadenichek. 2006*b*. "Pre-Electoral Coalition Formation in Parliamentary Democracies." *British journal of political science* 36(2):193–212.

Hirano, Shigeo, James M Snyder Jr and Michael M Ting. 2009. "Distributive politics with primaries." *The Journal of Politics* 71(4):1467–1480.

Hortala-Vallve, Rafael and Hannes Mueller. 2015. "Primaries: the Unifying Force." *Public Choice* 163:289–305.

Ibenskas, Raimondas. 2016. "Marriages of Convenience: Explaining Party Mergers in Europe." *The Journal of Politics* 78(2):343–356.

Invernizzi, Giovanna M. 2022. "Antagonistic Cooperation: Factional Competition in the Shadow of Elections." *American Journal of Political Science* .

Invernizzi, Giovanna Maria. 2023. "Why do Parties Merge? Electoral Volatility and Long-Term Coalitions." *Journal of Politics* .

Invernizzi, Giovanna Maria and Carlo Prato. 2024. "Bending the Iron Law." *Forthcoming, American Journal of Political Science* .

Izzo, Federica. 2023. "With Friends Like These, Who Needs Enemies?" *Forthcoming, Journal of Politics* .

Key, V. O. 1949. *Southern politics.* New York: Knopf.

Kselman, Daniel M, Eleanor Neff Powell and Joshua A Tucker. 2016. "Crowded Space, Fertile Ground: Party Entry and the Effective Number of Parties." *Political Science Research and Methods* 4(2):317–342.

Levy, Gilat. 2004. "A Model of Political Parties." *Journal of Economic theory* 115(2):250–277.

Lijphart, Arend. 1994. *Electoral Systems and Party Systems: a study of twenty-seven democracies, 1945-1990.* Oxford University Press.

Lo, James, Sven-Oliver Proksch and Jonathan B Slapin. 2016. "Ideological Clarity in Multiparty Competition: A New Measure and Test Using Election Manifestos." *British Journal of Political Science* 46(3):591–610.

Mershon, Carol and Olga Shvetsova. 2013*a*. "The Microfoundations of Party System Stability in Legislatures." *The Journal of Politics* 75(4):865–878.

Mershon, Carol and Olga Shvetsova. 2013*b*. *Party System Change in Legislatures Worldwide: Moving Outside the Electoral Arena.* Cambridge University Press.

Ordeshook, Peter C and Olga V Shvetsova. 1994. "Ethnic Heterogeneity, District Magnitude, and the Number of Parties." *American journal of political science* pp. 100–123.

Pedersen, Mogens N. 1979. “The Dynamics of European Party Systems: Changing Patterns of Electoral Volatility.” *European Journal of Political Research* 7(1):1–26.

Persico, Nicola, José CR Pueblita and Dan Silverman. 2011. “Factions and Political Competition.” *Journal of Political Economy* 119(2):242–288.

Powell, Robert. 2004. “The Inefficient Use of Power: Costly Conflict with Complete Information.” *American Political Science Review* 98(2):231–241.

Rokkan, Stein and Seymour Martin Lipset. 1967. “Cleavage Structures, Party Systems and Voter Alignments: An Introduction.” *Party Systems and Voter a—Slignments* pp. 1–63.

Singer, Matthew M. 2013. “Was Duverger correct? Single-member district election outcomes in fifty-three countries.” *British Journal of Political Science* 43(1):201–220.

Snyder, James M and Michael M Ting. 2002. “An Informational Rationale for Political Parties.” *American Journal of Political Science* pp. 90–110.

Taagepera, Rein and Bernard Grofman. 2003. “Mapping the Indices of Seats–Votes Disproportionality and Inter-Election Volatility.” *Party Politics* 9(6):659–677.

Appendix

Table of Contents

Appendix A: Proofs of Main Results

Proof of Proposition 1. We proceed by backward induction, analyzing factions' behavior in the second period.

Second Period. Suppose there is a split in $t = 1$. In $t = 2$, factions re-merge in equilibrium if and only if both factions prefer running in the same party to staying split, i.e.,:

$$(\sigma^a)^2 < \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right], \tag{A-1}$$

and

$$(\sigma^b)^2 < \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^b - \sigma^a)\right]. \tag{A-2}$$

Notice that, under $\phi < \frac{1}{2\bar{\sigma}^2} < 1$, these conditions establish upper bounds on σ^a and σ^b, respectively.

Suppose instead there is no split in $t = 1$. Factions remain merged in $t = 2$ if and only if a prefers staying merged to splitting:[17]

$$\sigma^a < 1 + \frac{1}{2}\left(\alpha_m r^2\right). \tag{A-3}$$

First period. Moving to the first period, we must consider four cases, depending on the behavior anticipated in the second.

Case 1: $(\sigma^a)^2 < \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right]$**,** $(\sigma^b)^2 < \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^b - \sigma^a)\right]$ **and** $\sigma^a < 1 + \frac{1}{2}(\alpha_m r^2)$.

In this case, factions always merge/remain merged in the second period. Thus, at $t = 1$, faction a compares the payoff from a stable merger to the payoff from a split-merger cycle.

[17]Recall that in this baseline model faction b is not allowed to initiate a split. We relax this assumption in Proposition 4.

On the equilibrium path, we have a stable merger if and only if

$$2+\frac{\alpha_m r(1+r)}{2}>\sigma^a+\left(\sigma^a+\sigma^b+\alpha_m r^2\right)\left[\frac{1}{2}+\phi(\sigma^a-\sigma^b)\right], \tag{A-4}$$

which rearranged yields

$$2+\frac{\alpha_m r(1+r)}{2}-\sigma^a-\left(\sigma^a+\sigma^b+\alpha_m r^2\right)\left[\frac{1}{2}+\phi(\sigma^a-\sigma^b)\right]>0. \tag{A-5}$$

Differentiating the LHS with respect to σ^a we obtain

$$-\frac{3}{2}-2\phi\sigma^a-\phi\alpha_m r^2<0.$$

Condition (A-5) is always satisfied at $\sigma^a=0$. Therefore, there must exist a value of σ^a below which (A-5) holds. Solving for σ^a yields the following upper bound on σ^a:

$$\tilde{\sigma}^a=\frac{2+\frac{\alpha_m r(1+r)}{2}-\frac{1}{2}\sigma^b-\frac{1}{2}\alpha_m r^2+\phi(\sigma^b)^2+\phi\alpha_m r^2\sigma^b}{\frac{3}{2}+\phi\sigma^a+\phi\alpha_m r^2}. \tag{A-6}$$

Thus, if

$$\sigma^a<\frac{2+\frac{\alpha_m r(1+r)}{2}-\frac{1}{2}\sigma^b-\frac{1}{2}\alpha_m r^2+\phi(\sigma^b)^2+\phi\alpha_m r^2\sigma^b}{\frac{3}{2}+\phi\sigma^a+\phi\alpha_m r^2}, \tag{A-7}$$

then we have a stable merger equilibrium.

Otherwise, if

$$\sigma^a>\frac{2+\frac{\alpha_m r(1+r)}{2}-\frac{1}{2}\sigma^b-\frac{1}{2}\alpha_m r^2+\phi(\sigma^b)^2+\phi\alpha_m r^2\sigma^b}{\frac{3}{2}+\phi\sigma^a+\phi\alpha_m r^2}, \tag{A-8}$$

then we have a split-merger cycle.

Case 2: $(\sigma^a)^2>\left(\sigma^a+\sigma^b+\alpha_m r^2\right)\left[\frac{1}{2}+\phi(\sigma^a-\sigma^b)\right]$ **and/or** $(\sigma^b)^2>\left(\sigma^a+\sigma^b+\alpha_m r^2\right)\left[\frac{1}{2}+\phi(\sigma^b-\sigma^a)\right]$ **and** $\sigma^a>1+\frac{1}{2}(\alpha_m r^2)$.

In this case, factions always split/remain split in the second period. Thus, at $t = 1$, faction a compares the payoff from a stable split to the payoff from a merger-split. The payoff from a merger-split is higher if and only if

$$1 + \frac{\alpha_m r}{2} + \sigma^a > \sigma^a + (\sigma^a)^2. \tag{A-9}$$

However, given $r > 1$, this contradicts $\sigma^a > 1 + \frac{1}{2}(\alpha_m r^2)$. Thus, under the conditions in Case 2, we always have a stable split in equilibrium.

Case 3: $(\sigma^a)^2 > \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right]$ **and/or** $(\sigma^b)^2 > \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^b - \sigma^a)\right]$ **and** $\sigma^a < 1 + \frac{1}{2}(\alpha_m r^2)$.

In this case, factions remain split at $t = 2$ if there is a split at $t = 1$, but remain merged at $t = 2$ if there is a merger at $t = 1$. For faction a, the payoff from a stable merger is higher than the payoff from a stable split if and only if

$$2 + \frac{\alpha_m r(1+r)}{2} > \sigma^a + (\sigma^a)^2, \tag{A-10}$$

which establishes the following upper bound on σ^a:

$$\sigma^a = \frac{\sqrt{9 + 2\alpha_m r(1+r)} - 1}{2}. \tag{A-11}$$

Thus, if

$$\sigma^a < \frac{\sqrt{9 + 2\alpha_m r(1+r)} - 1}{2}. \tag{A-12}$$

then we have a stable merger in equilibrium. Otherwise, if

$$\sigma^a > \frac{\sqrt{9 + 2\alpha_m r(1+r)} - 1}{2}. \tag{A-13}$$

then we have a stable split.

Case 4: $(\sigma^a)^2 < \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right]$ **and** $(\sigma^b)^2 < \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^b - \sigma^a)\right]$ **and** $\sigma^a > 1 + \frac{1}{2}(\alpha_m r^2)$.

In this case, factions split after a merger in $t = 1$, but re-merge after a spilt at $t = 1$. At $t = 1$, faction a's payoff from splitting is lower than the payoff from staying merged if and only if

$$\sigma^a + \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right] - 1 - \frac{(\alpha_m r)}{2} - \sigma^a < 0. \tag{A-14}$$

Recall that, in Case 4, we have $(\sigma^a)^2 < \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right]$. Thus, condition (A-14) requires that $(\sigma^a)^2 < 1 + \frac{\alpha_m r}{2}$. However, this contradicts $\sigma^a > 1 + \frac{1}{2}(\alpha_m r^2)$. Thus, under the conditions in Case 4, we always have a split-merger cycle in equilibrium.

To pull these cases together, first suppose that $(\sigma^b)^2 > \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^b - \sigma^a)\right]$. Then, depending on the values of the other parameters, we must be in either Case 2 or Case 3. It follows immediately that there exists a threshold in σ^a, function of the other model primitives, such that in equilibrium we have a stable split if σ^a is above this threshold, and a stable merger otherwise.

Suppose instead that $(\sigma^b)^2 < \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^b - \sigma^a)\right]$, and denote by $\hat{\sigma}^a$ the value of σ^a that solves

$$(\sigma^a)^2 = \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right].$$

The value of $\hat{\sigma}^a$ can be simplified as follows:

$$\hat{\sigma}^a = \frac{\left(\frac{1}{2} + \phi\alpha_m r^2\right) + \sqrt{\left(\frac{1}{2} + \phi\alpha_m r^2\right)^2 + 4(1-\phi)\left(\frac{1}{2}\sigma^b + \frac{1}{2}\alpha_m r^2 - \phi(\sigma^b)^2 - \phi\alpha_m r^2\sigma^b\right)}}{2(1-\phi)}.$$

Recall our assumption that $\phi \in [0, \frac{1}{2\bar{\sigma}^2}]$. Then, it is easy to show that $\hat{\sigma}^a > 1 + \frac{1}{2}(\alpha_m r^2)$. To see why, notice that, for $\phi < \frac{1}{2\bar{\sigma}^2} < \frac{1}{2}$:

$$\frac{\frac{1}{2} + \phi\alpha_m r^2}{2(1-\phi)} > 1 + \frac{1}{2}\left(\alpha_m r^2\right),$$

therefore

$$\hat{\sigma}^a > \frac{\frac{1}{2} + \phi\alpha_m r^2}{2(1-\phi)} > 1 + \frac{1}{2}(\alpha_m r^2). \tag{A-15}$$

This implies that, if $(\sigma^b)^2 < \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^b - \sigma^a)\right]$, then we must be either in Case 1, 2 or 4, depending on the value of σ^a.

Then, the on-the-path behavior is as follows:

- If $\sigma^a > \hat{\sigma}^a$, then the factions split at $t = 1$ and remain split at $t = 2$;
- If $\sigma^a \in \left[\min\left\{1 + \frac{1}{2}(\alpha_m r^2); \tilde{\sigma}^a\right\}, \hat{\sigma}^a\right]$, then the factions split at $t = 1$ and re-merge at $t = 2$;
- If $\sigma^a \in \left[\min\left\{1 + \frac{1}{2}(\alpha_m r^2); \tilde{\sigma}^a\right\}\right]$, then the factions remain merged in both periods. □

Proof of Corollary 1. Follows from the conditions identified above. □

Proof of Proposition 2. Suppose that r is sufficiently large that (A-1), (A-2) and (A-3) hold, so that we are in Case 1 from the proof of Proposition 1. Then, in equilibrium we have a stable merger if

$$\sigma^a + \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right] - 2 - \frac{1}{2}\left(\alpha_m r + \alpha_m r^2\right) < 0, \tag{A-16}$$

and a cycle if

$$\sigma^a + \left(\sigma^a + \sigma^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right] - 2 - \frac{1}{2}\left(\alpha_m r + \alpha_m r^2\right) > 0. \tag{A-17}$$

Differentiating the LHS with respect to α_m we get

$$r\phi(\sigma^a - \sigma^b) - \frac{1}{2}, \tag{A-18}$$

which is positive for $\sigma^a > \sigma^b$ and a sufficiently large r. □

Proof of Proposition 3. Follows from inspection of the existence conditions from the proof of Proposition 1, noting that, if $\sigma^a > \sigma^b$, then $\hat{\sigma}^a$ is increasing in ϕ and $\tilde{\sigma}^a$ is decreasing in ϕ. Otherwise, if $\sigma^a < \sigma^b$, then $\hat{\sigma}^a$ is decreasing in ϕ and $\tilde{\sigma}^a$ is increasing in ϕ. □

Proof of Proposition 4. Intuitively, if we consider parameter values for which we would have a stable split in equilibrium in the baseline, we continue to have a stable split in this case. Similarly, because a re-merger always requires b's consent, under the conditions sustaining a cycle in the baseline we continue to have a cycle here. Thus, here we will focus on the case in which $\sigma^a < \underline{\sigma}^a$, i.e., in the baseline we have a stable merger in equilibrium.

Here, we will show that while allowing b to split erodes the stable-merger region, this kind of equilibrium continues to arise for some parameter values.

First, notice that the conditions identified in Proposition 1 remain necessary for the existence of a stable merger equilibrium, ensuring that faction a has no profitable deviation. Consider instead faction b's behavior.

By backward induction, consider faction b's choice to stay merged in $t = 2$, if factions were merged in $t = 1$. This choice is optimal if and only if

$$1 + \frac{(\alpha_m r^2)}{2} > \sigma^b. \tag{A-19}$$

Moving to the first period, we need to consider the possible off-path conditions. First, suppose that the equilibrium prescribes factions to remain split after a split in $t = 1$. That is, either a wants to remain split in $t = 2$:

$$(\sigma^a)^2 > \left(\sigma^a + \sigma^b + \alpha_m r^2\right) \left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right], \tag{A-20}$$

or a wants to merge but b does not: i.e., Equation A-20 does not hold and

$$(\sigma^b)^2 > \left(\sigma^a + \sigma^b + \alpha_m r^2\right) \left[\frac{1}{2} + \phi(\sigma^b - \sigma^a)\right]. \tag{A-21}$$

In this case, in the first period, faction b does not split if and only if:

$$\frac{1}{2}(2 + \alpha_m r) + \frac{1}{2}(2 + \alpha_m r^2) > \sigma^b + (\sigma^b)^2, \tag{A-22}$$

which establishes an upper bound on σ^b.

Suppose instead that factions re-merge after a deviation, which requires:

$$\left(\sigma^a + \sigma^b + \alpha_m r^2\right) \left[\frac{1}{2} + \phi(\sigma^a - \sigma^b)\right] > (\sigma^a)^2, \tag{A-23}$$

and

$$\left(\sigma^a + \sigma^b) + \alpha_m r^2\right) \left[\frac{1}{2} + \phi(\sigma^b - \sigma^a)\right] > (\sigma^b)^2. \tag{A-24}$$

In this case, in $t = 1$, faction b does not split if and only if:

$$\frac{1}{2}(2 + \alpha_m r) + \frac{1}{2}(2 + \alpha_m r^2) > \sigma^b + \left(\sigma^a + \sigma^b) + \alpha_m r^2\right) \left[\frac{1}{2} - \phi(\sigma^a - \sigma^b)\right]. \tag{A-25}$$

Which again establishes an upper bound on σ^b. Thus, depending on the parameter values either (A-19), (A-22) or (A-25) will be binding, and there exists a unique $\underline{\sigma}^b$ s.t. a stable merger equilibrium exists if and only if $\sigma^b < \underline{\sigma}^b$ and $\sigma^a < \underline{\sigma}^a$ (where $\underline{\sigma}^a$ is as characterized in the proof of Proposition 1). □

Proof of Proposition 5. Proceeding by backward induction, we characterize the equilibrium of the second-period subgame. Let $\beta^i_t(\mathcal{B}^i_{t-1}, \mathcal{B}^j_{t-1})$ be the probability that i is selected as the proposer in period t, as a function of the factions' brands at the beginning of the period (i.e., the brands inherited from $t-1$). Further, denote as x_j the share of the pie which the proposer allocates to factions j, and $1 - x_j$ the share allocated to faction i.

Suppose i gets recognized as the proposer in period 2. Then, assuming an interior solution, if i proposes a merger/staying merged, it will propose an x_j s.t. j is indifferent between being together or apart. This solves

$$(\mathcal{B}^i_1 + \mathcal{B}^j_1 + \alpha_m r^\gamma) x_j = \mathcal{B}^j_1 \sigma^j, \tag{A-26}$$

which yields

$$x_j^* = \frac{\mathcal{B}_1^j \sigma^j}{\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma}. \tag{A-27}$$

Notice that the solution is always strictly larger than 0 and, for a sufficiently high γ, smaller than 1.

Given (A-27), i chooses to propose a merger/staying merged rather than splitting/remaining split if and only if

$$\Big(1 - \frac{\mathcal{B}_1^j \sigma^j}{\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma}\Big)(\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma) > \mathcal{B}_1^i \sigma^i. \tag{A-28}$$

Which is again always satisfied for a sufficiently high γ.

A similar analysis establishes that if faction j is selected in period 2 and proposes a merger, the proposed allocation will be

$$1 - x_j = \frac{\mathcal{B}_1^i \sigma^i}{\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma}. \tag{A-29}$$

Then, j prefers to propose a merger/staying merged rather than than splitting/remaining split if and only if

$$\Big(1 - \frac{\mathcal{B}_1^i \sigma^i}{\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma}\Big)(\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma) > \mathcal{B}_1^j \sigma^j. \tag{A-30}$$

Suppose then that γ is sufficiently large that (A-27) and (A-29) are interior, and (A-28) and (A-30) are satisfied.

Moving backwards, a sufficient condition to ensure that a split emerges in equilibrium in period 1, regardless of which faction is selected as the proposer, is that one faction always prefers to split even if the other is willing to offer the entire first-period pie. Consider faction i. Denote by $\Gamma(\mathcal{B}_1^i, \mathcal{B}_1^j)$ the expected second period payoff of faction i, given the brands inherited from the first period. In equilibrium:

$$\begin{aligned}\Gamma(\mathcal{B}_1^i, \mathcal{B}_1^j) = \beta_2^i(\mathcal{B}_1^i, \mathcal{B}_1^j)\Big((\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma)(1 - \frac{\mathcal{B}_1^j \sigma^j}{\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma})\Big) \\ + (1 - \beta_2^i(\mathcal{B}_1^i, \mathcal{B}_1^j))\Big((\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma)\frac{\mathcal{B}_1^i \sigma^i}{\mathcal{B}_1^i + \mathcal{B}_1^j + \alpha_m r^\gamma}\Big),\end{aligned} \tag{A-31}$$

where the values of $\mathcal{B}_1^i$ and $\mathcal{B}_1^j$ depend on the factions' decision to split or merge in the first period.

Then, sufficient condition to ensure that i prefers a split in the first period is

$$\mathcal{B}_0\sigma^i + \Gamma(\mathcal{B}_0\sigma^i, \mathcal{B}_0\sigma^j) > 2\mathcal{B}_0 + \alpha_m r + \Gamma(\mathcal{B}_0, \mathcal{B}_0), \tag{A-32}$$

where the RHS is i's dynamic payoff from remaining merged in the first period and obtaining the entirety of the pie in that period.

Substituting in the value of $\Gamma(\mathcal{B}_1^i, \mathcal{B}_1^j)$, notice that we can rewrite $\Gamma(\mathcal{B}_0\sigma^i, \mathcal{B}_0\sigma^j) - \Gamma(\mathcal{B}_0, \mathcal{B}_0)$ as $\Big(\beta_2^i(\mathcal{B}_0\sigma^i, \mathcal{B}_0\sigma^j) - \beta_2^i(\mathcal{B}_0, \mathcal{B}_0)\Big)\alpha_m r^\gamma + \Delta$, where Δ is not a function of γ. Then, A-32 reduces to

$$\begin{aligned}\Big(\beta_2^i(\mathcal{B}_0\sigma^i, \mathcal{B}_0\sigma^j) - \beta_2^i(\mathcal{B}_0, \mathcal{B}_0)\Big)\alpha_m r^\gamma + \Delta > \\ 2\mathcal{B}_0 + \alpha_m r - \mathcal{B}_0\sigma^i.\end{aligned} \tag{A-33}$$

Suppose $\sigma^i > \sigma^j$. Then, $\beta_2^i(\mathcal{B}_0\sigma^i, \mathcal{B}_0\sigma^j) - \beta_2^i(\mathcal{B}_0, \mathcal{B}_0) > 0$, by assumption, and the condition is always satisfied for a sufficiently high γ.

Proceeding in the same way we can show that, for a sufficiently high γ, j always prefers a split in the first-period if $\sigma^j > \sigma^i$.

Appendix B: Time-dependent Shifts in Factional Brands

In the baseline model the evolution of factions' brands is dictated by the value of σ^i: if $\sigma^i > 1$ then a split helps faction i to build its own brand, while if $\sigma^i < 1$ then a split damages the faction's individual brand. In what follows we relax the simplifying assumption that σ^i is constant over time.

Formally, σ_1^i denote the shift in faction i's brand following a split when factions start-off together. The shift σ_2^i in turn characterizes faction i's brand evolution when factions are already split. In other words, the subscript refers to the number of periods in which factions are split (not to the time period in which the split occurs). We allow for $\sigma_1^i \neq \sigma_2^i$ for both factions, and we assume $\sigma_t^i \in [0, \overline{\sigma}]$ for $i = a, b$ and $t = 1, 2$.

To exclude mechanical cycles that arise from static incentives, we impose the following:

Assumption B-1. *For $i = a, b$:*
If $\sigma_1^i > 1$, then $\sigma_2^i > 1$.
If $\sigma_1^i < 1$, then $\sigma_2^i < 1$.

We proceed as for the proof of Proposition 1.

Second period. Beginning by backward induction, consider faction's behavior in the second period. Suppose there is a split in $t = 1$. In $t = 2$, factions re-merge in equilibrium if and only if both factions prefer running in the same party to staying split, i.e.,:

$$\sigma_2^a \sigma_1^a < \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right) \left[\frac{1}{2} + \phi(\sigma_1^a - \sigma_1^b)\right], \tag{B-1}$$

and

$$\sigma_2^b \sigma_1^b < \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right) \left[\frac{1}{2} + \phi(\sigma_1^b - \sigma_1^a)\right]. \tag{B-2}$$

Thus, factions re-merge in $t = 2$ if and only of both σ_2^a and σ_2^b are below the following thresholds:

$$\sigma_2^a < \frac{\left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)}{\sigma_1^a} \left[\frac{1}{2} + \phi(\sigma_1^a - \sigma_1^b)\right],$$

and

$$\sigma_2^b < \frac{\left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)}{\sigma_1^b} \left[\frac{1}{2} + \phi(\sigma_1^b - \sigma_1^a)\right].$$

Suppose instead there is no split in $t = 1$. Factions remain merged in $t = 2$ if and only if a prefers staying merged to splitting:

$$\sigma_1^a < 1 + \frac{1}{2}\left(\alpha_m r^2\right). \tag{B-3}$$

First period. Moving to the first period, we must consider four cases, depending on the behavior anticipated in the second.

Case 1: $\sigma_2^a \sigma_1^a < \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right) \left[\frac{1}{2} + \phi(\sigma_1^a - \sigma_1^b)\right]$**,** $\sigma_2^b \sigma_1^b < \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right) \left[\frac{1}{2} + \phi(\sigma_1^b - \sigma_1^a)\right]$ **and** $\sigma_1^a < 1 + \frac{1}{2}(\alpha_m r^2)$.

In this case, factions always merge/remain merged in the second period. Thus, at $t = 1$, faction a compares the payoff from a stable merger to the payoff from a split-merger cycle.

On the equilibrium path, we have a stable merger if and only if

$$2 + \frac{\alpha_m r(1+r)}{2} > \sigma_1^a + \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma_1^a - \sigma_1^b)\right], \tag{B-4}$$

Solving the quadratic inequality for σ_1^a, we obtain that if

$$\sigma_1^a < \frac{\left(1 + \frac{1}{2}\alpha_m r^2\right) + \sqrt{\left(1 + \frac{1}{2}\alpha_m r^2\right)^2 + 4(1-\phi)\left(2 + \frac{\alpha_m r(1+r) - \alpha_m r^2}{2} - \frac{\sigma_1^b}{2} - \phi((\sigma_1^b)^2 - \alpha_m r^2 \sigma_1^b)\right)}}{2(1-\phi)}, \tag{B-5}$$

then we have a stable merger equilibrium.

Otherwise, if

$$\sigma_1^a > \frac{\left(1 + \frac{1}{2}\alpha_m r^2\right) + \sqrt{\left(1 + \frac{1}{2}\alpha_m r^2\right)^2 + 4(1-\phi)\left(2 + \frac{\alpha_m r(1+r) - \alpha_m r^2}{2} - \frac{\sigma_1^b}{2} - \phi((\sigma_1^b)^2 - \alpha_m r^2 \sigma_1^b)\right)}}{2(1-\phi)}, \tag{B-6}$$

then we have a split-merger cycle.

Case 2: $\sigma_2^a \sigma_1^a > \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma_1^a - \sigma_1^b)\right]$**, and/or** $\sigma_2^b \sigma_1^b > \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma_1^b - \sigma_1^a)\right]$ **and** $\sigma_1^a > 1 + \frac{1}{2}(\alpha_m r^2)$.

In this case, factions always split/remain split in the second period. Thus, at $t = 1$, faction a compares the payoff from a stable split to the payoff from a merger-split. The payoff from a merger-split is higher if and only if

$$1 + \frac{\alpha_m r}{2} + \sigma_1^a > \sigma_1^a(1 + \sigma_2^a). \tag{B-7}$$

However, given $r > 1$ and Assumption B-1, this contradicts $\sigma_1^a > 1 + \frac{1}{2}(\alpha_m r^2)$. Thus, under the conditions in Case 2, we always have a stable split in equilibrium.

Case 3: $\sigma_2^a \sigma_1^a > \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma_1^a - \sigma_1^b)\right]$ **and/or** $\sigma_2^b \sigma_1^b > \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma_1^b - \sigma_1^a)\right]$ **and** $\sigma_1^a < 1 + \frac{1}{2}(\alpha_m r^2)$.

In this case, factions remain split at $t = 2$ if there is a split at $t = 1$, but remain merged at $t = 2$ if there is a merger at $t = 1$. For faction a, the payoff from a stable merger is higher than the payoff from a stable split if and only if

$$2 + \frac{\alpha_m r(1 + r)}{2} > \sigma_1^a(1 + \sigma_2^a), \tag{B-8}$$

Rearranging, if

$$\sigma_1^a < \frac{4 + \alpha_m r(1 + r)}{1 + \sigma_2^a}. \tag{B-9}$$

then we have a stable merger in equilibrium. Otherwise, if

$$\sigma_1^a > \frac{4 + \alpha_m r(1 + r)}{1 + \sigma_2^a}. \tag{B-10}$$

then we have a stable split.

Case 4: $\sigma_2^a \sigma_1^a < \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma_1^a - \sigma_1^b)\right]$ **and** $\sigma_2^b \sigma_1^b < \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma_1^b - \sigma_1^a)\right]$ **and** $\sigma_1^a > 1 + \frac{1}{2}(\alpha_m r^2)$.

In this case, factions split after a merger in $t = 1$, but re-merge after a split at $t = 1$. At $t = 1$, faction a's payoff from splitting is lower than the payoff from staying merged if and only if

$$\sigma_1^a + \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma_1^a - \sigma_1^b)\right] - 1 - \frac{(\alpha_m r)}{2} - \sigma_1^a < 0. \tag{B-11}$$

Recall that, in Case 4, we have $\sigma_2^a \sigma_1^a < \left(\sigma_1^a + \sigma_1^b + \alpha_m r^2\right)\left[\frac{1}{2} + \phi(\sigma_1^a - \sigma_1^b)\right]$. Thus, condition (B-11) requires that $\sigma_2^a \sigma_1^a < 1 + \frac{\alpha_m r}{2}$. However, under Assumption B-1, this contradicts $\sigma_1^a > 1 + \frac{1}{2}(\alpha_m r^2)$. Thus, under the conditions in Case 4, we always have a split-merger cycle in equilibrium.

To pull these cases together, first, suppose that σ_2^a and σ_2^b are low enough so that both factions want to re-merge following a split, that is conditions (B-1) and (B-2) hold (so we are in either in Case 1 or Case 4). Then, if σ_1^a is sufficiently high (i.e., above the minimum between $1 + \frac{1}{2}(\alpha_m r^2)$ and the threshold defined in (B-5)) we have a split-merger cycle in equilibrium. If instead σ_1^a is sufficiently low (i.e., below the minimum between $1 + \frac{1}{2}(\alpha_m r^2)$ and the threshold defined in (B-5)) we have a stable merger equilibrium.

Second, suppose that σ_2^a and/or σ_2^b are high enough so that factions always remain split in $t = 2$, i.e., condition (B-1) and/or (B-2) doesn't hold (so we are in either in Case 2 or Case 3). Then, if σ_1^a is sufficiently high (i.e., above the minimum between $1 + \frac{1}{2}(\alpha_m r^2)$ and the threshold defined in (B-9)) we have a stable split equilibrium. If instead σ_1^a is below the minimum between $1 + \frac{1}{2}(\alpha_m r^2)$ and the threshold defined in (B-5), we have a stable merger equilibrium. □

Appendix C: Micro-Foundation of Factions' Payoffs

In this section, we present a possible electoral micro-foundation for the factions' payoffs, and for how these are affected by a split through its impact on their relative brand.

As in the baseline model, each faction's payoff is given by

$$u_t^a = \begin{cases} \left(\mathcal{S}_t^a(\mathcal{B}_t^a) + \mathcal{S}_t^b(\mathcal{B}_t^b) + \alpha_m r^t\right) \left[\frac{1}{2} + \phi(\mathcal{S}_t^a(\mathcal{B}_t^a) - \mathcal{S}_t^b(\mathcal{B}_t^b))\right] & \text{if in the same party in } t \\ \\ \mathcal{S}_t^a(\mathcal{B}_t^a) + \alpha_s r^t & \text{if running alone in } t \end{cases} \tag{8.1}$$

In this micro-founded model, $\mathcal{S}_t^i$ is the number of votes attracted by faction i at time t. Thus, when factions are together in the same party, the total number of votes for the party is given by the votes attracted by the two factions, plus the votes of the ideological voters that are committed to supporting a left-wing party. Each faction then obtains a share of the spoils proportional to the number of votes it brings to the party.

Here, we micro-found how the factions' relative brands determine the values of $\mathcal{S}_t^a$ and $\mathcal{S}_t^b$. We consider a mass of voters of size 2. We model their choice as sequential. First, each voter v compares the two factions in the left-wing camp, a and b, given their brands $\mathcal{B}_t^a$ and $\mathcal{B}_t^b$, and a valence shock. Then, they decide whether to vote for their preferred left-wing faction, abstain, or vote for a right-wing party. Formally, in the first stage voter v prefers faction a iff

$$\mathcal{B}_t^a + \varepsilon_t^v > \mathcal{B}_t^b, \tag{8.2}$$

where ε_t^v is drawn from a uniform distribution on $[-\frac{1}{2\psi}, \frac{1}{2\psi}]$.

This gives us a pool of *potential* voters for faction a of size $1 + 2\psi(\mathcal{B}_t^a - \mathcal{B}_t^b)$.

In the next stage of the decision process, each *potential* voter of faction a decides whether to vote for a, abstain, or vote for a right-wing party. Specifically, v votes for a iff

$$\mathcal{B}_t^a > K_v \tag{8.3}$$

where $K_v = \max\{c_v, \widetilde{B}_v\}$, where c_v represents the cost of voting for voter v, and $\widetilde{B}_v$ is the brand of the most attractive right-wing faction or party for the voter.[18] For simplicity, let $K_v \sim U[0, 1]$.

This then leaves us with the following number of votes for faction a:

$$\mathcal{S}_t^a = \mathcal{B}_t^a \Big(1 + 2\psi(\mathcal{B}_t^a - \mathcal{B}_t^b)\Big). \tag{8.4}$$

We then use (8.4) to map the parameters in the baseline model to the results of this micro-foundation.

As in the baseline model, let $\mathcal{B}_0^a = \mathcal{B}_0^{\mathcal{B}} = 1$. This gives us that, *absent a split* in the first period

$$\mathcal{S}_1^a = \mathcal{S}_1^b = \mathcal{B}_0^a = \mathcal{B}_0^{\mathcal{B}} = 1. \tag{8.5}$$

Suppose instead there is a split in the first period. Assume the split influences each faction i's brand, which evolves by a factor $\chi_1^i \in [0, \bar{\chi}]$.[19] Then, faction a's votes are

$$\mathcal{S}_1^a = \chi_1^a \mathcal{B}_0^a \Big(1 + 2\psi \mathcal{B}_0^a(\chi_1^a - \chi_1^b)\Big). \tag{8.6}$$

[18] In the background, the voter compares the various right-wing factions or parties in a way analogous to the way she compares the left-wing ones, as described above.

[19] To ensure vote shares are always interior, we assume $\psi < \frac{1}{2\bar{\chi}^2}$.

Comparing (8.6) to (8.4), we obtain that the effect of a first-period split on faction i's votes, i.e., the parameter σ_1^a in the baseline model,[20] solves

$$\sigma_1^a \mathcal{B}_0^a\Big(1 + 2\psi(\mathcal{B}_0^a - \mathcal{B}_0^{\mathcal{B}})\Big) = \chi_1^a \mathcal{B}_0^a\Big(1 + 2\psi\mathcal{B}_0^a(\chi_1^a - \chi_1^b)\Big). \tag{8.7}$$

Plugging in $\mathcal{B}_0^a = \mathcal{B}_0^{\mathcal{B}} = 1$, this yields

$$\sigma_1^a = \chi_1^a\Big(1 + 2\psi(\chi_1^a - \chi_1^b)\Big). \tag{8.8}$$

We can similarly derive

$$\sigma_1^b = \chi_1^b\Big(1 - 2\psi(\chi_1^a - \chi_1^b)\Big). \tag{8.9}$$

Recall that $\psi < \frac{1}{2\bar{\chi}^2}$, therefore both $(1-2\psi(\chi_1^a-\chi_1^b))$ and $(1-2\psi(\chi_1^a-\chi_1^b))$ are between 0 and 1. Thus, for appropriate values of ψ, χ_1^a and χ_1^b, we can sustain all four possible configurations of parameters from the baseline model, i.e., σ_1^a and σ_1^b both above or below 1, or one above and one below.

[20] Notice that in this micro-foundation we allow the effect of a split to change across periods, hence the subscript t. This is in line with our generalization of the baseline model analyzed in Appendix B.