On the Workings of Tribal Politics*

Assaf Patir Bnaya Dreyfuss Moses Shayo

January 26, 2021

Abstract

This paper tries to understand the workings of economies in which an (endogenous) subset of voters support certain ("tribal") candidates regardless of their policies; and politicians choose whether or not to run on a tribal ticket. Two political regimes emerge. Non-tribal politics is characterized by centrist policies. Tribal politics produces extreme policies, typically from the right, despite the fact that the tribal base is from the lower middle class. Allowing policy in one period to determine the income distribution in the next, the economy either converges to a steady state or cycles between tribal and non-tribal regimes, depending on the vote share of the minority group, the scope for redistributive policy, and the salience of inter-group disparities.

*Patir: assaf.patir@mail.huji.ac.il ; Dreyfuss: bdreyfuss@g.harvard.edu; Shayo: mshayo@huji.ac.il

1 Introduction

> *What Brexiteers, Catalan separatists, Russian nationalists and Islamic fundamentalists all have in common is that their politics are all about identity. India's Modi and Israel's Netanyahu have both found political profit in pitting one local identity against another. And what is the massive backlash against immigration if not the assertion of one identity over another?*
>
> –Andrés Velasco (2020)

The resurgence of tribal politics in recent years appears to shape a wide range of policies, from immigration and international trade, to income redistribution, the rule of law, and even the responses to a pandemic and to climate change. But how does tribal politics work? Our main departure from conventional political economy models is the existence of a set of voters whose overriding concern is "who is with us and who is against us" and who support candidates representing their ethnic, religious, or national group whatever the policy they promote. Building on insights from identity economics, we argue that this set is endogenous. Depending on the state of the economy, two classes of political regimes can emerge: tribal and non-tribal. Non-tribal regimes are characterized by centrist policies, catering to the median voter in society as a whole, whereas tribal regimes are typically characterized by more extreme policies. However, despite the fact that the tribal vote base tends to come from the middle and lower ranks of the socioeconomic distribution, policy in tribal regimes is generically (though not always) from the right, catering to the better-off segments of the majority group, and effectively ignoring the distribution of preferences among the minority group.

The conditions for the emergence of tribal politics include three sets of factors. First, "economic" factors that affect the distribution of income or socioeconomic status, and in particular the concentration of voters in the lower-middle parts of the distribution of the majority group. Second, "cultural" factors, specifically the salience of inter-group cleavages in society. Third, institutional and demographic factors, specifically the share of the minority group in the population that is eligible to vote. These forces interact with each other in concrete ways. An economic shock has different implications depending on the salience of, say, inter-ethnic differences and tensions. And a xenophobic message is more likely to have political impact when a larger share of the majority is concentrated at the lower rungs of the income distribution.

All three factors can be endogenous to policy (and possibly to political rhetoric). Analytically, the most intricate is the distribution of income-related attributes, whose dynamics we analyze at some detail. In particular, we consider a dynamic model in which voters at each period determine the distribution for the next period from a three-parameter family of distributions and subject to a resource constraint. This endogenously divides the electorate into three classes, based on the type of distributions that maximize their material welfare.

We find that the policy chosen during periods of non-tribal politics is the one that is preferred by the middle class. However, if there are sufficiently many tribal voters, the economy enters a period of tribal politics, and the policy chosen could be the one that is preferred by the rich, the middle class or even the poor for some parameter values. Over time, the system can converge to a steady state or cycle between different distributions and between periods of tribal and non-tribal politics. We systematically classify all possible asymptotic behaviors as a function of the fundamentals of the economy. We also find that for some parameter values the system exhibits hysteresis, that is, the asymptotic behavior depends on the initial conditions chosen.

The analysis of the dynamic model illustrates how tribal politics can sow the seeds of its own undoing: tribal politics typically caters to voters who prefer higher inequality, which can in turn deplete the set of tribal voters. Therefore, tribalism is more likely to end itself the more public policy is effective in changing the distribution of socioeconomic status.

The paper builds on—and helps tie together—a wealth of recent empirical work on the rise of identity politics, populism, polarization, and social identity. For reviews, see Guriev and Papaioannou (2020); Iyengar et al. (2019); Noury and Roland (2020); Shayo (2020). Recent theoretical contributions include Bernhardt et al. (2020); Besley and Persson (2019); Gennaioli and Tabellini (2019); Grossman and Helpman (2020). The existence of voters who support their party independently of issue preferences has also been extensively studied. See Green et al. 2002; Huddy et al. 2015.[1] We discuss relevant findings from these literatures in relation to the concrete assumptions and results of the model. Note that while our analysis pertains to the study of populism, our focus here is on the tribal nature of politics and on equilibrium policies, and less on the anti-elite rhetoric that is typical of many populist leaders.

2 Baseline Model

Consider an economy with a continuum of agents and two groups or *tribes* (e.g. ethnic groups). Each agent i belongs to either the majority group, M, or the minority group, m. Agent i's group is denoted $g_i \in \{M, m\}$. We abstract from the possible endogeneity of the franchise and assume that the ratio between the sizes of the minority and the majority in the voting population is $\eta < 1$. Each agent i is endowed with an attribute y_i, representing a desirable feature such as socioeconomic status, productivity, or income. Within each group $g \in \{M, m\}$, y_i is distributed according to a continuous CDF F_g with PDF f_g. The attribute y_i induces policy preferences over some policy space X. For now, assume that the policy space is continuous and unidimensional (section 4 analyzes a multi-dimensional policy space), and that i's ideal policy is a monotonic function of y_i. Without loss of generality, we can then set i's ideal policy to be equal to y_i. Finally, i's "material payoff" from policy x is $\pi_i(x) = -(x - y_i)^2$. We shall sometimes refer to x as the "economic policy": we think of it as any policy that directly affects socio-economic variables.

[1] As Huddy et al. (2015) point out, the view of partisanship as an identity rather than as an instrument can account for the stability of partisan attachments, their relative immunity to short-term economic and political fluctuations, the influence of partisanship on vote choice independently of issue preferences, and the power of partisan elites to influence rank-and-file partisan opinion, evidence of which is difficult to reconcile with the instrumental model (see Huddy et al. 2015 for references).

2.1 Tribal politics

Political candidates can come from any tribe. A candidate c commits to a policy pair $\{x_c, I_c\}$. The first argument $x_c \in X$ is directly related to voters' material payoff.[2] The second argument, $I_c \in \{0, 1\}$, represents the candidate's stance regarding their tribal affiliation. We call candidates with $I_c = 1$ "tribal candidates." We can think of I_c as being purely rhetorical: Employing "us against them" terminology and portraying oneself as the group's defender against outsiders. But it may also represent policies that, while having little direct effect on one's personal economic condition, may affect the perceived honor and recognition accorded to the group (respecting our group's traditions, honoring our flag, controlling our borders).

The key feature of tribal politics is the existence of individuals whose political behavior is primarily driven by loyalty to their group: They vote for their tribe regardless of economic policy. Some of these voters might be willing to "swallow" economic policies (or candidates) they otherwise dislike, in order to support their group against the out-group. Others might adjust their attitudes and policy preferences to match those of their tribe's candidate (on this phenomenon see Cohen 2003; Druckman et al. 2013; Petersen et al. 2013; Christenson and Kriner 2017; Bisgaard and Slothuus 2020). Either way, they vote for the tribal candidate.[3] To make this feature as stark as possible, we make the following assumption.

Assumption 1. *Agents who identify with their tribe always vote for a tribal candidate from their group whenever such a candidate exists. If there is more than one tribal candidate from their group, they randomize between them.*

Below we exploit insights from identity economics to characterize those agents who choose to identify with their tribe. For now notice that to the extent that such agents exist, this generates an incentive for politicians to play the identity card. Such a move, however, entails an important tradeoff: by adopting a tribal stance you gain the votes of your tribal supporters, but you lose those of the other group. Formally:

Assumption 2. *An agent from group g never votes for a candidate from group $g' \neq g$ if that candidate is tribal (i.e. if $c \in g'$ and $I_c = 1$).*

Assumptions 1-2 can be jointly interpreted as follows: people from g who identify with their group primarily care about I_g, and will always support a candidate from g that promotes it. On the other hand, voters from g' are harmed, alienated, or offended by I_g, and therefore never support candidates who promote it.

Agents vote for the candidate whose policy maximizes their material payoffs, subject to Assumptions 1-2, with ties broken randomly.

[2] In the dynamic analysis (section 4) we explicitly restrict the set of economic policies to those that candidates can credibly commit to. For the baseline case this does not meaningfully affect the analysis, hence we allow candidates to commit to any policy in X.

[3] This is an extreme variant of social identification as defined by Shayo (2009): When it comes to voting, the main thing that matters is that your group wins and that you behave like the prototypical member of your group.

2.2 Identity

While it is well established that humans have a strong tendency to associate themselves with groups, individuals do not identify with a group simply because they belong to it. People are more likely to identify with their group if it provides them with a sense of pride or "status", and if it is perceived as similar to them along a set of salient attributes (see Shayo 2020 for a review of the evidence). Agents in our model have two attributes: their socioeconomic status or income y_i, and their group membership g_i. Let $\hat{y}_g$ be some statistic which describes the typical or central value in group g, e.g. its mean, median or mode. We make $\hat{y}_g$ more explicit in the dynamic analysis, but for now we only assume $\hat{y}_g$ is strictly in the interior of the support of y_i in group g. We assume the status of group g to be simply $\hat{y}_g$. The perceived distance between agent i and group g is given by

$$d_{i,g} = \omega \cdot \mathbf{1}\left(g \neq g_i\right) + (1-\omega)\left(y_i - \hat{y}_g\right)^2 \tag{1}$$

where the weight parameter $\omega \in [0,1]$ captures the (potentially endogenous) salience of distinctive group attributes such as mother-tongue, religion, or skin color relative to socioeconomic differences. The value for agent i of identifying with group g is then

$$V_{ig} = \hat{y}_g - \alpha d_{i,g} \tag{2}$$

where $\alpha > 0$. In words, individuals take pride in seeing themselves as members of a high-status group, but pay a dissonance cost for identifying with a group of people that are very different from them.

We assume that each agent i either identifies with their tribe g_i, or they can simply identify as i. The latter case can be interpreted as identifying with a narrow set of agents that have the same traits as i (i.e. your immediate social group). Alternatively, it can be interpreted as having an *individualistic* identity, therefore caring only about one's personal payoffs and status. We simply refer to these agents as "non-tribal". From equations 1 and 2 we have:

$$V_{ig} = \begin{cases} \hat{y}_{g_i} - \alpha(1-\omega)\left(y_i - \hat{y}_{g_i}\right)^2 & \text{if } g = g_i \text{ (tribal)} \\ y_i & \text{if } g = i \text{ (non-tribal)} \end{cases}. \tag{3}$$

Note that identifying with oneself entails no cognitive cost ($d_{ii} = 0$), and the status of non-tribal agents is simply their individual status y_i.

Finally, we assume that agents choose the identity with a higher value. Specifically, agent i identifies with group g_i if and only if $V_{ig_i} \geq V_{ii}$. While identity choices may not be conscious and deliberate, we employ an optimization assumption to capture the idea that people are more likely to identify with those groups that give them higher status and that are more similar to them. See Atkin et al. (Forthcoming); Grossman and Helpman (2020); Shayo (2020).

For simplicity we ignore non-socioeconomic drivers of group status. It is straightforward, however, to see how adding an exogenous shifter to group status would expand or shrink the set of tribal voters. For instance, a group with a high status due to its glorious past (or some other factor not captured by its socioeconomic characteristics), might attract members with high y_i or with y_i quite different from the typical member of their group. The qualitative results, however, would be similar, except for the case where the majority group's status is so high that *everyone* in the majority is tribal. In this case, the winning policy would not be unique: any arbitrary whim of the tribal politician can be supported.

3 Static Analysis

We concentrate on the triple (c, I_c, x_c) which is a Condorcet winner, if one exists. That is, we are looking for elements in the core of the simple majority voting rule. Specifically, we look for the combination of candidate c, tribal stance I_c, and policy x_c, that would receive at least half the votes in a two-candidate election against *any* other combination $(c', I'_{c'}, x'_{c'})$. This allows us to abstract from the specifics of the political process within and across parties, and focus on characterizing the political regime that is likely to emerge given the state of the economy.

Denote the set of tribal voters in group g by Γ_g. Furthermore let $\tilde{\omega} \equiv \frac{1}{\alpha(1-\omega)}$ be the "effective salience" of the tribal identity: it is increasing in the salience of the tribal attributes and decreasing in the dissonance cost α of identifying with dissimilar others. Theorem 1 characterizes the conditions for tribal vs. non-tribal politics, and the nature of policy in the two regimes.

Theorem 1.

1. *If* $F_M(\hat{y}_M) - F_M(\hat{y}_M - \tilde{\omega}) \geq \eta$ *then a Condorcet winner* $(c^*, I^*_{c^*}, x^*_{c^*})$ *exists. Furthermore,*
   - (a) $c^* \in M$ *(the winner is from the majority group);*
   - (b) $I^*_{c^*} = 1$ *(the winner is a tribal candidate);*
   - (c) $x^*_{c^*} = med\{y_i | i \in M \setminus \Gamma_M\}$ *(the winning policy is the ideal policy of the median non-tribal member of the majority group).*
2. *If* $F_M(\hat{y}_M) - F_M(\hat{y}_M - \tilde{\omega}) < \eta$ *then a Condorcet winner may or may not exist, depending on the distributions. For any Condorcet winner* $(c^*, I^*_{c^*}, x^*_{c^*})$*:*
   - (a) $I^*_{c^*} = 0$ *(any Condorcet winner is non-tribal);*
   - (b) $x^*_{c^*} = med\{y_i\}$ *(the winning policy is the ideal policy of the median voter in the population).*

Below we discuss the intuition for these results, and show how they can help explain established empirical patterns. All proofs are in the appendix. The proof also provides the conditions on F_M, F_m and η such that a Condorcet winner exists in part 2 of the theorem. Appendix B also shows that for the cases in which a Condorcet winner does not exist, the non-tribal candidate proposing the median voter's preferred policy is the unique winner of a simple two-stage game.

3.1 The conditions for tribal politics

Start with the set of tribal voters. Agents with high socioeconomic status ($y_i > \hat{y}_g$) have little to gain from adopting a tribal identity: it does not enhance their status and entails some dissonance cost. At the same time, people very distant from their group's prototype—possibly including the "very poor"—may not identify with it either. This can happen if the effective salience $\tilde{\omega}$ of the tribal identity is low, since in this case the status gain may not offset the added dissonance cost. Thus, the set Γ_g of tribal voters in group g is its *"lower middle class":* a set of people situated just below g's typical member. This is illustrated in Figure 1 by the shaded areas in the distributions. Notice that according to this analysis, the tribal voters are not necessarily the poorest of the poor. Consistent with our model, Gidron and Hall (2020) report that the relationship between individual economic conditions and voting for the radical right is not monotonic: People facing the most difficult economic circumstance are most likely to vote for the radical left (consistent with their ideal economic policy), but support for the radical right is strongest among those whose economic situation is slightly better.

Now, tribal politicians can ensure winning only if Γ_g is large enough to overcome the loss of voters from the other group. This requires that the majority's share of the voting population is large enough. Therefore, the condition for the emergence of tribal politics is:

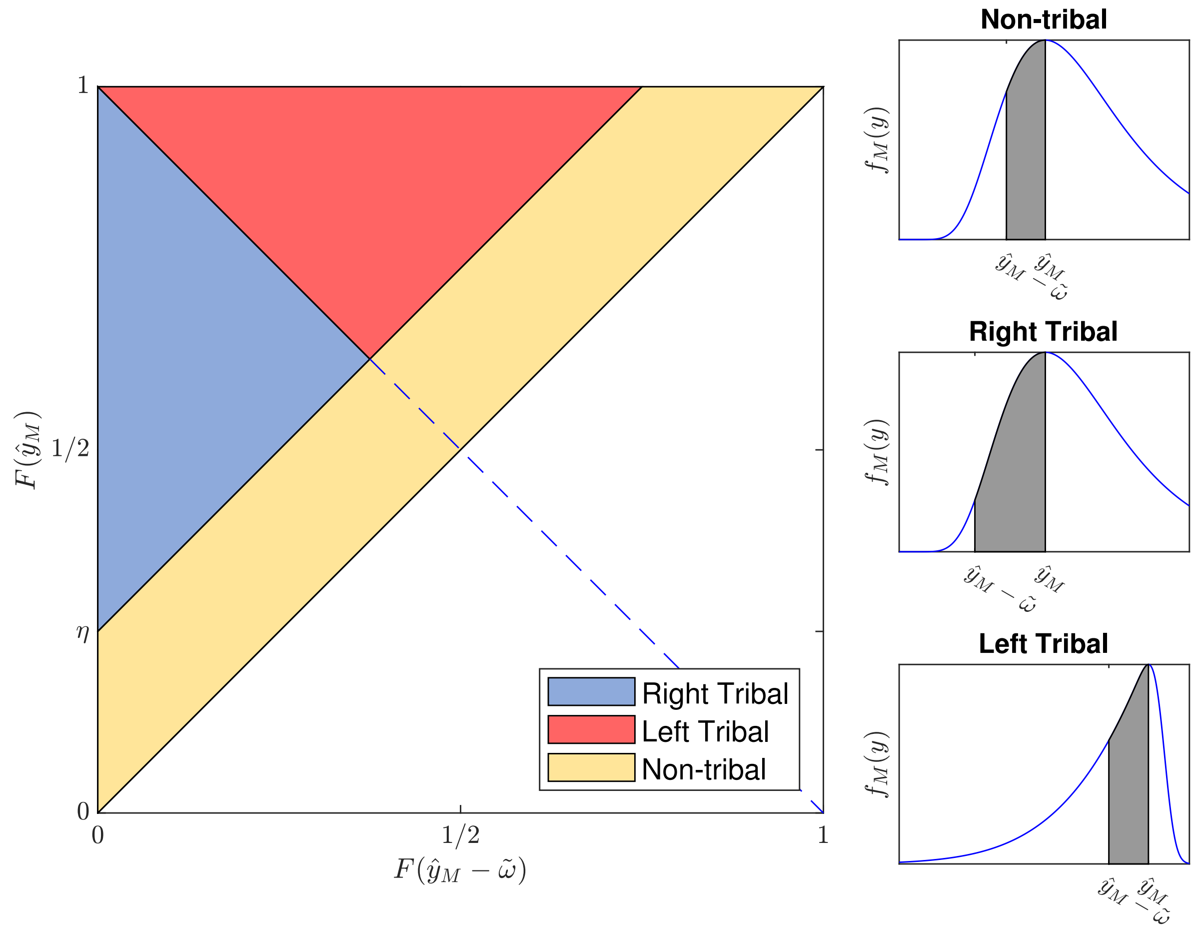
Condition 1. $$\underbrace{F_M(\hat{y}_M) - F_M(\hat{y}_M - \tilde{\omega})}_{\substack{\text{Share of tribal voters} \\ \text{within majority}}} \geq \underbrace{\eta}_{\substack{\text{Ratio of minority} \\ \text{to majority voters}}}$$

The left-hand side of Condition 1 highlights the interaction between two sets of drivers of tribal politics, sometimes referred to in the literature as "economic" and "cultural" (Noury and Roland, 2020). The first includes changes in F_M, the socioeconomic distribution of the majority group. The second includes changes in $\tilde{\omega}$, the effective salience of the tribal identity.

Consider first shocks to F_M, that push formerly middle-class occupations into the lower parts of the distribution. For example, Autor and Dorn (2013) and Goos et al. (2014) argue that falling costs of offshoring and automating routine tasks have led, in both Europe and the US, to relative employment declines in occupations at the middle of the skill distribution, and to slower wage growth in these occupations. The expansion of international trade has also contributed to the decline in manufacturing employment and to lower incomes for affected workers (Autor et al., 2013; Pierce and Schott, 2016; Caliendo et al., 2019). In turn, these shocks are associated with increased support for nationalistic or nativist parties and politicians (e.g., Colantone and Stanig, 2018; Autor et al., forthcoming; Dippel et al., forthcoming. See Guriev and Papaioannou, 2020 for an extensive review). Notice however that according to Condition 1, tribal politics is not driven by economic hardship per se, but rather by the *relative* position of members of the majority group, and could thus show up even if the overall economy is doing well.[4]

[4]Autor et al. (forthcoming) find that the rightward shifts in ideological affiliation and voting patterns in the US

Figure 1: Three Condorcet Winners



Note: Non-tribal politics can only emerge when Condition 1 is violated, i.e., $F_M(\hat{y}_M) - F_M(\hat{y}_M - \tilde{\omega}) < \eta$ (the yellow region). Otherwise, we have tribal politics. Right-wing tribal politics emerges if, in addition, $1 - F_M(\hat{y}_M) > F_M(\hat{y}_M - \tilde{\omega})$. Left-wing tribal politics emerges if, in addition to Condition 1, $F_M(\hat{y}_M - \tilde{\omega}) > 1 - F_M(\hat{y}_M)$. The three figures on the right illustrate socioeconomic distributions of the majority group that support these outcomes, with the tribal voters in the shaded areas.

Second, consider variations in $\tilde{\omega}$. These may stem from cultural and historical forces that affect present-day perceptions of inter-group differences and rivalry (e.g. Voigtländer and Voth 2012; Acharya et al. 2016), sometimes with the help of tribal politicians who stir up old hatreds or facilitate their expression (e.g. Ochsner and Roesel 2019; Cantoni et al. 2020). But salience of inter-group rivalry is also affected by contemporary events, policies, and political campaigns. Perhaps the most conspicuous shocks to $\tilde{\omega}$ relate to immigration (Halla et al., 2017; Dustmann et al., 2019; Tabellini, 2020) and to inter-group conflict (Wilkinson, 2004; Shayo and Zussman, 2011; Iyer and Shrivastava, 2018; Wasow, 2020; Atkin et al., Forthcoming).

due to import competition from China, are more concentrated among or driven by the majority, non-Hispanic whites. Indeed, in districts with a non-white majority, rising trade exposure was associated with higher odds of electing more liberal democrats. Gidron and Hall (2017) show that across developed democracies, lower levels of *subjective* social status are associated with support for right populist parties. And as Mutz (2018) argues, "status threat, not economic hardship, explains the 2016 presidential vote."

Note the interaction between cultural and economic factors in Condition 1: If $\tilde{\omega} \approx 0$ then shocks to the income distribution do not lead to tribal politics. Income shocks can facilitate tribalism especially in those places where $\tilde{\omega}$ is already high. Indeed, Doerr et al. (2018) find that economic repercussions of the German 1931 banking crisis increased Nazi support especially in cities with well-established anti-Semitism. Conversely, a local shock to ethnic salience is more consequential when $f_M(\hat{y}_M - \tilde{\omega})$ is high, that is, when a large part of the majority group is concentrated in the lower rungs of society. This might help explain the finding of Funke et al. (2016) that, between 1870 and 2014, parties on the far right have been the most successful in attracting voters following financial crises. To the extent that these crises are associated with decline in economic status among the majority middle class, they should also be associated with a larger impact of political rhetoric that attributes blame to minorities or foreigners.

A third set of factors contained in Condition 1 includes institutional and demographic determinants of the share of minority voters in the population that is eligible to vote. Tribal politics is more likely to emerge when the vote share of the minority is low, i.e., when η is small.

In the next section we take η and $\tilde{\omega}$ as exogenous, and concentrate on understanding the endogenous evolution of the socioeconomic distribution F_M, which is analytically much harder. It should be clear, however, that tribal candidates of the majority group do better if they succeed in decreasing η or increasing $\tilde{\omega}$. The former might be achieved, for example, by suppressing minority voting, by facilitating majority voting, or by denying minorities entry or citizenship. Increasing $\tilde{\omega}$ can be achieved by accentuating economic, cultural or social threats from "outsiders", and by allowing or even encouraging inter-group conflict. Fighting immigration may be particularly effective as it can *both* limit η and increase $\tilde{\omega}$. Recalling the discussion in section 2.2, if perceptions of group status can also be manipulated, then tribal candidates do better if they manage to increase the perceived status of their group.

We now turn from the conditions for tribal politics, to the nature of policy under the different regimes.

3.2 Policy

A non-tribal political regime is characterized by centrist policies, coinciding with the material interests of the median voter in society (Theorem 1, part 2). In contrast, tribal politics can produce extreme policies, either from the left or, more likely, from the right. The logic is simple but, we believe, compelling. Since tribal agents vote inelastically for their tribal candidate, a winning tribal candidate must choose policies that cater to the swing non-tribal voters. Thus, in a tribal regime, the policy serves the independent voters from the majority group. Typically (though, as we shall see, not always) this means a policy that serves the better-off segments of society. This observation can help explain why tribal politicians often seem to pursue policies that are not in fact supported by their base. For example, nationalistic parties often pursue tax cuts coupled with reductions in social programs that tend to serve much of their base.

Even when tribal candidates are not Condorcet winners, the second part of Theorem 1 implies that there may be situations in which tribal candidates can propose extremist policies and destabilize a non-tribal regime. The precise conditions are provided in the appendix. In essence what is required is that the majority group includes a large set of tribal voters situated on one side of the population median, alongside a large enough block of non-tribal voters located on the opposite side of the population median. By proposing a policy that attracts the extreme non-tribals (e.g. the right tail in case the block of non-tribals is to the right of the population median), the tribal candidate can form a non-convex coalition that defeats the non-tribal centrist candidate. This coalition, however, is not robust and can be defeated by an even more extreme tribal candidate. Eventually, the tribal candidate is so extreme that they are defeated by a non-tribal centrist candidate.

A final implication of Theorem 1 worth noting is that tribal politics renders the distribution of policy preferences *within* the minority irrelevant for policy making. Intuitively, in a tribal regime the minority voters vote for a non-tribal candidate, for a minority tribal candidate, or they do not vote at all. Thus, when forming policies, the majority tribal candidate ignores the distribution of policy preferences within the minority. And while other candidates may compete over minority voters, they do not set the policy.

3.3 Summary

Qualitatively, there are four possible outcomes: a non-tribal regime (with policy coinciding with the material interests of the median voter); a "right-tribal" regime where policy coincides with the ideal point of some individual from the right tail of the majority distribution ($y_i > \hat{y}_M$); a "left-tribal" regime with policy coinciding with the ideal point of someone from the left tail of the distribution ($y_i < \hat{y}_M - \tilde{\omega}$); and a situation without a Condorcet winner, where majorities cycle between centrist non-tribal candidates and increasingly-extreme tribal candidates. Figure 1 illustrates the first three. The left panel shows the Condorcet winner that emerges under different parameters and socioeconomic distributions. On the right we illustrate distributions of the majority group that can sustain each of the possible outcomes, with shaded areas depicting tribal voters.

Intuitively, if F_M is symmetric or right-skewed, then the bulk of the majority's non-tribal voters are from the right tail of the distribution (as illustrated in the top two distributions in Figure 1). Now, if $\tilde{\omega}$ is sufficiently large (producing a "thick" set of tribal voters), and η is sufficiently low, then we have tribal politics with policies serving voters from the right tail of the majority distribution. Note also that, unless the minority is relatively rich (i.e., so long as $\hat{y}_M$ is not too far below the population median), we can expect the pivotal voter from the majority group to be relatively well-off and, importantly, to the right of the population median. Since income distributions are typically right-skewed, we would expect this to be the generic outcome of tribal politics.

However, if F_M is sufficiently left-skewed, as illustrated in the bottom distribution, then the bulk of non-tribal voters are from the left tail of the distribution. In this case, the policies under tribal regimes would serve voters from the left-tail of the majority distribution. Thus, we could have nationalistic leaders that promote redistribution of wealth.

4 A Dynamic Model of Tribal Politics and Redistribution

While economic shocks to the income distribution of the type discussed in the previous section (due to trade and automation) are crucial for understanding tribal politics, endogenous political developments are also important. In the remainder of the paper we analyze a dynamic version of our model of tribal politics, in which the winning policy today affects the distribution in the next period.

For tractability and to make the underlying mechanisms transparent, we shall assume that status or income in both groups is identically distributed according to a triangular distribution.[5] Triangular distributions, as the name implies, are distributions for which the probability density function is triangular, and therefore are fully described in terms of the location of the minimum value of the support, a, the length of the support Δ, and the quantile of the mode p^c. Figure 3 below provides some examples.[6] The PDF is given by:

$$f(y;a,\Delta,p^c)=\frac{2}{\Delta^2}\begin{cases}\frac{y-a}{p^c} & y\in[a,a+\Delta p^c]\\ \frac{a+\Delta-y}{1-p^c} & y\in[a+\Delta p^c,a+\Delta]\\ 0 & \text{otherwise}\end{cases}.$$

Triangular distributions are useful for our purposes, since they allow for non-zero skewness and yet are simple enough so that we can obtain analytical results, and transparently explain the underlying mechanisms.

Let the prototypical member in each group be the individual at the mode, which is the individual at quantile p^c, and whose income is $\hat{y}=a+\Delta p^c$. Thus, the tribal-voters in each group are those whose income is between $\max\{a,a+\Delta p^c-\tilde{\omega}\}$ and $a+\Delta p^c$. The non-tribal voters consist of the individuals that are to the right of the mode, whose share within the tribe is $1-p^c$; and the non-tribal left whose share is:

$$S_L(\Delta,p^c)=p^c\left(\max\left\{0,1-\frac{\tilde{\omega}}{\Delta p^c}\right\}\right)^2. \tag{4}$$

The share of tribal voters within each group is $TV=p^c-S_L(\Delta,p^c)$. The mean of a triangular distribution is given by $\mu=a+\Delta(1+p^c)/3$.

Rather than defining the particulars of the redistribution mechanism, we assume that at each period the elected government can implement policies that determines the income distribution for the next period subject to the following constraints:

[5]The assumption that the majority and minority are identically distributed is needed because a redistribution scheme that only takes into account the individual's income (but not one's group) cannot, in general, keep the distributions of both tribes triangular. An alternative approach that also allows some tractability is to assume that the majority's distribution is triangular and the minority is an atom. This also allows for analysis of how inter-group inequality interacts with tribal politics. We leave this for future research.

[6]The standard parameterization of triangular distributions (see, e.g. Kotz and Dorp, 2004), in which the parameters are the values of the support, $[a,b]$, and the location of the mode, c, maps to our parameterization via $\Delta=b-a$, and $p^c=(c-a)/(b-a)$.

1. The ranking of individuals is unchanged, i.e. individuals remain at their original quantile over time.
2. The total (and mean) income remains unchanged.
3. The resulting distribution is triangular (a, Δ, p^c), with $a \geq 0$, $\Delta \geq \underline{\Delta} > 0$, $p^c \in \left[\underline{p}^c, \overline{p}^c\right]$, and $0 < \underline{p}^c < 1/2 < \overline{p}^c < 1$.

Additionally, we assume that

$$\mu > \frac{2 - \underline{p}^c}{2 - \overline{p}^c} \cdot \frac{1 + \overline{p}^c}{3} \underline{\Delta}, \tag{5}$$

which is sufficient to ensure a non-empty core in some situations as we will discuss.

The requirement that the mean and the ranking remain unchanged is sufficient to assure that such a redistribution can be achieved through a monotonically increasing tax schedule combined with a universal transfer. The restrictions on Δ and p^c should be thought of as constraints stemming from the fundamentals of the economy and its institutions, that make it infeasible to implement certain distributions such as perfect equality ($\Delta = 0$) or extremely skewed distributions (e.g. $p^c = 0$, i.e. the mode is at the bottom percentile).

4.1 The space of policy alternatives

We restrict attention to policies that are the ideal point of *some* agent in the economy. One way to think about this is that the political candidates are themselves citizens, and the only policy they can credibly commit to is their bliss point (e.g. Besley and Coate, 1997). Therefore, our first task is to find the optimal distribution that an individual at quantile p will prefer. Formally,

$$\max_{a, \Delta, p^c} F^{-1}(p; a, \Delta, p^c) \text{ s.t.} \left\{ \begin{array}{c} a + \frac{\Delta}{3}(1 + p^c) = \mu \\ \Delta \geq \underline{\Delta},\ a \geq 0 \\ p^c \in \left[\underline{p}^c, \overline{p}^c\right] \end{array} \right\}, \tag{6}$$

where $F^{-1}(p; a, \Delta, p^c)$ is the inverse CDF of a triangular distribution with parameters (a, Δ, p^c), and the first constraint says that mean (or equivalently total) income is fixed. The inverse CDF, i.e. the income of an individual at quantile p, is given by

$$F^{-1}(p; a, \Delta, p^c) = a + \Delta \begin{cases} 1 - \sqrt{(1-p)(1-p^c)} & 0 \leq p^c \leq p \leq 1 \\ \sqrt{p p^c} & 0 \leq p \leq p^c \leq 1 \end{cases}.$$

The solution to the problem is described by the following theorem:

Theorem 2. *There exist two constants $0 < p^P < 1/2 < p^R < 1$, which depend on $\left(\underline{p}^c, \overline{p}^c\right)$, and divide individuals into three classes, such that the solution to (6), i.e. the optimal distribution (a^*, Δ^*, p^{c*}) chosen by an individual at quantile p, is as follows:*

- *All individuals with $p \in (p^R, 1]$ (hence, the "rich") choose $a^* = 0$, $p^c = \underline{p}^c$, and $\Delta^* = 3\mu/(1 + \underline{p}^c)$. That is, set a and p^c to their minimal value and Δ to its maximal value.*

- *All individuals with $p \in (p^P, p^R)$ (hence, the "middle-class") choose $a^* = 0$, $p^c = \overline{p}^c$, and $\Delta^* = 3\mu/(1+\overline{p}^c)$. That is, set a to its minimal value, p^c to its maximal value, and Δ to its maximal value given $p^{c*} = \overline{p}^c$.*
- *Individuals with $p \in [0, p^P)$ (hence, the "poor") choose $p^c = \min\left\{\overline{p}^c, \max\left\{\underline{p}^c, \frac{9}{4}p\right\}\right\}$, $a^* = \mu - \underline{\Delta}(1+p^{c*})/3$, and $\Delta^* = \underline{\Delta}$. That is, set Δ to its minimal value, p^c to $\frac{9}{4}p$ or the closest allowed value, and a to its maximal value given p^{c*}.*

The (compact) set of feasible solutions implied by the constraints is plotted in Figure 2, which also shows the bliss points of the three endogenous classes. The bliss point of the rich is marked R^*, and that of the middle class is marked M^*. The bliss point of the poor always has Δ kept at its minimum, but the ideal p^c varies as a function of the individual's quantile p. To understand the intuition, note that increasing a acts like a universal transfer: it increases everyone's income by the same amount. Increasing Δ, however, increases one's income by more the higher one's quantile is. Finally, increasing p^c is more beneficial the closer one's quantile is to p^c (intuitively, for both the richest and poorest individuals, increasing p^c while holding a and Δ constant makes no difference). Therefore, there is a hierarchy in how different individuals prioritize using the resources of the economy to invest in a, p^c, and Δ. The rich benefit most from Δ, and therefore prefer increasing the gap between the richest and poorest individuals to spending any resources on either a or p^c. The middle-class prioritize p^c over Δ, and Δ over a. Finally, poor individuals do not want any resources spent on Δ, and choose to use the resources on a and p^c in some combination that depends on the individual's quantile.

In the proof of Theorem 2 we derive explicit expressions for the boundaries between classes: p^P and p^R. Since these turn out to be quite cumbersome, we leave them in the appendix. However, it can be shown that $p^P \in (4/9, 1/2)$, and $p^R \in (3/5, 7/8)$, which provides a sense of where these classes are located. This also implies that the endogenous middle class always includes the voter with median income.

We refer to the bliss point of the voter at quantile p by $(a^*, \Delta^*, p^{c*})(p)$, and also introduce the following shorthand notation:

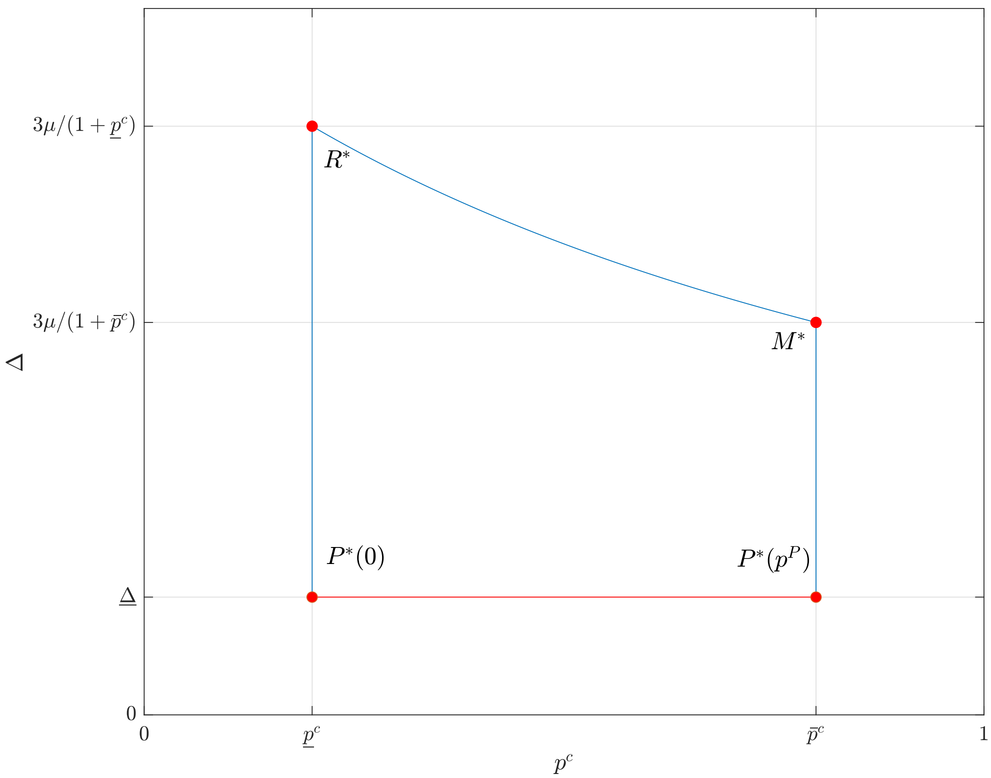
$$(a^*, \Delta^*, p^{c*})(p) = \begin{cases} P^*(p) & p \in \left[0, p^P\right) \\ M^* & p \in \left(p^P, p^R\right) \\ R^* & p \in \left(p^R, 1\right] \end{cases},$$

i.e. R^* is the preferred distribution of the rich voters, namely those with $p > p^R$; M^* is the preferred distribution of the middle class, and $P^*(p)$ is the preferred distribution of a poor voter in quantile p.[7]

Figure 3 plots the ideal probability density functions $P^*(0.2)$,$P^*(0.4)$, M^*, and R^* for some

[7]The individuals at $p = p^P$ are exactly indifferent between $P^*(4/9)$ and M^*, and, similarly, the ones at $p = p^R$ between $M^*$and R^*. However, since these are sets of zero measure, whichever methods they use to break ties is inconsequential to our analysis.

Figure 2: The Policy Space

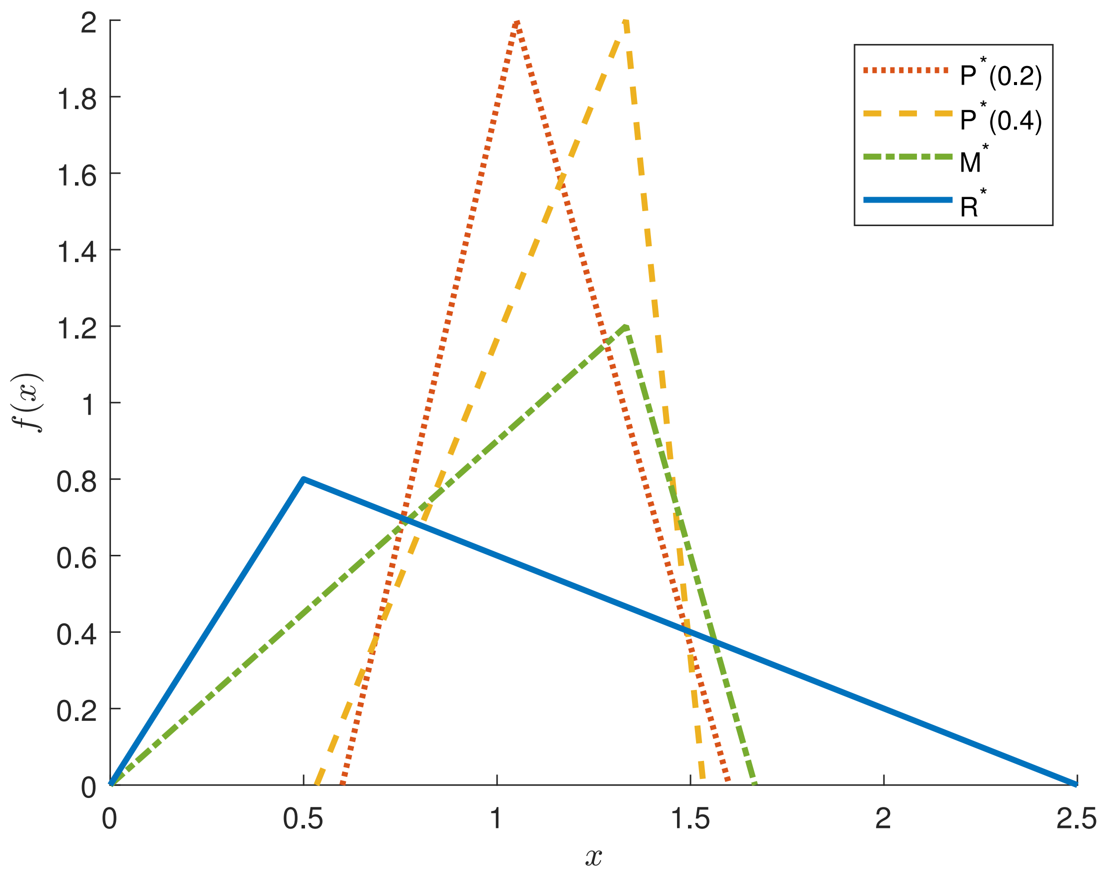


Note: The constraints in (6) generate the compact set depicted above in the (p^c, Δ) plane. Red circles mark the bliss points of the rich (R^*), and the middle class (M^*). The poor's bliss points are along the $\Delta = \underline{\Delta}$ line, and the red points mark the two extreme cases $P^*(0)$ and $P^*(p^P)$.

particular choice of parameters. The poor prefer a distribution with the lowest possible disparities between rich and poor (low Δ), but they differ in the precise shape of their ideal distribution. The poorer of the poor want a distribution with a relatively high minimal income (a), whereas those who are somewhat better-off (at the 40th percentile in this example), prefer a distribution with more people close to the top, at the cost of a somewhat lower minimal income. In contrast, both the rich and the middle-class prefer to keep the poorest individual at the lowest possible value $a = 0$. However, while the middle-class prefer a distribution with relatively small disparities and a large mass close the top of the distribution, the rich prefer a distribution that is highly skewed to the right.

It is worth noting that while we derive this form of inter-class conflict of interests from the analysis of triangular distributions, it is probably much more general. For example, think of a as a direct transfer to households, and interpret a policy that increases p^c as an improvement to a public service that individuals who are higher up the social ladder are better able to take advantage of (e.g., education). The poor don't stand to gain enough from the public service, so

Figure 3: Optimal Distributions



Note: Probability density functions for the preferred distributions of individuals at quantiles: 0.2, 0.4, 0.6 and 0.8. The parameters are $\underline{\Delta}$=1, $[\underline{p}^c, \overline{p}^c]$=$[0.2, 0.8]$, $\mu = 1$.

they prefer the transfer. The middle-class and rich stand to gain most from the program, but only the middle-class support it, because the rich expect to bear too much of the cost.

The fact that the very poor and the rich share an interest in low p^c may in general lead to an empty core under simple majority voting. It is also the case that the preferences of the rich and of some of the middle-class are not single peaked. However we do have the following two results which allow for the existence of a non-empty core under fairly broad conditions.

Lemma 1. ***Single-Peakedness of the poor's preferences.*** *Define the following order on the set of policy alternatives:*

$$\left(P^*(0), \ldots, P^*(p^P), M^*, R^*\right).$$

The preferences of any poor voter $p \in (0, p^P)$ *are single-peaked with respect to this order, that is,*

$$(p_1 < p_2 \leq p) \vee (p \leq p_2 < p_1) \Rightarrow F^{-1}(p; (a^*, \Delta^*, p^{c*})(p_2)) > F^{-1}(p; (a^*, \Delta^*, p^{c*})(p_1)).$$

Lemma 2. ***Non-poor prefer M^* to any ideal policy of the poor.*** *For all voters at quantiles $p \in (p^P, 1]$, $F^{-1}(p; M^*) > F^{-1}(p; P^*(q))$ for all $q \in [0, p^P]$.*

4.2 Dynamics with high effective salience of the tribal identity

Analyzing the dynamics of the model is a matter of determining for each of the possible distributions whether it leads to tribal politics, and what the electoral outcome is in terms of the distribution for the next period.

Recall that tribal politics emerges when the mass of majority tribal voters is larger than that of the minority. Under triangular distributions this means that $p^c - S_L(\Delta, p^c) > \eta$, where, recalling equation 4, S_L is the mass of the non-tribal left, i.e. the voters that are too poor to identify with the majority group.

During periods of *non*-tribal politics, Lemmas 1 and 2 together imply a familiar "centrist" policy. Since $p^P < 1/2 < p^R$, the median always belongs to the middle-class, and the chosen distribution for the next periods will be M^*. The reason is that middle-class voters can form a majority with the rich to vote for M^* against any P^*-type distribution, and with the poor to win against R^*.

During periods of tribal politics, however, the median non-tribal voter can in principle belong to any of the three classes. Furthermore, if this voter belongs to either the rich or the middle class, then the policy R^* or M^*, respectively, is the unique Condorcet winner. In contrast, if the median non-tribal voter during a period of tribal politics belongs to the poor class, then the fact that the preferences of some of the middle-class and rich voters are not single-peaked over P^*-type distributions may generate Condorcet cycles and an empty core.

In the remainder of this subsection, we focus on the case where the effective salience of the tribal identity, $\tilde{\omega}$, is sufficiently high such that $S_L(\Delta, p^c) < 1 - p^c$ for any distribution (a, Δ, p^c) that satisfies the constraints. That is, the non-tribal right is always larger than the non-tribal left. (We return to the issues that arise when this does not hold in the next subsection). In Table 1 we fully characterize the asymptotic dynamics of the model as a function of the parameters of the model, for any initial conditions.

Start by noticing that under the above restriction on $\tilde{\omega}$, the only possible policy outcomes are M^* and R^*. Hence we can concentrate only on these two distributions. Using Theorem 2 and equation 4, the share of tribal voters for these distributions is

$$TV(p^c) = p^c - S_L\left(\frac{3\mu}{1+p^c}, p^c\right) = p^c - p^c\left(\max\left\{0, 1 - \frac{\tilde{\omega}}{3\mu} \cdot \frac{1+p^c}{p^c}\right\}\right)^2, \tag{7}$$

which is a monotonically increasing function of p^c. Therefore, we have three possibilities:

(i) $\eta < TV(\underline{p}^c)$ where politics are tribal under both M^* and R^*;

(ii) $TV(\underline{p}^c) < \eta < TV(\overline{p}^c)$ where politics are tribal under M^*, but not under R^*;

(iii) $\eta > TV(\overline{p}^c)$ where politics are non-tribal under both M^* and R^*.

Table 1: Dynamics under large $\tilde{\omega}$

<table>
<tr><th></th><th>$\eta < TV(\underline{p}^c)$</th><th>$TV(\underline{p}^c) < \eta < TV(\overline{p}^c)$</th><th>$TV(\overline{p}^c) < \eta$</th></tr>
<tr><td>$2p^R - 1 < TV(\underline{p}^c)$</td><td>(Tribal, R^*)</td><td rowspan="2">Cycle:
(Tribal, R^*) /
(Non-tribal, M^*)</td><td rowspan="3">(Non-tribal, M^*)</td></tr>
<tr><td>$TV(\underline{p}^c) < 2p^R - 1 < TV(\overline{p}^c)$</td><td>Cycle:
(Tribal, M^*) /
(Tribal, R^*)</td></tr>
<tr><td>$TV(\overline{p}^c) < 2p^R - 1$</td><td colspan="2">(Tribal, M^*)</td></tr>
</table>

Note: Each box in the table represents the election outcome for the conditions described in the relevant column and row. A single outcome represents a steady-state, e.g. the top-left entry "(Tribal, R^*)" represents a steady-state where politics are always tribal and the winning policy is R^*. "Cycle" implies that the economy fluctuates between the two outcomes in alternating periods, and for each period we detail the political regime and the election outcome.

Next, notice that during tribal periods, the pivotal voter is at quantile $(1 + p^c - S_L)/2$. As a result, the pivotal voter will be rich if $(1 + TV(p^c))/2 > p^R$, and from the middle class otherwise. This implies three categories, depending on how p^R compares with $TV(\underline{p}^c)$ and $TV(\overline{p}^c)$. Intersecting all of these categories provides a complete characterization of the dynamics, described in Table 1.

The asymptotic dynamics are determined uniquely by η, $\tilde{\omega}/\mu$, $\underline{p}^c$, and $\overline{p}^c$ (this follows from equation 7 and the solution for p^R, provided in equation 11 in the appendix). Since we place no restrictions on the magnitude of distributional changes that can be achieved in any given period, the system converges to its asymptotic dynamics immediately after the initial period, regardless of initial conditions. These dynamics are characterized either by a fixed steady-state, or by a two-period cycle.

The main takeaway from this analysis is that tribal politics can be its own undoing: tribal politics leads to policies that are skewed toward the rich, which in turn lead to more unequal distributions. This reduces the mass of individuals that identify with their tribe. This section shows that whether or not this happens depends strongly on the boundaries, $[\underline{p}^c, \overline{p}^c]$, that is, the extent to which public policy can affect the skewness of the distribution. Thus, if the cultural, institutional, and economic fundamentals in society allow the government to quickly increase inequality ($\underline{p}^c$ is low), then periods of tribal politics are more likely to end quickly.

We can use Table 1 to analyze shocks to the relative share of minority voters (η), to the effective salience ($\tilde{\omega}$), and to the constraints constraints on the skewness of the distribution $\underline{p}^c$ and $\overline{p}^c$. In the table, η only enters in the inequalities that determine which column the economy is in, so changing η will move the economy between the columns while keeping the row fixed. As

η increases and crosses the thresholds of $TV(\underline{p}^c)$ and $TV(\overline{p}^c)$, the economy moves from tribal to non-tribal equilibria. This simply reflects the fact that tribalism is more costly the larger the minority share of the vote.

Turning to the effective salience of tribal identity, note that the expression for $TV(p^c)$ is weakly increasing in $\tilde{\omega}/\mu$, and is also supermodular as a function of $(p^c, \tilde{\omega}/\mu)$. Increasing $\tilde{\omega}/\mu$ has the effect of increasing $TV(\underline{p}^c)$, and $TV(\overline{p}^c)$, as well as the difference between them. Therefore, if we begin in one of the equilibria of Table 1, an increase in $\tilde{\omega}/\mu$ will eventually move the economy up the rows and toward the left column. This reflects the fact that, *ceteris paribus*, higher salience means more tribal voters, which makes tribal politics more likely, and also makes it more likely that the pivotal voter belongs to the rich class during tribal periods.

Finally, increasing $\underline{p}^c$ or $\overline{p}^c$ has two effects: first, it respectively increases $TV(\underline{p}^c)$ or $TV(\overline{p}^c)$, so just like $\tilde{\omega}/\mu$ it moves the economy toward the left column. However, since p^R is a (complicated) function of $\underline{p}^c$ and $\overline{p}^c$ (see 11), the effect in the direction of the rows is ambiguous. Higher p^c always means more tribal voters, which is similar to the effects of increasing $\tilde{\omega}/\mu$. However, for small $\tilde{\omega}/\mu$, when the marginal effect of p^c on $TV(p^c)$ is small, the effect on p^R dominates.

4.3 Dynamics with low effective salience

In the previous subsection we analyzed the model under the restriction that $\tilde{\omega}$ is sufficiently high, so that there are always more non-tribal voters on the right side of the distribution than on the left, i.e. $1 - p^c > S_L(\Delta, p^c)$. When this is not the case, the median non-tribal voter's quantile is $\hat{p} = (S_L(\Delta, p^c) + 1 - p^c)/2$. As long as $\hat{p} > p^P$, i.e. the median non-tribal voter does not belong to the poor class, then we still have a unique Condorcet winner which is now necessarily M^*. However, if for some distribution we have both tribal-politics

$$p^c - S_L(\Delta, p^c) > \eta, \tag{8}$$

and a poor median non-tribal voter

$$1 - p^c < S_L(\Delta, p^c) < 2p^P - (1 - p^c), \tag{9}$$

then the simple majority voting rule typically has an empty core. To understand why this happens, recall that $P^*(q)$ distributions all have $\Delta^*(q) = \underline{\Delta}$, and only differ in that higher q implies higher $p^{c*}(q)$ and lower $a^*(q)$, and both functions are piecewise linear in q. For all individuals, the marginal benefit of increasing a is constant, but the marginal benefit of increasing p^c is highest for individuals at the quantile $p = p^c$ and drops in both directions as we move away from p^c. Consequently, when contemplating two distributions $P^*(q_1)$ and $P^*(q_2)$ with $q_1 < q_2$, there is always a mass of individuals on the right-side of the distribution that prefers $P^*(q_1)$.

Therefore, when both 8 and 9 are satisfied, determining the electoral outcome typically requires making additional assumptions about the nature of the political process. Nonetheless, even without adding more structure to the voting process, we can analyze particular situations where a unique electoral outcome exists despite the lack of single-peakedness, and discuss the dynamics that can arise in the general case.

In Section 4.2 we found a unique Condorcet winner despite non-single-peaked preferences by setting $\tilde{\omega}$ to be large. Under these conditions, with few non-tribal poor voters, distributions of type P^* become irrelevant, so the fact that some voters have non-single-peaked preferences over them does not matter. Another way to obtain a unique winner is to effectively exclude the *voters* with non-single-peaked preferences. In particular, for distributions with large p^c (close to unity) the share of the non-tribal-right is small, and it is possible to construct examples where the simple majority rule leads to a unique winner.

By combining large p^c with small effective salience we can construct economies in which a distribution of the type $P^*(p)$ is a stable steady-state.[8] Such economies display a new possible asymptotic behavior, namely, convergence to a $P^*(p)$ distribution with $2/9 < p \leq 4/9$. Such steady-states cannot exist for $p < 2/9$, since the non-tribal left is no longer a majority in this case and the outcome should be M^* or R^*.

We can only enter a $P^*(p)$ from a state where the poor are a majority, which means either another P^* distribution or M^*, and it must also be the case that politics are tribal in the preceding period. Therefore, overall small effective salience leads to three additional types of possible dynamics: a steady-state $P^*(p)$, a cycle of the type $M^* \to \{P^*\} \to M^*$, and a cycle of the type $M^* \to \{P^*\} \to R^* \to M^*$. In the cases of cycles, the $\{P^*\}$ denotes one or more periods of different P^*-type distributions, and politics must be tribal in all periods except perhaps the R^* period.

Another important difference from the large $\tilde{\omega}$ analysis is that we can now have multiple steady-states and, more generally, the asymptotic dynamics may be history-dependent. [9]

5 Conclusions

A key observation of this paper is that tribal politics is closely tied to the shape of the socioeconomic distribution. This is consistent with a series of empirical studies documenting the effects of distributional changes triggered by trade and automation, on support for nativist politicians. In a dynamic setting, however, the income distribution is itself the outcome of public policies chosen in previous periods. As we have seen, allowing for a multi-dimensional policy space endogenously generates a 3-class structure, with inherent conflict of interests between different social classes. Importantly, as the income distribution changes from period to period, so does the prototypical tribe-member, and, consequentially, the set of voters who choose to identify with their tribe.

[8]For example, consider the parameter values $\left[\underline{p}^c, \overline{p}^c\right] = [0.1, 0.9]$, $\underline{\Delta} = 1$,$\mu = 2$, $\tilde{\omega} = 0.9 - 0.3\sqrt{7.9}$, and $\eta < 0.098$. The distribution $P^*(4/9)$ is a steady-state for this distribution, that is, it is the unique winner when the distribution is $P^*(4/9)$. It wins over M^* and R^* since the poor (0.45) form a majority among non-tribal voters (0.89); and it also wins over $P^*(q)$ despite the support of the rich and some of the poorest voters.

[9]For example, for the parameter values if footnote 8, if $\eta > 0.042$, then M^* is also a steady state: due to the larger Δ of M^* there are fewer tribal voters, and thus the distribution leads to stable non-tribal politics as in the leftmost column of table 1.

One consequence of this conflict of interests is that tribal-politics can, under some circumstances, be its own undoing. During periods of tribal-politics, politicians cater their policy toward non-tribal voters. In the typical cases, where the majority of non-tribal voters are on the upper side of the distribution, this increases inequality. As long as the salience of the tribal cleavage is not too high, this eventually leads to fewer voters identifying with their tribe. In the model this happens in a matter of one period, but it is easy to imagine how this would be a more gradual process if one assumes that changing the distribution is costly.

Whether or not tribalism ends itself depends on the distributions generated during tribal periods, which in turn depends on the extent to which policy can influence the distribution (specifically the parameters$\underline{p}^c$ and $\overline{p}^c$). For example, if the fundamentals of the economy or the strength of its institutions prevent a “tribal” government from implementing a policy that generates very high inequality (that is, $\underline{p}^c$ is high), then tribalism is less likely to end.

We also found that a tribal-left steady-state is in principle possible. In such economies the politicians entice the middle- and high-middle-classes by appealing to their tribal identity, and apply very egalitarian policies to appeal to the poor. However, this requires a delicate balance of the parameters. For other parameter values, a non-tribal-left majority will only lead to temporary redistribution.

References

Acharya, Avidit, Matthew Blackwell, and Maya Sen, “The political legacy of American slavery,” *The Journal of Politics*, 2016, *78* (3), 621–641.

Atkin, David, Eve Colson-Sihra, and Moses Shayo, “How Do We Choose Our Identity? A Revealed Preference Approach Using Food Consumption,” *Journal of Political Economy*, Forthcoming.

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi, “Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure,” *American Economic Review*, forthcoming.

Autor, David H and David Dorn, “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market,” *The American Economic Review*, 2013, pp. 1553–1597.

Autor, David H., David Dorn, and Gordon H. Hanson, “The China Syndrome: Local Labor Market Effects of Import Competition in the United States,” *American Economic Review*, October 2013, *103* (6), 2121–68.

Bernhardt, Dan, Peter Buisseret, and Sinem Hidir, “The Race to the Base,” *American Economic Review*, March 2020, *110* (3), 922–42.

Besley, Timothy and Stephen Coate, “An economic model of representative democracy,” *The Quarterly Journal of Economics*, 1997, *112* (1), 85–114.

_ **and Torsten Persson**, “The Rise of Identity Politics,” 2019.

Bisgaard, Martin and Rune Slothuus, “How Political Parties Shape Public Opinion in the Real World,” *American Journal of Political Science*, 2020.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "Trade and labor market dynamics: General equilibrium analysis of the china trade shock," *Econometrica*, 2019, *87* (3), 741–835.

Cantoni, Davide, Felix Hagemeister, and Mark Westcott, "Persistence and Activation of Right-Wing Political Ideology," 2020.

Christenson, Dino P and Douglas L Kriner, "Constitutional qualms or politics as usual? The factors shaping public support for unilateral action," *American Journal of Political Science*, 2017, *61* (2), 335–349.

Cohen, Geoffrey L, "Party over policy: The dominating impact of group influence on political beliefs.," *Journal of personality and social psychology*, 2003, *85* (5), 808.

Colantone, Italo and Piero Stanig, "The trade origins of economic nationalism: Import competition and voting behavior in Western Europe," *American Journal of Political Science*, 2018, *62* (4), 936–953.

Dippel, Christian, Robert Gold, Stephan Heblich, and Rodrigo Pinto, "The Effect of Trade on Workers and Voters," *The Economic Journal*, forthcoming.

Doerr, Sebastian, Stefan Gissler, José-Luis Peydró, and Hans-Joachim Voth, "From finance to fascism: The real effect of Germany's 1931 banking crisis," 2018.

Druckman, James N, Erik Peterson, and Rune Slothuus, "How elite partisan polarization affects public opinion formation," *American Political Science Review*, 2013, pp. 57–79.

Dustmann, Christian, Kristine Vasiljeva, and Anna Piil Damm, "Refugee Migration and Electoral Outcomes," *The Review of Economic Studies*, 09 2019, *86* (5), 2035–2091.

Funke, Manuel, Moritz Schularick, and Christoph Trebesch, "Going to extremes: Politics after financial crises, 1870–2014," *European Economic Review*, 2016, *88*, 227–260.

Gennaioli, Nicola and Guido Tabellini, "Identity, beliefs, and political conflict," *Available at SSRN 3300726*, 2019.

Gidron, Noam and Peter A Hall, "The politics of social status: economic and cultural roots of the populist right," *The British Journal of Sociology*, 2017, *68*, S57–S84.

_ **and Peter A. Hall**, "Populism as a Problem of Social Integration," *Comparative Political Studies*, 2020, *53* (7), 1027–1059.

Goos, Maarten, Alan Manning, and Anna Salomons, "Explaining job polarization: Routine-biased technological change and offshoring," *American economic review*, 2014, *104* (8), 2509–26.

Green, Donald P, Bradley Palmquist, and Eric Schickler, *Partisan hearts and minds: Political parties and the social identities of voters*, Yale University Press, 2002.

Grossman, Gene and Elhanan Helpman, "Identity Politics and Trade Policy," *Review of Economic Studies*, 2020.

Guriev, Sergei and Elias Papaioannou, "The political economy of populism," 2020.

Halla, Martin, Alexander F Wagner, and Josef Zweimüller, "Immigration and voting for the far right," *Journal of the European Economic Association*, 2017, *15* (6), 1341–1385.

Huddy, Leonie, Lilliana Mason, and Lene Aarøe, "Expressive partisanship: Campaign involvement, political emotion, and partisan identity," *American Political Science Review*, 2015, *109* (1), 1–17.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J Westwood, "The origins and consequences of affective polarization in the United States," *Annual Review of Political Science*, 2019, *22*, 129–146.

Iyer, Sriya and Anand Shrivastava, "Religious riots and electoral politics in India," *Journal of Development Economics*, 2018, *131*, 104–122.

Kotz, Samuel and Johan Rene Van Dorp, *Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications*, Singapore ; Hackensack, NJ: World Scientific Publishing Company, December 2004.

Mutz, Diana C, "Status threat, not economic hardship, explains the 2016 presidential vote," *Proceedings of the National Academy of Sciences*, 2018, *115* (19), E4330–E4339.

Noury, Abdul and Gerard Roland, "Identity Politics and Populism in Europe," *Annual Review of Political Science*, 2020, *23*, 421–439.

Ochsner, Christian and Felix Roesel, "Mobilizing History," 2019.

Petersen, Michael Bang, Martin Skov, Søren Serritzlew, and Thomas Ramsøy, "Motivated reasoning and political parties: Evidence for increased processing in the face of party cues," *Political Behavior*, 2013, *35* (4), 831–854.

Pierce, Justin R and Peter K Schott, "The surprisingly swift decline of US manufacturing employment," *American Economic Review*, 2016, *106* (7), 1632–62.

Shayo, Moses, "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution," *American Political Science Review*, 2009, *103* (02), 147–174.

_ , "Social identity and economic policy," *Annual Review of Economics*, 2020, *12.*

_ **and Asaf Zussman**, "Judicial ingroup bias in the shadow of terrorism," *The Quarterly Journal of Economics*, 2011, *126* (3), 1447–1484.

Tabellini, Marco, "Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration," *The Review of Economic Studies*, 2020, *87* (1), 454–486.

Velasco, Andrés, "Populism and identity politics," *LSE Public Policy Review*, 2020.

Voigtländer, Nico and Hans-Joachim Voth, "Persecution perpetuated: the medieval origins of anti-Semitic violence in Nazi Germany," *The Quarterly Journal of Economics*, 2012, *127* (3), 1339–1392.

Wasow, Omar, "Agenda Seeding: How 1960s Black Protests Moved Elites, Public Opinion and Voting," *American Political Science Review*, 2020.

Wilkinson, S.I., *Votes and Violence: Electoral Competition and Ethnic Riots in India*, Cambridge, UK: Cambridge University Press, 2004.

Appendix

A Proof of Theorem 1

From equation 3 we have the set of tribal voters in group g given by$\Gamma_g \equiv \{i|g_i = g \text{ , } y_i \in [\hat{y}_g - \tilde{\omega}, \hat{y}_g]\}$, where $\tilde{\omega} \equiv \frac{1}{\alpha(1-\omega)}$. Denote $\rho = (c, I_c, x_c)$ and similarly $\rho^* = (c^*, I^*_{c^*}, x^*_{c^*})$. Denote the share of the tribal voters within group g by $TV_g = F_g(\hat{y_g}) - F(\hat{y}_g - \tilde{\omega})$. We now prove the two cases separately.

1. Assume $TV_M \geq \eta$, and consider the triplet $\rho^* = (c \in M, 1, y_{TV})$ where $y_{TV} \equiv med\{y_i|i \in M \setminus \Gamma_M\}$. By way of contradiction assume that there exists a different triplet ρ' that wins more than half of the votes. First, Assumption 2 implies that a tribal candidate c' from the minority will get no more than $\eta/(1+\eta) < \frac{1}{2}$ of the votes against ρ^*. Next, Assumption 1 implies that a tribal candidate c' from the majority will get exactly half of the votes of the tribal voters in the majority against ρ^*. Now, if $x'_{c'} = x^*_{c^*}$ then all participating voters are indifferent between the two candidates (notice that in this case people from the minority do not vote). If $x'_{c'} \neq x^*_{c^*}$ then by a standard median voter argument c^* gets more than half the votes of the non-tribal voters from the majority (notice that since $\hat{y}_M < max\{y_i|i \in M\}$ non-tribal majority voters always exist), since preferences over policies are single-peaked. Last, a non-tribal candidate c' (from either group) will get at most all the minority group, and half the non-tribal voters from the majority, i.e., $\frac{\eta}{1+\eta} + \frac{(1-TV_M)}{2(1+\eta)}$. But $\frac{2\eta+(1-TV_M)}{2(1+\eta)} > \frac{1}{2}$ implies $TV_M < \eta$, a contradiction.

2. Assume $TV_M < \eta$, and consider the set $R := \{\rho|\rho = (c, 0, y_{50}), c \in \{M, m\}\}$ where we denote the population median by $y_{50} := med\{y_i\}$. By way of contradiction assume that there exists $\rho' \notin R$ that is a Condorcet winner (CW). First, assume $\rho' = (c', 1, y'_i)$ with some arbitrary y'_i, and consider a candidate $\rho'' = (M, 0, y'_i)$. If $c' \in m$, ρ'' gets all the votes from the majority, and wins. If $c' \in M$, then ρ' gets the tribal voters from the majority, ρ'' gets all the voters from the minority, and the independent voters from the majority are equally split between the two candidates. Since $TV_M < \eta$, in this case ρ'' gets more than half the votes, a contradiction. Now assume $\rho' = (c', 0, y'_i)$ where $y'_i \neq med\{y_i\}$. A standard median-voter argument ensures that in this case $\rho \in R$ wins more than half the votes against ρ', a contradiction. Therefore, if a CW exists, it must be some $\rho \in R$. We now describe the conditions for a CW to exist when $TV_M < \eta$.

Existence of a Condorcet Winner When $TV_M < \eta$

Define the share of the majority $\lambda := \frac{1}{\eta+1}$. Consider again a triple $\rho \in R := \{\rho|\rho = (c, 0, y_{50}), c \in \{M, m\}\}$. We are looking for necessary and sufficient conditions for ρ to be a CW. Recqll that by Theorem 1, if a CW exists, then it must be some $\rho \in R$.

First, notice that as noted above, by standard median-voter arguments, any $\rho \in R$ gets at least half the votes against any non-tribal candidate, so we only need to consider tribal opponents. Moreover, by single-peakedness of preferences, it is sufficient to look at opponents proposing policies in an arbitrarily small neighborhood of y_{50}. We start by stating a fact about the distribution of y_i:

Fact. *For any distributions of y and any $\lambda \in (0.5, 1)$, we have $\lambda F_M(y_{50}) < 0.5$ and $\lambda(1 - F_M(y_{50})) < 0.5$.*

For intuition, notice that it cannot be the case that the share of the whole *population* on a segment is smaller than the share of the *majority* on the same segment.

Denote the set of voters from the majority that are above y_{50} by $M_+ := \{i | i \in M, y_i > y_{50}\}$, and equivalently the set of voters from the majority that are below y_{50} by $M_- := \{i | i \in M, y_i < y_{50}\}$. In general, we need to consider two types of opponents' policies against $\rho \in R$: tribal candidates with policies just above ("right tribals"), and just below ("left tribals") the population median. Notice that the set of voters for right-tribal opponents is bounded from above by $A_R := \Gamma_M \cup M_+$, and similarly, the set of voters for left-tribal opponents is bounded above by $A_L := \Gamma_M \cup M_-$. For convenience, for any $B \subseteq M$, denote the corresponding share of voters from the majority by $\mu(B) = \int_{i \in B} dF_M(y_i)$. We consider three separate cases.

Case 1: $y_{50} > \hat{y}_M$**.** First, consider left-tribal opponents. Notice that in this case we have $\Gamma_M \subset M_-$ and therefore $A_L = M_-$, and so $\lambda\mu(A_L) = \lambda F_M(y_{50}) < 0.5$, which implies that $\rho \in R$ wins against tribal opponents from the left. For right-tribal opponents, we have $\Gamma_M \cap M_+ = \emptyset$, and so $\lambda\mu(A_R) = \lambda(1 - F_M(y_{50}) + TV_M)$. Therefore, since any CW is in R, we must have $\lambda(1 - F_M(y_{50}) + TV_M) \leq 0.5$ in order to have a CW. Notice that this condition is both necessary an sufficient: if it does not hold, i.e., if $\lambda(1 - F_M(y_{50}) + TV_M) > 0.5$ a right-tribal opponent can choose a policy arbitrarily close to y_{50} and get more than half the votes against $\rho \in R$.

Also notice that such tribal candidate can be defeated by a tribal candidate that proposes yet even more extreme policy, e.g., y_{TV}. However, a tribal candidate proposing y_{TV} is defeated by $\rho \in R$, and therefore we get a Condorcet cycle.

Case 2: $y_{50} < \hat{y}_M - \tilde{\omega}$**.** Just like case 1, we have $\Gamma_M \subset M_+$ and therefore since $A_R = M_+$,we have $\lambda\mu(A_R) = \lambda(1 - F_M(y_{50})) < 0.5$, so $\rho \in R$ wins against tribal opponents from the right. For left tribal opponents, we have $\Gamma_M \cap M_- = \emptyset$, and so $\lambda\mu(A_L) = \lambda(F_M(y_{50}) + TV_M)$. Therefore we must have $\lambda(F_M(y_{50}) + TV_M) \leq 0.5$ in order to get a CW, and as before the condition is both necessary and sufficient, with the same Condorcet cycle (albeit in the opposite direction).

Case 3: $y_{50} \in [\hat{y}_M - \tilde{\omega}, \hat{y}_M]$**.** Since $y_{50} \in [\hat{y}_M - \tilde{\omega}, \hat{y}_M]$ left tribal opponents get the set of tribal voters plus all majority voters located to the left of the tribal segment $[\hat{y}_M - \tilde{\omega}, \hat{y}_M]$, i.e., all voters to the left of $\hat{y}_M$. Similarly, right tribal opponents get the set of tribal voters plus all majority voters located to the right of the tribal segment, i.e., all voters to the right of $\hat{y}_M - \tilde{\omega}$. To preclude Condorcet cycles where such triples defeating $\rho \in R$, we must require $\max\{\lambda F_M(\hat{y}_M), \lambda(1 - F_M(\hat{y}_M - \tilde{\omega}))\} \leq 0.5$. As before, this condition is both necessary and sufficient.

B A Two-Stage Political Competition

In this section we construct a simple game that allows us to characterize the winner even when a CW does not exist. Consider the following two-stage, two-player game:

- There are two parties, 1 and 2.
- At period 1, parties choose simultaneously a candidate $c \in \{M, m\}$, and a tribal stance $I_c \in \{0, 1\}$.
- At period 2, parties choose simultaneously a policy in X.
- Voters then cast their voters according to the assumptions stated in section 2.
- The payoff is 1 for the party that wins, 0 for the party that loses, and $\frac{1}{2}$ for each party if the votes are equally split.

Proposition 1. *In any pure SPE, the equilibrium result is such that:*

1. If $TV_M > \eta$, both parties choose $\rho^* = (c \in M, 1, y_{TV})$ (i.e., the winner is always a tribal candidate from the majority offering the policy preferred by the median non-tribal voter from the majority).
2. If $TV_M < \eta$, both parties choose (ρ, ρ'), with $\rho, \rho' \in R$ (i.e., the winner is always non-tribal candidate offering the centrist policy).
3. If $TV_M = \eta$, any choice in for both parties $R \cup \rho^*$ can be a part of a SPE (i.e., the winner and the policy are undetermined and can be a tribal candidate as in part 1, or a non-tribal candidate as in part 2).

Proof. The strategy space of each party i is $S_i = (z_i, x_i(z_j, z_i))$ where $z_i \in (M, m) \times \{1, 0\}$, and $x_i \in X$. As above, we denote the tribal stance of a party by I. First, note that choosing $(m, 1)$ is never a best response, and so we can WLOG eliminate this strategy from the tree. Second, by a standard median-voter argument, conditional on z_i, z_j,both parties choose the same policy in the second stage. Specifically, in this modified game tree, in any pure SPE we must have

$$x_i(z_j, z_i) = \begin{cases} y_{50} & I_i = I_j = 0 \\ y_{TV} & \text{otherwise} \end{cases}$$

that is, both parties choose the policy preferred by the pivotal voter. We can therefore restrict our attention to each party's first-stage strategy.

The following matrix lists each party's vote share in the second stage as a function of first-stage actions:

	1	0
1	$\frac{1}{2}, \frac{1}{2}$	$\lambda TV_M + \frac{1}{2}\lambda\,(1 - TV_M)$, $(1-\lambda) + \frac{1}{2}\lambda\,(1 - TV_M)$
0	$(1-\lambda) +$ $\frac{1}{2}\lambda\,(1 - TV_M)$, $\lambda TV_M +$ $\frac{1}{2}\lambda\,(1 - TV_M)$	$\frac{1}{2}, \frac{1}{2}$

(Notice that we merge $\{m, 0\}$ and $\{M, 0\}$ into a single action 0, since both are payoff equivalent.)

We then have the following three obvious conclusions:

1. Assume $\lambda TV_M > 1 - \lambda$ (i.e., $TV_M > \eta$), then the unique SPE strategy profile is $(1, 1)$. The policy chosen in this case is y_{TV}.

2. Assume $\lambda TV_M = 1 - \lambda$ (i.e., $TV_M = \eta$), then any $\{I_i, I_j\} \in \{1, 0\}^2$ is a SPE. The policy depends on the equilibrium, but must be either y_{TV} or y_{50}.

3. Assume $\lambda TV_M > 1 - \lambda$ (i.e., $TV_M < \eta$), then the unique SPE strategy profile is $(0, 0)$. The policy chosen in this case is y_{50}.

□

C Proof of Theorem 2

By substituting the constraint $a + \Delta(1 + p^c)/3 = \mu$ into (6), the maximization problem can be rewritten

Proof.

$$\max_{\Delta, p^c} \Delta \cdot L(p^c, p) \text{ s.t. } \begin{array}{c} \underline{\Delta} \leq \Delta \leq 3\mu/(1+p^c) \\ p^c \in \left[\underline{p}^c, \overline{p}^c\right] \end{array},$$

where

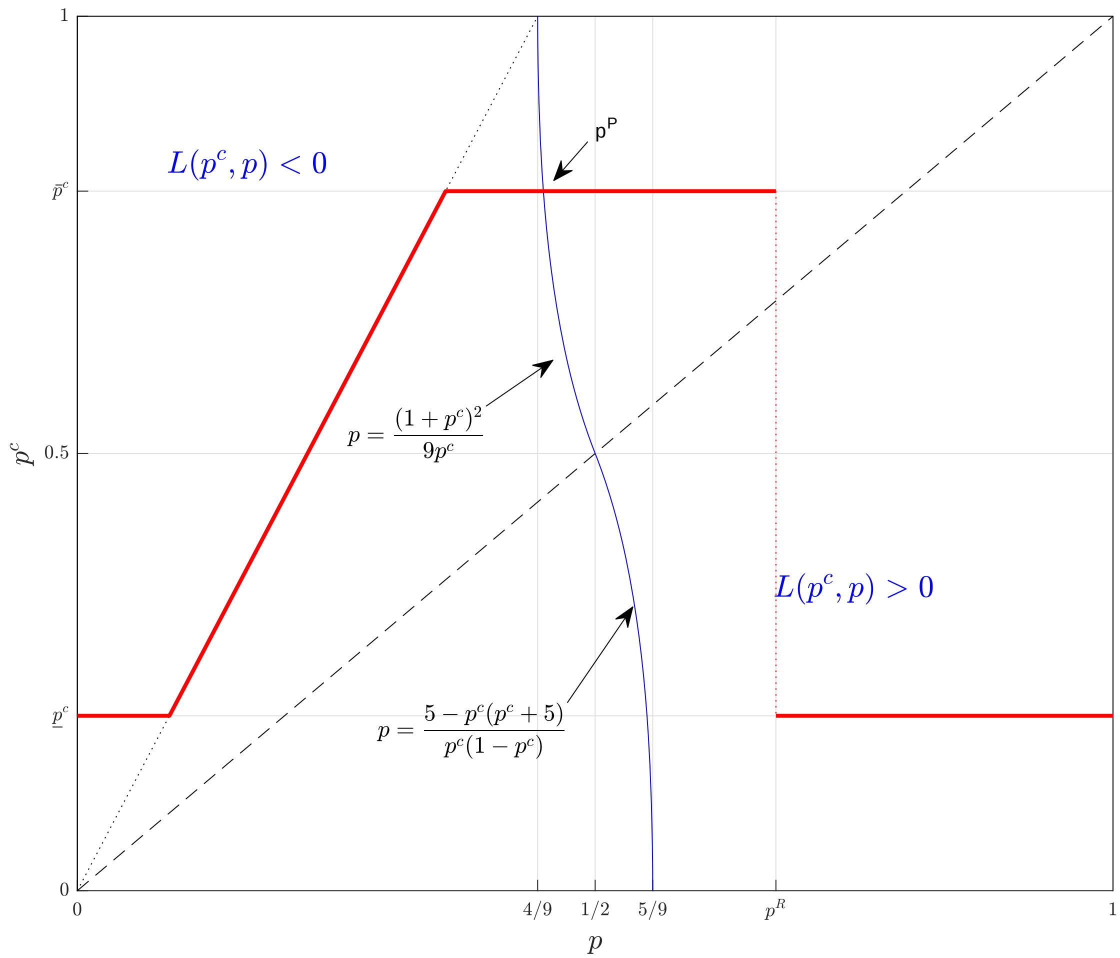
$$L(p^c, p) = -\frac{1+p^c}{3} + \begin{cases} 1 - \sqrt{(1-p)(1-p^c)} & p^c \in [0, p] \\ \sqrt{pp^c} & p^c \in [p, 1] \end{cases}. \tag{10}$$

□

The domain of the function $L(p^c, p)$ is $[\underline{p}^c, \overline{p}^c] \otimes [0, 1]$, and it is strictly negative for all p^c when $p < p^P$, where $p^P = (1 + \overline{p}^c)^2/(9\overline{p}^c)$ (see Figure C.1). Since individuals with $p < p^P$ always have $L < 0$,[10] they maximize $\Delta \cdot L$ by choosing Δ at its minimal value, i.e. $\Delta = \underline{\Delta}$. These individuals are thus left with the simple problem of maximizing $L(p^c, p)$ over $p^c \in [\underline{p}^c, \overline{p}^c]$, which is solved by choosing $p^c = \min\left\{\overline{p}^c, \max\left\{\underline{p}^c, \frac{9}{4}p\right\}\right\}$.[11] This establishes the result for the "poor" individuals.

[10]The fact that $L < 0$ for individuals with $p < p^P$ leads to another definition of the poor class: these are the

Figure C.1: The function $L(p^c, p)$ and the optimal choice of $p^{c*}(p)$.



Notes: The bold red line represents $p^{c*}(p)$. The function $L(p^c, p)$ is positive to the right of the solid thin blue line, and negative to the left. The class boundary $p^P = (1+\bar{p}^c)^2/(9\bar{p}^c)$ is the value of p at which the blue and the red line cross.

Individuals with $p > p^P$ can choose values of p^c for which $L(p^c, p) > 0$, which is always strictly preferred (since $\Delta > 0$) . This also means that the upper constraint on Δ will be binding: $\Delta = 3\mu/(1+p^c)$. The problem of these individuals can therefore be restated as:

$$\max_{p^c \in [\underline{p}^c, \bar{p}^c]} \frac{L(p^c, p)}{1+p^c}.$$

It is straightforward to verify that the expression to be maximized has no local maxima, and therefore, we are just left to compare the values at the two boundaries: $L(\underline{p}^c, p)/(1+\underline{p}^c)$ and

individuals whose status will be below average for any allowed redistribution.

[11]To show this, first note that $\partial L(p^c, p)/\partial p^c < 0$ for all $p < 5/9$ and $p^c < p$, next note that $L(p^c, p)/\partial p^c$ is concave for $p^c > p$, and finally that a local maximum exists at $p^c = 9p/4$.

$L(\overline{p}^c, p)/(1+\overline{p}^c)$. Thus, define

$$H(p) = \frac{L(\overline{p}^c, p)}{1+\overline{p}^c} - \frac{L(\underline{p}^c, p)}{1+\underline{p}^c},$$

and note that:

- $L(\overline{p}^c, p^P) = 0 > L(\underline{p}^c, p^P) \Rightarrow H(p^P) > 0$.
- $H(1) = (1+\overline{p}^c)^{-1} - (1+\underline{p}^c)^{-1} < 0$.
- $H'(p) < 0$.

It follows that there exists $p^R \in (p^P, 1)$ such that $H(p) > 0$ for $p \in (p^P, p^R)$, and $H(p) < 0$ for $p \in (p^R, 1]$. This proves that $p^{c*}(p) = \overline{p}^c$ for $p \in (p^P, p^R)$, and $p^{*c}(p) = \underline{p}^c$ for $p \in (p^R, 1]$, as needed. The rest follows from substituting this into the constraints.

Finally, it is not needed for the proof, but the value of p^R can be found by solving $H(p^R) = 0$, and the solution is:

$$p^R = \begin{cases} \left(\frac{\underline{p}^c(1+\overline{p}^c)}{(1+\underline{p}^c)\sqrt{\overline{p}^c}-\sqrt{(1-\underline{p}^c)(\overline{p}^c-\underline{p}^c)(1-\underline{p}^c\overline{p}^c)}}\right)^2 & \overline{p}^c \geq \frac{\sqrt{5-4\underline{p}^c}-1}{2(1-\underline{p}^c)} \\ \frac{32-(5-\underline{p}^c)(5-\overline{p}^c)+\underline{p}^c(1-\underline{p}^c)\overline{p}^c(1-\overline{p}^c)-2(1+\underline{p}^c)(1+\overline{p}^c)\sqrt{(1-\underline{p}^c)(1-\overline{p}^c)}}{(3+\underline{p}^c+\overline{p}^c-\underline{p}^c\overline{p}^c)^2} & \text{otherwise} \end{cases}. \tag{11}$$

D Proof of Lemmas 1 and 2

Lemma 1 states the the preferences of the poor are single-peaked.

Proof. Consider a voter at quantile $p \in (0, p^P)$, and let $q \in (0, p^P)$. From theorem 2 we have that

$$F^{-1}(p; P^*(q)) = \mu + \underline{\Delta} L(p^{c*}(q), p),$$

where $L(p^c, p)$ is as defined in 10. It is straightforward to show that for all $p \in (0, p^P)$ the function $L(p^c, p)$ has a unique maximum at $p^c = 9p/4$, and that it is monotonically increasing on $[0, 9p/4]$ and decreasing on $[9p/4, p^P]$. Therefore, the preferences of the voter at quantile p are single-peaked on $[0, p^P]$. Next, we must show that $P^*(p^P) \succ_p M^* \succ_p R^*$. Since $p < p^P < 1/2 < \overline{p}^c$,

$$F^{-1}(p; P^*(p^P)) = \mu + \underline{\Delta}\left[\sqrt{p\overline{p}^c} - \frac{1+\overline{p}^c}{3}\right],$$

$$F^{-1}(p; M^*) = \frac{3\mu}{1+\overline{p}^c}\sqrt{p\overline{p}^c},$$

$$F^{-1}(p; R^*) = \frac{3\mu}{1+\underline{p}^c}\begin{cases} 1 - \sqrt{(1-p)(1-\underline{p}^c)} & \underline{p}^c \leq p \\ \sqrt{p\underline{p}^c} & p \leq \underline{p}^c \end{cases}.$$

We calculate

$$F^{-1}(p; P^*(p^P)) - F^{-1}(p; M^*) = \left[\mu - \underline{\Delta}\frac{1+\overline{p}^c}{3}\right]\left[1 - 3\frac{\sqrt{p\overline{p}^c}}{1+\overline{p}^c}\right] > 0,$$

where the term in the first bracket is positive due to Assumption 5 on μ, and the second term is positive for all $p < p^P = (1-\overline{p}^c)^2/(9\overline{p}^c)$. Also, for $p \leq \underline{p}^c$ we have $F^{-1}(p; M^*) > F^{-1}(p; R^*)$, since the function $\sqrt{pp^c}/(1+p^c)$ is an increasing function of $p^c \in (0,1)$. For $p \in (\underline{p}^c, p^P)$,

$$F^{-1}(p; M^*) = \frac{3\mu}{1+\overline{p}^c}\sqrt{p\overline{p}^c} \geq \mu\sqrt{2p},$$
$$F^{-1}(p; R^*) = \frac{3\mu}{1+\underline{p}^c}\left(1 - \sqrt{(1-p)(1-\underline{p}^c)}\right) \leq 3\mu\frac{p}{1+p},$$

and it is straightforward to show that $\sqrt{2p} > 3p/(1+p)$ for all $p < 1/2$.

Lemma 2 states that the non-poor prefer M^* to $P^*(q)$ for any $q \in [0, p^P]$. This is trivial for the middle-class, since M^* is their bliss-point, therefore we must only prove this for the rich.

Proof. Let $p \in [p^R, 1]$, and note that $F^{-1}(p; P^*(q)) = \mu + \underline{\Delta} L(p^{c*}(q), p)$. By differentiating $L(p^c, p)$ with respect to its first argument, we find that $L(p^{c*}(q), p)$ is either decreasing in q for all $q \in (0, p^P)$,[12] or decreasing in q for $q \in (0, p-5/9)$ and increasing for $q \in (p-5/9, p^P)$.[13] In either cases, $L(p^{c*}(q), p)$ has no local maxima, so it is sufficient to verify that $F^{-1}(p; M^*) > F^{-1}(p; P^*(p^P))$ and $F^{-1}(p; M^*) > F^{-1}(p; P^*(0))$.

Proof. Comparing M^* and $P^*(p^P)$, we find

$$F^{-1}(p; M^*) - F^{-1}(p; P^*(p^P)) = L(p, \overline{p}^c)\left[\frac{3\mu}{1+\overline{p}^c} - \underline{\Delta}\right] > 0,$$

since for all $p > p^R$, $L(p, p^c) > 0$, and the term in brackets is positive by the assumption on μ. In order to compare M^* and $P^*(p^P)$, define:

$$G(p) = F^{-1}(p; M^*) - F^{-1}(p; P^*(0)) = \frac{3\mu}{1+\overline{p}^c} L(\overline{p}^c, p) - \underline{\Delta} L(\underline{p}^c, p).$$

The function $G(p)$ is continuously differentiable on $p \in (p^P, 1)$, and the derivative can only vanish for $p < \overline{p}^c$. However, it is also the case that $G''(p)$ exists and is negative for $p \in (1/2, \overline{p}^c)$, therefore the minimum of $G(p)$ on $p \in [p^R, 1]$ must occur at one of the boundaries. We therefore verify:

$$G(1) = \frac{2-\overline{p}^c}{1+\overline{p}^c}\mu - \underline{\Delta}\frac{2-\underline{p}^c}{3} > 0,$$

due to 5. Also, by the definition of p^R,

$$G(p^R) = \frac{3\mu}{1+\overline{p}^c} L(\overline{p}^c, p^R) - \underline{\Delta} L(\underline{p}^c, p^R) = \left[\frac{3\mu}{1+\underline{p}^c} - \underline{\Delta}\right] L(\underline{p}^c, p^R) > 0.$$

Thus, $G(p) > 0$, and $M^* \succ_p P^*(0)$ as needed. □

[12] This is the case when $p \in [(4\overline{p}^c + 5)/9, 1]$.
[13] This is the case when $p \in [(4\underline{p}^c + 5)/9, (4\overline{p}^c + 5)/9]$.