On the Theory of Ethnic Conflict

Francesco Caselli and Wilbur John Coleman II[1]

First Draft: December 2001; This Draft: September 2011

[1]London School of Economics, CEPR and NBER (f.caselli@lse.ac.edu); Duke University (coleman@duke.edu). Alberto Alesina, Robert Barro, Paul Collier, Lisa Cook, Joan Esteban, Raquel Fernandez, Ben Friedman, Nicola Gennaioli, Seema Jayachandran, Peter Klenow, Michael Kremer, Krishna Kumar, Ben Olken, Dominic Rohner, Silvana Tenreyro, Jaume Ventura, and Romain Wacziarg provided valuable comments.

Abstract

Why are ethnic groups sometimes in conflict and soemtimes at peace? We present a theory of ethnic conflict in which coalitions formed along ethnic lines compete for the economy's resources. The role of ethnicity is to enforce coalition membership: in ethnically homogeneous societies members of the losing coalition can more easily pass themselves as members of the winning group, and this reduces the chances of conflict as an equilibrium outcome. We derive a number of implications of the model relating social, political, and economic indicators such as the incidence of conflict, the distance among ethnic groups, group sizes, income inequality, and expropriable resources. One of the insights is that the incidence of ethnic conflict is non-monotonic in expropriable resources as a fraction of total resources, with a low incidence for either low or high values. We use the model's predictions to interpret historical examples of conflict associated with skin pigmentation, body size, language, and religion.

1 Introduction

Ethnic conflict is frequently in the news, and pervasive throughout history. In many countries ethnic groups are or have been visiting violence on each other, sometimes on a horrific scale (the word "genocide," by definition, refers to a type of ethnic conflict). Fearon and Laitin (2003) identify no less than 58 ethnic civil wars between 1945 and 1999, constituting 51% of the total number of civil wars.[1] Less visible and newsworthy, but quite possibly much more pervasive, is non-violent ethnic conflict. Non-violent ethnic conflict can take multiple forms. In some countries ethnic groups compete through overtly ethnic parties, vying for power. In others a dominant group discriminates against and exploits the others. As Esman (1994) succinctly puts it "when an ethnic group gains control of the state, important economic assets are soon transferred to the members of that community" (p. 229).

Yet, and crucially, ethnic conflict is by no means universal in ethnically heterogenous societies: in many countries ethnic groups coexist peacefully. Nor is it constant over time: many ethnically heterogenous societies experience long (sometimes very long) periods of fairly harmonious ethnic relations before or after periods of conflict. Why do some countries experience ethnic conflict and others don't? Why does ethnic conflict wax and wane over time in the same country? The goal of this paper is to contribute to the effort to answer these questions.

We view each society as being endowed with a set of wealth-creating assets, such as land and mineral resources. There is therefore an incentive for a subset of agents to form a group to wrest control of these assets from the rest of the population. Once a group has won control over the country's riches, however, it faces the task of enforcing the exclusion of non-members. In particular, agents not belonging to the winning group will attempt to infiltrate it, so as to participate in the distribution of the spoils. For example, they will apply for land titles, or for government jobs. This infiltration defeats the winning group's purpose, as it dilutes the "dividend" each original member receives. In large communities of millions of citizens it can be quite costly to keep track of the genuine members so as to successfully discriminate against the non-members. Our key idea is that, when groups can be formed along ethnic lines, ethnic identity can be used as a marker to recognize potential infiltrators. By lowering the cost of enforcing membership in the winning group, ethnic diversity makes it less susceptible to *ex-post* infiltration by members of the losing one. Hence, for group that expects to prevail in a conflict, a bid for a country's resources is an *ex-ante* more profitable proposition if this bid occurs along ethnic lines than if it occurs along non-ethnic (and therefore more porous) lines.

[1] Of the remaining 56 civil wars an additional 20 is classified as "ambiguous," in the sense that Fearon and Laitin are not sure whether it was faught along ethnic lines or not (their definition of "ethnic" war).

An important implication of this idea is that not all ethnic distinctions are equally effective ways of enforcing group membership. In particular, it is possible that some types of ethnic identities are harder to shed than others. One reason for this is that some ethnic identities are more easily observed by members of other groups. The clearest case of this is the case of skin color, or other physical characteristics that differ markedly among ethnic groups. Ceteris paribus ethnic boundaries based on physical differences should be easier to police than boundaries based on non-visible differences. Another reason why not all ethnic cleavages are equally resistant to passing is that the psychic costs of giving up one's ethnic identity may vary with the nature of that identity. For example, in some cases passing from one group to the other may require religious conversion, while in others both origin and destination groups have the same religion. Abandoning one's religious identity may be more costly psychologically than abandoning other traits of one's cultural identity. Furthermore some religions create physical markers, such as circumcision or scarring, that further increase the cost of passing.[2]

To capture this heterogeneity, we build on the notion of *ethnic distance*.[3] In our model ethnic distance is the cost to be born by a member of one group to successfully pass himself as a member of the other group. In general, we would expect ethnic distance to be maximal when there are differences in skin color and other physical characteristics that make passing all but impossible. Distance may be fairly high in the case of religious differences among groups. Language barriers could plausibly be argued to be a somewhat weaker source of distance. Potential infiltrators can assimilate through learning the language, or more realistically through having one's children do so. Finally, ethnic cleavages that are only marked by a shared sense of identity or history, unsupported by additional differences of color, religion, language, or other observable characteristics, should be associated with the lowest levels of ethnic distance.[4]

Having established the concept of ethnic distance we can give a preview of how our

[2]Maimonides in the late 12th century explains the practice of circumcision as a way of preventing "strangers" from saying they are members of the faith. "For sometimes people say so for the purpose of obtaining some advantage." Needless to say the practice has backfired when circumcision has been used to identify Jews for the purposes of persecution, as depicted most memorably in Luis Malle's *Au Revoir les Enfants* (1987).

[3]A notion of group distance is also important in Esteban and Ray's (1994, 1999) measures of polarization and their subsequent work on ethnic conflict (reviewed below). However in their context distance is best interpreted as distance in preference or income space, not in terms of ease of migration among groups.

[4]Needless to say, ethnic bounaries can be and often are multi-dimensional, involving various combination of physical, religious, linguistic, and other cultural differences. Ethnic distance is the cumulative effect of these differences. Also, ethnic distance is not country specific but is specific to each pair of ethnic groups within a country. Group A could be "close" to group B but distant from C, etc. While in the paper we mostly focus on the case of just two groups, we also discuss a multiple-group extension.

model answers the questions that motivate the paper. In particular, the probability of observing ethnic conflict is:

1. Increasing in ethnic distance. The further the distance, the most limited the passing from the losing/exploited group into the winning/dominant group, and hence the greater the reward from conflict behavior for the latter.

2. Inverted-U shaped in expropriable assets – resources that can be captured through conflict behavior – as a share of overall income. In our model an increase in the share of expropriable assets has two opposing effects on the intensity of conflict. It increases the "prize" to be gained by the winning group, and hence its incentive to seek conflict. But it also increases the incentive for the losers to pass into the winning group, enhancing the dilution effect from infiltration, and thus reducing the incentive for conflict by the prospective dominant group. In our model the net effect turns out to be inverted-U shaped, with conflict intensity being maximized for intermediate levels of the expropriable-resource share in total wealth.

3. Increasing in the wealth of the losing/exploited group. This is because passing often involves having to give up part of one's assets. For example, starting a life under a new ethnic identity may require moving to a different region where one is not previously known, which in turn may require abandoning assets such as land, a shop, or one's own residence. One's human capital may also be location specific. Hence, the cost of passing is increasing in one's wealth, so the richer the "victim" group the less concerned with infiltration and dilution the dominant group will be.[5]

4. The effect of pre-conflict groups sizes on the likelihood of conflict is complex. In a baseline model where the only possible form of ethnic conflict is exploitation of a weak group by a strong group, conflict is monotonically decreasing in the pre-conflict size of the strong group. This is because the larger the strong group the smaller the per-capita gain from exploitation. However on an extended model where the weaker group can also "fight back" to protect some of its own access to the country's resources, or even exploit the stronger group's passivity to appropriate them to itself, the probability of conflict may become U shaped in the stronger group's size. This is because as the weaker group becomes extremely small it has extremely large per-capita benefits from

[5] Clearly there are other possible reasons why conflict is more likely when the side which is expected to lose is rich. For example, the winner may expect to appropriate some of the losing side's private assets. But the mechanism we highlight, the effect of the losers' wealth on the incentive to pass, is unique to our framework and, more importantly, it applies even to the component of private wealth that cannot be expropriated (for example human capital).

exploiting the fact that group A has correspondingly low incentives for conflict, and it becomes possible to observe forms of exploitation of the majority by the minority.

5. Decreasing in the destructiveness of conflict, and decreasing in the wealth of the prevailing group (because conflict destroys some of this wealth).

The mechanisms in points (1)-(3) are consequences of our notion of ethnic distance and therefore unique to our framework. The U shape in relative size result in point (4) is not particularly dependent on our notion of distance but as far as we know is novel in the literature. The results in (5) would operate in a very large class of models of civil conflict.[6]

Given these results, cross-country differences in proneness to civil conflict would result from differences in all the determinants of conflict just listed, and transitions from conflict to peace would equally be driven by changes over time in these determinants. Particularly likely seem changes in the share of expropriable assets in total wealth, and we discuss below a number of historical examples where we conjecture such changes may have led to long-run changes in ethnic relations. Changes in ethnic distance due to changes in the perceived psychic costs of passing seem also possible, and so are of course changes in the relative wealth of the groups, changes in relative group size (for example due to migration or differential population growth), and of course changes in conflict technology that may make conflict more or less destructive.

Another source of distance is of course geography. Our model applies equally well to groups that form based on the geographical base of their membership. When one group's army enters a city in enemy territory, its soldiers can be pretty confident that the overwhelming majority of the civilians they encounter belong to the enemy group. Hence, our theory of conflict among geographically separated groups is isomorphic to our theory of ethnically distant groups, and one may therefore be able to use our model, together with other relevant state variables identified in this paper, to explain changes over time in the intensity of inter-regional (and perhaps even international) conflict.

The rest of the paper is organized as follows. In the next section we discuss related literature. This discussion is structured along the following lines. First, we discuss literature that provides empirical and historical underpinnings to our key notion of ethnic distance, and how it relates to "passing" between ethnic groups. Since passing is the central strategic

[6] Obviously we are not saying that conflict will *only* arise in societies with deep ethnic divides. If the benefits of conflict are large enough, a group aiming to exclude the rest of the population may arise even in relatively homogenous societies: this group will tolerate a certain amount of leakage and/or will be willing to pay relatively large costs to set up artificial methods to enforce membership (e.g. party affiliation). We are merely saying that, *ceteris paribus*, distance increases the likelihood for conflict, particularly if the other conditions listed in the text are satisfied.

concern of the dominant group in our model, this discussion serves to highlight the relevance of the ideas developed in the paper. Other parts of the literature review section deal with alternative theories of ethnic conflict, existing empirical evidence on ethnic conflict, and (briefly) the various ways in which the notion of "ethnicity" has been conceptualized in the literature.

Our theory is developed formally in Sections 3-4. Section 3 sets out a benchmark model of exploitation with two ethnic groups that captures the basic idea of ethnic distance and highlights the non-monotonic role of expropriable resources. Section 4 develops a number of extensions. In one extension we distinguish between two types of conflict: exploitation of one group at the hands of the other, and open conflict between the two groups. We show that our insights are robust to this richer description of ethnic relations, and we make predictions as to when conflict will take the form of exploitation and when it will be bilateral. In a second extension we allow for a distinction between "leaders" and "followers" within each group. Once again our main insights go through. A third extension considers the possibility of multi-dimensional ethnic cleavages, and hence multiple groups. Finally we look at a simple dynamic extension of our baseline static model. The results from the dynamic analysis are broadly in line with the baseline. However one new insight is that when factoring the possibility of passing the dominant group may optimally manipulate the timing at which it begins and cease to exploit the other. As a result, changes in ethnic relations can occur well before or long after changes in the exogenous state variables, and ethnic relations can change multiple times between peace and conflict even if the state of economy changes only once.

Section 5 discusses how our model can contribute to the understanding a number of historical examples that we think our theory sheds some light on, including Black-White relations in the United States and South Africa; Hutu-Tutsi relations in Rwanda and Burundi; Muslims and Hindus in India; and others.

2 Related Literature

2.1 Passing and Distance

At least since Barth's (1969) classic book social scientists have been aware of overwhelming evidence that individuals change their ethnic affiliation in response to external circumstances. Barth also stressed that ethnic categories (or labels) can endure despite the porosity of ethnic boundaries.

An often-cited case of passing is represented by light-skinned African-Americans who "passed" and "lived on the other side" [to use the language of the New York Times, 9/7/2003], albeit at the cost of severing all ties with their families and childhood friends, as poignantly depicted in Roth (2000).[7] The porosity of ethnic boundaries is also evident in the widely commented swings in self-reported ethnic identification in censuses. Nagel (1995) argues that the more than three-fold increase in the American Indian population between the 1960 and 1990 U.S. censuses is due to massive ethnic switching from American Indians who previously identified themselves as non-Indian, but returned to an American-Indian identity. Evans et. al. (1993) argue that the 46 percent increase in Aboriginal and Torres Strait Islanders between the 1981 and 1986 Australian censuses is due to "an increase in the propensity to identify themselves as such in the Census." Lieberson and Waters (1993) present Census-based evidence of significant ethnic switching during a person's lifetime even for Whites in the U.S. Jeganathan (1997) reports that Tamil families living near Colombo give Sinhalese names to their children and teach them Sinhalese cultural practices to help them escape identification in case of riots. We will discuss several further examples of passing in Section 5.

A path breaking experimental study by Habyarimana et al. (2007) highlights the porosity of ethnic boundaries as well as the potential for individuals to manipulate their identity. The authors asked people from diverse ethnic groups in Uganda to view digital images of other Ugandans and categorize them into the appropriate ethnic group. They found that, for example, members of the largest ethnic group (the Baganda, which comprised about 45 percent of the survey), correctly identified a Bagandan approximately 70 percent of the time, and incorrectly identified a non-Bagandan as Bagandan approximately 20 percent of the time. They also found that, given the right incentives, members of some groups can send signals to members of other groups that fool them into mistakenly accepting them as members of their own. Lastly, they found that the degree of ethnic identifiability (or the ability to pass oneself off as a member of another ethnic group) varies across group pairs, indicating that some bilateral cleavages are more porous than others.[8]

Experimental evidence also confirms humans' reliance on visual cues in group formation: Kurzban, Tooby, and Cosmides (2002) found that experimental subjects tend to classify individuals by race when there are no other visual markers to rely on, but when other markers are added, such as, for example, one of two basketball team's jerseys, observers become equally likely to switch to the new visual markers as to continue using race. Of course from our point of view it is critical that outside of the experiment people can change shirt color but not skin color.

[7] Joseph Roth is not the only novelist who saw the literary potentital of passing. Other examples include Naipaul (1990, India), Danticat (1998, Haiti), Akpam (2008, Nigeria).

[8] Interestingly, there are no systematic differences in skin color among Ugandan ethic groups so evidently Ugandans are using cues other than skin color in making these judgments.

2.2 Explanations for Ethnic Conflict

The social-science literature on ethnic conflict and ethnic politics is vast and impossible to comprehensively review here. We limit our review to some classic contributions and those contributions that are most closely related to ours. Broadly speaking, theories of ethnic conflict rely on some combination of two broad categories of motives: "instrumentalist" (sometimes also called "rationalist") and "primordialist" (or "consummatory").[9] Instrumentalist explanations emphasize the fact that participants in conflict hope to derive some material benefit from the conflict, such as jobs, wealth or power. Primordialist views focus on the visceral dimension of conflict, which they interpret as an eruption of mutual antipathy. To some extent it is possible to think of these two views as, respectively, conceiving of ethnic conflict as a "technology," as it modifies the constraints and opportunities faced by individuals, or putting it directly into agents' preferences.[10] Our contribution is clearly in the instrumentalist tradition.

The classic statement of the instrumentalist view is in Bates (1973, 1982). Bates' foremost point is that ethnic conflict is conflict among rational agents over scarce resources. He buttresses this claim by organizing an astounding wealth of case-studies from Sub-Saharan Africa. Taking Bates' view of the reasons for conflict as our starting point, we formalize the reasons why ethnicity is a rational basis for coalition building and provide a characterization of some of the conditions that make ethnic conflict more likely.[11] Many subsequent scholars have identified numerous further examples where leaders favor their own ethnicity when allocating resources [see e.g. Posner (2005) for Africa.]

Horowtiz (1985) combines primordialist and instrumentalist elements. A central element of his analysis is the role of self-esteem that individuals derive from seeing members of their ethnic group succeed in business and, especially, in politics. This allows group leaders (who view ethnicity instrumentally) to mobilize coethnic support whether in the form of votes for ethnic parties or participation in violent confrontations. That group self-esteem considerations are important in ethnic conflict seems extremely plausible, and our model could easily be extended to add potential self-esteem benefits to engaging in conflict. As long

[9] This is a very rough and overly rigid classification. See Cederman (2002) for a more subtle and articulated taxonomy. Cederman also draws the important distinction between nation and ethnicity.

[10] Hence, to an economist, primordialist behavior does not need to be irrational, if it leads to higher utility for the perpetrator. Identifying instrumentalist with rational is therefore somewhat problematic.

[11] Bates' own view of the reason why coalitions form along ethnic lines is somewhat different from ours. He stresses a combination of factors, chief among which is that ethnic groups are associated with geographic regions, and the allocation of scarce resources and administrative instituions are also often organized along geographic criteria. He also argues that to the extent that there are pre-existing political and social networks within ethnic groups, as well as a common language, the costs of intragroup organizing are lower than coalition-building across groups.

as there are *also* material benefits in terms of redistributed appropriable resources our main results would not change. Similarly, Horowtiz is surely right that the benefits of conflict are very unequally distributed between group leaders and the masses, and this is a point that has been greatly emphasized by much of the subsequent literature [Fearon and Laitin (2000) have a particularly lucid discussion of this issue]. We discuss a possible extension of our model where conflict disproportionately benefits group leaders (and provide further references to the literature) in Section 4.

Hardin (1995) revisits several of these issues giving many examples where individual participation in ethnic conflict can be construed as an individually rational response given agents' information sets, potential or perceived material gains from conflict, and concerns about the intentions of members of other ethnic groups.[12]

More recently Chandra (2004) asks why and when is it rational for voters to support ethnic parties and when is it rational for political entrepreneurs to target the votes of coethnics. Her answer is that voters find collecting information on candidates' background and intentions costly, while ethnicity is readily observable. Hence, they use ethnicity has a noisy but low-cost signal of candidates propensity to favor them in allocating public goods and transfers. Given this behavior by voters, it can be rational for parties to organize along ethnic lines. There is some connection between Chandra's use of ethnicity as a low-cost signal of intentions and our use of ethnicity as a low-cost technology to police coalition boundaries, and once again our analyses are somewhat complementary. The closest antecedent to our work, however, is Fearon (1999), who asks why ethnic politics and "pork" politics often tend to go together, and conjectures informally that allocating pork according to ethnicity (or other features that are not easily chosen or changed by individuals) is a way of preventing political losers from attempting to enter the winning group.

None of the work surveyed above uses formal models. There is a tradition of formal models of social and international conflict, featuring among others contributions by Grossman (1991, 1999), Hirshleifer (1995), Azam (1995, 2001), McDermott (1997), Esteban and Ray (1994, 1999), Gershenson and Grossman (2000), and Grossman and Mendoza (2003). However few of these models are explicitly designed to specifically capture conflict among ethnicities. The only clear antecedent we are aware of is Robinson (1991), whose focus is in predicting whether class or ethnic conflict should be expected to be more virulent.

More recently Esteban and Ray (2011) have proposed a model where, as in Horowtiz, conflict is driven both by instrumentalist and primordialist motives. The main focus is the role of within- and between-group income differences. However, unlike in our model,

[12] Hardin also devotes considerable attention to the mechanism leading to group identity and cohesion, an issue on which considerable progress has recently been made by Berman (2000, 2009), and Berman and Laitin (2008).

Esteban and Ray assume that ethnic relations are inherently conflictual (the model does not feature a peaceful outcome) and the composition of the groups is fixed and immutable. Our analyses are therefore complementary. Esteban and Ray (2008a), taking fundamentally an instrumentalist view, study a society with both class and ethnic cleavages and ask when one should expect to see ethnic as opposed to class conflict (or no conflict). Esteban and Ray (2008b) focus on the (conflicting) effects of polarization on conflict onset and conflict intensity. Esteban and Ray (forthcoming) investigate the relation between the intensity of conflict and various measures of heterogeneity used in the empirical literature. Fernandez and Levy (2008) focus on the interaction between income and preference heterogeneity in a redistribution game. Rohner (2010) and Rohner, Thoenig and Zilibotti (2011) model the two-way interaction between conflict and trust. In none of these contributions are changes in group composition allowed.[13]

2.3 Empirics of Ethnic Conflict

Several authors have tried to identify a causal link between a society's ethnic structure and measures of (typically violent) conflict.[14] Generally speaking, the traditional measure of "ethnolinguistic fractionalization" (the probability that two randomly drawn individuals will belong to different ethnic groups) is not a very good predictor of conflict [Fearon and Laitin (2003)], while more sophisticated constructs used by Montalvo and Raynal-Querol (2005) [building on Esteban and Ray (1994, 1999) and Reynal-Querol (2002)] and Cederman and Girardin (2007) do appear to successfully predict civil wars or inter-group conflict. A simple dummy variable taking the value of one when the largest ethnic group accounts for between 45 and 90% of the population has also been found to predict conflict [Collier and Hoeffler (2004)].

These findings are important and interesting, and on the whole suggest that certain ethnic structures are more conducive to conflict than others. However, they are not particularly informative in the specific context of the present paper. One reason is that these studies take the existing ethnic structure of the population as exogenous. But our model predicts that relative group sizes change in response to conflict, so regressing conflict outcomes on statistics that depend on the size-distribution of groups is very close to getting the direction

[13] Our understanding is that Esteban and Ray (2011) was first written roughly contemporaneously and independently of our contribution, while the other contributions cited in this paragraph are subsequent to ours.

[14] This is just part of the broader literature on the effects of ethnolinguistic fractionalization. Other outcomes studied in this literature include growth, redistribution, corruption, trust, and the provision of public goods. Examples include Mauro (1995), Easterly and Levine (1997), Miguel and Gugerty (2005), Alesina, Baqir, and Easterly (1999), Alesina and La Ferrara (2000), Alesina et al. (2003) and Desmet et al. (forthcoming).

of causality wrong. The exogenous variable in our model is the *initial* group size distribution, but this is not what typical data sets measure.[15] Second, and perhaps more importantly, our theory implies that the distribution of group sizes is an insufficient predictor of conflict. At a minimum, information on relative group sizes should be complemented with information on ethnic distance, as the same group-size structure could be associated with very different conflict outcomes depending on inter-group distance patterns. Furthermore our model implies a complex set of interactions between group size, distance, and other variables, and particularly the share of appropriable assets in the economy. Simple linear specifications are unlikely to capture these.

Very recently some progress has been made in measuring distance. In particular, building on work by Laitin (2000) and Fearon (2003), Desmet et al. (forthcoming) and Esteban et al. (2010) bring to bear evidence on linguistic distances. In particular, Desmet et al. (forthcoming) report that only when linguistic fractionalization is computed on the basis of language groups that have split a long time ago is such fractionalization a significant predictor of conflict. This is consistent with our model as older language cleavages likely correspond to the languages being more different, making assimilation and passing more arduous. Esteban et al. (2010) show that a measure of polarization constructed using linguistic distances is a robust predictor of conflict. While they interpret linguistic distance as proxying for differences in preferences, we think that an equally likely interpretation is that linguistic distance directly measures (a dimension of) the costs of passing among groups.[16]

As already mentioned, distance in space is another possible dimension of ethnic distance in the sense of our model. If groups are spatially clustered, jobs, subsidies, and other benefits of being the dominant group can be effectively targeted using geographic criteria - and practices of this kind are abundantly documented (see, e.g., the Bates papers). Passing becomes correspondingly more costly, as it requires moving to a different region of the country. Hence, our theory also has the implication that geographically clustered and isolated ethnic groups are more likely to find themselves parties to conflicts. Matuszeski and Schneider (2006) present evidence that geographical clustering of ethnic groups is significantly related to the incidence, duration, and severity of civil war. Similarly, Cederman et al. (2009), Weidman (2009) and Weidman et al. (2010) find that groups that are more clustered, tend to live in mountainous areas (and are therefore more costly to reach/move away from), and are further away from the national capital are more likely to be in conflict with the central government [see also Toft (2003) for similar results].

[15] See Ahlerup and Olsson (2011) for another model of endogenous ethnic structure formation, and Fletcher and Iygun (2010) for empirical evidence that indeed ethnic structure is a function of past conflict outcomes.

[16] It could also be that linguistic distance proxies for other types of distances, e.g. visible physical differences. This would still be consistent with our model.

Further progress on distance has also recently been made by Guiso, Sapienza, and Zingales (2009), and Spolaore and Wacziarg (2009), who look at some economic consequences of "genetic distance," as measured by the frequency of certain alleles in various populations. While genetic distance is not the same thing as ethnic distance (most genetic differences do not lead to visible differences), this is certainly a step in the right direction.[17] Bossert et al. (forthcoming) have followed a different path trying to construct measures of distance that take into account income, education, age, etc.

Our paper is also related to empirical studies of the role of economic factors, such as natural-resource wealth or income shocks, on civil conflict [e.g. Miguel et al. (2004), Ciccone (2011), Besley and Persson (2011)]. A particularly interesting finding due to Collier and Hoeffler (2004) is that the probability of conflict is inverted-U shaped in the fraction of primary commodities in total exports. If the latter is a proxy for the resources whose control the conflict is about, this is exactly as predicted by our model.

2.4 Construction and Salience of Ethnicity

Two closely-related propositions enjoy near-universal consensus in the literature on ethnicity. The first proposition is that ethnicity's "salience" changes over time, both within the lifetime of individuals and in terms of wider societal perceptions. In other words individuals and communities ascribe to ethnic identities more importance in certain periods than in others (and sometimes no importance at all). This view is entirely consistent with our framework. Indeed, our model offers an explanation for why ethnicity's salience varies across time and space. In the model periods of harmonious relations may be interpreted as periods where ethnicity is not salient, while periods where conflict or exploitation take place are periods where ethnicity has become salient. As discussed, such transitions from non-salience to salience can be triggered by changes in macro-economic conditions, changes in the wealth status of certain groups, or changes in the perceived social cost of conflict.

The second widely held view is that ethnic identity is a "social construct," in the sense that it results from social "discourses" that end up conditioning individuals to identify with particular groups. This idea seems implicit in Barth (1969) and has been extensively elaborated. A famous application is in Anderson (1983). Social constructivism is in opposition to an alternative approach that views ethnic identity as an immutable feature of human nature.[18] Once again our contribution is fully consistent with the social-constructivist posi-

[17] The distinction between genetic and ethnic distance may become less relevant in the future due to changes in biotechnology. Dando (2004) argues that RNA technology can now potentially be used militarily to shut down specific mutations of important genes that are known to be prevalent in certain populations.

[18] Confusingly, this latter view is also called "primordialist" in the literature, perhaps because those who believe that ethnicity is an immutable factor also tend to believe that ethnic conflcit arises from feeling of

tion. In our framework, like in much instrumentalist writing on ethnic conflict, ethnic groups are socially constructed to build winning coalitions.[19] Our twist on social constructionism is to point out that such discourses are easier to make where there exist markers (of color, or language, or religion, etc.) around which the discourses can be organized. In other words it is easier to create a social construction of identity when this identity can be pegged on the hook of, say, skin color than when such a hook is absent.

In this respect our approach shares some features with van den Berghe's (1978, 1981, 1995) theory of ethnic identification and racism. Building on evolutionary psychology, Van den Berghe argues that agents are strongly motivated by "nepotism," an evolutionary-driven tendency to seek to benefit individuals who are more likely to share a larger proportion of one's genotypes. This induces agents to look for cues that can provide some information on common ancestry, such as skin color and *visible* physical features (leading to racism) or cultural markers (leading to ethnic identification). Like us, van den Berghe stresses the gradient among possible markers of ancestry: "where physical, genetic markers do a reliable job of differentiating between groups they *are* used," but "most ethnic groups *look* so much like their neighbors that they *must* rely on cultural markers of distinction" (1995, p. 361). He then goes on to discuss the relative effectiveness of dress, cultural markers which permanently change physical appearance (such as scarification), language, etc. The difference between van den Berghe and us is that we do not require nepotism for agents to be interested in identifying markers that lower the cost of policing group boundaries - our agents are purely selfish. Furthermore, van den Berghe's analysis does not directly address fluctuations over time in the salience of racial and ethnic identities.

3 A Model of Exploitation

3.1 Assumptions

We study a society populated by a continuum of individuals of measure 1. Each individual is initially assigned to one of two ethnic groups, A or B (we discuss multiple groups later). The initial size of group A is n (so the initial size of B is $1-n$). Within each group, all individuals are identical (we discuss within-group heterogeneity later). Each member of group A (B) has

mutual anthipaty. However, there is no logical reason for expecting the primordialist view of ethnicity to imply the primordialist view of ethnic conflict. One can certainly believe ethnicity to be immutable and still propound an instrumentalist interpretation of ethnic conflcit. See again Fearon and Laitin (2000) for an excellent discussion of the relation between social constructivist views of ethnicity and theories of ethnic conflict.

[19] Harris and Sim (2002) say "... advocates of this social constructionist perspective on race maintain that the function of race is to reinforce and perpetuate social differences."

an initial exogenous income stream y_A (y_B) from assets that cannot be expropriated. One may loosely think of y_A as human capital. In addition, society is endowed with aggregate resources that generate an income stream of z, that must somehow be distributed among the population. z could be the rental value of land, mineral resources, or any other endowment that is valuable to a country.

We will assume that one of the two groups is "stronger" and can set up an exploitation regime. We have in mind that one of the two groups has greater fire power and can largely impose its will. In many cases the stronger group will be the numerical ethnic majority. However, in some cases ethnic minorities may be stronger if they can mobilize greater resources per capita, or equivalently have greater human capital (e.g. South Africa during Apartheid). Without loss of generality we assume that A is the stronger group.[20]

Group A then chooses between two actions: C (for conflict) or P (for peace). We don't model the specific mechanism through which this collective decision is taken, but we assume that the choice maximizes the utility of agents who start out as members of group A.[21] If group A chooses C, it takes hold of the common resource z, to the exclusion of the members of the other group from enjoyment in it. Exploitation is costly. If group A decides to seize control, a fraction δ of all the country's resources is lost. There are several possible interpretations of the conflict cost δ. It could represent the cost of the repressive apparatus needed to enforce the exploitation of group B. It can also represent the deadweight cost of discrimination. For example, exploitation may call for excluding talented members of group B from administrative and managerial posts (and having to search further down the talent distribution of group A to replace them). Net of this cost, conflict results in a reallocation of the common resource z to group A, with the *ex-post* (i.e. end-of-game) members of the group sharing equally in it. If group A chooses action P, z is divided equally among all citizens.[22]

Group A's conflict or peace decision takes up the first stage of the game. In the second stage, members of the weaker group decide whether to keep their ethnic identity, or

[20] We could formalize the definition of stronger by saying, for example, that group A is stronger if its aggregate wealth is greater, i.e. $y_A n > y_B(1-n)$, but since the formal definition of "stronger" plays no role in the subsequent analysis we leave other possibilities open.

[21] Because all of the members of group A are identical, almost all rules to aggregate preferences will give rise to the same decision as of whether to exploit or not to exploit group B, as long as the spoils are shared equally among group A's members. In turn, the equal-sharing option would be the natural choice on a "behind the veil of ignorance" basis. As already mentioned, we return to within-group heterogeneity later.

[22] There is here, and even more clearly in the extension of Section (4.1), where we look at the possibility of Group B fighitng back, an implicit assumption that groups cannot precommit to act cooperatively. In the present context, group B could agree to cooperate to its own exploitation (saving the economy the cost δ) while still allowing A to take all of the pie. We implicitly assume that if A lets down its guard, say by eschewing a repressive apparatus or by allowing talented members of groups B to take on important jobs (but none of the benefits) group B will then have an incentive to renege and try to keep some of Z for itself.

to "pass" and join the majority.[23] Passing involves a proportional income loss of ϕ. There are many interpretations for ϕ. At the simplest level, changing ethnic group may involve considerable loss of ethnicity-specific human capital. For example, one may have to sacrifice business contacts, or leave a profession that has an ethnic connotation to it. Changing identity may also involve geographical relocation to an area where one's ancestry is not known, with attendant further loss of business contacts or location-specific human capital. It may also involve some kind of primitive surgery, the payment of bribes to counterfeit identification documents or change names, payments to families of other groups in order to marry (one's children) into them, etc. All these costs are likely to have a component that is proportional to one's income. Finally, there are the obvious psychic costs. A key idea in the paper is that all of these costs vary depending on the nature of the ethnic distinction (race, religion, skin color, etc.). For example, it is far more costly for a person with very dark skin to pass himself off as white than for a low-caste Hindu to become Catholic. We therefore assume that ϕ can vary continuously from zero (to capture a completely homogenous country) to infinity.[24]

Identity switchers cannot be separately identified from original members of the group. The number of ex-post members of group A is denoted n', and is equal to n plus the number of initial members of group B who switched identity. After individuals have made (and executed) their ethnic identity decision, resources are allocated based on all prior decisions and characteristics of the society. Individuals derive utility exclusively from consumption, and consumption equals income.

Society can be characterized by the initial relative group-size n, non-expropriable endowments y_A and y_B, aggregate resources z, switching cost ϕ, and exploitation-cost parameter δ. Given these characteristics, group A decides collectively whether or not to engage in conflict, and individuals of group B choose their ethnic identity, giving rise to n'.

3.2 Equilibrium

Consider the first-stage decision by group A whether or not to exploit group B. If A decides for peace (action P) its per-capita payoff is simply

$$U_A^P = y_A + z. \tag{1}$$

[23] It will be obvious below that members of the stronger group never pass in this simple version of the model. We explore later a richer model where group B can "fight back," or even attempt to exploit group A. In that model, members of group A may also wish to pass.

[24] This is a shortcut. A more rigorous formalization would involve a fixed cost ϕ_0 and a proportional cost ϕ_1, so that the overall cost of passing is $\phi_0 + \phi_1 y_B$. We would then let ϕ_0 vary between 0 and infinity and ϕ_1 between 0 and 1. This more rigourous version would be somewhat unwieldy and it is hard to imagine that the qualitative insights of the model would change.

I.e., members of group A have complete access to their initial endowment, as well as to the common resource z, which is divided equally among all members of society. If instead, they decide to seize control of z (action C) their payoff is

$$U_A^C = (1-\delta)\left[y_A + \frac{z}{n'}\right]. \tag{2}$$

Hence, conflict leads to the loss of δy_A units of the individual endowment as well as δz units of the collective good. On the other hand, through action C group A obtains full control of the natural resource. This amount is divided equally among the final membership of group A, n'.

It is clear by comparing the last two expressions that group A's decision as to whether or not to play C depends on the equilibrium response of n' if it does so: the greater the expected ex-post size of group A in the event of a conflict, the less likely group A is to seek it. For example, it is immediately apparent that there will be no equilibria where a conflict induces *all* of the members of group B to switch identity: with $n' = 1$ we have $U_A^C = (1-\delta)[y_A + z]$, which is certainly less than U_A^P. More generally, by comparing eqs. (2) and (1), we see that group A will seek to exploit group B if and only if $n' < \tilde{n}$, where

$$\tilde{n} \equiv \frac{(1-\delta)z}{\delta y_A + z}.$$

This "exploitation threshold" is increasing in z, falling in the cost of exploitation δ, and falling in the income of the victorious group y_A: the richer group A is, the more it is concerned about the destructive effects of exploitation. A very rich group has much to lose from engaging in conflict. Note that $\tilde{n} < 1$.

In case of conflict each member of group B decides his ethnic identity.[25] If he passes to group A he receives utility

$$U_B^S = (1-\delta)\left[(1-\phi)y_B + \frac{z}{n'}\right],$$

where the first term in the square bracket reflects the cost of changing identity and the second term is the gain represented by access to resources seized by group A. Since there is exploitation all resources are net of the cost δ. If he sticks to his original identity his utility is

$$U_B^{NS} = (1-\delta)y_B.$$

The pro of passing is that it allows the passer to retain access to the common resource. The con is that one has to pay the switching cost.

[25] It should be obvious that there is no switching by members of group B if there is no conflict (they would pay the switching cost, but gain nothing).

Note that the gain from switching is decreasing in n'. For low values of n' the gains from defecting to the winners are relatively large, as the spoils of exploitation are divided among few people. As n' increases an infiltrator's share falls, and so does the incentive to pass. Hence, passing by some reduces the incentive for further passing by others. Indeed, for n' large enough gaining access to z is not a sufficient compensation for the switching cost, and the net incentive to pass may become negative. In particular, we have that members of group B pass as long as $n' < \overline{n}$, where

$$\overline{n} \equiv \frac{z}{\phi y_B}.$$

The "switching threshold" $\overline{n}$ is increasing in the spoils of conflict z (the bigger the pie, the larger the number of people one is willing to share it with), and decreasing in the cost of switching ϕy_B. Note that it is possible for $\overline{n}$ to be larger than 1. These are cases in which, in the event of conflict, members of the weak group have an incentive to defect at all values of n' (the pie to share is just too large relative to the cost of changing sides).

The equilibrium value of n' in the event of a conflict depends on the relative positions of the initial group size n and the switching threshold $\overline{n}$. If $n < \overline{n}$, and a conflict occurs, citizens of group B will start switching to A. If $\overline{n} < 1$ the flow of defectors will stop when no further incentives to switching are left, i.e. the equilibrium value of n' is $\overline{n}$. If $\overline{n} > 1$ the flow of defectors will stop when all members of group B have switched sides, i.e. $n' = 1$. On the other hand, if $n > \overline{n}$ there are already "too many" people in group A to start with, and no member of group B wishes to switch. The equilibrium in this case features $n' = n$. In summary, if the dominant group A seeks to exploit group B, we have $n' = \max[n, \min(1, \overline{n})]$.

Recall now that group A seeks to exploit group B if it does not expect too much switching in response, i.e. if $n' < \tilde{n}$, where $\tilde{n}$ is the "conflict threshold." We therefore have conflict if $\max[n, \min(1, \overline{n})] < \tilde{n}$. Recall also that $\tilde{n} < 1$ (A never engages in conflict if, in the event of conflict, everyone switches to A), so there can never be conflict if $\bar{n} \geq 1$. This allows us to simplify the condition for conflict to

$$\max(n, \overline{n}) < \tilde{n}. \tag{3}$$

We summarize this discussion with the following

Proposition: *Group A exploits group B if and only if (3) holds. If, furthermore, $n < \bar{n}$, then there is switching from B to A, and $n' = \bar{n}$. Otherwise $n' = n$.*

If $n < \overline{n} < \tilde{n}$, then there is conflict, and the equilibrium value of n' is $\overline{n}$. The size of the dominant group is sufficiently small that members of group B switch, but not in large enough numbers to make conflict unprofitable for the dominant group. For $\overline{n} < n < \tilde{n}$ there is still conflict, but no switching. The exclusionary benefits of conflict are large enough for the dominant group to seek conflict, but not large enough for members of the weak group to incur the switching cost ϕ. For $n > \tilde{n}$ it is never worth it for the dominant group to exploit the small minority in B. Finally, if $n < \tilde{n}_A < \overline{n}$, group A would benefit from taking control of z if its ex-post size was the same as its ex-ante one, but it expects too much switching in equilibrium, so it does not attempt it.

3.3 Comparative Statics

Depending on the configuration of parameters ϕ, δ, n, z, y_A, and y_B, a country will or will not experience an ethnic conflict. We want to know how the "exploitation" v. "no exploitation" status changes as these 6 parameters vary. Substituting the expressions for $\bar{n}$ and $\tilde{n}$ in 3 we easily get that conflict occurs if and only if (i) $n < (1-\delta)$ and, (ii)

$$\frac{\delta y_A n}{(1-\delta)-n} < z < (1-\delta)\phi y_B - \delta y_A. \tag{4}$$

The comparative static properties of the model follow immediately from (4). First, conflict will occur only for relatively high values of the passing cost ϕ. Second, conflict will occur only for intermediate values of the resource-rent flow z. Third, conflict occurs only for relatively low values of the income of the stronger group A, and, fourth, for relatively high values of the income of the weaker group B. Fifth, conflict will only occur for relatively small relative sizes of the dominant group A. Sixth, conflict will only occur if the cost of conflict δ is relatively small.

It is also of interest to ask which parameters are associated with more passing in case of conflict. In case of conflict, passing occurs when $\bar{n} > n$, or

$$\frac{z}{\phi y_B} > n.$$

Hence, it is immediate that more passing (conditional on conflict) is associated with smaller ϕ, y_B and n, and with larger z.

We further illustrate and discuss our results with the help of Figure 1, which measures z on the horizontal axis, and ϕ on the vertical axis. The figure features a large triangle denoted "conflict." This is the set of (z, ϕ) combinations that satisfy condition (4), and hence give rise to exploitation of B by A (holding constant the other parameters). Outside of this triangle A does not attempt to gain control. The "conflict" region is further divided into two triangles. The "no switch" triangle corresponds to combinations of parameters such that all the members of group B stay in group B, while the "switch" triangle features some switching from B to A.

The figure shows a (weakly) positive relationship between conflict and ethnic distance, ϕ. For given z/y, there is no conflict if ϕ is very low, and there is conflict if ϕ is high enough.

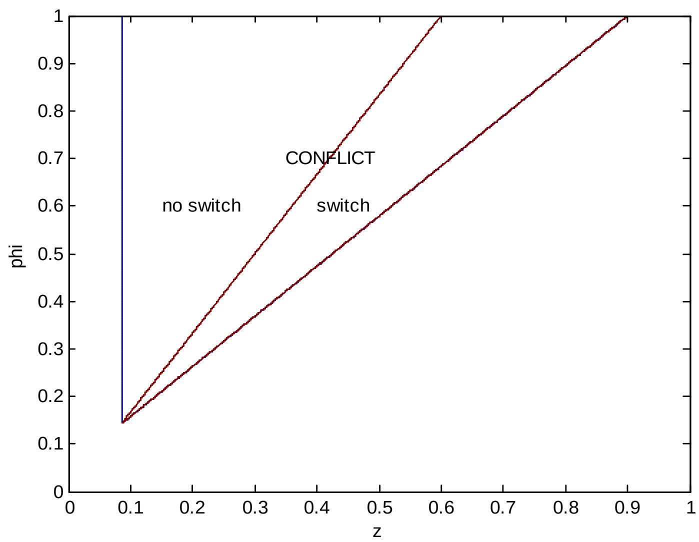


Figure 1: Exploitation v. peace as functions of z and ϕ

Hence, ethnic proximity acts as a deterrent to conflict: the dominant group eschews any attempts at exploitation when it expects a large inflow of group B members should it try to do so. A low ϕ allows for such a massive switching.

The figure also shows an "inverted-U shaped" relationship between z/y and conflict. Moving from left to right for a given (sufficiently high) value of ϕ, we see that there is no exploitation for z low - it does not pay. However, conflict also disappears as an equilibrium for z large. The reason is that the larger is z the larger is the number of B members who switch to A in case of conflict. Anticipating this massive switching, group A backs off. Hence, A exploits B only if z is large enough to make for an appealing booty, but not so large that it triggers a massive switching from B to A. However, the existence of a "switch" sub-region in the conflict region shows that A can tolerate a moderate amount of infiltration and still pursue exploitation. As mentioned in Section 2, Collier and Hoeffler (2004) find an inverted-U shaped relationship between conflict and natural resource intensity.

Figure 1 also highlights the interaction between ethnic distance ϕ and abundance of resources z. In particular, the greater the ethnic distance the larger the set of values of z such that exploitation occurs. The intuition is immediate from the previous discussion: the more costly it is to switch, the smaller the elasticity with respect to z of inter-group migration in response to conflict. Hence, the greater the ethnic distance, the more aggressive group A can be in appropriating large amounts of riches.

Finally, the figure shows that, not surprisingly given the discussion above, switching occurs for relatively low ϕ and relatively high z, with a similar interaction between these variables as found in the conflict decision of A.

Changes in group incomes can also be illustrated with reference to Figure 1. An increase in y_A causes the vertical line to shift right and the diagonal line to shift left, shrinking the conflict region. As group A becomes richer (relative to the resource endowment) we move from conflict to peace. This is the standard insight that the stronger group is more interested in conflict when the resources at stake are abundant, relative to the cost of conflict (which is indexed by the group's human capital). An increase in y_B causes the diagonal line to rotate clockwise, thereby expanding the conflict region. The reason for this is more specific to our model: high income weak groups have more to lose from switching identity. Hence, we can conclude that conflict is more likely when the stronger group has low per-capita income and the weaker group has high per-capita income (always relative to the resource endowment).

An increase in the dominant-group size n shifts the vertical line Figure 1 to the right, so that the conflict region shrinks. In particular, there are now fewer values of ϕ and fewer values of z/y for which conflict occurs. A larger initial size of the stronger group implies a smaller per-capita gain in the amount of natural resources appropriated through conflict, and hence a smaller incentive. Indeed, as per condition (i) above, there always are values of n that are large enough that no conflict occurs (the conflict region disappears). This particular result will receive some qualification in Section 4.1, when we allow group B to "fight back."

Increases in δ have very similar effects as declines in ϕ. Increases in δ tend to reduce the set of other parameter values such that there is conflict (the area of conflict with a larger δ is always a subset of the area of conflict with a smaller δ). For δ large enough we are always in the no-conflict region. Destructive conflicts are in nobody's interest. Indeed, there is always a neighborhood of $\delta = 1$ such that conflict does not take place, irrespective of other parameters' values.[26]

3.4 Summing Up

In sum, if group A is the stronger group, we are more likely to observe exploitation of group B by group A if: (i) The ethnic distance between A and B is large; (ii) the country's endowment of expropriable resources is neither too small nor too large; (iii) group B has high per-capita income; (iv) group A has low per-capita income; (v) group A is small; and (vi) the efficiency costs of exploitation are modest.

It is very important to stress that for *all* variables the threshold values that trigger conflict are defined in terms of the other variables in the model. For example, the lower

[26] As a technical detail, unlike in the case for ϕ, the opposite *is* true: when $\delta = 0$ conflict is a weakly dominant strategy for the stronger group: they can do no worse than with peace, and we should therefore always observe exploitation. This is a discontinuity, however: for any $\delta > 0$, if, say, ϕ is low enough war is no longer an equilibrium. Clearly $\delta > 0$ is the empirically relevant case.

δ the lower the required threshold for ϕ. This has important empirical implications. For example, consider the potential inverted-U shaped pattern that the theory predicts for the effect of variation in z on the peace-conflict status of a country. The upper threshold is clearly increasing in ϕ and, indeed, if $\phi = \infty$ then the relationship between z and conflict status becomes monotonic: since switching identity is prohibitively expensive, the deterrent effect of switching does not counter-balance the incentive to fight for a larger z. Hence, the model predicts that the width of the U shape depends on the value of ϕ.

4 Extensions

4.1 Exploitation v Conflict

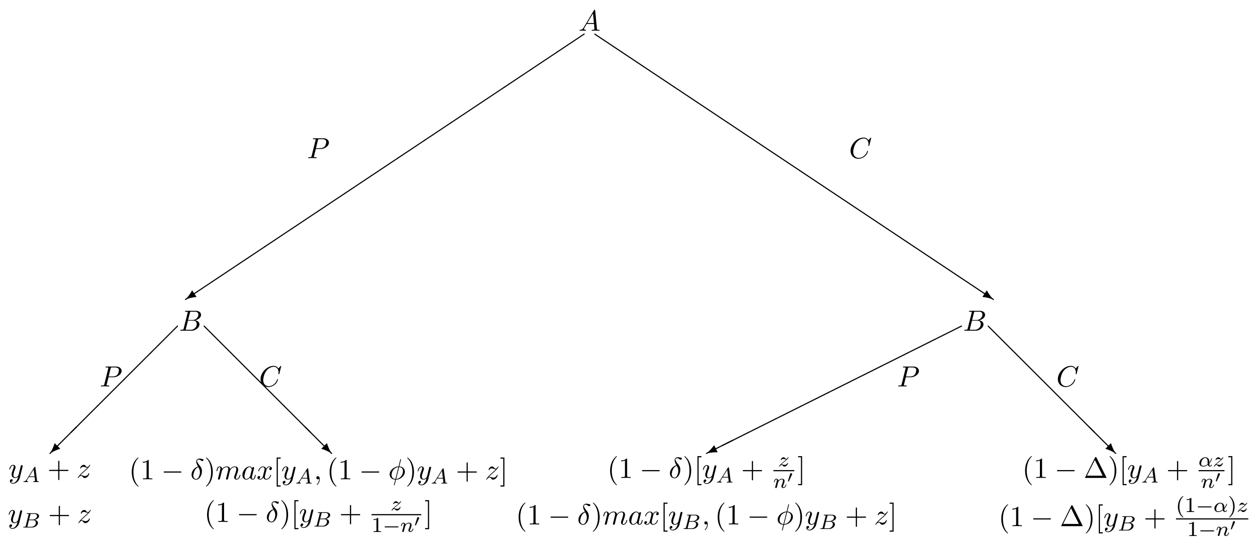
In the model of the previous section, when group A goes on the offensive and decides to appropriate the resource z, the only choice open to members of group B is whether or not to pass themselves off as members of the dominant group. The model does not distinguish between situations in which the losers "surrender," and give the winners free reign on the country's resources – a situation we have termed "exploitation" – and one where the losers "fight back," and try to retain control over at least some share of the country's resources – a situation for which we now specialize the meaning of the word "conflict." We now turn to a simple extension that accommodates a distinction between these two outcomes.

We continue to assume that, realistically, the stronger group, group A, moves first, and chooses between a "conflict action," C, and a "peace action," P. However, we now introduce a new second stage where group B can also respond with a C action or a P action. Furthermore, in the third stage we now explicitly consider not only the possibility of switching from B to A, but also from A to B.

The consequences of various series of actions are as follows. If both groups have played P, peace prevails, and each group i receives $y_i + z$, i.e. their "inalienable"endowment y_i plus an equal stake in the country's natural resources. This is the same as the no-exploitation equilibrium in the previous section's model. If one of the two groups has played C, and the other group has played P, we are in a situation where the C–playing group is exploiting the P–playing group, which acquiesces. In this case, the $C-$playing group gains control of all the natural resources z, which are then shared among the *ex-post* members of this group. Exploitation has enforcement costs and/or introduces distortions that reduce all incomes by a fraction δ. This is analogous to the "exploitation" scenario of the previous section, except that we leave open the possibility that group B exploits A, and not only A exploits B.

The more radically new type of scenario that is possible in this extension pertains to the outcome when both groups play C. We now assume that in this case the stronger group,

Figure 2: The 3-Stage Game



group A, receives a fraction α of the natural resource, with $\alpha > 0.5$, while the weaker group, say B, receives the remaining $(1-\alpha)$. Hence, relative to acquiescing to being exploited by A, and losing all control over z, B can "fight back" and retain some fraction, albeit less than its "fair share," of the country's endowment. However, this fighting-back option comes at a cost. We assume that open conflict causes greater social losses than exploitation. The destruction rate of output in the CC equilibrium is $\Delta > \delta$.

The extended form of the game is (partially) depicted in figure 2, where at each final node the payoff of A is listed first and the payoff of B second. The interpretation of the payoffs is straightforward in the PP case, where peace prevails. In the cases of exploitation (PC or CP) the exploiting group receives its own endowment y plus z divided by the number of ex-post group members, both depreciated at rate δ. The exploited group's payoff depends on this group's passing behavior. Non-passers receive only their individual endowment y. Hence if there is no passing, or if passing occurs until members of the exploited group have become indifferent between switching and maintaining their identity, the payoff for members of the exploited group is $(1-\delta)y$. On the other hand, if all the members of the exploited group pass over to the exploiting group, their welfare is $(1-\delta)[(1-\phi)y+z]$. In other words they pay the switching cost but recover access to their share of the country's resources. Universal passing of the group occurs when this last quantity exceeds $(1-\delta)y$, which explains the formula for the exploited group's payoff.

The payoffs in case CC, or open conflict, also depend on switching behavior. We show later that only members of group B switch to A, if at all. In equilibrium, members of group B prefer to remain in their original group, or are indifferent between switching and not switching. (Note that since stayers get some positive amount of the natural resource, there is no possibility that the entire membership of the group will switch identity.) Hence, the utility of members of group i in case CC is y_i plus the per ex-post member amount of natural resource that the group manages to preserve in the conflict. This payoff is now discounted at the higher rate Δ.

Solving this version of the model is conceptually straightforward. For each of the four final nodes PP, PC, CP, and CC one needs first to determine the equilibrium ex-post group sizes, or n'. Given n' one can determine whether B prefers PP or PC, and whether it prefers CP or CC. This provides A with B's response function to its actions. Given that, A chooses its best option between P and C.[27] The formal analysis, which is quite tedious, can be found in the online appendix to the paper.[28] Here we describe the properties of the equilibrium and the comparative statics.

The general structure of the model of this section can best be discussed with reference to Figure 3. In the figure we hold constant y_A, y_B, α, n, δ, and Δ, and study how the nature

[27] As already discussed in footnote 22 we implicitly rule out side deals. For example, an interesting variant of this model would give group A the option of offering to group B a division of Z which is more favorable to A than under the PP equilibrium, but not as favorable as under the CP or CC equilibrium. While such arrangements are sometimes observed in reality, they do heavily depend on both parties being able to make binding commitments. For example typically partial exploitation will require that the dominant group controls all the resources, and hands out group B's agreed share voluntarily and on an ongoing basis. It may be very difficult for B to monitor that this is appropriately done, particularly when the government's budget accounting is murky. It also requires B to commit not to take advantage of situations in which A has lowered its guard. In practice, inability to commit seems likley to be a frequent situation.

[28] The appendix studies in detail the model under the following restrictions on the parameters: $(1-\Delta)\alpha < n$ and $\alpha \geq 2n - n^2$. Exploring other regions of the parameter space would not materially change the qualitative insights from the model.

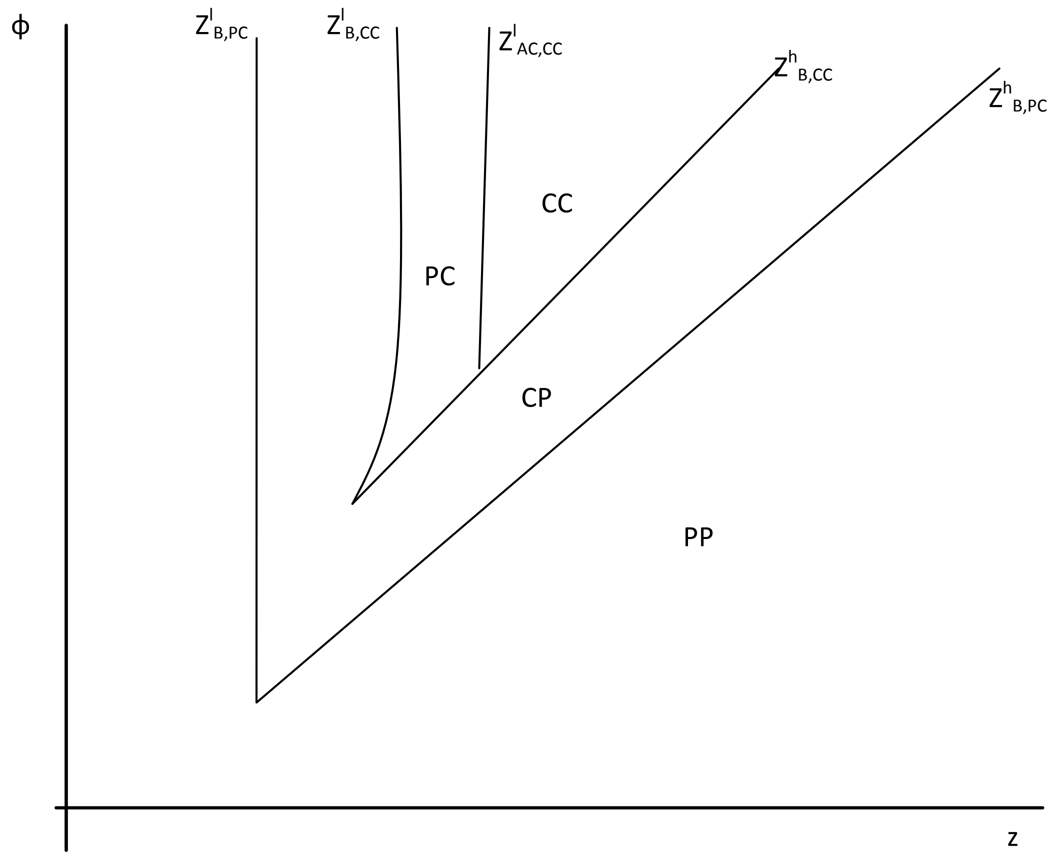


Figure 3: Regions of Peace, Exploitation, and Open Conflict

of the equilibrium vary as we vary ϕ and z. Each type of equilibrium is identified by the final node reached in the game between ethnic groups.[29] As in the benchmark model, there is a broadly triangular region featuring some type of conflict, while the complement features peace. Hence, peace prevails for low values of the passing costs ϕ, and for values of z that are neither too small nor too large. We also see again the important interaction between z and ϕ: as ϕ increases conflict occurs for a larger range of values of z.

One new feature of the equilibrium is that attempts to capture the resource z can now result in either exploitation or open conflict. In particular, there is an "inner triangle" featuring open conflict, CC, and outside "corridors" featuring exploitation by A on B, CP.

[29] For example CP is the region of the parameter space where in equilibrium A plays C and B plays C. The lables $z^l_{B,PC}$, $z^l_{B,CC}$, etc. are explained in the online appendix.

Finally, between the left CP corridor and the inner CC triangle there can be a region, PC, where B exploits A, rather than the other way around. Hence, for ϕ sufficiently large, as z increases from a sufficiently low value, the economy potentially transitions from peace, to exploitation of B by A, of A by B, to open conflict, back to exploitation by A, and finally back to peace However it is important to note that not all these regions necessarily exist. The only two regions that always exist (i.e. for all combinations of values of α, δ, Δ, y_A, y_B, and n) are PP and CC (we return to comparative statics with respect to these parameters below).

The intuition for how ethnic relations change with z is as follows. For z very low neither party wishes to disturb the peace, as the pie to fight over is too small. When z is larger both A and B become interested in exploiting each other, but not yet willing to engage in full-scale conflict, as the costs of the latter are still too large compared to the benefit. Hence, it is possible for A to play C without B fighting back. As z rises further, B begins to fight back when A plays C. Whether we enter directly the CC region, or we first transit through a PC region (as depicted in the figure), depends on parameter values. The reason why there may be a PC region is that, if z is not large enough, A may prefer to be exploited rather than bear the very large costs of an open conflict. Further increases in z beyond the CC region can bring about a new region of exploitation of B by A. The intuition for this region is interesting. For z large enough, the entire B population passes into group A under exploitation, so the payoff of group B under exploitation becomes increasing in z. In particular, it can exceed the payoff of open conflict, leading the group as a whole to choose to be exploited rather than fight back. However in the same region B would respond to P with C, as it is attractive to exploit the other group. A is therefore faced with a choice between being exploited or exploit, and obviously chooses the latter, even if the benefit is nil. Finally, for z very large again neither group wishes to exploit the other, as the volume of passing (both ways) would nullify the benefits.

The model's comparative statics with respect to the income of the stronger group, y_A, are slightly richer than in the benchmark model of exploitation. As in that model, an increase in y_A makes group A generally more peace oriented, as it increases the cost of conflict (of any type). In particular, this results in a shift to the right of the lower bounds of both the CP and the CC region, meaning that the onset of conflict is generally for higher values of z. However, an increase in y_A also makes group A less mobile, thereby increasing the region in which B responds to P with C. As a consequence, A is forced more often to play C to preempt being exploited by B, resulting in shifts to the right also of the upper bounds of the CP and the CC region, meaning that conflict generally persists for higher values of z as well. Another effect is that the PC region widens. In sum, an increase in y_A leads to shifts to the right of both the inner and outer conflict triangles, as well as an expansion of the inner PC region.[30]

Similarly, increases in y_B tend to increase the cost of conflict for B, which generally tend to shrink the CC area. The one countervailing force comes about, once again, when B must choose between CC and a CP situation where all the members of B pass into A. Since an increase in y_B increases the cost of passing, the net effect may be that B chooses CC more often in this region. Increases in B also have the effect of moving the outer edge of the CP region inward: the reason is that the higher y_B implies that B does not try to exploit A as often. As a result, A is forced to preempt less often. Thus, contrary to the benchmark model, increases in y_B could shrink, rather than expand, the overall conflict area. On the whole, however, the insight is similar: increases in one group's income make that group less aggressive, and the other group more aggressive.

The effect of an increase in n on the overall conflict area is also more ambiguous than in the benchmark model. The main conflicting forces are that: A has less to gain from conflict, which is the only force in the benchmark model, but B has now more to gain from conflict. This reflects the asymmetric nature of conflict: it does not depend on the average gain from conflict between the two groups, but on the maximum gain between the two groups. The maximum gain occurs when a small minority exploits a large majority. In this case, very low values of n tend to be associated with exploitation of B by A, and extremely large values of n may be associated with exploitation of A by B. In between, we tend to have either peace, or open conflict. Hence, the model accommodates the frequent pattern of powerful minorities exploiting the weaker majority (e.g. Sunnis exploiting Shias in pre-war Iraq), but also the seemingly puzzle fact that sometimes what look like small and weak minorities enjoy a seemingly privileged status. In these cases the strong majority (group A) prefers to entirely acquiesce to group B's voracity. Perhaps the current treatment of the surviving "American-Indians" in the US, and the Indian Tribes' fierce policing of their ethnic boundaries against (what they consider to be) infiltrators, may resemble this situation.

Many of the results in this sub-section highlight an important tension: the larger the group, the greater its power, but the less its incentive to engage in exploitation. This result may explain why the persecution of minorities is often accompanied and fueled by accusations that the minority is conspiring against the majority. It is true that in open conflict the minority stand to obtain a relatively minor share of the country's resources, but it is also true that if the majority lowers its guard and opens itself to exploitation by the minority the latter has enormous incentives to seize the opportunity.

[30]The parameter y_A, together with n, discussed below, is also the most important determinant of whether some of the regions exist at all. In particular, for y_A sufficiently small both the PC and the right corridor of the CP region disappear.

4.2 Leaders and Followers

As mentioned in Section 2.2 the literature on ethnic conflict has emphasized the unequal gains from ethnic competition. Leaders of ethnic groups stand to gain large amounts of wealth and power, so their behavior is easy to explain. But what about the masses? On this question there is some disagreement.

Some point out that the unequal distribution of material benefits does not imply that there is no benefit for the masses (or, at least, no *expected* material benefit). The elite may share (or promise to share) enough of the cake as to make participation or acquiescence to other group's exploitation worthwhile even for the foot soldiers. In the case of ethnic politics this will take the form of public-sector jobs, handouts, subsidies, location of public projects and infrastructure, or a law-enforcement system skewed in favor of coethnics. Bates and Posner, among others, take this view and provide many examples. In the case of open conflict the evidence is less systematic, but it seems clear that the masses of followers tend to enjoy freedom to loot, which could be a significant reward. Another benefit is that followers may use the open conflict to eliminate creditors, or take over property like land, cattle, and housing, which used to belong to members of the losing group. And of course there is the expectation of further benefits from the group's political control of the state once the conflict is over.

Other authors are more skeptical that the masses are in it for the material benefits. We have already seen that Horowtiz stresses individuals' self-esteem from seeing coethnics in positions of power. Others focus on elite manipulation of coethnic primordial feelings [e.g. Brass (1997), Woodward (1995), Glaeser (2005)] and/or information.[e.g. de Figuereido and Weingast (1999)].

In this sub-section we return to the baseline model of exploitation and sketch an extension where the masses follow for the material benefits. In this extension the stronger group, group A, has an elite of size ν, and the reminder $n - \nu$ are "the masses." The elite moves first and chooses between a C and a P action. If the elite chooses P the rest of the game is exactly as in the baseline model. If the elite chooses C there is a "spoil-sharing" rule which determines that a fraction β of the resources appropriated through conflict gets equally divided among the masses, while the remaining $1 - \beta$ is divided equally among the elite. For simplicity we treat β as an exogenous parameter. Next, the masses decide whether to support the elite in the conflict decision or to abstain. If the masses abstain the outcome is once again the peace outcome, with no social costs and equal society-wide division of the country's resources. This captures the idea that the elite needs the support of the mass of its coethnics to implement an exploitation strategy. If the masses cooperate there is an exploitation equilibrium, in which members of the elite receive a per-capita fraction $(1-\beta)/\nu$

of the country's appropriable resources (net of exploitation costs) and members of the masses receive a fraction $\beta/(n' - \nu)$.[31] Finally, members of group B decide whether to pass or not. Realistically, we assume that members of group B who pass will be part of group A's masses, i.e. it is impossible to pass oneself as a member of A's elite.

The passing decision in this extended model leads to a solution for the equilibrium size of group A in the case of conflict similar to the baseline case, namely $n' = \max[n, \min(1, \nu + \beta\bar{n})]$, where $\bar{n}$ was defined in the baseline section as $z/(\phi y_B)$. Hence the only difference is that the passing threshold $\beta\overline{n}$ now depends on the spoil-sharing rule. Exploitation occurs if both elite and masses of group A are better off under conflict. This results in the following two conditions:

$$\frac{\beta}{(n' - \nu)}(1-\delta)z + (1-\delta)y_A > z + y_A, \tag{5}$$

for the masses, and

$$\frac{1-\beta}{\nu}(1-\delta)z + (1-\delta)y_A > z + y_A, \tag{6}$$

for the elite. Note that (6) is always satisfied for ν sufficiently small, so we assume this constraint is never binding. In other words the elite of the dominant group always gains from ethnic politics, which seems realistic. The question is whether the masses go along.

From (5) the masses will go along if

$$n' < \nu + \frac{\beta(1-\delta)z}{z + \delta y_A} = \nu + \beta\tilde{n}.$$

Once again this expression is closely reminiscent of the baseline model of Section 3. Indeed as the size of the elite goes to 0, both the passing threshold and the conflict threshold converge to the baseline expressions, except that both thresholds are multiplied by β. Aside from this rescaling, therefore, all the comparative static results with respect to the baseline model parameters are exactly the same as in the homogenous case, and the model delivers the same messages.

4.3 Multiple Groups

So far we have focused on countries with only one (potential) ethnic cleavage. In many countries there are multiple politically-relevant ethnic groups. Furthermore, in many countries there are multiple dimensions along which ethnicity can become politically salient. For example there could be cross-cutting religious and skin-color cleavages, and the relevant dimension for group action could turn out to be religion (giving rise to groups that are heterogenous in

[31] Once again because all members of the masses are identical their colllective decision would be the same under virtually any mechanism to aggregate preferences.

skin color) or racial (giving rise to groups that are heterogenous in religion). In this section we sketch how our model could be extended to account for the multiplicity of potential ethnic categories.

Suppose that there are I "ethnically homogenous" groups, in the sense that all the members of each groups have the same physical, religious, linguistic, and cultural features. In the example above with two skin colors and two religions $I = 4$. Each group is characterized by its per-capita income, y_i, and relative group size, n_i. Furthermore, each *pair* of groups i, j is characterized by a switching cost $\phi_{i,j}$, which is the cost of switching identity from i to j and j to i.

Assume next that there is a nonempty set of "potentially winning coalitions." A potentially winning coalition is a coalition of groups that has the capability of imposing an exploitation equilibrium on the groups who are not in it. Naturally a potentially winning coalition could be made of a single group. Denote by W both the set and the number of potentially winning coalitions. Next assume that there is a "natural order" among the potentially winning coalitions. A coalition in W gets to decide whether to impose an exploitation equilibrium only if none of the previous coalitions in the natural order has decided to exploit. If a coalition gets an opportunity to decide, it imposes an exploitation equilibrium if and only if all the groups in the coalition play the same action C. If any group in the coalition plays P the decision passes on to the next coalition in W.[32] There is an exploitation equilibrium if one of the coalitions in W decides to exploit the other groups. Without loss of generality coalitions in W are indexed by their natural order (i.e. coalition 1 is the first in the natural order, etc.)

If no coalition exploits, then each member of group i receives $y_i + z$, for every i. If coalition $t \in W$ exploits, then each member of group $i \in t$ receives $(1-\delta)(y_i + z/n')$, where $n' = \sum_{s \in t} n'_s$, i.e. n' is the ex-post sum of members of the groups in the exploiting coalition.[33] Each member of group $i \notin t$ receives $(1-\delta)y_i$ if he does not pass, and $(1-\delta)((1-\tilde{\phi}_{it})y_i + z/n')$, where $\tilde{\phi}_{it} = \min\{\phi_{is}, s \in t\}$. In other words members of each exploited group will pass, if they pass at all, into the group in the exploiting coalition that is less distant from them. In order to avoid possible indeterminacies we assume that passing also occurs sequentially among members of the exploited groups. In particular, the group whose members have most to gain from passing passes first. If passing makes residual members of this first group indifferent between passing or staying then there is no further passing. However if all the members of the first group have passed then the opportunity moves to members of the group with the

[32] The "natural order" assumption is a simplification. It would be possible to use results in the literature on coalition formation to endogenize the order in which coalitions decide.

[33] Hence we assume that members of the exploiting coalition share equally. This could be extended by introducing a within-coalition bargaining stage.

highest gains from passing among the remaining exploited groups, etc. Notice that unlike in the baseline case there can be exploitation in equilibrium even when an entire group passes - provided there are sufficiently many remaining people in other exploited groups.

There is no way to solve this model in closed form, but it would be easy to do so numerically. One would start by computing for each group the value of the game if the decision reached coalition W, i.e. the last potentially winning coalition. This would depend on the decision of the members of coalition W to play C or P, as well as the amount of equilibrium passing (in case of conflict) from the groups not in W. One would then move to coalition $W-1$ (the preceding one in the natural order) and compute for each group in $W-1$ the relative payoff of exploiting as members of $W-1$ or let the game move on to coalition W. One would then proceed recursively backward all the way to coalition 1 in W. Note that some coalitions in W may pass on the opportunity to exploit even if they prefer conflict to peace, if some of the groups in that coalition prefer to be part of a smaller exploiting coalition further down the natural order (and predict that the game will arrive to that coalition). Using this algorithm, for each vector of incomes y_i, ex-ante group sizes n_i, matrix of bilateral passing costs ϕ_{ij}, resource rents z, and exploitation cost δ one can predict whether an exploitation equilibrium will prevail.

Despite the elusiveness of general closed-form results, it should be easy at this point to see that the model will share qualitative features of the baseline model. For example, there will be no conflict if all the ϕ_{ij}s are close to zero, as all potentially winning coalitions will be infiltrated to the point of making exploitation pointless. By the same token, at least some of the ϕ_{ij}s need to be reasonably large to make exploitation worthwhile. Hence, the model still implies that ethnic distance is a key determinant of conflict. Similarly, very large values of z will trigger more passing and thus discourage conflict. The model will thus generate a similar inverted-U relation between z and conflict (holding constant the ethnic structure) as the baseline model.

4.4 Dynamics

The final extension we discuss is an infinite-horizon version of our static model. The primary advantage of looking at an explicit dynamic formulation is to make more precise some of our relatively informal discussion of the model's implications for changes over time in some of the exogenous parameters of the model - most notably the flow of resource rents z. However a general treatment of the dynamic properties of the model is beyond the scope of the present contribution so we limit the discussion to some specific examples for the time path of z.

Time is discrete and all individuals are infinitely lived (or belong to infinitely-lived, altruistic dynasties). In each period there is a continuum of agents of mass 1. In period t

a measure n_t of these agents belong to group A, while a measure $1 - n_t$ belong to group B. Of the agents in group A, some have always been in A, while others have passed from B to A sometimes in the past. No one can tell the original members of A from the passers. Each agent receives an inexpropriable per-period endowment: y_A is the per-period endowment of agents who have always been in group A, and y_B is the endowment of agents of group B, and of agents who have switched to A sometimes in the past. Both y_A and y_B are constant over time.[34] In every period there also is a flow of expropriable resources z_t. z_t follows an exogenous, deterministic time-series path, to be specified below.

At the beginning of each period members of group A choose between two actions: C or P. We assume that the choice maximizes the present value of utility of agents who have always been in group A. This assumption could be motivated in a number of ways. Perhaps group A is ruled by a chief or a council of elders with an old lineage. Or perhaps the decision is taken collectively, but if any member of group A voted or acted in a way inconsistent with the interests of original members he would give himself away as an infiltrator.[35] The rest of the period game is as in the baseline model. If group A has played the P action, peace ensues, while if A has played the C action there is exploitation, with period payoffs specified as in the baseline model. Finally, some members of B (who have not yet passed in previous periods) may decide to pass. The post-passing group size is n_{t+1}.

There is no technology for saving, so once total incomes have been determined, each agent simply consumes his income. Utility is linear in consumption, so all decisions are taken in order to maximize

$$\sum_{t=0}^{\infty} \beta^t c_t.$$

The exogenous variables are the initial group sizes, n_0 and the history of z, $\{z_t\}_{t=0}^{\infty}$. The state variables at time t are n_t and the remaining history of z, $\{z_\tau\}_{\tau=t}^{\infty} \equiv \zeta_t$.The exogenous parameters are y_A, y_B, ϕ, δ, Δ and β.

We look for *Markov perfect* equilibria, which in this context implies among other things that we let each agent's action in period t depend only on the state variables in period t (and not, for example, on the specific history of the game up to period t, nor on communication occurring among players in period t), and that we specify strategies that are utility maximizing both on and off the equilibrium path. In the online appendix we

[34] To seme extent this is without loss of generality, as the main driver of our results is the ratio of the ys to z, and we will let z vary. However a truly satisfying dynamic extension would let the ys change *endogenously* over time. It is possible to speculate that such an extension could generate conflict-underdevelopment traps, but we leave this to future work.

[35] The assumption simplifies the model considerably, because it allows us to ignore potential conflicts of interest between original members and passers in deciding whether to play C or P. The conflict of interests could arise because the individual incomes y^A and y^B are potentially different.

prove that for any history ζ_0 such that $\lim_{t \to \infty} z_t$ exists and is finite, there exists a unique Markov-perfect equilibrium.

The equilibrium dynamics of the model clearly depend on the particular path z_t follows. We focus on simple, tractable cases that deliver sharp predictions. The simplest case is obviously the case where z is constant: $z_t = z$, for all t. Not surprisingly in this case the economy is either permanently in conflict or permanently at peace. The comparative statics of conflict with respect to the model's parameters are identical to those in the baseline model. The proof of this is part of the proof of existence and uniqueness, referenced above.

The next-simplest case is one where z_t changes only once, from a high to a low value, at some date T. In particular,

$$z_t = \begin{cases} z_h & \text{for} \quad t < T \\ z_l & \text{for} \quad t \geq T, \end{cases}$$

with $z_h > z_l$. This example is meant to capture in stylized fashion the structural transformation of the economy from a resource-intensive mode of production, particularly one based on heavy use of land as an input, as in agriculture, to one where (domestic) natural resources play a smaller role, such as manufacturing and services.

This case is studied formally in the online appendix. Despite its extreme stylized nature, this example gives rise to a surprisingly rich variety of possible dynamic paths for ethnic relations. In particular, depending on parameter values, there are five possible "histories" of ethnic relations. Naturally it is possible to have permanent peace, for example when both z_h and z_l are very low, so that conflict is never a very attractive option, or when ϕ is very low, so that irrespective of z conflict triggers too much passing. It is possible to have permanent conflict, when ϕ is sufficiently high that there is little passing and both z_h and z_l represents sufficiently attractive booties. Another fairly natural dynamic path for conflict that is possible in this example is to have conflict for $t < T$ and peace for $t \geq T$. This would arise, for example, when passing is fairly costly, z_h is large enough to make conflict worthwhile to group A, but z_l is too small. In this case, group A exploits group B while z is high, and shifts to peace when z is low. In terms of Figure 1, this case can be interpreted as crossing from the conflict into the peace region as z falls from an intermediate value to a low value. Perhaps more surprisingly the structural-transformation example can also generate the reverse pattern: peace for $t < T$ and conflict for $t \geq T$. Intuitively this is analogous in Figure 1 to moving from the peace region to the right of the conflict area, where peace is "enforced" by the expectation of large amounts of passing, into the conflict region as z falls from a high to an intermediate value.

But the most surprising new insight from looking at this dynamic example is that the conflict-peace outcome can change more than once. In particular, there exist combinations of parameters where we observe *two changes* in conflict status. Specifically, there is a date $T_0 < T$ such that we have: peace for $t < T_0$; conflict for $T_0 \leq t < T$, and peace for $t > T$. The intuition for this result is that the amount of passing from B to A depends on the number of periods that A will exploit B. If A starts exploiting "too early" it will trigger excessive passing, as passers will be able to recoup their passing cost over many periods. It may therefore be optimal for A to wait and begin exploiting only when only few (if any) members of B decide that it is worthwhile to pass. This insight is important because observers of ethnic conflict have often remarked that ethnic groups that have peacefully coexisted for many years seem to suddenly begin clash even when there are have been no obvious economic shocks to trigger the fighting. Our model can rationalize these observations.

Another simple case is the case where z_t changes only once but, in this case, it changes up rather than down. I.e.

$$z_t = \begin{cases} z_l & \text{for} \quad t < T \\ z_h & \text{for} \quad t \geq T. \end{cases}$$

This could be interpreted as capturing in stylized fashion the discovery of new oil fields or other mineral natural resources. This example turns out to be consistent, depending on parameter values, with the exact same possible set of histories of conflict as the structural-transformation example. It is straightforward to see that the cases of permanent peace, permanent conflict, one-time switch from peace to conflict, and one-time switch from conflict to peace all can arise. More surprising is perhaps that in the case of two switches the pattern is again <<peace - conflict - peace>>. However the intuition is straightforward. This outcome arises if z_h is so high that it triggers too much passing. Hence, group A chooses peace after T. Before T, z_l is still large enough that A would like to exploit B, and it turns out that if the period of exploitation is short enough only few or no members of B pass. There is therefore an optimal *stopping* time before the windfall occurs, which leads to outcomes observationally similar to the optimal *waiting* time in the structural-transformation example. One interesting implication of the last two examples, therefore, is that, if a country has only monotonic changes in resource endowments, then it is difficult to get a <<conflict-peace-conflict>> pattern, while it is relatively easy to generate a <<peace-conflict-peace>> pattern.

The last case we examine is the case of a temporary windfall. In particular,

$$z_t = \begin{cases} z_l & \text{for} \quad t < T_l \text{ and for } t \geq T_h \\ z_h & \text{for} \quad T_l \leq t < T_h. \end{cases}$$

This case requires (even more) lengthy calculations that we have not performed, largely because the analysis of the previous examples allows us to predict with considerable confidence the kind of patterns that arise in this case. All the possible qualitative patterns of conflict, peace, and passing of the previous two cases are again possible here, and are supported by the same reasoning. However in this case it is also possible to observe <<conflict-peace-conflict.>> If group A follows a strategy of sharing peacefully during periods when z is high, and limits exploitation to periods when z_l is low, it can prevent passing from B to A altogether. If z_l is large enough, this may be an optimal strategy.

5 Historical Examples

5.1 Pigmentation

In the United States no other ethnic group stands out for its troubled relationships with the white majority (and other groups, for that matter), and for its persistently disadvantaged socioeconomic status, as the African-Americans. Our theory suggests this may in part be due to the fact that African-Americans are also the ones who most stand out visually: they are "black," as opposed to "white." Hence, the greatest amount of conflict is associated with the greatest ethnic distance, ϕ, as suggested by our theory.

Of course African-Americans come in different shades of black, so ϕ varies within this group as well. Consistent with our theory, an increasing body of work shows that light-skinned African-Americans have persistently (i.e. from the pre-civil war era to the present day) had better outcomes (nutrition, education, income, wealth) than darker-skinned ones [Bodenhorn and Ruebeck (2007) survey this literature]. In a striking recent contribution, Goldsmith et al. (2006) have shown that the light-skin premium is non-monotonic: light-skin blacks have wages that are essentially the same as whites (controlling for the usual demographic characteristics), while medium and dark-skinned blacks have wages that are indistinguishable from each other and significantly below those of whites and light-skinned blacks. As the authors conclude, this suggests that << Employers ... in many cases, ... believe they are hiring someone who is just as white as they are themselves.>> (p. 245). Another striking finding by Gymah-Brempong and Price (2006) is that blacks with darker hues receive longer prison sentences than light-skin ones for the same crimes. A light-skin premium has also been documented for Mexican-Americans [Murguia and Telles (1996), Mason (2004)].[36]

The black-white conflict in America is particularly striking because there would have been no shortage of alternative (or additional) minorities to discriminate and exploit: Irish,

[36] The other group that is both distant from the white majority and historically greatly exploited is of course the Native Americans. As argued in Section 4, our model can explain both their tragic experience in the 19th century, and their currently privileged status. Asians – another ethnically distant group – have also been singled out, witness for example the detention camps during World War II. But their "luck" was to arrive in the US mostly during the industrialization phase, when the incentives for exploitation had already declined considerably.

Italians, Jews, Poles, and other migrant communities could have been equally attractive objects. Why haven't they been targeted in the way blacks have? According to our theory, this is simply because continued exclusion of these white immigrants would have been too costly to enforce given the close physical proximity, or low ϕ, with the Anglo elite. Had the latter tried to perpetuate such discrimination, there would now be many more Americans with names like Coleman, and many fewer with names like Caselli, as the holder of the latter would have switched in mass to the former. Hence, the "Anglo" majority refrained from a systematic attempt to disenfranchise the white immigrants - who have therefore been able to preserve their ancestral identity.[37]

It is not that these immigrant communities did not suffer their own share of initial discrimination and exploitation, but that the "Anglos" have "backed off" fairly soon, say within one or two generations. One or two generations is probably the time required for the newcomers (i.e. their descendents) to learn the language well enough, and to overcome the physical baggage of pre-migration malnutrition, that they would be able to disguise their ancestry – if necessary. Of course in equilibrium this is not necessary. Also, it is not that it would have been impossible for the Anglos to set up a vast bureaucracy keeping track of everyone's ancestry, but in the case of physically similar individuals it was evidently too costly.[38]

It is now increasingly widely recognized that discrimination against blacks has been slowly but steadily declining over the last century in the US. Over the same span of time, the economy has undergone a huge structural transformation from largely agricultural to industrial and then service-based. This transformation has meant that land and other natural-resource rents have become an increasingly insignificant share of total income. In our model, this is equivalent to a decline in z/y. Hence, the model does indeed provide a possible interpretation for the gradual and ongoing phasing out of discrimination against blacks.

The South-African case presents of course many analogies with the US case, and our model describes it even better, if one identifies the dominant group as the one that has greater total resources. While whites are a numerical minority in South-Africa, their per-capita resources so dwarf those of the black majority that their "firepower" is greater. This allowed them to establish the apartheid regime. The rich mineral resources of the country, coupled with the small number of whites to divide them, provided the incentive. In other

[37] Another distinguishing factor for blacks vis-a-vis other groups is that their ancestors came to the US as slaves. But it is not entirely clear why, after the abolition of slavery, people of anglosaxon descent would want to specifically target descendants of slaves (rather than descendandants of voluntary immigrants) for exploitation and discrimination. The evidence on light-skin premia discussed above also seem hard to reconcile with the view that blacks are differentially discriminated solely because of slavery.

[38] Imagine enforcing a policy of separate water fountains for Italian-Americans!

words South Africa has historically been a high ϕ, low n, and high z/y country, making it "ideally suited" for exploitation. Over time, as the economy grew and diversified away from the primary sector, and the sanction regime against the white government became increasingly aggressive, z/y fell, and the cost of maintaining the regime became too large relative to the benefits.[39] The whites decided then to start a transition to the "no conflict" equilibrium. The model of Section 4.1 suggests that the nature of the apartheid regime may have changed from "by choice" to "preemptive" before further changes in the state variables made it safe enough for the whites to relinquish power.

One could keep going with examples of conflict or exploitation where differences in skin color plays a critical role in enabling members of one group to pinpoint members of the "other" group. The Dominican police openly uses skin complexion and texture as a criterion for identifying "Haitians" to be mass deported from the country.[40] Humphreys and ag Mohamed (2005) compare Southern Senegal and Northern Mali, and argue that in the former ethnic tensions are much less severe than in the latter – despite broadly similar socioeconomic conditions – because in Mali the minorities (Tuareg and Maures) are more readily physically distinguished from the majority than in Senegal (Diola).

5.2 Body size

The black-white gradient is of course an important physical source of ethnic distance, but by no means the only one. An illustration of this is provided by the Rwandan case, where so-called "Hutus" and "Tutsis" have been in extremely bloody – if somewhat intermittent – conflict since the end of he colonial era. Much has been written about the artificial birth of the Hutu-Tutsi split as part of the divide-and-conquer strategy of Belgium, the colonial power. For us, what is notable is the rich anecdotal evidence that physical attributes play a critical role in the conflict. On average, "Tutsis" are taller and more slender, they have somewhat lighter skin, and thinner noses. Indeed, the Belgian colonists classified a person as Tutsi if they had a long nose (or ten cows). During the genocidal campaign that led to the death of more than one half of a million people in 1994, "Hutus" reportedly made use of these visual cues to identify potential victims. This of course implies that many "Hutus"

[39] Mineral Sales as a fraction of GDP for South Africa declined from 25 percent in 1980 to 11 percent in 1994 (the end of apartheid).

[40] According to Human Rights Watch (2002) "the Dominican authorities have conducted mass expulsions of Haitians and Dominico-Haitians. ... Snatched off the street, dragged from their homes, or picked up from their workplaces, 'Haitian-looking' people are rarely given a fair opportunity to challenge their expulsion during these wholesale sweeps. Questioned by Human Rights Watch as to how undocumented Haitians are identified, the subdirector for Haitian affairs of the Dominican government's migration department insisted that they can be spotted ... Noting that Haitians also have 'rougher skin,' the subdirector declared that 'they're much blacker than we are. They're easy to recognize.' "

were also victimized, as they did not fit the stereotypical description (for example they were too tall or too thin). To us, the willingness of the genocide's perpetrators to commit such "type I" errors strongly supports the "group enforcing" interpretation of ethnic conflict over explanations based on hatred or within-group altruism.[41] To put it crudely, pre-genocide Rwanda was a country on the verge of an impending famine, mainly due to excess population pressure on the land. A genocide was one way to relieve such pressures, and targeting Tutsis, or rather – as it turned out – the tall and thin, assured that the designated victims could not infiltrate the dominant group (i.e., in this case, escape the killers).[42]

The use of height in the Rwandan case raises the interesting question of why is height not used more systematically around the world as a boundary-enforcing marker. In particular, it would seem that in ethnically-homogeneous countries one should observe winning groups of individuals below or above a certain height threshold. We speculate that the typical shape of the height distribution makes it unsuitable to the purpose of boundary enforcement. In particular, within ethnic groups (and gender) height distributions are known to be normal (and thus are unimodal and with thin tails). This implies that any group boundary that makes conflict worthwhile must be drawn at a point which leaves large masses of people on both of its sides. Because height is not easily measured perfectly, this means that the number of type I and type II errors is vast, and the scheme may become unworkable.[43]

5.3 Language

Another way this is done is through language. Examples of this go literally back to biblical times – with tales of warring tribes using the pronunciation of certain words to establish who should be slaughtered [Judges 12:4-6] – and stretch to 21st century Northern Ireland, where, as reported by *The Economist* of June 15th, 2002, "a group of masked men [entered a school and] demanded that students produce identification or repeat the alphabet. Many Catholics

[41] The killers also targeted so-called "moderate Hutus," i.e. Hutus who did not cooperate in the genocide.

[42] The infamous Radio Mille Collines broadcast: "Those of you who live along the road, jump on the people with long noses, who are tall and slim, and want to dominate us." (Peterson, 2000, p. 327). Very similar considerations, only in reverse, apply to Burundi, where the tall and thin Tutsis dominate the Hutus. There, too, physical characteristics play an explicit role. For example, the army has a "height-by-girth" requirement that so happens to exclude from the ranks the average Hutu. And there, too, changing economic circumstances affect the incentive of the dominant group to tighten the exploitation equilibrium: when coffee prices (the export crop) fall, the relative return to government jobs increase, and the Tutsis fight Hutu "infiltration" more fiercely (Gurr, 2000).

[43] The (gender-specific) Rwandan height distribution is a mixture of the Tutsi distributuion and the Hutu distribution, which have different means. Hence it is conceivable that the resulting overall distribution is bimodal, and the valley between the two modes could conceivably be sufficiently deep such that drawing the boundary near the trough minimizes the number of type I and type II errors. It is also important to remember that height was only one of several physical markers used during the genocide.

pronounce the letter "h" differently to Protestants, with an aspiration influenced by the Irish language. Students were evacuated before it became clear what was planned for people with the wrong accent."[44] Another example is provided by the 1937 massacre of Haitians in the Dominican Republic, where victims were identified by their inability to pronounce the word *perejil* (parsley) "correctly" [e.g. Danticat (1998), who also highlights the occurrence of type I errors.]

5.4 Religion

Religion is often cited as a conflict-inducing cleavage. The effectiveness of religion as a signpost for group boundaries depends on the psychic costs involved in conversion, and is therefore likely to vary considerably across religions, across pairs of religions (namely conversion to religion A may be psychologically less costly than conversion into religion B, for someone originally raised in religion C), and across people of the same original group. It is therefore unlikely that religious differences will be found *systematically* to relate to conflict. Indeed, Alesina et al. (2003) find that religious fractionalization does not significantly predict the rent-seeking policy distortions usually associated to other types of ethnic fractionalization. Similarly, examining a large cross-section of conflicts, Fox (1997) finds that in only a small minority do religious issues play more than a marginal role.

Anecdotal evidence confirms that in at least some cases individuals are able and willing to shed their religion to respond to external circumstances, particularly discrimination against one's group. In post-Reform Europe entire populations switched back and forth between Catholicism and Protestantism, as the political alliances of their princes switched back and forth between the Pope, the Emperor, and other potentates.[45] In Fascist Italy many Jews converted to Catholicism to escape discrimination. In modern-day India it is common for lower-caste Hindus to convert to the Muslim or Catholic faiths, which are relatively less discriminated against.

Generally speaking, therefore, while some individuals seem clearly unwilling to shed their religious affiliation at any cost, our theory suggests that physical differences should be more systematically related to conflict than religious ones. A stark example of color working better than religion as a group enforcing mechanism is recounted by Horowitz (1985,

[44] We pointed out above that language-based markers can be overcome over a couple of generations, so at first sight it may seem unlikely that they would sustain a multi-generational conflict such as the Northern Irish one. However, Northern Ireland may be a special case of our model where the two groups have virtually equal strength, so that there is no clear winner or loser. As we show in Section 4 in this case conflict and persistent ethnic differentiation can coexist.

[45] And the so-called "religious wars" were mostly international wars that happened to involve the Papacy as one of the territorial contenders.

p.43): "In seventeenth century North-America, the English were originally called 'Christians,' while the African slaves were described as 'heathens.' The initial differentiation of groups relied heavily on religion. After about 1680, however, a new dichotomy of 'whites' and 'blacks' supplanted the former Christian and heathen categories, for some slaves had become Christians. If reliance had continued to be placed mainly on religion, baptism could have been employed to escape from bondage. Color provided a barrier seemingly both 'visible and permanent.' " [46]

5.5 No conflict

So far our examples have involved cases of conflict, and we have asked whether our model can shed light on these episodes. In principle, we would like to offer examples were there is no conflict because there is insufficient distance. Doing so is difficult, however, because such examples in the limit become tautological: there is no ethnic conflict in Sweden because the ethnic distance among all Swedes is virtually zero! Nevertheless, we venture here that the model may be useful in explaining Norway's escape from the "natural resource curse." Because of its rich oil reserves Norway is probably a high z/y for the purposes of our model. While most countries with a high share of natural resources in income seem to have fraught social relations and poor economic outcomes, Norway has neither. Perhaps its high degree of ethnic homogeneity is the key to this success. A similar example may be Botswana, where the physical similarity of different groups is cited by Acemoglu et al. (2003) as a possible reason why conflict over natural resources has not erupted there.[47]

A more subtle example of ethnic proximity leading to relatively peaceful ethnic relations may perhaps be found in the Indian case.[48] In a world were all ethnic cleavages are equally important, for a very poor, over-populated country such as India, the 13% Muslim minority should constitute an attractive target for massive exploitation, if not for Rwandan-style elimination. Instead, Muslims have for the most part equal economic and political rights. Our

[46] An argument could probably be made that a similar shift occurred at various times from religious to racial anti-Semitism, for example after the expulsion of Jews from Spain.

[47] The only shadow on Botswana's reputation as a model of ethnic harmony is cast by the advocacy group Survival International's claim that the government is mistreating the San, a tribe of Bushmen. Surprise surprise, the Pigmy-sized Bushmen have very high ϕ vis-a-vis other Southern Africans.

[48] There seemingly is a lot of communal violence in India, so some readers may find it paradoxical to treat India as a case of relative ethnic harmony. The fact, is, however, that *relative to the size of the population*, ethnic violence in India is actually fairly trivial. For example, Varshney (2002) estimates that between 1950 and 1995 there was a total of 7,173 deaths caused by communal rioting, which leads to an average of 155.9 deaths per year for those 46 years. In contrast, Pakistan seems to be engaged in repression of the Hindu minority. Although objective evidence is difficult to obtain, anecdotal evidence suggests that passing seems to be taking place, especially in the form of switching last names from Hindu to Muslim.

speculation is that India enjoys this relative harmony precisely because the ethnic distance between Muslims and Hindus is quite modest: too oppressive an exploitation equilibrium by the Hindu majority would be unsustainable in the face of mass ethnic switching by the Muslims.

6 Conclusions

In this paper we attempted to developed a new, simple explanation for the salience of ethnicity in exploitation and conflict around the world. Ethnicity provides a technology for group membership and exclusion which is used to avoid indiscriminate access to the spoils of conflict. Without such a technology groups become porous and the spoils of conflict are dissipated. In relating the incidence of ethnic conflict to variables such as group size and the share of expropriable assets in overall wealth, we were able to derive various implications that seem to shed light on a wide variety of historical episodes of conflict (and lack thereof).

It is natural to try to use the insights of the model to suggest policy recommendations to minimize the incidence of conflict along ethnic lines. The model suggests that economic development alone will remove the incentives for ethnic conflict, particularly if it is accompanied, as it often is, by a structural transformation where control over natural (expropriable) resources plays a smaller and smaller role The paper therefore adds to the list of good things that come with growth, beyond higher consumption. It also offer a foundation for the assertion by Habyarimana, et. al., (2008) that "modernization may be the antidote to ethnic nationalism rather than its cause."

Secondly, the model of Section 4 suggests that ethnic conflict is sometimes preemptive, in that the stronger group preempts with conflict to protect itself from aggression by a smaller group. If the smaller group could commit to no conflict, then the larger group would feel no need for preemption. This is certainly not a paper about how to form institutions that facilitate commitment, but it highlights the role of such institutions in avoiding conflict.[49]

Perhaps most interestingly, the paper suggests that any policy that blurs sharp distinction between groups will reduce the incidence of ethnic conflict. One such policy is the promotion of intermarriage. Policies such as tax breaks for interracial couples (which may increase with the number of children) and affirmative action programs for mixed-race individuals may help achieve such a goal. Policies to encourage interracial adoption could also be

[49] Policies that increase transparency on the magnitude and destination of natural-resource export revenues, such as the Extractive Industries Development Initiative (EITI), in which participating governments and oil companies agree to disseminate detailed information on quantities extracted, revenues, and royalties paid to the government, also find support in our model. Likewise for certification processes that keep conflict diamonds out of rich-country markets, as was done for the fighting in Sierra Leone.

justified along the lines of our model. The surest path to a world without racism is a world without races.

There are also policies our analysis does not provide support for. One is to resort to drastic measures of creating nation states along ethnic lines to avoid ethnic conflict, as Muller (2008) has recently argued. This may work in some cases, but our results suggest that the creation of such a state may even lead to the endogenous creation of new ethnic groups that compete for the economy's resources. Recent tensions in South Sudan among groups who previously fought together for independence underscores this concern. Our paper also provides very little *prima facie* support for increased aid flows to countries in or at risk of conflict. This is because aid flows are very similar to increases of other appropriable resources, and may therefore increase the incentive of groups to fight over them. Peterson (2000) makes a very compelling case that aid exacerbated the conflict in Southern Sudan.

Although we presented several historical examples of ethnic conflict that are consistent with the premise of this paper, there is a clear sense in which the data required to fully test the implications of the model are not yet available. Our theory highlights the role of ethnic "distance" in leading to ethnic conflict: ceteris paribus, ethnic groups are more likely to clash the more pronounced the differences that mark the ethnic cleavage. Systematic data on ethnic distance has not yet been collected. Extending the empirical results of Habyarimana, et. al. (2007) for Uganda to many other countries would be a great start in this direction. Given the importance of ethnic conflict in the world, we hope that research such as ours would motivate the collection of this type of data.

7 References

Acemoglu, Daron; Simon Johnson and James Robinson (2003): "An African Success Story: Botswana,", in Dani Rodrik (ed.), *In Search of Prosperity: Analytic Narratives on Economic Growth*, Princeton; Princeton University Press.

Ahlerup, Pelle and Ola Olsson (2011): "The Roots of Ethnic Diversity," working paper, University of Gothenburg.

Akpam (2008): *Say You're One of Them,* Little, Brown, and Company.

Alesina, Alberto; Baqir, Reza, and William Easterly (1999): "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, Vol. 114 (4). p 1243-84. November 1999.

Alesina, Alberto, and Eliana La Ferrara (2000): "Participation in Heterogeneous Communities," *Quarterly Journal of Economics*, Vol. 115 (3). p 847-904, August 2000.

Alesina, Alberto; Devleeschauwer, Arnaud; Wacziarg, Romain; Easterly, William, and Sergio Kurlat (2003): "Fractionalization," *Journal of Economic Growth*, vol. 8, no. 2, June, pp. 155-194..

Azam, Jean-Paul (1995): “How to Pay for the Peace?” *Public Choice*, 83(April), 173-184.

Azam, Jean-Paul (2001): “The Redistributive State and Conflicts in Africa,” *Journal of Peace Research*, 38(4), 429-444.

Anderson, B. (1983): *Imagined Communities*, Verso, London.

Bates, Robert H. (1974): "Ethnic Competition and Modernization in Contemporary Africa." *Comparative Political Studies* (January).

Bates, Robert H. (1982): “Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa,” in *State versus Ethnic Claims: African Policy Dilemmas*, Donald Rothchild and Victor A. Olunsorola (eds.), Boulder, CO: Westview Press.

Barth, Fredrik (1969): *Ethnic groups and Boundaries*, Universitetforlaget.

Besley and Persson (2011): “The Logic of Political Violence,” *Quarterly Journal of Economics*, forthcoming.

Berman, Eli (2000): “Sect, Subsidy, and Sacrifice: An Economist’s View of Ultra-Orthodox Jews,” *Quarterly Journal of Economics*, 115(3).

Berman, Eli (2009): *Radical, Religious and Violent: The New Economics of Terrorism.* MIT Press

Berman, Eli and David D. Laitin (2008): “Religion, Terrorism and Public Goods: Testing the Club Model,” *Journal of Public Economics*, 92, 942-1967.

Bodenhorn, Howard, and Christopher S. Ruebeck (2007): “Colorism and African-American Wealth: Evidence from the Ninetheenth Century South,” *Journal of Population Economics*, 20(3) 599-620.

Bossert, Walter; D’Ambrosio, Conchita, and Eliana La Ferrara (forthcoming): “A Generalized Index of Ethno-Linguistic Fractionalization,” *Economica.*

Brass, Paul R. (1997): *Theft of an Idol: Text and Context in the Representation of Collective Violence.* Princeton University Press.

Cederman, Lars-Erik (2002): “Nationalism and Ethnicity,” in Carlsnaes, Walter, Thomas Risse, and Beth Simmons (eds.), *Handbook of International Relations*, 409-428. Sage.

Cederman, Lars-Erik, and Luc Girardin (2007): “Beyonf Fractionalization: Mapping Ethnicity into Nationalist Insurgencies,” *American Political Science Review*, 101, 1, 173-185.

Cederman, Lars-Erik, Halvard Buhaug, and Jan Ketil Rod (2009): “Ethno-Nationalist Dyads and Civil War: A GIS-Based Analysis,” *Journal of Conflict Resolution*, 53, 4, 496-525.

Chandra, Kanchan (2004): *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India*, Cambridge University Press

Ciccone (2011): “Economic Shocks and Civil Conflict: A Comment,” *American Economic Journal: Applied Economics*, forthcoming.

Collier, Paul. and Anke Hoeffler (2004): “Greed and Grievance in Civil War,” *Oxford Eco-*

nomic Papers, 56, 563-595.

Dando, Malcolm (2004): *Biotechnology, Weapons, and Humanity II*, British Medical Association.

Danticat, Edwige (1998): *The Farming of Bones*, Penguin Books.

de Figuereido, Rui J. P. Jr and Barry R. Weingast (1999): “The Rationality of Fear,” in *Civil Wars, Insecurity and Intervention*, edited by Barbara F. Walter and Jack Snyder, Colu,bia University Press. Pp. 261-302.

Desmet, Klaus, Ignacia Ortuno-Ortin, and Romain Wacziarg (forthcoming), “The Political Economy of Linguistic Cleavages,” *Jorunal of Development Economics*.

Easterly, William and Levine, Ross (1997): “Africa’s growth tragedy: policies and ethnic divisions.” *Quarterly Journal of Economics*, 112, November, 1203-1250.

The Economist (2002): “How do you pronounce hate?” 15 June, 56.

Esman, Milton (1994): *Ethnic Politics*, Cornell University Press.

Esteban, Joan-María, and Debraj Ray (1994): “On the Measurement of Polarization,” *Econometrica*, 62(4), 819-851.

Esteban, Joan and Debraj Ray (1999): “Conflict and Distribution,” *Journal of Economic Theory*, 87, 379-415.

Esteban, Joan and Debraj Ray (2008a). “On the Salience of Ethnic Conflict." *American Economic Review*, 98(5): 21852202.

Esteban, Joan and Debraj Ray (2008b). “Polarization, Fractionalization and Conflict." *Journal of Peace Research*, 45(2): 163-182.

Esteban, Joan and Debraj Ray (2011). “A Model of Ethnic Conict.” *Journal of the European Economic Association*.

Esteban, Joan and Debraj Ray (forthcoming). “Linking Conflict to Inequality and Polarization." *American Economic Review*.

Esteban, Joan, Laura Mayoral and Debraj Ray (2010). “Ethnicity and Conflict: An Empirical Study," working paper, Institut d’Analisi Economica.

Evans, Julie, Dietmar Kahles and Catriona Bate (1993): “1991 Census Data Quality: Aboriginal and Torres Straight Islander Countes,” Australian Bureau of Statistics, Population Census Evaluation, Census Working Paper 93/6.

Fearon, James (2003): “Ethnic Structure and Cultural Diversity by Country,” *Journal of Economic Growth*, 8, 2, 195-222.

Fearon, James (1999): “Why Ethnic Politics and “Pork” Tend to Go Together,” unpublished, Stanford University.

Fearon, James and David Laitin (2000): “Violence and the Social Construction of Ethnic Identity,” *International Organization,* 54, 4, Autumn, 845-877.

Fearon, James and David Laitin (2003): "Ethnicity, Insurgency, and Civil War, *American Political Science Review,* 91, 1, February, 75-90.

Fernandez, Raquel and Gilat Levy (2008): "Diversity and Redistribution," *Journal of Public Economics*, 92, 925-43.

Fletcher, Erin and Murat Iygun (2010): "The Clash of Civilizations: A Cliometric Investigation," working paper, University of Colorado.

Fox, Jonathan (1997): "The Salience of Religious Issues in Ethnic Conflicts: A Large-N Study," *Nationalism & Ethnic Politics,* 3, 3, Autumn, 1-19.

Gershenson, Dmitriy, and Herschel I. Grossman (2000): "Civil Conflict: Ended or Never Ending?" *Journal of Conflict Resolution*, 44, 6, December, 807-821.

Glaeser (2005): "The Political Economy of Hatred," *Quarterly Journal of Economics*, CXX(1), 45-86.

Goldsmith, Arthur H., Darrick Hamilton, and William Darity Jr. (2006): "Shades of Discrimination: Skin Tone and Wages," *American Economic review*, 96 (2), 242-45.

Grossman, Herschel I (1991): "A General Equilibrium Model of Insurrections," *American Economic Review*, 81, 4, 912-921.

Grossman, Herschel I. (1999): "Kleptocracy and Revolutions," *Oxford Economic Papers*, 51, April, 267-283.

Grossman, Herschel I. and Juan Mendoza (2003): "Scarcity and Appropriative Competition," *European Journal of Political Economy*, 19, 4 (November), 747-758.

Guiso, Luigi, Paola Sapienza and Luigi Zingales (2009): "Cultural Bis in Economic Exchange," *Quarterly Journal of Economics*, 124(3), 1095-1131, August.

Gurr, Ted Robert (2000): *Peoples Versus States: Minorities at Risk in the New Century*, United States Institute of Peace.

Gymah-Brempong, Kwabena, and Gregory N. Price (2006): "Crime and Punishment: And Skin Hue Too?" *American Economic review*, 96 (2), 246-50.

Gurr, Ted Robert (2000): *Peoples Versus States: Minorities at Risk in the New Century*, United States Institute of Peace.

Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein (2007): "Placing and Passing: Evidence from Uganda on Ethnic Identification and Ethnic Deception," unpublished, Georgetown University.

Habyarimana, James, Macartan Humphreys, Daniel N. Posner, Jeremy M. Weinstein, Richard Rosecrance, Arthur Stein, and Jerry Z. Muller (2008): "Is Ethnic Conflict Inevitable?," *Foreign Affairs*, July/August.

Hardin, Russel (1995): *One For All: The Logic of Group Conflict*, Princeton University Press.

Harris, David R., and Jeremiah J. Sim (2002) "Who is Multiracial? Assessing the Complexity of Lived Race," *American Sociological Review*, 67, 614-627.

Hirshleifer, J. (1995) "Anarchy and its Breakdown," *Journal of Political Economy*, 103, February, 26-52.

Horowitz, Donald L. (1985): *Ethnic Groups in Conflict*. Berkeley, CA: University of California Press.

Humphreys, Macartan and Habaye ag Mohamed (2005): "Senegal and Mali: A Comparative Study of Rebellions in West Africa" in Collier and Sambanis (eds.) *Understanding Civil War Africa: Africa Evidence And Analysis.*

Human Rights Watch (2002): *Dominican Republic*, available at http://www.hrw.org/reports/2002/domrep/index.html.

Jeganathan, Pradeep (1997): *After a Riot: Anthropological Locations of Violence in an Urban Sri Lankan Community*. PhD dissertation, University of Chicago.

Kurzban, Tooby, and Cosmides (2002), *Proceedings of the national academy of sciences.*

Laitin, D. (2000). "What is a language community?" *American Journal of Political Science*, 44, 142–54.

Lake, David A. and Donald Rothchild (1996): "Containing Fear: The Origins and Management of Conflict," *International Security*, 21(2), 41-75.

Lieberson, Stanley and Mary Waters (1993): "The Ethnic Responses of Whites: What Causes Their Instability, Simplification, and Inconsistency?" *Social Forces*, 72(2), December, 421-450.

Mason, Patrick L. (2004): "Annual Income, Hourly Wages, and Identity Formation among Mexican Americans and other Latinos," *Industrial Relations*, 43, 4, 817–834, October.

Matuszeski, Janina and Frank Schneider (2006): "Patterns of Ethnic Group Segregation and Civil Conflict," unpublished, Harvard University.

Mauro, Paolo (1995): "Corruption and Growth", Quarterly Journal of Economics, CX, 3, 681–712, August.

McDermott, John (1997): "Exploitation and Growth," *Journal of Economic Growth*, 2, September, 251-278.

Miguel, Edward and Mary Key Gugerty (2005): "Ethnic diversity, social sanctions, and public goods in Kenya ," *Journal of Public Economics*, 89, 11-12, December, 2325-2368.

Miguel, Edward, Shankar Satyanath, and Ernest Sergenti (2004): "Shocks and Civil Conflict: An Instrumental Variable Approach," *Journal of Political Economy*, 112(4), 725-753.

Murguia, Edward and Edward E. Telles (1996): "Phenotype and Schooling among Mexican Americans," *Sociology of Education*, 69(4), 276-289

Montalvo, Jose G. and Marta Reynald-Querol (2005): "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review.*

Muller, Jerry (2008): "Us and Them," *Foreign Affairs*, March/April.

Nagel, Joane (1995): "American Indian Ethnic Renewal: Politics and the Resurgence of Identity," *American Sociological Review*, 60, 947-965.

Naipaul, V. S. (1990): *India: A Thousand Mutinies Now*, Deutsch.

Peterson, Scott (2000): *Me Against My Brother : At War in Somalia, Sudan and Rwanda*, Routledge.

Posner, Daniel N. (2005): *Institutions and Ethnic Politics in Africa*, Cambridge University Press, New York.

Reynal-Querol, M. (2002). Ethnicity, political systems and civil war. *Journal of Conflict Resolution*, 46, 29–54.

Robinson, James A. (2001): "Social Identity, Inequality, and Conflict," *Economic of Governance*, 2, 85-99.

Rohner, Dominic (2010): "Reputation, Group Structure, and Social Tensions," *Journal of Development Economics*.

Rohner, Thoenig and Zilibotti (2011): "War Signals: A Theory of Trade, Trust and Con‡ict," working paper, University of Zurich.

Roth, Philip (2000): *The Human Stain*.

Spolaore, Enrico, and Romain Wacziarg (2009): "The Diffusion of Development," *Quarterly Journal of Economics*, 124(2), pp. 469-530.

Toft, Monica Duffy (2003): *The Geography of Ethnic Violence: Identity, Interests, and the Indivisibility of Territory*, Princeton University Press.

van den Berghe (1978): "Race and Ethnicity: A Sociobiological Perspective," *Ethnic and Racial Studies*, 1, 4, 401-11.

van den Berghe (1981): *The Ethnic Phenomenon*, Elsevier.

van den Berghe (1995): "Does Race Matter?" *Nations and Nationalism*, 1 (3), 357-88.

Varshney, Ashutosh (2002): *Ethnic Conflict and Civil Life*, Yale University Press.

Weidman, Nils B. (2009): "Geography as a Motivation and Opportunity: Group Concentration and Ethnic Conflict," *Journal of Conflict Resolution*, 53 (4): 526-543.

Weidman, Nils B., Jan Ketil Rod (2010): "Representing Ethnic Groups in Space: A New dataset," *Journal of Peace Research*, 47 (4), 491-99.

Woodward, Susan L. (1995): *Balkan Tragedy: Chaos and Dissolution After the Cold War*. Brookinggs Institution.

Unpublished Appendix for "On the Theory of Ethnic Conflict"

Francesco Caselli (LSE) and John Coleman (Duke)

September 2011

1 Unpublished Appendix 1. Analysis of Model with open Conflict

1.1 Equilibrium group sizes

1.1.1 At node CC

At node CC in principle there could be passing from either group.

No passing from A to B We first show that members of group A never pass in CC equilibria. Define $\bar{n}_{CC}^{A}$ the hypothetical ex-post size that would make members of A indifferent between staying and passing. $\bar{n}_{CC}^{A}$ solves

$$(1-\Delta)\left(y_A + \frac{\alpha z}{\bar{n}_{CC}^{A}}\right) = (1-\Delta)\left((1-\phi)y_A + \frac{(1-\alpha)z}{1-\bar{n}_{CC}^{A}}\right)$$

so

$$\bar{n}_{CC}^{A} = \frac{-\left[(1-\alpha)z - \phi y_A\right] \pm \sqrt{\left[(1-\alpha)z - \phi y_A\right]^2 + 4\alpha z \phi y_A}}{2\phi y_A}$$

It can be shown that the "-" root is always less than 0, while the "+" root is greater than 1 (using $\alpha > 0.5$). We conclude that there is never passing from A to B at the CC node.

Equilibrium passing from B to A Now define $\bar{n}_{CC}^{B}$ the ex-post size that makes members of B indifferent between staying and passing at node CC. The condition for $\bar{n}_{CC}^{B}$ is

$$(1-\Delta)\left(y_B + \frac{(1-\alpha)z}{1-\bar{n}_{CC}^{B}}\right) = (1-\Delta)\left((1-\phi)y_B + \frac{\alpha z}{\bar{n}_{CC}^{B}}\right)$$

so

$$\bar{n}_{CC}^{B} = \frac{(z+\phi y_B) \pm \sqrt{(z+\phi y_B)^2 - 4\alpha z \phi y_B}}{2\phi y_B}$$

While the "+" root is always greater than 1, the "-" root is always between 0 and 1. We conclude that $\bar{n}^B_{CC}$ is strictly between 0 and 1 and is given by:

$$\bar{n}^B_{CC} = \frac{(z+\phi y_B) - \sqrt{(z+\phi y_B)^2 - 4\alpha z \phi y_B}}{2\phi y_B}.$$

Passing behavior in the CC node can then be summarized as follows. If $n < \bar{n}^B_{CC}$ there will be passing from B to A. If $n > \bar{n}^B_{CC}$ there will be no passing. Hence, if the node is CC we define $n'_{CC} = \max\{n, \bar{n}^B_{CC}\}$.

For future reference, we note that $\bar{n}^B_{CC}$ is an increasing and concave function of z, which passes through the origin and asymptotes to α for $z \to \infty$. Hence, n'_{CC} is a constant equal to n if $\alpha < n$ (which makes sense). Instead, if $\alpha > n$, n'_{CC} is a constant through n up to a "kink," and then it becomes increasing and concave, and converges to α. Assumptions we make below for a variety of reasons imply that we focus on cases where $\alpha > n$. The kink is at z^k defined by

$$n = \frac{(z+\phi y_B) - \sqrt{(z+\phi y_B)^2 - 4\alpha z \phi y_B}}{2\phi y_B}$$

so

$$z^k = \frac{\phi y_b n(1-n)}{(\alpha - n)}$$

1.1.2 At other nodes

The analysis of equilibrium population size at nodes CP and PC is identical to the case of exploitation in the baseline model. In particular, at node CP we have $n'_{CP} = \max\{n, \min[1, \bar{n}^B_{CP}]\}$, where

$$\overline{n}^B_{CP} \equiv \frac{z}{\phi y_B}.$$

At node PC we have $n'_{PC} = \min\{n, \max[0, \bar{n}^A_{PC}]\}$, where

$$\overline{n}^A_{PC} \equiv 1 - \frac{z}{\phi y_A}.$$

Finally, at node PP we obviously have $n' = n$.

1.2 Group B's Decision

1.2.1 At node C

We can now look at B's strategic decisions. When A has played C, B has to choose between acquiescence and fight back. B chooses fight back if

$$(1-\Delta)\left[y_B + \frac{(1-\alpha)z}{1-n'_{CC}}\right] > (1-\delta)\max\left[y_B, (1-\phi)y_B + z\right]$$

or

$$1 - \frac{(1-\Delta)(1-\alpha)z}{\max\left[0, (1-\delta)(z-\phi y_B)\right] + (\Delta - \delta) y_B} < n'_{CC}.$$

We already know the RHS is constant at n, has a kink at z^k, and then is increasing and concave asymptoting to α.

If $\phi \leq (\Delta - \delta)/(1-\delta)$ the left side as a function of z starts at 1 and decreases monotonically (though with a kink at $z = \phi y_B$) asymptoting to $1-(1-\Delta)(1-\alpha)/(1-\delta)$. Since this is always greater than α the left side and right side never intersect so B never responds to C with C.

If $\phi > (\Delta - \delta)/(1-\delta)$ the left side as a function of z starts at 1 and decreases until $z = \phi y_B$, after which it turns increasing, asymptoting (from below) to $1-(1-\Delta)(1-\alpha)/(1-\delta)$.

Assumption: $\alpha \geq 2n - n^2$

With this assumption, the kink in LHS is always to the right of z^k. Note that this assumption implies $\alpha > n$. We make this assumption exclusively for convenience so as not to have to consider too many different cases. However the assumption does not affect the qualitative results in any significant way. In particular the discussion in the text is entirely unaffected.

The assumption implies that we can ignore the horizontal segment of the RHS. We can then reformulate the problem by saying that B responds with C if

$$1 - \frac{(1-\Delta)(1-\alpha)z}{\max\left[0,(1-\delta)(z-\phi y_B)\right] + (\Delta-\delta)y_B} < \bar{n}_{CC}^B.$$

This equation has either two solutions or no solution. The solution exists if, at the kink $z = \phi y_B$ the right hand side exceeds the left hand side, or

$$1 - \frac{(1-\Delta)(1-\alpha)\phi y_B}{(\Delta-\delta)y_B} < \frac{(\phi y_B + \phi y_B) - \sqrt{(\phi y_B + \phi y_B)^2 - 4\alpha\phi y_B \phi y_B}}{2\phi y_B}$$

or

$$\phi > \frac{1}{\sqrt{1-\alpha}}\frac{(\Delta-\delta)}{(1-\Delta)}$$

Now note that the condition $\phi > \frac{1}{\sqrt{1-\alpha}}\frac{(\Delta-\delta)}{(1-\Delta)}$ also implies $\phi > \frac{(\Delta-\delta)}{(1-\delta)}$.

Now we look at the two solutions. The first one is the solution to:

$$1 - \frac{(1-\Delta)(1-\alpha)z}{(\Delta-\delta)y_B} = \frac{(z+\phi y_B) - \sqrt{(z+\phi y_B)^2 - 4\alpha z \phi y_B}}{2\phi y_B}$$

Define

$$q = \frac{(1-\Delta)(1-\alpha)}{(\Delta-\delta)}$$

then the solution is

$$z = \frac{(1-\alpha)y_B + q y_B \phi}{(q^2\phi + q)} \equiv z_{B,CC}^l$$

The other solution is implicitly given by

$$1 - \frac{(1-\Delta)(1-\alpha)z}{(1-\delta)(z-\phi y_B) + (\Delta-\delta)y_B} = \frac{(z+\phi y_B) - \sqrt{(z+\phi y_B)^2 - 4\alpha z \phi y_B}}{2\phi y_B}, \tag{1}$$

and we denote it by $z_{B,CC}^h$.

In conclusion, if $\alpha \geq 2n - n^2$ then group B responds to C with C if and only if:

$$\phi > \frac{1}{\sqrt{1-\alpha}} \frac{(\Delta - \delta)}{(1 - \Delta)}$$

and

$$z \in (z^l_{B,CC}, z^h_{B,CC}).$$

The subscript "B, CC" is a mnemonic for "B responds to C with C," and the superscripts l and h denote lower- and upper-bound values.

It is important for future reference to characterize the behavior of $z^l_{B,CC}$ and $z^h_{B,CC}$ as functions of ϕ.

By construction, we have $z^l_{B,CC} = z^h_{B,CC} = \phi y_B$ when $\phi = \frac{1}{\sqrt{1-\alpha}} \frac{(\Delta-\delta)}{(1-\Delta)}$.

$z^l_{B,CC}$ is strictly increasing in ϕ and converges to y_B/q.

$z^h_{B,CC}$ is strictly increasing in ϕ and grows without bound.

1.2.2 At node P

Group B's decision if A has played P is isomorphic to A's decision in the baseline model. In particular, B plays C if

$$1 - n'_{PC} < \frac{(1-\delta)z}{\delta y_B + z},$$

or

$$\max\{1 - n, \min[1, \frac{z}{\phi y_A}]\} < \frac{(1-\delta)z}{\delta y_B + z}$$

The left side starts out flat at $1 - n$, then it increases linearly, then it becomes flat again at 1. The right side increases from 0 in concave fashion and converges to $(1 - \delta)$. The two sides either never intersect or they intersect twice, once in the flat part of the left side and one in the increasing part of the left side. They intersect if and only if $n - \delta > \delta y_B/(\phi y_A)$

In conclusion, (i) if $\phi \leq \delta y_B/(y_A (n - \delta))$ then B never responds to P with C; (ii) if $\phi > \delta y_B/(y_A (n - \delta))$ then B responds to P with C if $z^l_{B,PC} < z < z^h_{B,PC}$, where

$$z^l_{B,PC} = \frac{(1-n)\delta y_B}{n - \delta}$$
$$z^h_{B,PC} = \phi y_A(1 - \delta) - \delta y_B,$$

which, not surprisingly, describes a triangle much like the one in figure 1.

1.3 Decision by A

We can now examine the behavior of A. There are four cases, depending on the combination of actions chosen by B.

1.3.1 When B plays C at both nodes

If B plays C at both nodes, A plays C if

$$(1-\delta)y_A < (1-\Delta)\left(y_A + \frac{\alpha z}{n'_{CC}}\right)$$

$$\max\left[n, \frac{(z+\phi y_B) - \sqrt{(z+\phi y_B)^2 - 4\alpha z\phi y_B}}{2\phi y_B}\right] < \frac{\alpha z}{(\Delta-\delta)\, y_A}$$

(note that if B plays C at P we know that $n'_{PC} > 0$, so we don't need to worry about the case where all the As pass.) There always is one and only one solution to this equation. Call this solution $z^l_{AC,CC}$ (for "A plays C when B plays C at both nodes"). The solution is either the solution to

$$n = \frac{\alpha z}{(\Delta-\delta)\, y_A}$$

$$z = \frac{(\Delta-\delta)\, y_A n}{\alpha} \tag{2}$$

or the solution to

$$\frac{(z+\phi y_B) - \sqrt{(z+\phi y_B)^2 - 4\alpha z\phi y_B}}{2\phi y_B} = \frac{\alpha z}{(\Delta-\delta)\, y_A}$$

$$z = \frac{(\Delta-\delta)\, y_A\left[\phi y_B - (\Delta-\delta)\, y_A\right]}{\left[\phi y_B \alpha - (\Delta-\delta)\, y_A\right]} \tag{3}$$

The solution to (2) is the "global" solution if it lies to the left of the kink in the LHS, z^k,or if

$$\frac{(\Delta-\delta)\, y_A n}{\alpha} < \frac{\phi y_B n(1-n)}{(\alpha-n)}$$

$$\frac{(\alpha-n)(\Delta-\delta)\, y_A}{\alpha y_B(1-n)} < \phi \equiv \phi_{C,CC}$$

Therefore: we conclude that, for $\phi \le \phi_{C,CC}$, A plays C if z exceeds the expression in (3), while if $\phi > \phi_{C,CC}$ A plays C if z exceeds the expression in (2).[1]

Now we characterize how $z^l_{AC,CC}$ varies with ϕ. It can be shown that this begins at $(\Delta-\delta)y_A$ for $\phi = 0$ and decreases over the interval $[0, \phi_{C,CC})$, after which it becomes constant at $(\Delta-\delta)y_A n/\alpha$.

[1] Note that the solution in (3) exists if and only if z

$$\phi \le \frac{(\Delta-\delta)y_A}{y_B}.$$

However, this is always greater than $\phi_{C,CC}$ so this constraint is never binding.

1.3.2 When B plays P and C

If B plays P at node P and C at node C, A plays C if

$$y_A + z < (1-\Delta)\left(y_A + \frac{\alpha z}{n'_{CC}}\right) \tag{4}$$
$$n'_{CC} < \frac{(1-\Delta)\alpha z}{z + \Delta y_A}.$$

The right hand side starts at 0 and then increases, converging to $(1-\Delta)\alpha$. Since as we know RHS asymptotes to α, there are either 0 or two solutions for z. There are no solutions if $(1-\Delta)\alpha < n$. If instead $(1-\Delta)\alpha \geq n$ there still could be either two or no solution.

When there are two solutions one solves

$$n = \frac{(1-\Delta)\alpha z}{z + \Delta y_A}$$

and the solution is

$$z = \frac{\Delta y_A n}{(1-\Delta)\alpha - n} \equiv z^l_{AC,PC},$$

where the subscript stands for "A plays C when B plays P and C."

The other solution solves

$$\frac{(z + \phi y_B) - \sqrt{(z+\phi y_B)^2 - 4\alpha z \phi y_B}}{2\phi y_B} = \frac{(1-\Delta)\alpha z}{z + \Delta y_A}$$

so

$$\begin{aligned} z &= \frac{-\left\{(1+\Delta)\Delta y_A + (1-\Delta)\left[(1-\Delta)\alpha - 1\right]\phi y_B\right\} +}{2\Delta} \\ &\quad + \frac{\sqrt{\left\{(1+\Delta)\Delta y_A + (1-\Delta)\left[(1-\Delta)\alpha - 1\right]\phi y_B\right\}^2 - 4\Delta^2 y_A\left[\Delta y_A - (1-\Delta)\phi y_B\right]}}{2\Delta} \\ &\equiv z^h_{AC,PC}. \end{aligned}$$

The condition for having two solutions is that $z^l_{AC,PC}$ is to the left of the kink in the left side of (4), z^k. I.e. the condition is

$$\frac{\Delta y_A n}{(1-\Delta)\alpha - n} < \frac{\phi y_b n(1-n)}{(\alpha - n)}$$
$$\phi > \frac{\Delta y_A(\alpha - n)}{\left[(1-\Delta)\alpha - n\right] y_b(1-n)} \equiv \phi_{C,PC} \tag{5}$$

In conclusion, A plays C when B plays P, C if and only if $(1-\Delta)\alpha \geq n$, (5) is satisfied, and $z^l_{AC,PC} < z < z^h_{AC,PC}$.[2]

As functions of ϕ, the two bounds share the same value $z^l_{AC,PC}$ at $\phi_{C,PC}$. As ϕ increases beyond this value, $z^l_{AC,PC}$ is constant while $z^h_{AC,PC}$ increases without bound.

[2] Note that the slope condition for a solution to the last equation exist is that the right hand side is steeper at the origin. But this is satisfied whenever $\phi \geq \phi_{C,PC}$, which in turn holds when $(1-\Delta)\alpha < n$. We focus on this case below.

1.3.3 When B plays C and P

If B plays C at node P and P at node C, A plays C if

$$(1-\delta)y_A < (1-\delta)\left(y_A + \frac{z}{n'_{CP}}\right),$$

so A always plays C in this case.

1.3.4 When B plays P and P

If B always plays P (this is the case of the baseline model) A plays C if

$$y_A + z < (1-\delta)\left(y_A + \frac{z}{n'_{CP}}\right)$$

$$\max\{n, \min[1, \frac{z}{\phi y_B}]\} < \frac{(1-\delta)z}{z+\delta y_A}$$

As in the benchmark model, we focus on the case where $(1-\delta) > n$. Then if $\phi \leq \delta y_A/(y_B\,(1-n-\delta))$ A never plays C, while if $\phi > \delta y_A/(y_B\,(1-n-\delta))$ A plays C if $\frac{\delta y_A n}{(1-\delta)-n} < z < (1-\delta)\phi y_B - \delta y_A$.This is of course the same condition as in the baseline model of exploitation.

1.4 Regions where different equilibria prevail

We now bring it all together and characterize the regions of the parameter space where the various equilibria obtain.

1.4.1 Equilibria of type CC

Equilibria where both groups play C occur in two scenarios. (i) B plays C at both nodes, and A plays C. Or (ii) B plays P at node P and C at node C, and A plays C. To limit the number of cases we focus on situations where $(1-\Delta)\alpha < n$. As we have seen this means that A never fights a war of choice: faced with a choice of peace and open conflict, it always chooses peace. Open conflict only arises when B plays C in both nodes, i.e. when A's choice is between being exploited and being the exploiter. Again this assumption only serves to reduce the number of appendix pages, without materially affecting the insights of the model.

Assumption: $(1-\Delta)\alpha < n$

Let's collect the conditions for CC to happen.

(i) B plays C at node P

$$\phi > \frac{\delta y_B}{y_A\,(n-\delta)}$$

$$z^l_{B,PC} < z < z^h_{B,PC}$$

(ii) B plays C at node C

$$\phi > \frac{1}{\sqrt{1-\alpha}}\frac{(\Delta-\delta)}{(1-\Delta)}$$

$$z^l_{B,CC} < z < z^h_{B,CC}$$

(iii) A plays C when B plays C, C

$$z^l_{AC,CC} < z$$

Therefore we observe open conflict if and only if

$$\max\left\{z^l_{B,PC}, z^l_{B,CC}, z^l_{AC,CC}\right\} < z < \min\left\{z^h_{B,PC}, z^h_{B,CC}\right\}$$
$$\max\left\{\frac{\delta y_B}{y_A(n-\delta)} \frac{1}{\sqrt{1-\alpha}} \frac{(\Delta-\delta)}{(1-\Delta)}\right\} < \phi.$$

Note that this region *always* exists. To see this, notice that both $z^h_{B,PC}$ and $z^h_{B,CC}$ grow without bound with ϕ, while all of $z^l_{B,PC}$, $z^l_{B,CC}$, and $z^l_{AC,CC}$ converge to finite constants.

An increase in y_A increases both $z^h_{B,PC}$ and $z^l_{AC,CC}$. The former reflects that in a possible PC equilibrium there will be less passing from A to B. This makes B more aggressive for a larger set of values of z, forcing A to choose C more often. The increase in $z^l_{AC,CC}$ reflects the fact that a higher y_A makes open conflict more costly for A, and thus increases the set of zs such that A is willing to let itself be exploited by B. Hence, an increase in the wealth of the stronger groups shifts the conflict region "to the right."

An increase in y_B increases $z^l_{B,PC}$ (exploiting A becomes more costly for B), $z^l_{B,CC}$ (open conflict becomes more costly), On the other hand an increase in y_B reduces $z^h_{B,PC}$ (again, the cost of exploiting A are greater), and have ambiguous effects on $z^h_{B,CC}$. Hence, the lower bound of the conflict region unambiguously increases, while the upper bound could either fall or increase.

An increase in n decreases $z^l_{B,PC}$ (exploiting A becomes more attractive when A is larger) and increases $z^l_{AC,CC}$ (A becomes more likely to acquiesce to being exploited). Hence the effect of n is ambiguous.

1.4.2 Equilibria of type CP

These equilibria emerge in two sets of circumstances. (i) When B plays P at both nodes, and A decides to exploit; and (ii) when B responds to P with C, and to C with P (we know that A always plays C in this case).

(i) B plays PP, A plays C (i.i) B plays P at P

$$\phi \leq \frac{\delta y_B}{y_A(n-\delta)}$$
$$\text{OR } z \leq z^l_{B,PC} = \frac{(1-n)\delta y_B}{n-\delta}$$
$$\text{OR } z \geq z^h_{B,PC} = \phi y_A(1-\delta) - \delta y_B,$$

(i.ii) B plays P at C

$$\phi \leq \frac{1}{\sqrt{1-\alpha}} \frac{(\Delta-\delta)}{(1-\Delta)}$$
$$\text{OR } z \leq z^l_{B,CC}$$
$$\text{OR } z \geq z^h_{B,CC}$$

(i.iii) A plays C when B plays PP

$$\phi > \frac{\delta y_A}{y_B(1-n-\delta)}$$
$$z^l_{AC,PP} \equiv \frac{\delta y_A n}{(1-\delta)-n} < z < (1-\delta)\phi y_B - \delta y_A \equiv z^h_{AC,PP}$$

Now notice that when all the other conditions in this case are satisfied, then $z^h_{B,CC} > z^h_{AC,PP}$. The reason is simple. If $z > z^h_{B,PC}$ it means that under CP the entire B population passes. But then it cannot be optimal for A to choose CP over PP. Hence this region is of the form

$$\frac{\delta y_A}{y_B(1-n-\delta)} < \phi \le \min\left\{\frac{\delta y_B}{y_A\,(n-\delta)}, \frac{1}{\sqrt{1-\alpha}}\frac{(\Delta-\delta)}{(1-\Delta)}\right\}$$
$$z \in \left[z^l_{AC,PP}, \min\left\{z^l_{B,PC}, z^l_{B,CC}\right\}\right]$$

(ii) B plays CP, A plays C (ii.i) B plays C at P

$$\phi > \frac{\delta y_B}{y_A\,(n-\delta)}$$
$$z^l_{B,PC} < z < z^h_{B,PC}$$

(ii.ii) B plays P at C

$$\phi \le \frac{1}{\sqrt{1-\alpha}}\frac{(\Delta-\delta)}{(1-\Delta)}$$
$$\text{OR } z \le z^l_{B,CC}$$
$$\text{OR } z \ge z^h_{B,CC}$$

Hence

$$\frac{\delta y_B}{y_A\,(n-\delta)} < \phi \le \frac{1}{\sqrt{1-\alpha}}\frac{(\Delta-\delta)}{(1-\Delta)}$$
$$z \in \left[z^l_{B,PC}, z^l_{B,CC}\right] \cup \left[z^h_{B,CC}, z^h_{B,PC}\right]$$

An increase in y_A increases $z^h_{B,PC}$. It also increases $z^l_{AC,PP}$ because an increase in y_A makes it more expensive for A to exploit B. Hence, the "bottom" corridor of the CP region narrows, while the "top" corridor may either narrow or widen. The reason for the difference in results with the baseline model is that now when y_A increases there is less passing from A to B when B exploits A. This makes B more likely to respond to P with C, and may force A to preemptively play C more often.

An increase in y_B increases $z^l_{B,PC}$ and $z^l_{B,CC}$, has ambiguous effects on $z^h_{B,CC}$, and reduces $z^h_{B,PC}$. Hence, the lower bound of the CP region unambiguously increases, while the upper bound may increase or decrease.

An increase in n decreases $z^l_{B,PC}$ and increases $z^l_{AC,PP}$ (lower benefits of exploitation by A). Hence the effect of n is ambiguous.

1.4.3 Equilibria of type PC

This equilibrium emerges only if B plays C at both nodes, and A prefers being exploited than engaging in open conflict. (We already know that when B plays C and O, A always plays C: better to exploit than being exploited.)

(i) B plays C at P

$$\phi > \frac{\delta y_B}{y_A (n-\delta)}$$
$$z^l_{B,PC} < z < z^h_{B,PC}$$

(ii) B plays C at C

$$\phi > \frac{1}{\sqrt{1-\alpha}} \frac{(\Delta-\delta)}{(1-\Delta)}$$
$$z^l_{B,CC} < z < z^h_{B,CC}$$

(iii) A plays P when B plays C and C

$$z^l_{AC,CC} \geq z.$$

Hence

$$\max\left\{z^l_{B,PC}, z^l_{B,CC}\right\} < z < \min\left\{z^l_{AC,CC}, z^h_{B,PC}, z^h_{B,CC}\right\}$$
$$\max\left\{\frac{\delta y_B}{y_A (n-\delta)} \frac{1}{\sqrt{1-\alpha}} \frac{(\Delta-\delta)}{(1-\Delta)}\right\} < \phi.$$

This converges to a sort of corridor, when it exists, whose upper bound is the limit of $z^l_{AC,CC}$ and the lower bound is the limit of either $z^l_{B,PC}$ or $z^l_{B,CC}$.

An increase in y_A increases both $z^h_{B,PC}$ and $z^l_{AC,CC}$. Hence, an increase in the wealth of the stronger groups unambiguously increases the region where the weaker group exploits the richer group.

An increase in y_B increases $z^l_{B,PC}$ and $z^l_{B,CC}$, has ambiguous effects on $z^h_{B,CC}$, and reduces $z^h_{B,PC}$. Hence, the lower bound of the PC region unambiguously increases, while the upper bound could either fall or increase.

An increase in n decreases $z^l_{B,PC}$ and increases.$z^l_{AC,CC}$ Hence an increase in n unambiguously increases the size of the PC region.

2 Unpublished Appendix 2. Dynamics

In a Markov perfect equilibrium group A's action at time t depends only on n_t and ζ_t. We represent this dependence by the function $\pi^A(n_t, \zeta_t)$, where $\pi^A(n_t, \zeta_t)$ is the probability that group A chooses C, and $1 - \pi^A(n_t, \zeta_t)$ is the probability that it chooses P. Each member of group B individually decides on passing having observed the collective action group A. Hence, we can specify the strategy of each member of group B as $[S(n_t, \zeta_t, P), S(n_t, \zeta_t, C)]$, where $S(n_t, \zeta_t, P)$ is the probability that this member passes when the state is (n_t, ζ_t) and A has played P, and $S(n_t, \zeta_t, C)$ is analogously defined for when A has played C. We also make the standard assumption that, whenever an agent observes an off-the-equilibrium path action by another (or more) agents, this agent assumes that all other agents will return to their equilibrium strategies in all subsequent subgames.

2.1 Existence and uniqueness for constant-z case

We begin by studying the special case where z is constant in all periods. The total per-period incomes flowing to different agents in the various scenarios are

$$\begin{aligned}
Y^A_P &= y^A + z \\
Y^A_C(n_{t+1}) &= (1-\delta)\left(y^A + \frac{z}{n_{t+1}}\right) \\
Y^B_{NS}(P) &= y^B + z \\
Y^B_S(P) &= (1-\phi)y^B + z \\
Y^B_{NS}(C) &= (1-\delta)y^B \\
Y^B_S(n_{t+1}, C) &= (1-\delta)\left((1-\phi)y^B + \frac{z}{n_{t+1}}\right) \\
Y^{BA}(P) &= y^B + z \\
Y^{BA}(n_{t+1}, C) &= (1-\delta)\left(y^B + \frac{z}{n_{t+1}}\right).
\end{aligned}$$

where the notation should be self-explanatory, except perhaps for BA which denotes agents who have switched before time t. Note that by our assumptions this group has no decisions to take. Also note that the relative size of the As and the BAs is irrelevant.

At the beginning of period t there are (at most) three types of agents. Agents who have always been (or whose lineage has always been) in group A, agents who have always been in group B, and agents who at some point in the past have switched from A to B. Define the respective maximized values of lifetime utility $V^A(n_t)$, $V^B(n_t)$, and $V^{BA}(n_t)$. It is useful to write down these "value" functions explicitly.

Writing down the value for group B is, in principle, somewhat cumbersome. However, we can make an observation that greatly simplifies the necessary notation.

Lemma 1 *For each member of group B, $S(n_t, P) = 0$ is a dominant strategy for every n_t.*

This is obvious. If group A plays P there are no current benefits from switching. If there are future benefits from switching, these can be realized by switching later. Since there is discounting switching later is preferable. This implies that $N(n_t, P) = n_t$, where N is the equilibrium mapping between the time t state variable and ethnic outcome and n_{t+1}. Given this, we can then write the value for members of B as

$$V^B(n_t) = \max\left[U^B_{NS,NS}(n_t), U^B_{NS,S}(n_t)\right]$$

where

$$\begin{aligned} U^B_{NS,NS}(n_t) &= [1 - C(n_t)]\left\{\log\left[Y^B_{NS}(P)\right] + \beta V^B(n_t)\right\} \\ &\quad + C(n_t)\left\{\log\left[Y^B_{NS}(C)\right] + \beta V^B\left[\nu(n_t)\right]\right\}, \\ U^B_{NS,S}(n_t) &= [1 - C(n_t)]\left\{\log\left[Y^B_{NS}(P)\right] + \beta V^B(n_t)\right\} \\ &\quad + C(n_t)\left\{\log\left[Y^B_S(\nu(n_t), C)\right] + \beta V^{BA}\left[\nu(n_t)\right]\right\}, \end{aligned}$$

where $v(n_t)$ is the equilibrium *unconditional* mapping from n_t to n_{t+1} (i.e. unconditional on the realization of C or P). Note that group B's decisions are atomistic (and there is perfect foresight) so agents take the function $\nu(n_t)$ as given.

For A we have:

$$V^A(n_t) = \max\left[U^A_P(n_t), U^A_C(n_t)\right],$$

where

$$\begin{aligned} U^A_P(n_t) &= \log(Y^A_P) + \beta V^A(n_t) \\ U^A_C(n_t) &= \log\left\{Y^A_C\left[N(n_t, C)\right]\right\} + \beta V^A\left[N(n_t, C)\right]. \end{aligned}$$

Finally, for people who switched in the past we have

$$\begin{aligned} V^{BA}(n_t) &= C(n_t)\left\{\log\left[Y^{BA}(\nu(n_t), C)\right] + \beta V^{BA}\left[\nu(n_t)\right]\right\} \\ &\quad + [1 - C(n_t)]\left\{\log\left[Y^{BA}(P)\right] + \beta V^{BA}(n_t)\right\}. \end{aligned}$$

Again note that these people don't take any decisions. However V^{BA} is important because it enters the decision of people who are still in B.

It is very useful to define the thresholds n^A and n^B as follows. n^A is the solution of

$$Y^A_P = Y^A_C(n^A).$$

Hence, if $N(n_t, C) = n^A$ group A receives the same income in period t if it plays P or C. n^B is the solution to

$$\sum_{t=0}^{\infty} \beta^t \log\left[Y^B_{NS}(C)\right] = \log\left[Y^B_S(n^B, C)\right] + \sum_{t=1}^{\infty} \beta^t \log\left[Y^{BA}(n^B, C)\right],$$

which has the following interpretation: if $N(n_t, C) = n^B$, and members of B expect A to play C in all periods following a period when they played C, then at time t members of group B are indifferent between switching and not switching in response to C. Note that both $n^A \in [0, 1]$ and $n^B \geq 0$ are well defined as functions of the model's parameters.

Lemma 2 *If A plays P in period t, A will play P in all subsequent periods.*

Proof. This is a consequence of the Markovian nature of the equilibrium coupled with Lemma 1. If P is dominant at time t, then the state n_t does not change. But then the same action by A must be dominant in the subsequent period. Hence, P is an absorbing state. ■

In other words the equilibrium can take only three possible forms: perpetual P, perpetual C, or a spell of C in the early part of the game followed by perpetual P.

Lemma 3 *If switching ever occurs, it occurs all in period* 0*.*

Proof. This is a consequence of Lemma 2, which implies that there are no alternating spells of P and C, and that if C occurs it occurs in a continuous spell starting at time 0. This is common knowledge, hence group B members either switch in period 0 or never switch at all. If any member of group B would switch at a later period of conflict, then he would have been strictly better off by switching at period 0, since the current period costs and benefits are the same, while there are more periods of expropriation left at period 0. ■

Lemma 4 *If A plays C in period* 0*, A will play C in all subsequent periods.*

Proof. This is an immediate consequence of Lemma 3 and the Markov nature of the equilibrium. ■

These last three lemmas narrow the possible set of equilibria down to those of perpetual peace and those of perpetual conflict. We define $\tilde{n} = \max\left\{n_0, \min(n^B, 1)\right\}$. Then we have:

Proposition 5 There exists a unique Markov Perfect equilibrium. If $\tilde{n} \geq n^A$ then the equilibrium features P always and $n_t = n_0$ for all t. If $\tilde{n} < n^A$ then the equilibrium features C always and $n_t = \tilde{n}$ for all $t > 0$.

Proof. Consider the following specification of strategies. For A, $C(n_t) = 1$ if $N(n_t, C) < n^A$, and $C(n_t) = 0$ otherwise. For members of B, play $S(n_t, C) = 1$ with probability $\max[0, \min(1, \left(n^B - n_t\right)/(1-n_t))]$ and play $S(n_t, P) = 0$ with probability 1. It is clear that if these are the strategies then the equilibrium is as described in the statement of the proposition. For, these strategies imply $N(n_0, C) = \tilde{n}$ and $N(n_0, P) = n_0$. So let's check that these strategies are mutually optimal responses. Consider the set of n_ts such that $N(n_t, C) \geq n^A$. In the hypothesized equilibrium A plays P at all these ts. Furthermore, n_t does not change, so A plays P in all subsequent periods as well. As a result, $V^A(n_t) = \log(Y_P^A) + \beta V^A(n_t) = \log(Y_P^A)/(1-\beta)$. Now consider a deviation in which for a set of these same ns A plays C instead. If the game reaches one of these ns n_t permanently jumps to (or stays at) $N(n_t, C)$. Because n_t is permanent then A must be playing C always. The value of this deviation therefore is $\log(Y_C^A(N(n_t, C)))/(1-\beta)$, and it is less than the value of the hypothesized strategy by definition of n^A. Now consider the set of n_ts such that $N(n_t, C) < n^A$. In the hypothesized equilibrium A plays C at all these ts. Furthermore, n_t jumps permanently to $N(n_t, C)$, so the value of this strategy is $\log(Y_C^A(N(n_t, C)))/(1-\beta)$. In a deviation, A plays P at n_t and then it plays P always. So the value of this deviation is $\log(Y_P^A)/(1-\beta)$. This is dominated by definition of n^A. The strategy hypothesized for individual members of B is obviously optimal as, given the behavior of A and the behavior of all other Bs, any agent's behavior does not change the outcome and, given the outcome, each agent is indifferent about the exact probability with which it switches in case of C. To conclude the uniqueness proof we must therefore just show that there is no equilibrium featuring perpetual P starting at $n^A \geq \tilde{n}$ and no equilibrium with perpetual C starting at $n^A < \tilde{n}$. But this follows immediately from the argument in the proof of existence. ■

2.2 Existence and uniqueness for general case

We now consider the case where z_t varies over time (in deterministic fashion), with the restriction that $\lim_{t\to\infty} z_t$ exists and is finite. We proceed in two steps: in the first step we show that the unique Markov perfect equilibrium after T is almost the same as in the case of constant z (where $\lim_{t\to\infty} z_t = z$), then we use the backward induction method to prove the existence of an equilibrium. All the functions defined in the previous sub-section are now also functions of the other state variable, z_t.

We start by redefining the previously used thresholds: let n_t^A be the solution of

$$Y_P^A(z_t) = Y_C^A(n_t^A, z_t).$$

Now, since $\lim_{t\to\infty} z_t = z$ it is easy to see that $\lim n_t^A = n^A$ where n^A is defined by $Y_P^A(z) = Y_C^A(n^A, z)$. It is immediate form this that if $z_t \in [z-\varepsilon, z+\varepsilon]$ then there exist boundaries such that $n_t^A \in [\underline{n}^A, \overline{n}^A] \subset \left[n_l^A, n_h^A\right]$ where n_l^A and n_h^A are defined by[3]:

$$Y_P^A(z-\varepsilon) = Y_C^A(n_l^A, z-\varepsilon),$$

$$Y_P^A(z+\varepsilon) = Y_C^A(n_h^A, z+\varepsilon).$$

Let n_t^B be the group size of A that makes members of group B indifferent between playing S in period t and playing NS in the future if they expect group A to play C in all periods s for $s > t$:[4]

$$\sum_{s=t}^{\infty} \beta^s Y_{NS}^B(C) = Y_S^B(n_t^B, C) + \sum_{s=t+1}^{\infty} \beta^s Y^{BA}(n_s^B, C).$$

As z_t varies the size of group A that makes members of group B exactly indifferent between S and NS will vary as well. However, the size of group A cannot decrease, so this equilibrium group-size will not always be realized. We know that there exist boundaries such that $n_t^B \in \left[\underline{n}^B, \overline{n}^B\right] \subset \left[n_l^B, n_h^B\right]$, where n_l^B and n_h^B are defined by:

$$\sum_{t=0}^{\infty} \beta^t Y_{NS}^B(C, z-\varepsilon) = Y_S^B(n_l^B, C, z-\varepsilon) + \sum_{t=1}^{\infty} \beta^t Y^{BA}(n_l^B, C, z-\varepsilon),$$

$$\sum_{t=0}^{\infty} \beta^t Y_{NS}^B(C, z+\varepsilon) = Y_S^B(n_h^B, C, z+\varepsilon) + \sum_{t=1}^{\infty} \beta^t Y^{BA}(n_h^B, C, z+\varepsilon).$$

Let us divide the parameter space the following way: when defining n^A and n^B as in Proposition (5) let Case 1 be when $n^B < n^A$ and Case 2 be when $n^B > n^A$ (the case where $n^A = n^B$ is zero measure, so we omit it).

[3]Note that $n_t^A = \frac{(1-\delta)z_t}{\delta y^A + z_t}$ is increasing in z_t, hence n_t^A will be the highest at the highest value of $z_t = z + \varepsilon$ and vice versa.

[4]Notice that this formula depends on future values of n_t^B, since due to the variation in z_t it might happen that in the future a larger group size will make members of group B indifferenet between switching and not switching.

2.2.1 Case 1: $n^B < n^A$

Lemma 6 *There exists an $\widehat{\varepsilon} > 0$ such that $\overline{n}^B < \underline{n}^A$.*

Proof. Recall that $n^A = \frac{(1-\delta)z}{\delta y^A + z}$ and $n^B = \frac{1}{1-\beta}\frac{z}{\phi y^B}$. Now $\overline{n}^B < n_h^B = \frac{1}{1-\beta}\frac{z+\widehat{\varepsilon}}{\phi y^B}$ and $\underline{n}^A > n_l^A = \frac{(1-\delta)(z-\widehat{\varepsilon})}{\delta y^A + z - \widehat{\varepsilon}}$. Since as $\lim_{\widehat{\varepsilon}\to 0} n_l^A = n^A$ and $\lim_{\widehat{\varepsilon}\to 0} n_h^B = n^B$ and $n^B < n^A$ for a small enough $\widehat{\varepsilon} > 0$ $n_h^B < n_l^A$ has to hold. This concludes our proof, since then $\overline{n}^B < n_h^B < n_l^A < \underline{n}^A$. ■

Denote by $\widehat{T} < \infty$ that guarantees that $z_t \in [z - \widehat{\varepsilon}, z + \widehat{\varepsilon}]$ for $t > \widehat{T}$.

Lemma 7 *There exists a $\widetilde{T} < \infty$ (and a corresponding $\widetilde{\varepsilon} > 0$) such that it is strictly not worth it for group A to postpone attacking in order to have a smaller group size forever:*

$$\sum_{t=0}^{\infty} \beta^{T+t} Y_C^A(n_{T+t+1}^B, z_{T+t}) > \sum_{t=0}^{S-1} \beta^{T+t} Y_P^A(n_{T+t}, z_{T+t}) + \sum_{t=S}^{\infty} \beta^{T+t} Y_C^A(n_{T+t+1}^B, z_{T+t}).$$

Proof. Notice that by substituting the highest possible value of n_t^B into the left hand side and the lowest possible value into the right hand side, the left hand side decreases, while the right hand side increases:

$$\sum_{t=0}^{\infty} \beta^{T+t} Y_C^A(n_{T+t+1}^B, z_{T+t}) > \sum_{t=0}^{\infty} \beta^{T+t}(1-\delta)\left(y^A + \frac{z_{T+t}}{n_h^B}\right), \text{ and}$$

$$\sum_{t=0}^{S-1} \beta^{T+t} Y_P^A(n_{T+t}, z_{T+t}) + \sum_{t=S}^{\infty} \beta^{T+t} Y_C^A(n_{T+t+1}^B, z_{T+t}) < \sum_{t=0}^{S-1} \beta^{T+t}(y^A + z_{T+t}) + \sum_{t=S}^{\infty} \beta^{T+t}(1-\delta)\left(y^A + \frac{z_{T+t}}{n_l^B}\right).$$

To conclude the proof it is therefore enough to show that there is an $\widetilde{\varepsilon} > 0$ such that

$$\sum_{t=0}^{\infty} \beta^{T+t}(1-\delta)\left(y^A + \frac{z_{T+t}}{n_h^B}\right) > \sum_{t=0}^{S-1} \beta^{T+t}(y^A + z_{T+t}) + \sum_{t=S}^{\infty} \beta^{T+t}(1-\delta)\left(y^A + \frac{z_{T+t}}{n_l^B}\right)$$

This inequality can be simplified to:

$$\sum_{t=0}^{S-1} \beta^{T+t}\left(z_{T+t}\left(\frac{1-\delta}{n_h^B} - 1\right) - \delta y^A\right) > \sum_{t=S}^{\infty} \beta^{T+t}(1-\delta) z_{T+t} (1-\beta)\, \phi y^B \frac{2\varepsilon}{(z+\varepsilon)(z-\varepsilon)}.$$

Now plugging in $z_{T+t} = z - \widetilde{\varepsilon}$ to the RHS, and $z_{T+t} = z + \widetilde{\varepsilon}$ we decrease the RHS and increase the LHS:

$$\beta^T \frac{1-\beta^S}{1-\beta}\left((z-\varepsilon)\left((1-\beta)\, \phi y^B \frac{1-\delta}{(z+\varepsilon)} - 1\right) - \delta y^A\right) > \beta^{T+S}(1-\delta)\phi y^B \frac{2\varepsilon}{(z-\varepsilon)}.$$

By letting $S = 1$ the LHS decreases and the RHS increases and by letting $\varepsilon \to 0$:

$$\lim_{\varepsilon\to 0, S\to 1} \frac{1-\beta^S}{1-\beta}\left((z-\varepsilon)\left((1-\beta)\, \phi y^B \frac{1-\delta}{(z+\varepsilon)} - 1\right) - \delta y^A\right) = (1-\beta)\, \phi y^B (1-\delta) - z - \delta y^A = K > 0^5, \text{ while}$$

$$\lim_{\varepsilon\to 0, S\to 1} \beta^S (1-\delta)\phi y^B \frac{2\varepsilon}{(z-\varepsilon)} = 0$$

[5] This is the per period gain from C as opposed to P for group A, and hence it has to be positive.

This implies that there is an $\widetilde{\varepsilon} > 0$ (and the corresponding $\widetilde{T}$) for which the inequality holds. ■

Let $T = \max(\widehat{T}, \widetilde{T})$ which is equivalent to letting z_t vary at most $\min(\widehat{\varepsilon}, \widetilde{\varepsilon})$ around z for any $t \geq T$. Given this we can identify the unique Markov perfect equilibrium for $t \geq T$ similarly to Proposition (5). Let $\underline{n}^A$ and $\overline{n}^A$ denote the respective lower and upper bound for n_t^A for $t > T$. There are two cases:

Proposition 8 *(i) When $n_T < n^A$ then the equilibrium features C in all periods $t \geq T$ when $N(n_t, C) < n_t^A$ and group A size $n_{t+1} = \widetilde{n}_t = \max\left\{n_t, \min(n_t^B, 1)\right\}$, when $N(n_t, C) \geq n_t^A$ group A plays P and their group size does not change. (ii) When $n_T > n^A$ then the equilibrium features P in all periods $t \geq T$ when $n_t > n_t^A$ and C when $n_t \leq n_t^A$ and group A size $n_t = n_T$ for all $t \geq T$.*

Proof. The equilibrium strategies are for group A to play $C(n_t) = 1$ if $N(n_t, C) < n_t^A$ and to play $C(n_t) = 0$ otherwise, and for members of group B to play a mixed strategy $S(n_t, C) = 1$ with probability $\max[0, \min(1, \left(n_t^B - n_t\right)/(1 - n_t))]$ and play $S(n_t, P) = 0$ with probability 1. These strategies clearly lead to the above described equilibria. These strategies imply $N(n_t, C) = \widetilde{n}_t$ and $N(n_t, P) = n_t$. Let us check whether these strategies are mutually optimal responses. Group B members will play a (real) mixed strategy as a response to group A playing C only if $n_t < n_t^B$. However, notice that since $n_t^B < \overline{n}^B < \underline{n}^A$ this will only occur when the initial group size $n_T < \underline{n}^A$ (only possible in (i)). Hence, when $n_T \geq \underline{n}^A$ there is no trade-off for group A between attacking in the current period or later, since their group size will not change in either case. This implies that for group A it is a strictly dominant per period strategy to play C whenever $\overline{n}^B < n_t < n_t^A$. Moreover, the group size of A will at most be $\max(n_T, \overline{n}^B)$. Consider first case (ii). Since $n_T > n^A > \overline{n}^B$ it is strictly dominant for members of group B to play NS whatever group A plays. Hence the action of group A will not affect the continuation value of any player. This implies that it is enough to compare the payoff of action C, P in any given period. When $n_t^A > n_T$ then group A is strictly better off in period t by playing C as opposed to P, and when $n_t^A \leq n_T$ then they are better off by playing P instead of C in period t. Hence, in case (ii) the above described strategies form the unique Markov perfect equilibrium, since they are all dominant strategies. Now consider case (i), and the set of n_ts such that $N(n_t, C) < n_t^A$: in the above equilibrium A plays C at all these ts. As mentioned before this implies a change in n_t only if $n_t < \overline{n}^B < \underline{n}^A$, and in these cases the above strategies imply that n_t will change to $n_t^B = N(n_t, C)$, always non-decreasingly, and remaining between the boundaries $\left[n_T, \overline{n}^B\right]$. Group A could deviate by playing P in one period. This deviation is equivalent to postponing the attack as in Lemma (??), however, the choice of T guarantees that this is strictly dominated. Note moreover, that if the if this deviation is in a period when n_t would not change even if A played C, then the deviation to P is clearly dominated, since there is no gain from not-attacking. When $N(n_t, C) \geq n_t^A$, implying that $N(n_t, C) = n_T$, then playing P is a dominant strategy for group A. The strategy hypothesized for individual members of B is obviously optimal as, given the behavior of A and the behavior of all other Bs, any agent's behavior does not change the outcome and, given the outcome, each agent is indifferent about the exact probability with which it switches in case of C. We show the uniqueness of these strategies for case (i) by contradiction. Suppose there exist another Markov perfect equilibrium - this has to differ in the action group A takes at some point from the above described equilibrium. If this involves playing C in some period instead of P, then the deviation to P would provide a higher payoff to group A, and hence would destroy that equilibrium. If this would involve playing P in some period when $n_t^B < n_t < n_t^A$, then deviating to C would be worth it for group A in that period without affecting the continuation value, and hence destroy the equilibrium again. So this other Markov perfect equilibrium would involve the exact same actions from some period on, since the group size of A can be at most $\max(n_T, \overline{n}^B)$ which will be smaller then the lower bound of $\underline{n}^A$ from some period $\widetilde{t}$ on. But then this other Markov perfect equilibrium would mean postponing the attack as in Lemma (??), which would then be dominated by deviating to playing C from one period before. Hence, we have reached a contradiction. ■

2.2.2 Case 2: $n^B > n^A$

Lemma 9 *There exists an* $\varepsilon > 0$ *such that* $\overline{n}^A < \underline{n}^B$*.*

Proof. Recall that $n^A = \frac{(1-\delta)z}{\delta y^A + z}$ and $n^B = \frac{1}{1-\beta}\frac{z}{\phi y^B}$. Now $\underline{n}^B > n_l^B = \frac{1}{1-\beta}\frac{z-\varepsilon}{\phi y^B}$ and $\overline{n}^A < n_h^A = \frac{(1-\delta)(z+\varepsilon)}{\delta y^A + z + \varepsilon}$. Since as $\lim_{\varepsilon \to 0} n_h^A = n^A$ and $\lim_{\varepsilon \to 0} n_l^B = n^B$ and $n^B > n^A$ for a small enough $\varepsilon > 0$ $n_l^B > n_h^A$ has to hold. This concludes our proof, since then $\underline{n}^B > n_l^B > n_h^A > \overline{n}^A$. ■

Denote by $T < \infty$ that guarantees that $z_t \in [z - \varepsilon, z + \varepsilon]$ for $t > T$.

Proposition 10 *For any value of* n_T *the unique Markov perfect equilibrium will feature perpetual* P *and group* A *size* $n_t = n_T$ *for all* $t \geq T$*.*

Proof. The equilibrium strategies are for group A to play $C(n_t) = 1$ if $N(n_t, C) < n_t^A$ and to play $C(n_t) = 0$ otherwise, and for members of group B to play a mixed strategy $S(n_t, C) = 1$ with probability $\max[0, \min(1, (n_t^B - n_t)/(1 - n_t))]$ and play $S(n_t, P) = 0$ with probability 1. Since $\overline{n}^A < \underline{n}^B$, these strategies imply that group A will always play P, and hence for members of group B it will be dominant to play NS. Group A will never deviate to C since that will generate a loss in that period. These strategies clearly form the unique Markov perfect equilibrium. ■

2.3 Structural transformation example

We begin by defining a number of thresholds. The thresholds n_T^B and n_T^A are defined by

$$\sum_{t=0}^{\infty} \beta^t Y_{NS}^B(C) = Y_S^B(z^l, n_T^B, C) + \sum_{t=1}^{\infty} \beta^t Y^{BA}(z^l, n_T^B, C),$$

and n_T^A

$$Y_P^A(z^l) = Y_Z^A(n_T^A, z^l, C)$$

respectively. These should look familiar, They are the analogs of n^B and n^A from the constant-z game for the subgame starting at time T in the structural transformation example. In particular, at time T, n_T^B is the post-switching group size that makes members of B indifferent between switching and not switching, when they expect the rest of the game to be characterized by conflict. n_T^A is the post-switching group size that equalizes the returns from peace and conflict for group A for $t \geq T$.

Next we define thresholds $\underline{n}_{T-1}^B$ $\overline{n}_{T-1}^B$ and n_{T-1}^A by

$$Y_{NS}^B(C) = Y_S^B(z^h, \underline{n}_{T-1}^B, C),$$

$$\sum_{s=0}^{\infty} \beta^s Y_{NS}^B(C) = Y_S^B(z^h, \overline{n}_{T-1}^B, C) + \sum_{s=1}^{\infty} \beta^s Y^{BA}(z^l, \overline{n}_{T-1}^B, C)$$

and

$$Y_P^A(z^h) = Y_Z^A(n_{T-1}^A, z^h, C).$$

The threshold $\underline{n}^B_{T-1}$ makes members of B indifferent between switching and not switching in response to C at time $T-1$, *if they expect* P *in all periods* $t \geq T$. The threshold $\overline{n}^B_{T-1}$ makes members of B indifferent between switching and not switching in response to C at time $T-1$, *if they expect* C *and* $n_t = \overline{n}_{T-1}$ *in all periods* $t \geq T$. The threshold n^A_{T-1} is analogous to n^A_T. Note that since $z^h > z^l$, $\overline{n}^B_{T-1} > \underline{n}^B_{T-1}$, $\overline{n}^B_{T-1} > n^B_T$, and $n^A_{T-1} > n^A_T$.[6]

In analogy with the constant$-z$ example we also define $\tilde{n}_T = \max\left\{n_0, \min(n^B_T, 1)\right\}$ and $\underline{\tilde{n}}_{T-1} = \max\left\{n_0, \min(\underline{n}^B_{T-1}, 1)\right\}$, and $\overline{\tilde{n}}_{T-1} = \max\left\{n_0, \min(\overline{n}^B_{T-1}, 1)\right\}$. We then have six possible configurations of parameters:

$$
\begin{array}{cccc}
(i) & n^A_T \leq \tilde{n}_T & \text{and} & n^A_{T-1} \leq \underline{\tilde{n}}_{T-1} \leq \overline{\tilde{n}}_{T-1} \\
(ii) & n^A_T \leq \tilde{n}_T & \text{and} & \underline{\tilde{n}}_{T-1} \leq n^A_{T-1} \leq \overline{\tilde{n}}_{T-1} \\
(iii) & n^A_T \leq \tilde{n}_T & \text{and} & \underline{\tilde{n}}_{T-1} \leq \overline{\tilde{n}}_{T-1} \leq n^A_{T-1} \\
(iv) & n^A_T > \tilde{n}_T & \text{and} & n^A_{T-1} \leq \underline{\tilde{n}}_{T-1} \leq \overline{\tilde{n}}_{T-1} \\
(v) & n^A_T > \tilde{n}_T & \text{and} & \underline{\tilde{n}}_{T-1} \leq n^A_{T-1} \leq \overline{\tilde{n}}_{T-1} \\
(vi) & n^A_T > \tilde{n}_T & \text{and} & \underline{\tilde{n}}_{T-1} \leq \overline{\tilde{n}}_{T-1} \leq n^A_{T-1}
\end{array}
$$

Proposition 11 *[Cases (ii) and (iii)]. If* $\tilde{n}_T \geq n^A_T$ *and* $\underline{\tilde{n}}_{T-1} < n^A_{T-1}$ *there exists a* τ, $0 \leq \tau \leq T-1$, *such that the structural-transformation game features peace in all periods* $t < \tau$ *(if any) and in all periods* $t > T-1$*, and conflict in periods* $\tau \leq t \leq T-1$.

Proof. First note that the subgame beginning at T is entirely identical to the constant-z game. Furthermore, since $\tilde{n}_T \geq n^A_T$ the equilibrium of this game can only feature perpetual peace. Hence, in all periods $t < T$ all players expect perpetual peace beginning at time T. So our task is to identify equilibrium strategies for periods $t < T$.

Now define $\underline{n}^B_t$ as the solution to

$$\sum_{s=t}^{T-1} \beta^s Y^B_{NS}(C) = \beta^t Y^B_S(z^h, \underline{n}^B_t, C) + \sum_{s=t+1}^{T-1} \beta^s Y^{BA}(z^h, \underline{n}^B_t, C),$$

for all $t \leq T-1$ (note that this is consistent with the definition of $\underline{n}^B_{T-1}$ above). $\underline{n}^B_t$ is the time-t post-switching group size that makes members of B indifferent between switching and not switching, when they expect the rest of the game to be characterized by conflict up to time $T-1$ and peace from time T onwards. Notice that $\underline{n}^B_t$ is a strictly decreasing function of t.[7] Since $\underline{n}^B_t$ is strictly decreasing, if it is profitable for A to play C at some $t^* < T-1$, then it is profitable to play C at all τ between t^* and $T-1$. So the dominant strategy for A has to take the form of playing C for a set of contiguous periods before $T-1$. The period from which on group A should play C is identified by maximizing the discounted income of group A members for the first T periods of the game. This τ can be pinned down by solving:[8]

$$\max_{0 \leq \tau \leq T-1} \sum_{s=0}^{\tau-1} \beta^s Y^A_P(z^h) + \sum_{s=\tau}^{T-1} \beta^s Y^A_C(\underline{n}^B_\tau, z^h).$$

[6] Instead, we cannot sign the difference between $n^B_T = z^l/((1-\beta)\,\phi y^B)$ and $\underline{n}^B_{T-1} = z^h/\left[\phi y^B\right]$.

[7] While $\underline{n}^B_t$ is strictly decreasing, it might happen that $\underline{n}^B_t \leq n_0$ from some $\widetilde{t}$ on. In case the conflict starts after $t \geq \widetilde{t}$, the relevant post-switching group size will naturally be n_0, and this value does not decline as t increases.

[8] Substitute n_0 for values of $\underline{n}^B_\tau$ if $\tau \geq \widetilde{t}$.

It is clear that the first term is increasing in τ, simply because there are more periods of $Y_P^A(z^h)$ income. The second term however is subject to two effects: it is decreasing due to the fewer periods we sum across, while every period's utility increases due to the decrease in the equilibrium $\underline{n}_\tau^B$. The maximization boils down to the following: if $(1-\delta)\,\phi y^B - \delta y^A - z^h > (<)\, 0$, then the maximand is decreasing(increasing) in τ, hence $\tau = 0$ in the first case and $\tau = T-1$ in the second case.[9] The equilibrium strategies of the game are for group A to play C for all $\tau \le t < T$ and to play P in all periods from T on, while for members of B to play $S(n_t, C) = 1$ with probability $\max[0, \min(1, \left(\underline{n}_t^B - n_t\right)/(1-n_t))]$ and $S(n_t, P) = 0$ with probability 1.

It is instructive to compare the equilibria of the structural transformation example when the transformation is anticipated and unanticipated. It is clear that both cases feature the same post-T equilibrium, i.e. peace. We also know that the economy with unanticipated transformation features C before T if and only if $\tilde{n}_h < n_h^A$, where $\tilde{n}_h$ and n_h^A are the relevant thresholds of a constant-z game with $z = z^h$.

Proposition 12 *$\tilde{n}_h \ge \underline{\tilde{n}}_{T-1}$ and $n_h^A = n_{T-1}^A$*

This follows easily from the definitions of the various thresholds. The significance of this result is that, perhaps paradoxically, the likelihood of conflict in the pre-transformation part of the game is higher when the decline in z is anticipated. This is because the anticipated future decline in z reduces the return from switching, so $\underline{n}_{T-1}^B$ is lower than n_h^B. As members of B switch less in response to conflict when a future decline in z is anticipated, A has a stronger incentive to engage in conflict, and conflict in the early part of the game is more likely. Now also recall that in the case of unanticipated change the initial switching and conflict thresholds are the same as in the case where z is permanently constant at its initial value. This has a surprising implication: consider two economies that start out with the same initial level of z, z_0. In economy 1 $z_t = z^h$ for all t. In economy B z falls to a lower value at date T. If both these paths are fully anticipated economy B is more likely to experience conflict between dates 0 and T.

For the remaining cases we simply sketch the proof.

Proposition 13 *[Case (i)]. If $\tilde{n}_T \ge n_T^A$ and $\underline{\tilde{n}}_{T-1} > n_{T-1}^A$ the equilibrium features peace in all periods.*

It is clear that the post-T game features peace in all periods. It is also clear that at time $T-1$ the only possible outcome is peace. This argument can be extended backwards - since $\underline{\tilde{n}}_t > n_t^A$ is true also for all $t < T$ - to show that P must prevail in all periods.

Proposition 14 *[Cases (v) and (vi)]. If $n_T^A > \tilde{n}_T$ and $n_{T-1}^A > \underline{\tilde{n}}_{T-1}$ and furthermore*

- *$\underline{\tilde{n}}_{T-1} > n_T^A$ and either*

$$A.(1-\beta)^2\,(1-\delta)\,\phi y^B - (1-2\beta)\delta y^A \quad > \quad (1-\beta)z^h - \beta z^l \ \text{ or} \tag{6}$$

$$B.\,(1-\delta)\,(1-\beta)\left(1-\beta^T\right)\phi y^B - \delta y^A \quad > \quad (1-\beta^T)z^h - \beta^T z^l \tag{7}$$

then the equilibrium in case A features C at time $T-1$ and P in all other periods and in case B the equilibrium features C until period $T-1$ and P in all periods from T on.[10]

[9] In the zero measures case where $(1-\delta)\,\phi y^B - \delta y^A - z^h = 0$ group A would be exactly indifferent to when to attack, since the gain from smaller group size would be offset by the loss from fewer periods of expropriation.

[10] Case A prevails if $(1-\delta)\,\phi y^B - \delta y^A - z^h < 0$ while Case B is relevant if $(1-\delta)\,\phi y^B - \delta y^A - z^h > 0$.

- $\tilde{\underline{n}}_{T-1} > n_T^A$ *and both*

$$A.(1-\beta)^2(1-\delta)\phi y^B - (1-2\beta)\delta y^A \leq (1-\beta)z^h - \beta z^l \text{ and} \quad (8)$$

$$B.(1-\delta)(1-\beta)\left(1-\beta^T\right)\phi y^B - \delta y^A \leq (1-\beta^T)z^h - \beta^T z^l \quad (9)$$

hold, then the equilibrium features P up to time $T-1$ and C in all other periods.

- $\overline{\tilde{n}}_t \leq n_T^A$ *for all $t < T$, then there exists a τ, $0 \leq \tau \leq T$, such that the equilibrium features P for $t < \tau$ and C for $t \geq \tau$*

- $\tilde{\underline{n}}_t \leq n_T^A < \overline{\tilde{n}}_t$ *for all $t < \underline{t}$, and $\overline{\tilde{n}}_t \leq n_T^A$ for all $\underline{t} \leq t \leq T$ then the equilibrium features P in all $t < \tau$ for some $\tau \in [\underline{t}, T]$ and C for all $t \geq \tau$.*

In the first subcase, because members of B expect perpetual P after T they respond to C at time $t < T$ by jumping to $\underline{n}_t^B \geq \underline{n}_{T-1}^B$. Since $\underline{n}_t^B \geq \underline{n}_{T-1}^B > n_T^A$ group A indeed has no incentives to play C after time T. In this case group A's decision is identical to the one in Proposition 1, they will either start attacking in period 0 or wait until $T-1$ with playing C. However a possible deviation is for A to play P until $T-1$ and C in all future periods. Since $n_T^A > \tilde{n}_T$ playing C in the post-T subgame would indeed be optimal. Condition (6) insures this deviation is dominated.

In the second sub-case, because A has played P up to time $T-1$, n_t is $\tilde{n}_T$ when A begins playing C at T, and since $n_T^A > \tilde{n}_T$ playing C is indeed optimal. The deviation in which A plays C at $t < T$ and continues to play C even after cannot dominate, because $\overline{\tilde{n}}_{T-1} > \tilde{\underline{n}}_{T-1} > n_T^A$. The deviation in which A plays C at $T-1$ and P in subsequent periods is dominated if (6) does not hold.

The third case should be straightforward at this point. Since $\overline{\tilde{n}}_\tau \leq n_T^A$ it is dominant for group A to play C in all periods from T on. As for the periods before time T, similarly as in Proposition 1 group A maximizes the lifetime income of its members by choosing the τ period as the start of the conflict spell. Again, as in Proposition 1, there are two conflicting effects of choosing a bigger τ: the periods of expropriation declines, while the after-switching group size declines as well. Considering this, group A has two choices: start the conflict at $\tau < T$ or at $\tau = T$. This can be written formally as:

$$\max_{0 \leq \tau \leq T-1} \sum_{s=0}^{\tau-1} \beta^s Y_P^A(z^h) + \sum_{s=\tau}^{T-1} \beta^s Y_C^A(\overline{\tilde{n}}_\tau, z^h) + \sum_{s=T}^{\infty} \beta^s Y_C^A(\overline{\tilde{n}}_\tau, z^l).$$

and then compare the maximized utility to

$$\sum_{s=0}^{T} \beta^s Y_P^A(z^h) + \sum_{s=T}^{\infty} \beta^s Y_C^A(\overline{\tilde{n}}_T, z^l).$$

If the former is larger (the gain from more periods is bigger) then group A chooses a period before T to start the spell of conflict, if the latter is greater (the gains from smaller group to share with is bigger), then group A waits until T to start the expropriation.

In the fourth subcase group A cannot start the conflict too early (before period $\underline{t}$) in equilibrium. If it would start the conflict before period $\underline{t}$, then group B would expect P after time T (leading to a group size $\tilde{\underline{n}}_t \leq n_T^A$) which would make it profitable for group A to deviate to playing C even after time T. Hence, in any equilibrium group A must play P before period $\underline{t}$. From period $\underline{t}$ on the game is identical to subcase three. Again group A takes two steps: first it maximizes

$$\max_{\underline{t}\le\tau\le T-1}\sum_{s=0}^{\tau-1}\beta^s\log(Y_P^A(z^h))+\sum_{s=\tau}^{T-1}\beta^s\log(Y_C^A(\bar{\tilde{n}}_\tau,z^h))+\sum_{s=T}^{\infty}\beta^s\log(Y_C^A(\bar{\tilde{n}}_\tau,z^l)).$$

then compares the maximized utility to

$$\sum_{s=0}^{T}\beta^s\log(Y_P^A(z^h))+\sum_{s=T}^{\infty}\beta^s\log(Y_C^A(\bar{\tilde{n}}_T,z^l)).$$

Again, if the former is larger, then group A starts the conflict at period $\tau < T$, if the latter is larger, then group A starts the conflict at T and then continues playing C forever.

Proposition 15 *[Case (iv)]. If $n_T^A > \tilde{n}_T$ and $n_{T-1}^A \le \underline{\tilde{n}}_{T-1} \le \bar{\tilde{n}}_{T-1}$ then the equilibrium features P for all $t < T$ and C for all $t \ge T$.*

Since $n_{T-1}^A \le \underline{\tilde{n}}_{T-1} \le \bar{\tilde{n}}_{T-1}$ and $\underline{\tilde{n}}_t, \bar{\tilde{n}}_t$ are strictly decreasing with t while n_t^A is constant for $t < T$, $n_t^A \le \underline{\tilde{n}}_t \le \bar{\tilde{n}}_t$ holds for all $t < T$. This implies that playing C for any $t < T$ is a strictly dominated strategy for group A. From period T however, playing C is a dominant strategy for group A, since $n_T^A > \tilde{n}_T$.