Essays on Conflict

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Economics

By

Danielle Louisa Monnier, M.A.

Washington, DC
April 27, 2012

Essays on Conflict

Danielle Louisa Monnier, M.A.

Dissertation Advisor: Garance Genicot

Abstract

These two essays use game theoretic models to examine factors contributing to conflicts within and between countries. The first chapter models alliance formation and investment in conflict in a society which can be divided into either class or ethnic groups. Analysis of the model shows that countries always mobilize and invest in conflict resources, and that the excluded sector of the population (that sharing neither class nor ethnicity with the governing group) is decisive in determining whether class or ethnic alliances form. As income inequality within a group increases, mobilization on that dimension becomes more likely if the excluded subgroup will bear less of the cost, and vice versa. The second chapter uses a model of conflict between two countries with informed leaders to examine the influence of domestic political concerns on the likelihood of truthful revelation and the probability of conflict. A democratic leader gains credibility through the need to inform his citizens, in order to maintain sufficient public support to remain in office. However, he may want to manipulate the population, if his desired strategy differs from that of the median voter. Results show that a democratic leader is more likely to falsely state that his country is vulnerable to attack when facing a democracy, but is more likely to understate the country's vulnerability when facing an autocracy. Regime type has a greater influence than information revelation on the probability of conflict, which is highest for a pair of autocracies and lowest for a pair of democracies, thus reinforcing the 'democratic peace' hypothesis.

INDEX WORDS: asymmetric information, conflict, democracy, inequality

Acknowledgments

I thank my advisor Garance Genicot for her guidance and encouragement. I also received valued feedback from Axel Anderson, Arik Levinson, Luca Anderlini, Roger Lagunoff and Stergios Skaperdas, and from participants at the congress of the European Economic Association, the GCER Alumni Conference, GPPI Development Workshop, Georgetown University Graduate Student Seminar, and the Southern Economic Association Annual Meetings.

Table of Contents

List of Figures

Chapter 1

Inequality and Conflict Mobilization

1.1 Abstract

This chapter models alliance formation and investment in conflict in a society which can be divided into either class or ethnic groups. It examines the effects of income inequality on the salient cleavages in society and the destruction of resources through wasteful conflict expenditure. Conflict mobilization always takes place in this model, but whether class or ethnicity is the salient identity depends on the interaction between vertical and horizontal inequality. The mobilization process requires two subgroups in the population which share a characteristic to agree to unite, but the excluded sector of the population, sharing neither class nor ethnicity with the governing group, is decisive in determining which alliances form. The relative levels of inequality influence preferences through their effect on conflict investment. Waste through conflict is inevitable, and once an alliance has formed conflict expenditure is 'U-shaped' in income inequality, increasing as inequality increases without bound. The effect at any switch point between ethnic and class alliances is indeterminate: total conflict expenditure may be non-monotonic at this level of inequality.

1.2 Introduction

Inequality creates animosity in society, between the rich and the poor and between different ethnic groups. Conflict may then follow, requiring mobilization of the population and investment in conflict resources. This chapter studies the effects of income inequality on the salient divisions in society, mobilization for conflict, and the wasteful expenditure that results.

Any society has divisions on multiple dimensions. Socio-economic groupings, ethnicity, religion, and geographical location are among the markers which can create cleavages along which society mobilizes itself when conflict takes place within countries.

Traditionally class divisions were seen as the main cause of intra-country conflict, and for much of the twentieth century many such conflicts were framed as a struggle between rich and poor. Central America and South East Asia were regions that experienced a particular concentration of class conflict in the latter half of the century, fueling Cold War fears of communist revolution. However, in recent decades there has been increased ethnicization of domestic conflicts. Examples of countries where conflict has been defined along ethnic divides include Sri Lanka, Fiji, the countries of the former Yugoslavia and, in sub-Saharan Africa, Burundi, Rwanda and Sudan. Even in regions where such devastating violent conflict has not occurred, political discourse, itself a form of passive conflict, has frequently become polarized around ethnic identity.

Much of the rhetoric in intra-country conflicts has been focused on inequalities between these groups, both income inequality and social inequality, which can include variation in access to government controlled resources, services and public goods, as well as, in a more abstract sense, the prevailing attitudes and ideologies in society.

When a social group is disadvantaged it is natural for its members to feel antagonism towards the advantaged group, and such feelings may be aroused by inequalities in society, leading to discourse and, potentially, conflict. Nevertheless, such antagonistic sentiment is not sufficient to mobilize society for conflict. This chapter addresses the need for coordinated action in alliance formation in a society which has multiple potential cleavages. It focuses on how income inequality influences this process; social inequality is modeled as government allocation of public goods budgets, and as such is endogenous to the model.

As well as creating antagonism, income inequality can affect the level of investment in conflict resources, which can include the payment of activists and the purchase of physical capital, and thus may influence the expected outcome, should violence break out. This wasteful expenditure can be seen as a measure of the intensity of conflict, or of the destruction of society's resources. If a sector of society can choose to align itself with either class or ethnic allies the level of inequality within each potential alliance must therefore be considered, as this will affect the relative contributions of alliance partners. I examine the interaction between vertical and horizontal inequality in the process of mobilization and its impact on the cost of conflict.

I use a model in which society can be divided on two dimensions: ethnicity or socio-economic class. This gives four ethnic-economic subgroups characterized by their *per capita* income, allowing for parametrization of the income differential between rich and poor, and between the ethnic majority and minority. In addition, society has budgets for public goods with ethnic characteristics and public goods with class characteristics, which it is the responsibility of a governing subgroup to allocate between the majority and minority and between the rich and the poor respectively. Two subgroups of mutual

class or ethnicity may choose to ally themselves, invest in conflict, and seize a share of the contested public goods budget.

The existence of inequality ensures that society will always mobilize, generating waste and inefficiency through investment in conflict, as a self-interested government will allocate the public goods budgets to maximize its own gains. It follows that the socially excluded subgroup, which shares neither characteristic with the government, is therefore decisive in the mobilization of society. This subgroup - the poor ethnic minority in the case of elite government by the rich majority, or the rich ethnic minority if a populist, poor majority government - will always prefer conflict to peace, when it receives no benefit from either contested budget. Whether this sector of the population favors ethnic or class mobilization, it will always find a willing alliance partner, since any subgroup sharing one characteristic with the government will also prefer to mobilize on the dimension on which it is excluded.

The salient social cleavage thus depends on this excluded subgroup's preferences, a function of inequality, which acts through the share of total contributions that will be made in the possible alliances. In any alliance the subgroup with the higher aggregate income will bear the greater share of investment, an effect which is exacerbated as inequality increases. Under a mild symmetry condition the payoff from conflict is shown to equate to population share, so the contribution effect is crucial to the process of alliance formation. Increasing inequality has a monotonic effect on groupwise contributions, so as either vertical or horizontal inequality varies there will be at most a single switch point between class and ethnic alliances.

The dependence of the dimension of mobilization on the identity of the group in government is an important result. One might, for example, expect that as vertical inequality increases, class conflict would become more likely, due to greater tension

between rich and poor, but this is not necessarily the case. Sri Lanka is one country where ethnic tensions became increasingly heightened as vertical inequality grew. This model provides an explanation for such results. It shows that if a poor subgroup is in government class mobilization becomes more likely as vertical inequality grows, but if a rich subgroup is in government, as with the elite of the Sinhalese majority in Sri Lanka, ethnic mobilization becomes more likely as the decisive poor minority (the Tamils in Sri Lanka) will have to bear a declining share of conflict expenditure if ethnic mobilization occurs.

It is also not always the case that increasing horizontal inequality will increase the likelihood of ethnic conflict: this will only result if a subgroup from the economically disadvantaged ethnicity is decisive. This was the case in Burundi, where the 1966 Micombero coup established Tutsi hegemony over the majority Hutu population. In the decades of military government and one party rule which followed, horizontal inequality also favored the Tutsi minority which, for example, dominated the coffee trade, which brought most of Burundi's foreign exchange earnings. Economic inequality exacerbated ethnic tensions, which erupted into periodic violence and genocide.[1]

The effect of income inequality on conflict investment is less clear cut. Assuming either ethnic or class mobilization, investment in conflict resources is shown to be a U-shaped function of inequality within alliances. This result follows from the fact that the contributions of the richer subgroup in an alliance increase as inequality grows while the contributions of the poorer subgroup are decreasing, and is independent of inequality on the opposing dimension, which becomes irrelevant once mobilization has

[1]The model predicts that the high vertical inequality in Burundi may also have contributed to the salience of ethnic divisions, since the Tutsi government was dominated by the social elite, often drawn from the top ranks of the military.

taken place. However, the effect at a switch-point between class and ethnic alliances is indeterminate. Conflict investment may either increase or decrease, depending on the other parameters, since all are contributing to total expenditure but the excluded subgroup is decisive in inducing the switch. Social division and mobilization always leads to inefficiency through investment in conflict and the waste of resources, an effect which can be of a substantial magnitude: total conflict expenditure may be greater than the contested budget.

Consider the common case of elitist government by the rich majority, making the excluded poor minority decisive. As vertical inequality increases their share of contributions in an ethnic alliance declines, while their expected payoffs and contributions in a class alliance remain unchanged, so ethnic conflict becomes more likely. However, total conflict investment may either increase or decrease at the switch between class and ethnic mobilization depending on population shares, horizontal inequality and the size of the contested budget.

The level of destruction caused by conflict is another factor influencing preferences over the dimension of mobilization. If the decisive excluded subgroup has lower aggregate income than its potential class alliance partner and higher aggregate income than its potential ethnic alliance partner, it is more likely to prefer ethnic alliances as destructiveness decreases. When the potential prize from conflict is greater, all groups are willing to contribute more, but the relative share of the contributions of the richer subgroup is lower.

Much has been written about the causes of inefficient conflict.[2] In many cases imperfectly defined property rights are a contributing factor [Grossman and Kim (1995), Hirshleifer (1995)]. This chapter draws on that result by modeling conflict over

[2] Garfinkle and Skaperdas (2007) provide a good overview.

a public goods budget. Although in this model of complete information active conflict does not arise,[3] mobilization of resources, a form of passive conflict, will still take place when all parties are fully informed about the others' characteristics, allowing the focus here to be on the effects of inequality and the process of alliance formation. This relates to the work of Jackson and Morelli (2007), who connect incentives for war to the political bias of leaders and show that in many cases (provided that the bias is sufficiently small), active conflict can be avoided through the use of transfer payments.

This chapter draws most strongly on the work of Esteban and Ray (2008), in which they show that in a 'symmetric' world there is a bias in favor of ethnic conflict. I adapt their basic model, expanding it to allow for income inequality between ethnic groups as well as between the rich and the poor. Esteban and Ray assume unranked ethnic groups, but it is precisely the effect of inequality between them which is of interest to this chapter. Introducing greater variation, I am able to examine the process of alliance formation and investment in conflict, and how these are influenced by the interaction between horizontal and vertical inequality.

Other work by Esteban and Ray also informs the discussion on conflict within countries. Esteban and Ray (2010) is particularly relevant, as it addresses the question studied here of the effect of inequality on ethnic conflict. They use a different approach which directly models both the financial contributions and activists required for conflict, and show an increase in conflict following from an increase in within-group inequality. Esteban and Ray (2009) also studies the relationship between inequality and conflict, and demonstrates an approximate linear relationship between the level of conflict and the Gini coefficient, among other factors.

[3] Baliga and Sjöström (2004) and Bester and Wärneryd (2006), among others, have shown that incomplete information can be a significant cause of conflict.

The relationship between inequality and conflict has also been addressed in the political science literature. I make particular note of Østby (2008), an empirical study relating conflict to horizontal inequalities, which concludes that horizontal inequalities are positively associated with conflict, an effect which is particularly strong in democracies. Fearon and Laitin (2000) also provides a useful discussion of the salience of ethnic identity in violent conflict.

1.3 Model

Consider a society which exhibits two-way cleavages on two dimensions: it can be divided into either ethnic or economic groups.[4] The two ethnic groups are indexed by T, the ethnic majority, and D, the ethnic minority. The population is normalized to have unit measure, and $n_T > 0.5$ and $n_D = 1 - n_T$ denote the population shares of the ethnic majority and minority respectively. The socio-economic groups, or classes, are the poor, p, and rich, r, with population shares $n_p > 0.5$ and $n_r = 1 - n_p$ respectively. Thus there are four ethnic-economic subgroups in the population.

These ethnic-economic subgroups are differentiated by *per capita* income; an individual of class i and ethnicity j has income y_{ij} and there is homogeneity within each subgroup. Income inequality exists on both dimensions, between the rich and the poor (vertical inequality) and between the ethnic majority and minority (horizontal inequality). $\alpha > 1$ measures inequality within ethnic groups, giving the ratio of *per capita* income between the poor and the rich such that $y_{rj} = \alpha y_{pj}$. $\beta > 0$ is the corresponding measure for inequality between ethnic groups, $y_{iD} = \beta y_{iT}$, allowing either the majority or minority to have higher *per capita* income.

[4] This model draws on that of Esteban and Ray (2008). The most significant change is expansion of the model to allow for ranked ethnic groups.

In addition to the individual endowment, the population benefits from the consumption of public goods, which are produced or otherwise funded by the state. Public goods are assumed to have either socio-economic or ethnic characteristics. Thus, rather than perfectly pure public goods, from which no members of society can be excluded, the term here refers to collective goods which have particular implications for one class or ethnic group in the population, but from which the other does not benefit.

Public goods may represent physical infrastructure and services, or dominating attitudes and ideologies. As such non-tangibles are included, the budgets are taken to be fixed and exogenously defined.[5] C is the total class budget available to fund public goods with class specific characteristics, while E is the corresponding total ethnic budget.

Class specific public goods may include investment in public services which have particular benefits for specific socio-economic sectors of society. Examples include public health care or transportation, or a focus on primary or higher education. Attitudes towards foreign investment and international trade would also fall into this category.

Examples of ethnic public goods include support for religious structures and festivals or cultural events, as well as religion- or language-specific education and investment in arts or culture. Expenditure on infrastructure or public services in regions where a particular ethnic group is dominant would also fall into this category.

I assume that all such goods have a monetary value and the public goods budgets are perfectly divisible.

[5] An interesting extension could be to consider introducing taxation and endogenizing the public goods budgets.

The benchmark situation in society is one without conflict. In peacetime ethnic-economic subgroup (i, j) obtains a share s_i of the class budget C and a share s_j of the ethnic budget E, giving a peacetime payoff of:

$$U_{ij} = ln(y_{ij}) + s_i C + s_j E.$$

The peace shares are chosen by one ethnic-economic subgroup which is considered to be in a position of power ('in government').

Alliances and Conflict

Society may remain at peace or enter into conflict. Alliances must form between subgroups sharing an ethnic or class characteristic before conflict can break out. If ethnic alliances form, the rich and poor of each ethnicity unite but the divide between the ethnic majority and minority remains. Any conflict is then over the allocation of the ethnic budget, and the shares of the class budget remain as in peacetime. Likewise, if class alliances form, the ethnic subgroups unite but the socio-economic cleavage remains, and any conflict is over the allocation of the class budget, leaving the shares of the ethnic budget unchanged.

Mobilization on one dimension requires two subgroups sharing a characteristic to choose to form an alliance. For example, ethnic mobilization will occur if both the rich and poor of the ethnic majority want to join together. The remaining subgroups will then prefer to unite and mobilize too, as if they do not they will gain no share of the contested budget. Mobilization takes place and alliances remain stable if both subgroups in an alliance prefer joining together to both an alliance on the other dimension and no mobilization at all. It is not necessary for this to hold on both sides of the cleavage, as the presence of one stable alliance will force the other subgroups

to unite. Assume that if one or both subgroups are indifferent between forming an alliance or not, alliances will form, and if indifferent between ethnic and class alliances, ethnic alliances will form.

Once alliances have formed each subgroup contributes conflict resources, denoted A_{ij} in ethnic alliances and B_{ij} in class alliances. These could be physical capital such as weaponry, munitions and transportation, or human capital, militants or activists. Each unit of conflict resources has cost, w_j (ethnic alliances) or w_i (class alliances), proportional to the income of the poorer subgroup in the alliance, supposing that any activists will be taken from this group.

$$w_i = \begin{cases} \gamma y_{iT} & \text{if } \beta > 1 \\ \gamma y_{iD} & \text{if } \beta < 1 \end{cases}$$

$$w_j = \gamma y_{pj}$$

where $\gamma \in [0, 1]$.

It is necessary to differentiate between violence and fighting (referred to as 'active conflict') and the situation where society mobilizes and invests in conflict, but chooses to avoid violence ('passive conflict').

'Active conflict' takes place following the formation of alliances when at least one alliance prefers conflict to peace. Assume that if indifferent between peace and conflict, peace will be maintained. Following violent conflict, the shares of the public goods budgets, σ_i and σ_j, are determined by:

$$\sigma_j = \frac{A_j}{A_T + A_D}, \qquad \sigma_i = \frac{B_i}{B_p + B_r},$$

where σ is a function of the aggregate conflict resources in alliances. $A_j = A_{pj} + A_{rj}$, $B_i = B_{iT} + B_{iD}$ are the aggregates.

If active conflict obtains, a fraction $(1-\delta) \in (0,1)$ of the contested budget is destroyed in conflict. The shares of the uncontested budget are allocated as in peacetime.[6]

Thus an individual in subgroup (i,j) obtains a conflict payoff of

$$U_{ij} = ln\left(y_{ij} - \frac{w_j A_{ij}}{n_{ij}}\right) + s_i C + \sigma_j \delta E$$

in ethnic conflict, and

$$U_{ij} = ln\left(y_{ij} - \frac{w_i B_{ij}}{n_{ij}}\right) + \sigma_i \delta C + s_j E$$

in class conflict.

This supposes enforceability within ethnic-economic subgroups, due to within group homogeneity, but not within the alliance as a whole. Each subgroup must choose its contributions using an individual maximization problem.

If both alliances prefer peace to launching conflict the country is in a state of 'passive conflict', since mobilization and investment means that violence is now a credible threat. In this case peacetime shares will be allocated.[7] Passive conflict can be seen as a hostile stalemate, but it may also be a form of political engagement where parties are formed along ethnic divisions (as in Malaysia) or to represent class interests (including the many parties which have grown out of labor unions) and resources are required to promote political ends.

[6] If, instead, $(1-\delta)$ of both public goods budgets is destroyed if active conflict obtains, the main results should still hold, but the relative size of the budgets becomes significant in determining allocations. This is best avoided since, in the model, the choice of budget size is somewhat arbitrary.

[7] Passive conflict payoffs: $U_{ij} = ln\left(y_{ij} - \frac{w_j A_{ij}}{n_{ij}}\right) + s_i C + s_j E$ (ethnic conflict), $U_{ij} = ln\left(y_{ij} - \frac{w_i B_{ij}}{n_{ij}}\right) + s_i C + s_j E$ (class conflict).

Game Structure

A multistage model of strategic interaction is used:

1. *Alliance Formation:* Population subgroups choose whether to enter class or ethnic alliances. If alliances form, the game advances to stage 2; if not, it jumps to stage 3.

2. *Investment in Conflict:* Each subgroup chooses how much conflict investment to contribute to its alliance.

3. *Allocation of Shares:* The subgroup in government allocates the peacetime shares of the class and ethnic budgets for all groups in the population. If society has mobilized, the game advances to stage 4; if not, it jumps to stage 5.

4. *Hostility Decision:* Each alliance chooses whether to launch active conflict or to maintain a peaceful stance.

5. *Outcome:* Shares are allocated according to the conflict success function, σ, if active conflict obtains on that dimension, and according to the shares decided in stage 3 otherwise.

1.4 Mobilization of Society

The solution is obtained by backward induction.

The game ends with the allocation of shares of the public goods budgets, according to the shares chosen by the subgroup in government or, if active conflict has been launched, by the conflict success function. When making the decision between active conflict and maintaining a peaceful stance each alliance must compare the allocation

assigned to them to the share they expect to win in conflict. By this stage conflict is only possible on one dimension, since the population has already mobilized and conflict investment is a sunk cost. The preferences of both subgroups in any alliance will therefore be aligned, as it is not possible to change the shares of the non-contested budget and the allocation of the contested public good enters linearly into the utility functions.

Consider a situation where ethnic alliances have formed, the peacetime shares of the public goods have been decided and investment in conflict has been made, determining the value of the conflict success function. An ethnic alliance chooses to maintain a peaceful stance if $s_j \geq \delta\sigma_j$. Society will remain at peace if this holds for both ethnic alliances. Similarly, if class alliances have formed society will remain at peace if $s_i \geq \delta\sigma_i$ for both i. Since there is no gain from conflict ($\delta \leq 1$) and perfect information is assumed the constraint can be satisfied simultaneously for both alliances, and so there exist possible allocations of shares such that active conflict will be avoided.[8]

Avoiding Violent Conflict

Prior to the hostility decision the ethnic-economic sub-group in government chooses the peacetime shares. Denote this subgroup (I, J); subgroup $(I, -J)$ shares the class characteristic of the subgroup in government, $(-I, J)$ the ethnic characteristic, and $(-I, -J)$ neither characteristic.

Start by supposing that ethnic alliances have formed. The class public good cannot be contested so the government can use the entire class budget for the benefit of its

[8] If $\delta = 1$ there is no destruction in conflict and the constraint can only be weakly satisfied, since both sides will be indifferent between peace and active conflict when $s_j = \sigma_j$ $\forall\ j$ ($s_i = \sigma_i\ \forall\ i$) and $\sigma_T + \sigma_D = 1$ ($\sigma_p + \sigma_r = 1$).

own social class: $s_I = 1$, $s_{-I} = 0$. Assigning shares of the ethnic budget such that $s_J = 1 - \delta\sigma_{-J}$, $s_{-J} = \delta\sigma_{-J}$ gives the opposing alliance the smallest share possible to ensure that they prefer peace to active conflict. $1 - \delta\sigma_{-J} \geq \delta\sigma_J$ so (I, J) prefers this allocation to giving the opposing alliance a smaller share and provoking active conflict.

Likewise when class alliances have formed ethnic conflict is not possible, and the opposing alliance is given the smallest share of the class budget sufficient to prevent it from launching active conflict: $s_J = 1$, $s_{-J} = 0$, $s_I = 1 - \delta\sigma_{-I}$, $s_{-I} = \delta\sigma_{-I}$.

Shares are always allocated such that both alliances choose not to launch active conflict. Although there are, of course, many examples of violent conflict, it is not inconsistent with this model, as it abstracts from many of the other issues which may contribute to the outbreak of violence.[9] Non-violent conflict still causes inefficiency and waste in society due to investment in conflict resources. Hostility is observed, here termed 'passive conflict', due to mobilization and investment, allowing the opposing alliance to extract a share of the contested public good budget, and the formation of alliances exhibits the salient division in society.

Investment in Conflict

Each ethnic-economic subgroup (i, j) chooses its level of conflict resources to maximize its own payoff U_{ij}. Following ethnic mobilization this yields the maximization problem:

[9] For example, introducing some uncertainty to the model would ensure that active conflict sometimes obtains. While arguably more realistic in terms of the information structure and conflict outcome, this would not change the main results concerning the effects of inequality, but would make the explication less clear.

$$max_{A_{ij}} \begin{cases} ln\left(y_{ij} - \frac{w_j A_{ij}}{n_{ij}}\right) + \left(1 - \frac{\delta A_{-j}}{A_T + A_D}\right) E & \text{for subgroups in governing alliance,} \\ ln\left(y_{ij} - \frac{w_j Aij}{n_{ij}}\right) + \frac{\delta A_j}{A_T + A_D} E & \text{for subgroups in opposing alliance.} \end{cases}$$

This formulation ignores the share of the class budget because it is unaffected by ethnic conflict. As the resource constraint is omitted an interior solution is assumed; that is, the budget constraint is assumed to be non-binding.

Each of the four first order conditions has the form:

$$\frac{w_j}{n_{ij} y_{ij} - w_j A_{ij}} = \frac{\delta A_{-j}}{(A_T + A_D)^2} E. \tag{1.1}$$

These conditions, and therefore expenditure on ethnic conflict, are independent of the identity of the subgroup in government. The absolute value of income drops out of the expressions: wage is proportional to *per capita* income, and each subgroup income can be expressed in terms of any other using the inequality parameters. Contributions are not influenced by the absolute income of individuals, but by their relative income and thus the level of inequality within the alliance.

This system of four equations can be solved to express the conflict investment of each subgroup in terms of the parameters of the model. Investment in ethnic conflict is a function of vertical income inequality, α, but not horizontal inequality, β. Population inequality between ethnicities still features, as contributions are a function of the population shares of both socio-economic and ethnic groupings.[10] However, differences in *per capita* income between the opposing alliances does not affect investment in conflict; the share of contributions that an alliance must bear has greater influence than the direct effect on the conflict success function.[11] Income inequality within

[10] See the proof of Lemma 1 in the First Appendix for the closed form solution for contributions to conflict resources.

[11] To focus on the effects of horizontal versus vertical inequality, symmetry of inequality within alliances has been supposed, i.e. *per capita* income inequality between the rich

ethnic groups remains significant as the subgroups do not collaborate when choosing how much to invest.

Similarly, conflict expenditure for class alliances is obtained by solving the corresponding maximization problems obtained assuming class mobilization. There is variation in the first order conditions, and thus the expenditure functions, depending on whether the majority or minority has the higher *per capita* income, as this affects the cost of resources.[12] Class expenditure is likewise a function of inequality within the alliances (in this case β) but not vertical inequality between the opposing groups.

Lemma 1 *In any alliance the subgroup with the higher aggregate income contributes more.*

That is, in any alliance the subgroup with the lower *per capita* income will invest more if their advantage in population size is sufficient to make up for their disadvantage in income. Thus the ethnic minority will never contribute more in class alliances if $\beta < 1$ because in this case they are disadvantaged in both respects.

A proof of the lemma is contained in the First Appendix; the result follows from comparing the expressions for the contributions obtained from the first order conditions. For example, the poor contribute more than the rich in an ethnic alliance if $\frac{n_p}{n_r} > \alpha$; that is, if the poor subgroup has the higher aggregate income: $n_p y_{pj} > \alpha n_r y_{pj}$. A corner solution may exist in which one subgroup bears the entire cost of conflict for an alliance, if the difference in aggregate income is sufficiently large.

and poor of the ethnic majority is the same as the vertical inequality within the ethnic minority. This symmetry ensures that as there is no interaction between the two dimensions of inequality, contributing to this result.

[12] The expenditure function when the majority is richer is equal to that when the minority is richer scaled by a factor of β.

Relative shares of conflict expenditure for subgroups in an alliance are influenced by both factors contributing to groupwise inequality: population share and the *per capita* income ratio. In what follows, $\lambda_{ij} = \frac{A_{ij}}{A_j}$ will denote the share of investment made by economic group i within an ethnic alliance, $\mu_{ij} = \frac{B_{ij}}{B_i}$ will denote the share of investment made by ethnic group j within a class alliance.

Proposition 1 *An alliance's effectiveness in conflict corresponds to its population share:* $\sigma_i = n_i$, $\sigma_j = n_j$.

A proof of the proposition is contained in the First Appendix.

This result follows from the fact that the contributions of the poor in an alliance of the ethnic majority differ from the contributions of the poor in an ethnic minority alliance only by population shares, so $\lambda_{pT} = \lambda_{pD}$ and subsequently $\lambda_{rT} = \lambda_{rD}$.[13] The alliances have the same *per capita* contributions, since inequality within alliances is symmetrical. Success in conflict is therefore only a function of population share. Although active conflict does not occur in this model, the threat of conflict allows the opposing alliance to extract a share of the contested public good budget from the governing alliance. The share of the budget obtained is proportionate to population share, as this determines the anticipated gains should violent conflict occur, but remains independent of income inequality.

Mobilization

The first stage of the game involves the formation of alliances, revealing the salient cleavages in society and so determining the dimension for all succeeding interactions. An alliance forms when both subgroups sharing a characteristic (ethnic or economic)

[13] Similarly, $\mu_{pT} = \mu_{rT}$, $\mu_{pD} = \mu_{rD}$.

choose to unite, so it is first necessary to consider the preferences of the subgroups. A simple assumption is required in order to obtain a closed form solution. It does not change any of the results presented, while allowing for clearer explication of the main findings.

A1: *Contributions are small relative to income:* U' *is approximately constant across* per capita *income net of contributions for an individual in any ethnic-economic subgroup.*

This simplifying assumption requires that contributions are not constrained by the size of the endowment. As has been demonstrated, investment in conflict does not depend on the absolute value of income, and an interior solution has already been implicitly assumed by the construction of the maximization problem, so it is reasonable to allow this assumption.

For simplicity, the situation where no alliances form will be referred to as 'peace', to differentiate it from 'passive conflict', where society mobilizes, sustained by inefficient investment in conflict resources, but both alliances are satisfied by the government allocation of shares of the public goods budgets.

A first step towards determining how society mobilizes is to consider the preferences of each ethnic-economic subgroup over all possible alliance outcomes.

Lemma 2 *Under A1, subgroup* (i, j) *prefers*

a) Ethnic mobilization to no mobilization if $s_j^E - s_j^P \geq \lambda_{ij}\delta n_T n_D$*;*

b) Class mobilization to no mobilization if $s_i^C - s_i^P \geq \mu_{ij}\delta n_p n_r C$*;*

c) Ethnic alliance to class alliance if $(s_i^E - s_i^C)C + (s_j^E - s_j^C)E \geq \lambda_{ij}\delta n_T n_D E - \mu_{ij}\delta n_p n_r C$;

where s_k^E *is the share of the relevant public good budget obtained by alliance* k *under ethnic mobilization,* s_k^C *the corresponding share under class mobilization, and* s_k^P *when no alliances form (peace).*

A proof of this lemma is contained in the First Appendix; however, the intuition is straightforward. The left hand side of each inequality gives the payoff benefit, the difference in the shares of the public goods budgets that will be obtained in the situations being compared. When mobilization on a given dimension is compared to peace there is no difference in the anticipated shares of the uncontested budget. The right hand side is a measure of the cost of investing in conflict, incorporating the subgroup's share of contributions in an alliance. If society does not mobilize no expense is incurred. An alliance is preferred if the expected gain is greater than the cost of investing in conflict.

Mobilization on one dimension will occur if any two subgroups sharing a characteristic prefer to ally themselves than to remain alone or to form an alliance on the alternative dimension. The two remaining subgroups will then prefer to form an alliance as well.

Preferences differ between the four subgroups, but the structure of the alliance formation process will not affect the outcome: since a single stable alliance will always exist. There will always be two subgroups sharing either an ethnic or economic characteristic, for whom uniting is the preferred action. One of these will be the ethnic-economic subgroup which shares neither its class or ethnic characteristic with the

subgroup in government, from hereon referred to as the 'excluded' subgroup. It follows that a formal model of alliance formation is not required, since it is irrelevant to the outcome of the game.

Proposition 2 *Society always mobilizes. Alliances form on the dimension preferred by the 'excluded' subgroup.*

The proof of this proposition is intuitive. Clearly the subgroup in government will always prefer peace to either class or ethnic alliances, as in peacetime it is able to extract the full share of both public goods budgets. For the same reason a subgroup sharing one characteristic with the government will prefer peace to mobilization on the dimension of the shared characteristic, but will prefer mobilization on the other dimension to peace. For example, if the rich ethnic majority is in government, the poor ethnic majority will prefer peace to ethnic alliances: in peacetime the government will allocate the full share of the ethnic budget to the majority but if ethnic alliances form it will be forced to reduce this share in order to give the minority sufficient to prevent active conflict. In contrast, they will prefer class alliances to peace, as in this case they will be able to use the threat of active conflict to obtain a share of the class budget, of which they would receive nothing in peacetime. Likewise the rich ethnic minority will prefer ethnic alliances to peace which in turn is preferred to class alliances.

Therefore the two subgroups sharing one characteristic with the governing subgroup will each prefer conflict on the dimension in which they have the 'opposing characteristic'. To form an alliance each will need the support of the excluded subgroup. This group (the poor ethnic minority in the case where the rich ethnic majority is in government) always prefers mobilization on either dimension to peace, as whenever an alliance forms they will be able to extract a share of one public good budget, whereas

they receive no benefit from public goods in peacetime. The result is that alliances will always form. Whether the excluded subgroup prefers ethnic or class mobilization, it will find a willing alliance partner. Thus the preferences of the subgroup sharing neither characteristic with the group in government is decisive in determining on which dimension alliances form.

Note that assumption A1 is not required for this result.

Propositions 1 and 2, together with preliminary result that any opposing alliance will be given the smallest share of the contested budget sufficient to prevent it from launching active conflict, combine to show that, when subgroup (I, J) is in government, ethnic alliances will form if the following constraint is satisfied, and class alliances will form otherwise:

$$n_{-J}E - n_{-I}C \geq \lambda_{-I-J} n_T n_D E - \mu_{-I-J} n_p n_r C. \tag{1.2}$$

It follows that there will always be wasteful conflict expenditure. All non-governing subgroups are willing to invest in conflict resources to obtain a share of the public good budget from which they are otherwise excluded, and, since the excluded subgroup prefers mobilization on either dimension to peace, alliances will always form and investment take place.[14] Mobilization on the dimension chosen by the excluded subgroup is the least favored situation for both subgroups in the governing alliance but they are forced together by the formation of the opposing alliance. They then choose to invest, as otherwise the opposing alliance would be able to launch active conflict and win the entire share of the public good budget with minimal investment.

[14]It can be shown that there are circumstances where the waste from conflict is sufficiently extreme that total conflict investment (A or B) is greater than the size of the contested budget.

The preferences of the excluded subgroup are decisive in determining the dimension of alliances, as this subgroup gains no share of either public good budget in peacetime and so will never be satisfied with a peaceful outcome. This subgroup's preferences are therefore crucial in determining the salient cleavage of society, so to examine the effect of inequality on this outcome requires examining how inequality effects the excluded subgroup's preferences over class and ethnic mobilization.

The main result concerns the effects of income inequality on alliance formation, and how this varies depending on the identity of the subgroup in government. As vertical inequality, α, increases, the decisive subgroup is more likely to prefer ethnic alliances if a rich subgroup is in government, and is more likely to prefer class alliances if a poor subgroup is in government. These results are conditional on holding all other parameters constant.

Proposition 3 *If ethnic mobilization takes place at some level of inequality, $\widehat{\alpha}$, then there will be ethnic mobilization,*

a) At all higher levels of inequality, $\alpha > \widehat{\alpha}$, when a rich subgroup is in government;

b) At all lower levels of inequality, $\alpha < \widehat{\alpha}$, when a poor subgroup is in government.

If class mobilization takes place at some level of inequality, $\widetilde{\alpha}$, then there will be class mobilization,

c) At all lower levels of inequality, $\alpha < \widetilde{\alpha}$, when a rich subgroup is in government;

d) At all higher levels of inequality, $\alpha > \widetilde{\alpha}$, when a poor subgroup is in government.

There is at most a single point at which the excluded subgroup is indifferent between ethnic mobilization and class mobilization.

A proof of this proposition is contained in the First Appendix.

Vertical income inequality influences the choice between ethnic and class mobilization through its effect on the division of conflict expenditure between the rich and poor subgroups in an ethnic alliance.[15] As inequality increases the income of the rich rises, so they can afford to devote more of their endowment to conflict and they therefore take on a larger share of conflict expenditure.[16] The gains in the event of active conflict do not change and, more pertinently, neither do the gains induced by the threat of conflict. Consequently the net return to the rich from ethnic alliances declines, and, since α features in neither the contributions to nor the payoff from class conflict, shifts their preferences, decisive when a poor subgroup is in government, towards class alliances.

The poor bear a lower share of ethnic conflict expenditure as vertical inequality increases. The gains in the event of conflict remain constant, so the expected return to the poor from ethnic conflict increases, shifting their preferences, decisive when a rich subgroup is in government, in favor of ethnic alliances.

The effect of α on the preference inequality is weakly monotonic, so it follows that there exists at most a single switch-point between ethnic and class alliances. A switch-point may not exist, however, as there is no change in preferences in the regions in which one subgroup contributes the entire share of costs in an alliance, since clearing shares are constant in this region. Denote by $\overline{\alpha}$ the value of α above which the rich bear the entire cost of ethnic conflict, and by $\underline{\alpha}$ the value below which the

[15]Esteban and Ray (2008) also highlight the importance of expenditure effects in the salience of ethnic conflict in a society with unranked ethnic groups. It is hard for the poor to coordinate for class conflict as their opportunity cost of resources is so high, but the rich have much lower opportunity costs, favoring ethnic divisions.

[16]It is reasonable to assume that the absolute income of the poor remains constant when α increases, to avoid the complication of societal credit constraints when inequality increases without bound.

poor bear the entire cost. Thus any switch-point will be in the region $[\underline{\alpha}, \overline{\alpha}]$, where shares, and therefore preferences, vary. Nevertheless, it may be that the excluded subgroup prefers mobilization of the same dimension at both $\underline{\alpha}$ and $\overline{\alpha}$, in which case no switch in preferences will occur as income inequality changes. Thus no limiting situation exists: α is bounded below, and it is not the case that a switch in alliances will always occur when α increases without bound.

Proposition 4 *If ethnic mobilization takes place at some level of inequality, $\widehat{\beta}$, then there will be ethnic mobilization,*

a) At all $\beta > \widehat{\beta}$ when an ethnic majority subgroup is in government;

b) At all $\beta < \widehat{\beta}$ when an ethnic minority subgroup is in government.

If class mobilization takes place at some level of inequality, $\widetilde{\beta}$, then there will be class mobilization,

c) At all $\beta < \widetilde{\beta}$ when an ethnic majority subgroup is in government;

d) At all $\beta > \widetilde{\beta}$ when an ethnic minority subgroup is in government.

There is at most a single point at which the excluded subgroup is indifferent between ethnic mobilization and class mobilization.

A formal proof is omitted, as it follows the same procedure as the proof of Proposition 3.

As β increases horizontal inequality shifts from favoring the majority to favoring the minority, and so the minority take on a greater share of expenditure in any class alliance. The share of the class budget obtained remains constant, so the net return to the minority from class alliances declines, shifting their preferences, decisive when

an ethnic majority subgroup is in government, towards ethnic alliances. Likewise the ethnic majority's share of expenditure in any class alliance declines, but returns remain constant, shifting their preferences in favor of class alliances.

Again, it is possible that one subgroup will contribute the entire share of conflict investment in a class alliance, if its aggregate income is sufficiently greater. Define $\underline{\beta}$ and $\overline{\beta}$ such that if $\beta < \underline{\beta}$ the ethnic majority bears the entire cost of class conflict and if $\beta > \overline{\beta}$ the ethnic minority bears the entire cost.[17] Any switch between class and ethnic alliances as β varies will be in the region $[\underline{\beta}, \overline{\beta}]$, but again there is no limiting case where a switch in alliances can be induced when β is increased without bound, since the excluded subgroup may prefer mobilization on the same dimension at both $\underline{\beta}$ and $\overline{\beta}$.

The political situation in Malaysia since independence in 1957 illustrates some of these results.[18] The government has been dominated by the UMNO party, representing the interests of the Malay majority, while the minority Chinese population has been more economically successful. Vertical inequality has remained stagnant, but horizontal inequality initially increased, leading to heightened ethnic tensions, as the model predicts when an ethnic majority subgroup is in government. The government then introduced redistributional policies and quotas for access to public resources, intended to reduce inter-ethnic disparity, leading to an amelioration of ethnic tensions as horizontal inequality declined.

Sri Lanka is another country which has had an ethnic majority government since independence but where inequality favored the minority. The Sinhalese majority was

[17] It may be that $\underline{\beta} > 1$ or $\underline{\beta} < 1$ but $\overline{\beta} > 1$ always, since the minority is by nature at a population disadvantage.

[18] Of course, the model provides a highly stylized example and abstracts from other social tensions.

more powerful politically, with the government dominated by their elite. Legacies of colonialism meant that the minority Tamil population was richer and socially advantaged, although to a lesser extent than political rhetoric in the country implied. Vertical inequality, however, was high and increasing. Theory suggests that with a poor subgroup decisive, ethnic conflict becomes more likely as vertical inequality increases. This was the case in Sri Lanka, where the government restricting opportunities for Tamils (the minimum share of the ethnic public goods budget), leading to political tensions and eventually the outbreak of civil war in 1983.[19] Tamil activism was largely lead by young, lower class Tamils, the decisive excluded subgroup.

This discussion has focused on the effects of *per capita* income inequality. Measures of inequality such as the Gini coefficient also depend on population shares as they look at the distribution of income throughout society.[20] However, the comparative statics for changes in population shares, and therefore also of these aggregate measures, are inconclusive.

Increasing the population share of a subgroup increases its aggregate income, and therefore its share of contributions in any alliance,[21] in this respect shifting preferences towards alliances on the alternative dimension, the same way as changes in *per capita* inequality. However, this is not the only way in which population shares enter into the constraint determining whether ethnic or class mobilization takes place (1.2). As well as effecting the cost of conflict through the share of contributions, population share determines the payoff through the conflict success function, and influences the share on contributions on the opposite dimension.[22] The direction of the partial effect depends

[19] Although active conflict does not obtain in this model, adding some asymmetry of information would result in active conflict erupting in some circumstances.

[20] For example, the Gini coefficient for horizontal inequality: $\frac{n_T(1+\beta)-\beta n_T^2}{\beta-n_T(\beta-1)}$.

[21] $\frac{\partial \lambda_{pj}}{\partial n_p} > 0$, $\frac{\partial \lambda_{rj}}{\partial n_r} > 0$, $\frac{\partial \mu_{iT}}{\partial n_T} > 0$, $\frac{\partial \mu_{iD}}{\partial n_D} > 0$.

[22] λ_{ij} is also a function of n_j, and μ_{ij} a function of n_i.

on the other parameters, so the different ways in which the population shares enter into the constraints may be acting in different directions, and no general conclusions can be drawn about which effects outweigh the others.

1.5 Cost of Conflict

The measures of aggregate conflict resources - $A = A_T + A_D$ for ethnic alliances, $B = B_p + B_r$ for class alliances - can be interpreted as total conflict expenditure, or the total waste to society due to investment in conflict. Therefore, to assess the impact of inequality on the social cost of conflict it is useful to consider these measures.

Returning the focus to *per capita* income inequality, once an alliance has formed contributions are only a function of inequality on the dimension of mobilization. That is, total investment in ethnic conflict, A, and the contributions of each subgroup, are only a function of α, vertical inequality within each ethnic alliance, and not of β, horizontal inequality between the ethnic groups.[23] Likewise investment in class conflict, B, is only a function of horizontal inequality.

It is therefore useful to consider how expenditure in an alliance varies with income inequality before moving on to consider the possible effects of a change in the dimension of mobilization. This is also relevant because it should be easier to adjust the level of investment than the dimension of mobilization when faced with a change in the level of inequality. Not only will it be costly to move conflict resources and establish organizational structure, but it takes time for group members to form a new identity and develop hostility towards a different subject.

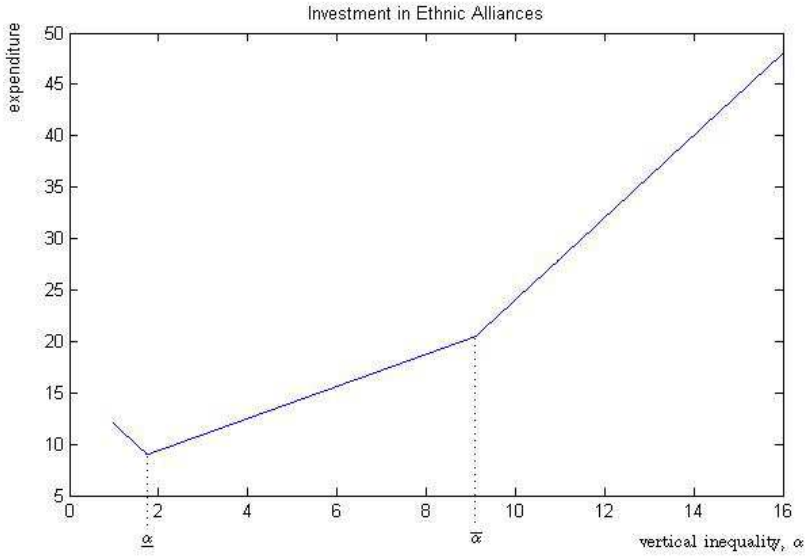


Figure 1.1: Expenditure with Ethnic Alliances

Expenditure with Ethnic Alliances

Conflict investment is not a function of absolute income, so total expenditure is only a function of the parameters of the model. As income inequality increases the rich have to pay more to sustain an ethnic alliance, as the poor can afford to contribute less due to their (relatively) lower income.[24] The overall effect of inequality on conflict expenditure will therefore depend on whether the increase in contributions of the rich outweigh the decline in contributions of the poor.

[23] See proof of Proposition 1

[24] Since it has been assumed that the budget constraint does not bind, an increase in income inequality should be interpreted as holding the income of one group constant while aggregate income varies, i.e. the rich get richer while the poor remain the same. Nevertheless, to ensure that the rate of growth of conflict investment does not outstrip that of total resources in the economy in the limiting case, it is necessary that a further parametric condition is satisfied.

Lemma 3 *Assuming ethnic mobilization, total investment in ethnic alliances is a 'U-shaped' function of vertical income inequality, α. [Figure 1.1]*

The proof, provided in the First Appendix, follows directly from the partial effects of aggregate contributions with respect to vertical inequality.

Across the entire range of α, the contribution of the rich subgroup in an ethnic alliance is weakly increasing in *per capita* income inequality, while the contribution of the poor subgroup is weakly decreasing. The poor can afford to contribute less, or choose to rely more on the rich due to their greater income. The rich must invest more to sustain the alliance, to maintain a sufficient level of resources to exert pressure on the group in government and maximize the return to the alliance.

Within the range of inequality for which both subgroups have non-zero contributions, $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, total investment increases as inequality within the ethnic alliances increases. Richer subgroups provide more resources while the poorer subgroup contributes less, but overall the greater contributions of the rich are the dominating effect and aggregate conflict investment increases.

When inequality is sufficiently high, $\alpha > \overline{\alpha}$, the contributions of the poor have fallen to zero and will remain at this level if inequality rises further. The rich are now wholly funding the alliances, and as inequality continues to increase so too will their conflict expenditure, as the relative cost of investment is lower when income is higher. In fact, in this range conflict investment is increasing more steeply than when $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, since the contributions of the rich are increasing while those of the poor are constant at zero.

When inequality is sufficiently low, $\alpha < \underline{\alpha}$, the poor bear the entire cost of ethnic conflict so contributions of the rich remain constant at zero. Therefore as inequality

increases conflict investment declines, until it reaches the point where the rich subgroup is forced to step in and contribute to the alliance. Depending on the levels of population inequality and other parameters, it may be that the poor never have sufficiently higher aggregate income to bear the entire cost of ethnic conflict; in this case, conflict investment in ethnic alliances will be strictly increasing across the range of α.

Returning to the example of Sri Lanka, the model predicts that when the poor minority is the decisive excluded subgroup, as with the Tamils in Sri Lanka increasing vertical inequality favors ethnic conflict, so once ethnic alliances have formed a change in the dimension of mobilization will not occur, but investment in ethnic conflict will increase. The intensification of ethnic conflict observed in Sri Lanka as the income divide between the rich and the poor grew is consistent with this result.

Expenditure with Class Alliances

Horizontal inequality has a similar effect on conflict expenditure taking class alliances as given.

Lemma 4 *Assuming class mobilization, total investment in class alliances is a 'U-shaped' function of horizontal income inequality, β. [Figure 1.2]*

The proof is omitted since it follows the same form as that of Lemma 3.[25]

When $\beta > 1$ the ethnic minority has higher *per capita* income than the ethnic majority. In this range inequality between ethnic groups increases as β increases, so its effect is similar to that of increasing inequality within ethnic groups. Rising

[25]It should be noted that the expenditure function varies slightly between $\beta < 1$ and $\beta > 1$.

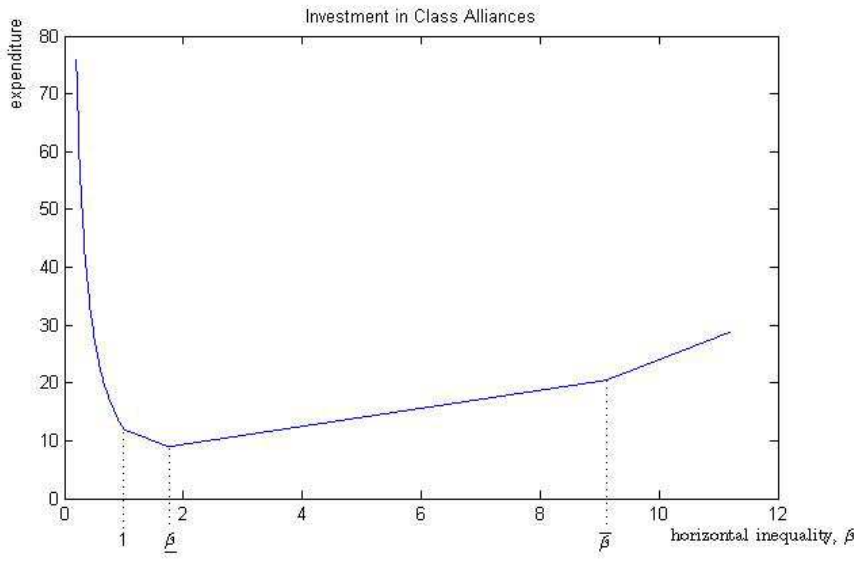


Figure 1.2: Expenditure with Class Alliances

inequality leads to higher investment by the richer minority while the contributions of the poorer majority decrease. If inequality between ethnic groups is sufficiently high, $\beta > \overline{\beta}$, the minority bears the entire cost of investment in ethnic conflict, and so total conflict expenditure is increasing. Below this level both subgroups are contributing to the alliance, but the overall effect remains that expenditure increases with inequality. It is possible that when inequality is sufficiently low the ethnic majority will bear the entire cost of the alliance due to their advantage in population share, even though the minority have higher *per capita* income.[26] If this is the case aggregate expenditure will be decreasing over the range $\beta \in [1, \underline{\beta}]$, as in Figure 2.

Inequality between ethnic groups takes on a larger range of values than intra-ethnic inequality, since it is also possible that the ethnic majority has higher *per*

[26] $\underline{\beta} = \frac{n_T}{n_D} \frac{1}{\delta n_p n_r C + 1}$ so can have $\underline{\beta} > 1$ or $\underline{\beta} < 1$.

capita income, $\beta < 1$. In this range inequality within class alliances increases as β decreases, but the general result still holds: contributions of the richer subgroup (in this case the majority) increase with inequality while contributions of the poorer subgroup (the minority) decrease. If inequality is sufficiently high - possible for the entire range of $\beta < 1$ (Figure 2) as the majority is advantaged in terms of population as well as *per capita* income - expenditure will increase as inequality increases, that is, as β decreases. If $\underline{\beta} < 1$ there is some level of inequality for which the minority will contribute a share despite the greater income and population share of the majority. For this range, $[\underline{\beta}, 1]$ as inequality increases the minority can afford to contribute less, so their investment declines while the majority invest more, but the overall effect is that total contributions increase. As the ethnic minority is disadvantaged in terms of both population share and *per capita* income there can be no situation where $\beta < 1$ but the minority bear the full cost of a class alliance.[27] Thus when the ethnic majority has a higher *per capita* income, aggregate expenditure on class conflict is always increasing with inequality (and therefore decreasing in β).

Combining these results gives a class conflict expenditure function which is 'U-shaped' in β. As β increases the situation described moves from high inequality in favor of the ethnic majority, through a point of income equality between ethnic groups, to high inequality in favor of the minority. The lowest level of expenditure cannot be at a point at which the majority have higher *per capita* income, i.e. it will exist at $\beta \geq 1$, since the minority will never bear the entire cost of conflict when it is disadvantaged in terms of both population and income.[28] Therefore for most of the range of horizontal inequality investment in class conflict increases as inequality increases;

[27] i.e. $\overline{\beta} > 1$

[28] Minimum expenditure is at $\beta = 1$ if $\underline{\beta} < 1$, and at $\beta = \underline{\beta}$ if $\underline{\beta} > 1$. The second case is depicted in Figure 2.

the only exception (which will not exist for all parameter combinations) being the case where the ethnic majority has the lower *per capita* income but is bearing the entire cost of class conflict due to its higher aggregate income.

Total Conflict Expenditure

There will be at most a single switch between ethnic alliances and class alliances as inequality on a single dimension varies, since the preferences of the decisive subgroup are weakly monotonic in inequality, as shown by Proposition 3 and Corollary 1.

Proposition 5 *Holding all other parameters constant, a sufficient condition for the existence of a switch in the dimension of mobilization as vertical inequality, α, varies is given by:*

$$n_{-J}E > (1 - \mu_{-I-J}n_I)n_{-I}C > n^2_{-J}E.$$

Holding all other parameters constant, a sufficient condition for the existence of a switch in the dimension of mobilization as horizontal inequality, β, varies is given by:

$$n_{-I}C > (1 - \lambda_{-I-J}n_J)n_{-J}E > n^2_{-I}C.$$

(I, J) is the ethnic-economic subgroup in government.

A proof of this proposition is contained in the First Appendix.

When this condition is satisfied and a switch between class and ethnic alliances occur there is a discontinuity in total conflict expenditure; expenditure may either increase or decrease at this point.

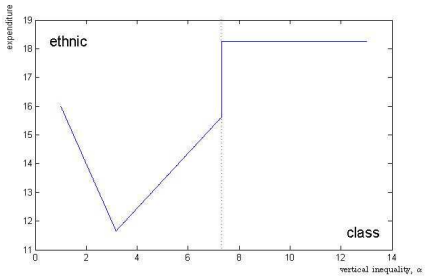
The excluded subgroup's preference over class and ethnic alliances (2) is determined not only by how much it will need to invest in each possible alliance but also

by the share of the contested public goods budget it expects to obtain in each case. It may be willing to invest more if it anticipates a higher payoff. For example, the poor ethnic minority will expect to obtain a larger share of the contested budget if class alliances are formed ($n_p > 0.5$ but $n_D < 0.5$). However, they may have to bear the majority share of the investment in a class alliance, while the rich will contribute the majority of ethnic investment. The trade off between cost and payoff will determine which alliance is preferred. In addition, the opposing alliance will also invest, even though they may have conflicting preferences over alliances, as otherwise they will forfeit all shares of the contested budget, adding to total conflict expenditure.

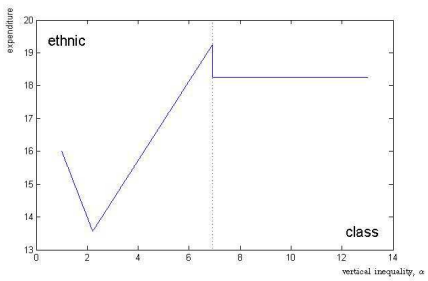
Any switch between class and ethnic alliances will take place in the region in which both subgroups contribute to an alliance. In this range investment in any alliance is strictly monotonic; that is, provided that all subgroups have non-zero contributions, total conflict expenditure is strictly increasing in inequality. When a switch in alliance dimension occurs there will be a discontinuity in total conflict expenditure. It may be that there is a 'jump up' in total expenditure at this point, thus ensuring that weak monotonicity holds throughout the region in which all contributions are non-zero, but this will not always be the case.

Figure 1.3 shows some examples of total conflict expenditure, demonstrating that monotonicity of conflict expenditure may either be maintained or fail at the switch point.[29] All of these examples show the effect of increasing inequality within ethnic groups, α, while holding all other parameters constant and imposing equality of the ethnic and class public goods budgets ($C = E$). This focuses the results on the effects of income and inequality, since the relative size of the budgets also influences preferences. It is intuitive that if there is a large difference between the size of the

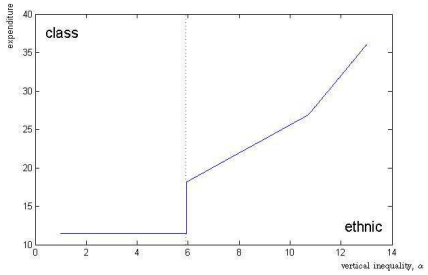
[29] Investment in conflict is expressed as an absolute value, to avoid introducing an additional parameter for the absolute value of *per capita* income.



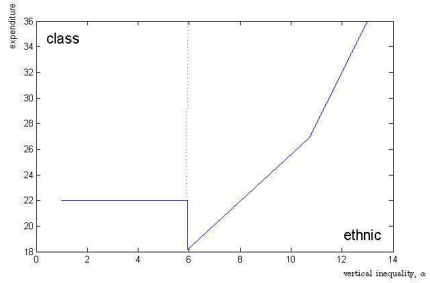
(a) Jump up in expenditure at switch point



(b) Jump down in expenditure at switch point



(c) Monotonicity maintained at switch point



(d) Monotonicity fails at switch point

Figure 1.3: Examples of the different forms that total conflict expenditure can take as vertical inequality varies.

public goods budgets mobilization on the dimension with the larger budget will be preferred. As the conflict success function determines the share of the budget obtained the same level of investment will produce a larger return simply due to the size of the contested budget.

Panels (a) and (b) show total expenditure when the poor ethnic majority is in government: the rich minority is decisive and its preferences shift from ethnic alliances to class alliances as inequality increases. The only difference between the two cases is the size of the contested budgets, all other parameters are held constant.[30] Expendi-

[30] $\beta = 0.8$, $n_p = 0.9$, $n_T = 0.85$, $\gamma = 0.05$, $\delta = 0.8$

ture on class conflict is not affected by variation in vertical inequality. Total expenditure on class conflict is greater than total expenditure on ethnic conflict at the point of indifference between ethnic and class mobilization when public goods budgets $C = E = 18$. The jump at the discontinuity is positive and total investment is weakly increasing over the range where all contributions are non-zero. However, when the contested budgets are increased ($C = E = 30$) the effect is reversed and total investment decreases at the switch in alliance dimension.

Panels (c) and (d) show situations with the rich ethnic minority in government, the poor majority decisive and a switch from class alliances to ethnic alliances. Again most of the parameters are held constant:[31] the difference here is the level of inequality between ethnic groups. When inequality is low ($\beta = 0.8$) an increase in total expenditure is observed at the switch point, but with a higher level of inequality ($\beta = 0.5$) ethnic expenditure is lower than class expenditure when the change in the dimension of mobilization occurs. This result follows from the fact that investment in class alliances is higher when there is higher inequality within class alliances, but this has no effect on investment in ethnic alliances.

In some cases, as in the examples given in panels (c) and (d), as inequality increases without bound, so too does total conflict expenditure. This is why an increase in α should be interpreted as an increase in inequality while holding constant the *per capita* income of the poor (or in the case of β, the poorer ethnic group). Otherwise investment would be credit constrained for sufficiently high levels of inequality.[32] It also supports the use of the assumption that contributions are small relative to income. A comparison of the rates at which investment and income are increasing

[31] $n_p = 0.8$, $n_T = 0.7$, $C = E = 10$, $\gamma = 0.05$, $\delta = 0.8$

[32] Aside from feasibility considerations in application, the choice of log utility prevents levels of investment greater than income in the context of this model.

depends on the absolute level of income,[33] but a large proportion of total resources may be devoted to conflict when inequality is high.

It follows that, in some circumstances, as inequality increases without bound, the limiting case is one where the proportion of resources in the economy devoted to conflict is minimal. Such a result would arise when total investment is constant in the limit (i.e when class mobilization takes place at $\overline{\alpha}$ as α varies or ethnic mobilization at $\overline{\beta}$ as β varies). However, it is likely that a significant proportion of resources is wasted on investment in conflict throughout the plausible range of parameters. To make a direct comparison it would be necessary to specify the absolute level of income,[34] since total conflict investment is only a function of levels of inequality, not absolute income. Any result would therefore depend on the choice of this parameter.

Another result demonstrated in Figure 1.3 is that the total conflict expenditure may be greater than the size of the contested budget, or indeed both public budgets combined. In panels (c) and (d) $C = E = 10$, which is lower than total expenditure throughout the range of α depicted here - and the parameters used in this example are not unrealistic.[35] Social division and conflict mobilization leads to inefficiency through waste of resources, an effect which can be of a significant magnitude. Inefficiency caused by the threat of conflict has a substantial effect on the resources available to society. Although investing in conflict ensures that both alliances in society will gain

[33] To ensure that conflict investment is not increasing faster than total resources in the economy, and thus that resources are not exhausted in the limit, *per capita* income must be such that $y_{pT} \geq \frac{1}{\gamma n_r(n_T+\beta n_D)}$. If the constraint holds with equality they will increase at the same rate.

[34] More specifically, y_{pT}, the *per capita* income of the poor ethnic majority, in terms of which all other income measures can be expressed.

[35] As a more specific example, for population parameters approximately equivalent to the population characteristics of India ($\alpha = 5$, $\beta = 0.7$, $n_T = 0.85$, $n_p = 0.7$) and public goods budgets $C = E = 10$. Conflict expenditure has a minimum value greater than $\frac{2}{3}$ of the contested budgets for all α and is greater than the contested budget at $\alpha > 3$. As β varies total conflict expenditure is greater than the contested budgets for all β.

a share of the public good budget, this gain in total utility will often be outweighed by the cost of investment of individual income.

1.6 Destruction of Resources

The indeterminancy of the effects of population shares, which also contribute to some measures of inequality, extends to their impact on investment in conflict. However, the effect of δ, the parameter measuring the level of destruction caused by conflict, is much clearer, and interacts with the income distribution in society.

δ is an inverse measure of destruction in conflict: if active conflict obtains $(1-\delta)$ of the contested budget is destroyed. This parameter features in both the constraint describing the preferences of the excluded subgroup over ethnic and class mobilization, and the level of investment in conflict resources.

The direct effect of δ is on the payoff from conflict, and thus from mobilization, but there is no difference between ethnic and class mobilization in the proportion of resources destroyed, and so it does not influence preferences in this respect. However, the level of destruction also factors into the level of conflict investment, and so the relative costs of ethnic and class mobilization.

If the excluded subgroup is from the class with higher aggregate income and the ethnic group with lower aggregate income,[36] then as the destructiveness of conflict decreases, the decisive excluded subgroup is more likely to prefer ethnic alliances. Likewise if the aggregate income condition is reversed the excluded subgroup is more likely to prefer class alliances as destructiveness decreases. This result is stated more formally in the following proposition.

[36] i.e. Its aggregate income is less than that of its potential class alliance partner and greater than that of its potential ethnic alliance partner.

Proposition 6 *When a subgroup (I, J) is in government, if ethnic mobilization takes place at some measure of destruction, δ', then,*

- *If $Y_{-I} > Y_I$ and $Y_{-J} < Y_J$, ethnic mobilization will take place at all $\delta > \delta'$.*
- *If $Y_{-I} < Y_I$ and $Y_{-J} > Y_J$, ethnic mobilization will take place at all $\delta < \delta'$.*

In all other cases the comparative statics in δ are indeterminate, and depend on the other parameters of the model. A proof of this proposition is contained in the First Appendix.

This result follows from the fact that when conflicts are less destructive the share of the public goods budgets to be gained from mobilization and the viable threat of conflict is larger, providing both subgroups in an alliance with a greater incentive to invest. When all contribute more the difference in the relative investment shares of the allied subgroups declines. Lemma 1 showed that the subgroup with the higher aggregate income invests more in any alliance. Therefore with less destructive conflict, if the excluded subgroup belongs to the poorer ethnic group and richer class, its share of contributions in a class alliance will increase while its share of contributions in an ethnic alliance will decrease, thus tending its preferences towards ethnic mobilization (and *vice versa* if it belongs to the richer ethnic group and aggregately poorer class, leading preferences to favor class mobilization).

As would be expected, when conflict is less destructive there is greater incentive to invest, and so the contributions of all subgroups, and therefore also aggregate contributions, increase. Proof of this result is omitted, as it is trivial and follows directly from the derivatives of contributions with respect to δ. This result holds weakly, since contributions beyond the range in which all subgroups have non-zero contributions are not a function of δ.

1.7 Conclusion

This chapter has developed and analyzed a model of alliance formation and investment in conflict in a society characterized by vertical and horizontal variation in *per capita* income. When a single population subgroup is in government, modeled as having responsibility for the allocation of public goods budgets, society always mobilizes, leading to inefficiency through investment in conflict, as a self-interested government will otherwise exclude those with opposing characteristics from benefiting from public resources.

The excluded subgroup, that which shares neither class nor ethnicity with the group in government, is decisive in determining whether class or ethnic divisions are salient in society. This group is disadvantaged with respect to both characteristics, so a society without mobilization is always its least favored outcome, as peacetime allocations give the excluded group no access to public resources. All else equal, this subgroup would prefer an alliance on the dimension on which has the greater economic disadvantage, as it will then bear a smaller share of the costs of the alliance.

These results hold regardless of the identity of the group in government, whether oligopolistic rule of the wealthy elite or a democratic government controlled by the poor majority through numerical dominance. However, the fact that there is a single subgroup is government is central to the results suggesting that the system of governance has a role in creating the environment which leads to conflict investment and inefficiency. It may therefore be useful to investigate the ways in which the structure of governance influences the outcomes, as this result suggests that limiting the exclusion of ethnic or economic groups may also limit the source of inefficiency. This is especially important when inequality is high as, in general, high inequality is associated with high levels of investment in conflict, and therefore greater waste of resources.

Chapter 2

The Art of the Possible: Information, Conflict and Political Power

"Politics is the art of the possible." Otto von Bismarck (1867)

2.1 Abstract

Can democratic leaders get away with lying? This chapter uses a game-theoretic model of conflict between two countries with informed leaders to examine the influence of domestic political concerns on the likelihood of truthful revelation and the probability of conflict. A democratic leader, who risks losing power if he lacks sufficient public support, may wish to reveal his private information about the costs of conflict. Democrats gain credibility through the need to inform their citizens, in contrast to autocrats, whose primary concern is to discourage their opponent from attacking. However, to gain support for his preferred strategy, a democratic leader may want to manipulate his population as well as his opponent, ensuring that there is not perfect information revelation. A democratic leader is more likely to overstate his country's vulnerability to attack when facing a democracy, but is more likely to understate vulnerability when facing an autocracy. Regime type has a greater influence than information revelation on the probability of conflict, which is highest for a pair of autocracies and lowest for a pair of democracies.

2.2 Introduction

Common opinion tends to suggest that politicians are not to be trusted, with a propensity to obfuscate, manipulate or hide information, or even lie outright.[1] Yet while democratically elected leaders may be considered untrustworthy, they have a legitimacy and credibility not afforded to autocrats. This chapter develops a game-theoretic signalling model in which democratic leaders can use their credibility, gained through the electoral process, to manipulate beliefs. It examines the effect of regime type on information transmission in the context of international conflict.

Bismarck, the Prussian Prime Minister and later first German Chancellor, was the first to describe politics as the 'art of the possible'. He was certainly a master of the art, overseeing the series of military conflicts culminating in the Franco-Prussian War, which brought about German unification. Military success can cement a leader's position of power, as was the case for Bismarck, but before launching conflict leaders must consider public opinion, knowing that defeat at the polls or removal from office can follow entry into an unpopular war. Anthony Eden was forced to resign as British Prime Minister following the debacle of the Suez Crisis in 1956; the backlash against the Vietnam War is attributed with destroying the career of President Lyndon Johnson.

Analysis of the model here demonstrates the trade-offs faced by democratic leaders, whose personal opinion about the desirability of war may not be shared by the electorate. It shows that a leader can increase public support for conflict by sharing information about the anticipated outcome, but if this knowledge will still leave a majority of the population opposed to war, he will be constrained by the

[1] For example, a 2010 survey by pollsters YouGov suggests that two thirds of the British public do not trust leading politicians to tell the truth even a "fair amount" of the time.

need to appeal to public support in order to remain in office. This is less of a concern to autocrats; although there are a variety of types of dictatorships, ranging from military rule to totalitarian regimes,[2] they share they the defining characteristic of not facing re-election, but can only be removed from office at extremely high cost. Modeling this feature reveals how democracy creates credibility: both democrats and autocrats have incentives to manipulate information to further their own ends, but unless there is also an incentive to tell the truth a leader will never be believed.

The restraining influence of democracy, which this chapter shows arising from the need to maintain electoral support, is also a factor in the 'democratic peace' hypothesis, the idea that democracy promotes peace. The theory dates back to the work of Kant and Paine in the late eighteenth century, and is supported by empirical evidence that conflict is more common between non-democratic states.[3] Yet democracies enter conflicts not only with autocracies but also with each other: notably the democratization of the former Yugoslavia led to years of brutal war. This chapter contributes to the literature on the democratic peace by showing that the effect of democracy dominates even when a leader is able to take advantage of his credibility to manipulate public opinion in favor of war.

The model combines a conflict game with private information with a signalling model, to examine the effect of democracy on information transmission and the likelihood of international conflict, and shows that information can only be credibly transmitted by a democratic leader. More specifically, the model analyzes the interaction between two countries assumes to have an inherently hostile relationship. The countries' leaders both choose between attacking and negotiating. Both citizens and

[2]Compare, for example, the totalitarian rule of Josef Stalin in the USSR, the oligarchic regimes of Ferdinand Marcos in the Philippines or Mobutu in Zaire, and half a century of military rule until 2011 in Burma.

[3]For example, Maoz and Russett (1993), Rousseau *et al* (1996), Leeds and Davis (1999).

potential leaders have a variety of preferences over the desirability of conflict, due to differences in the private cost of an attack. A leader also wants to remain in power. Simplifying from the extremely high cost and effort required to overthrow a dictator, the model uses an assumption that an autocrat will remain in office regardless of the outcome of the international conflict, while a democratic leader needs the support of at least half the population to do so.

Prior to choosing whether to attack or negotiate, a leader can send a public message about the country's vulnerability to attack, parameterized as the cost to all in a country when faced with a surprise attack, and assumed to be his private information. When vulnerability is high, attacking becomes more desirable. Vulnerability induces fear, as higher vulnerability entails a greater cost from failing to resist an attack. Citizens support war when they believe that their country is vulnerable, because they would rather make a preemptive strike than risk the costs of failing to resist a surprise attack. The signal allows a leader to inform his citizens about this vulnerability, but it is also observed by the opponent. A democratic leader therefore considers two audiences when choosing his signal.

The construction of the model illustrates the concept that hostility breeds hostility, as the gain from attacking is greatest when the opponent also attacks, due to the destruction and possible occupation resulting from facing an unresisted attack. Each country is therefore more likely to attack when its leader expects that his opponent also has a high probability of attacking. This feature demonstrates the way mutual fear can escalate into aggression, even between democracies.

Analysis of the model shows that an autocrat is never able to credibly transmit information. Although a democratic leader gains credibility, the temptation to manipulate both his population and the opponent ensures that there is not full revelation.

Whether he prefers to remain at peace or simply hopes to strike the first blow, any leader wants to deter his opponent from attacking, regardless of the truth of the statements he must make to do so. An autocrat has a single audience - the opponent - for any signal, and thus whatever the true vulnerability, he will always choose the same signal; therefore no information is transmitted.[4] A democratic leader, however, may face a conflict between ideology and pragmatism, between his personal preference about about the desirability of conflict and his wish to be re-elected. This provides a dual audience for any signal: the opponent and the domestic population, some of whose preferred strategy depends on the true vulnerability. A democratic leader may therefore want to inform citizens to increase support for his favored strategy, giving him credibility and allowing him to transmit information.

Observation of the persistence of incorrect beliefs, even in democracies, endorses the result that democratization is not sufficient to ensure that all information is revealed, and democratic leaders do not always tell the truth. The Soviet threat during the Cold War was persistently exaggerated; from the 1950s onwards the United States government had increasingly hard intelligence evidence that the USSR had no strategic weapons advantage, yet continued to imply the opposite. In more recent years, in the run up to the Iraq War the British and American governments manipulated evidence and relied on poor quality intelligence to exaggerate the threat of weapons of mass destructions (WMD). In both cases vulnerability was overstated to increase support for a hawkish strategy (military expenditure in the Cold War, war in Iraq).

A democratic leader needs public support; this may motivate him to reveal information about his country's strength or defenses, but it may also provide an incen-

[4]This analysis is purely in the context of conflict between pairs of countries and thus ignores strategic issues of interaction in a wider international community.

tive to lie or conceal information in an attempt to manipulate popular beliefs and increase support for his favored action. Analysis of the model shows that if a democratic leader's preferred strategy depends on the true vulnerability, as the median voter's does, he has an incentive to reveal this information, increase support for his choice of action. An incentive to lie develops when preferences are no longer aligned. A sufficiently hawkish leader always chooses to attack and thus, to increase support for this strategy, he will indicate that the country is vulnerable to a surprise attack, regardless of whether or not this is true. Others will give credence to this statement, because democratic leaders tell the truth more often than not.

The case of the Iraq War can be interpreted as a manipulation of information by intelligence agencies or government officials from democratic regimes, the United States and United Kingdom, both of which were inferred to have largely hawkish leaders. Although not a case of outright lying,[5] government actions lacked clarity, and used uncorroborated intelligence, despite a lack of hard evidence and the doubts of weapons inspectors and some in the intelligence community, to exaggerate the threat of WMDs and bolster wavering public support for the invasion. When it became clear that the WMDs did not exist at all, already wavering support for the war declined further.

An important result evaluates the influence of democratization of the opponent on the information revealed by a democratic leader. The presence of the domestic audience can induce truthful revelation, but still provides the possibility of false signals. A democratic leader may try to manipulate the population, but is also constrained by them; yet he must also pay attention to a second audience: the opponent. An

[5]The Hutton Inquiry cleared the British government of wrongdoing, but was itself met with scepticism.

autocratic opponent will be a more significant audience relative to the domestic population than a democratic opponent, as he cares little about his citizens' opinions and therefore is more susceptible to any signal received. This leads to the result that a democratic leader is more likely to understate his country's vulnerability when the opponent is an autocracy, to minimize the probability of being attacked, but has a higher probability of falsely stating that vulnerability is high when the opponent is also a democracy, to encourage public support for a war.

Comparison of equilibria for different country pairs reinforces the democratic peace hypothesis. Democratic leaders, like autocrats, may be hawks who always prefer to attack, but when a country democratizes, this has a direct regime effect reducing the likelihood of conflict, as the leader must consider the opinion of the population, the majority of whom are assumed not to be uncritically in favor of conflict. An information effect exists simultaneously: if it is revealed that the country is vulnerable to attack conflict becomes more likely, because public support for war is greater and, indirectly, because it encourages the opponent to attack. The regime effect dominates the information effect, since there is not complete information revelation, confirming the standard democratic peace result. The probability of conflict is greatest for a pair of autocracies and lowest for a pair of democracies.

2.3 Related Literature

There is a wide body of literature in both the economic and political science fields on the 'democratic peace' hypothesis, the observation that two democracies rarely fight each other. This phenomenon has quantified by, amongst others, Rousseau *et al* (1996) and Leeds and Davis (1999), who analyze foreign policy at both the monadic and dyadic level and show that democratic states behave less conflictually, an effect

that is compounded when democracies interact with each other. That observation is replicated in this chapter, which contributes to the analysis of the democratic peace by explicitly modeling the influence of domestic political concerns on the decision to go to war, and evaluating it in the context of information effects, which are shown to be dominated by the effects of regime type.

Baliga and Sjöström (2011) demonstrate that democracy acts as a check on leaders, using a model with complete information. The democratic process gives more weight to the opinion of the median voter, as a leader cares about his re-election prospects. The emphasis on the restraining influence of popular opinion is also present in the incomplete information models of Bueno de Mesquita *et al* (1999), which uses a model with rents from remaining in office, and Levy and Razin (2004), where a representative voter chooses the conflict strategy in a democracy, although these models depict societies with homogenous preferences. This chapter combines elements of the different models, using both incomplete information and heterogenous preferences to capture the diversity of opinions and uncertainly about the desirability of war and costs suffered in conflict. Uniting these features elucidates democratic leaders' incentives to both reveal and manipulate information, while demonstrating that the widely noted pacifying effects of democratization dominate even when a leader can take advantage of his credibility to manipulate public opinion.

Another explanation for the democratic peace is provided by Moaz and Russett (1993), who contrast the structural approach with a normative approach, suggesting that expectations of compromise and cooperation between democracies prevent conflicts of interest from erupting into violence. Tångeras (2009) posits the relative risk aversion of democratic leaders as an alternative explanation for the rarity of conflict between democracies. Democratic leaders are more reluctant to initiate wars, since

they may lose office following a bad conflict outcome. This argument is supported by the observation that democratic initiators are significantly more likely to win wars, as formalized by the analysis of Reiter and Stam (1998). Although this chapter assumes that leaders of all regimes have the same degree of risk aversion, it does not contradict that argument, as democratic leaders must also implicitly consider costs that will be born by citizens.

Also related is Jackson and Morelli (2007), which takes a different approach to analyzing the relationship between leaders and citizens, and show that only unbiased leaders are able to use commitment mechanisms to avoid war. This chapter examines the strategic interaction between countries to show that it is when a leader's preferences differ from those of the median voter that he has both the opportunity and incentive to manipulate the population. Baliga and Sjöström (2004) explore a different facet of this strategic interaction where similar modeling characteristics are used to illustrate how "fear and distrust" can spiral into an arms race situation.

This chapter also contributes to the literature on incomplete information and war. Incomplete or asymmetric information has frequently been modeled as a cause of conflict, including by Bester and Wärneryd (2006) and Kirshner (2000).[6] Riboni (2011) is particularly relevant, analyzing information transmission between an informed principal and a naive agent ahead of a conflict game, drawing on work on the significance of a dual audience in a cheap talk signalling game, first formalized by Farrell and Gibbons (1989). However, while Riboni assumes that the preferences of the principal and the agent are aligned, this chapter's most interesting results arise when the favored strategy of a democratic leader differs from that of his population.

[6] An overview is provided by Garfinkle and Skaperdas (2007).

2.4 Model

The model consists of a two stage game. A communication game, in which information concerning the payoffs from conflict may be revealed, is followed by a conflict game modeling the strategic interaction between the leaders of two countries. The conflict game draws on the model developed by Baliga and Sjöström (2004) and Baliga, Lucca and Sjöström (2011).

Formally, the game is divided into three periods. At $t = 0$ the state of each country, the level of vulnerability, is drawn by nature, and is observed only by the leader of that country. The leader of each country sends a public signal about its state at $t = 1$. The citizens and the opponent update their prior beliefs, then at $t = 2$ the leaders of the two countries play the conflict game. Finally the payoffs are allocated.

Conflict Game

The relationship between two countries is assumed to be inherently conflictual. The leaders of the two countries, indexed $i \in \{1, 2\}$, play a conflict game in which each chooses either an aggressive strategy, to *Attack*, or a conciliatory strategy, to *Negotiate*.

Each country has a continuum of citizens, who will each bear a private cost if the leader chooses to attack. This cost includes the risk of death if an individual is required to fight, as well as financial costs, either through direct contributions or increased taxation. It can also incorporate underlying attitudes about the desirability of war: aggressive individuals will have a low cost of conflict, but those who object to war on principle can be considered to have high cost. Cost type, c, therefore varies across the population: each citizen's type is drawn from a distribution F with support

$[\underline{c}, \overline{c}]$. It is assumed that F is strictly increasing and convex. The median cost type is denoted c^m, $F(c^m) = 0.5$.

There is no *ex ante* difference between the two countries; that is, both have the same distribution of cost types. In practice, of course, countries have a variety of underlying differences that may affect the probability of one attacking another - for example, relative size and GDP. The United States invaded Grenada in 1983, but it would have been impossible for Grenada to have the forces to invade the USA. The assumption of distributional symmetry allows this chapter to focus on the effects of regime type and information, abstracting from other factors influencing strategy choices in international relations.

The payoffs for a citizen of country i with cost type c are shown in the matrix below. The row gives the strategy choice of leader i and the column the choice of the leader of the opposing country, j.

i \ j	*Attack*	*Negotiate*
Attack	$-c$	$\mu - c$
Negotiate	$-d_i$	0

When leader i chooses to attack, each citizen of country i will always pay cost c (each according to his own draw from F). If i attacks but j does not each citizen also obtains benefit μ. This gain to being offensive can be interpreted as a 'first mover advantage' for the country; benefits include the element of surprise, better coordination of resources and military strategy, or keeping conflict on the opponent's territory, all net of any sanctions for aggression imposed by the international community.

If j attacks but i does not, each citizen suffers a common loss relating to the country's vulnerability, d_i.[7] This cost can take on two values, $d_i \in \{d^L, d^H\}$, where $d^H > d^L$. Vulnerability measures the ability to withstand costs and damages caused by an attack, and so depends on defenses and military resources. A country will be vulnerable to attack if, for example, it is lacking in defense infrastructure or has a poorly trained military. For simplicity of explication, a country will be referred to as 'vulnerable' if $d_i = d^H$ and 'not vulnerable' if $d_i = d^L$ for the remainder of this chapter.

The gain from unresisted offense and the country's vulnerability are the same for all citizens and the leader, but while μ is common knowledge, the realization of d_i is only observed by leader i. A country's leader will have complete knowledge of its defenses and military resources, not all of which is shared with the general population, as well as intelligence reports and communications from private negotiations which may provide information about how successfully the country will be able to respond to an attack.

The conflict game exhibits strategic complementarities: it is assumed that $0 < \mu < d^L < d^H$, so the gain from attacking is highest if the opponent also attacks. This feature demonstrates the propensity of conflicts to escalate, since fear that the opponent will attack makes aggression more desirable.

Attack is a dominant strategy for any citizen with cost type $c < \mu$, henceforth referred to as hawkish types (or 'hawks') since they always prefer to attack regardless

[7] It may be more realistic to assume that country i will suffer a loss whenever it is attacked, but this cost is smaller when country i also attacks. This could be added to the model by changing the payoff to a citizen of country i when both countries attack to $-\alpha d_i - c$, where $\alpha \in [0, 1]$. Provided that α is sufficiently small, $\alpha < \frac{d^H - \mu}{d^H}$, the results will be unchanged. Otherwise the strategic complementarity effect is reversed, so the basic information result remains but the direction of all effects is also reversed. Thus, for clarity, this additional cost is omitted, equivalent to setting $\alpha = 0$.

of the action of the opponent or the country's vulnerability. *Negotiate* is a dominant strategy for pure dovish types ('doves') with $c > d^H$, pacifists who prefer to negotiate even if the opponent attacks and the country is vulnerable to attack. Citizens without a dominant strategy are 'co-ordination types': attacking is their best response if the opponent attacks and negotiating their best response if the opponent negotiates. Citizens with cost type $\mu < c < d^L$ are always co-ordination types. This leaves those with private cost in the range $d^L < c < d^H$, who will be co-ordination types if $d_i = d^H$ but doves if $d_i = d^L$. The cost of aggression is high enough to dissuade them from unreservedly favoring attacking, and if the country is not vulnerable they will take the chance of being attacked by the opposing country, but if vulnerable they are unwilling to risk this chance and prefer to attack if the opponent also does.

The fraction of hawks in the population is always the same, $F(\mu)$, but if the country is vulnerable there are more coordination types and less doves: $1 - F(d^H)$ doves and $F(d^H) - F(\mu)$ coordination types, compared with $1 - F(d^L)$ doves and $F(d^L) - F(\mu)$ coordination types when it is not vulnerable. A greater proportion of the population support a coordination strategy (and therefore will support attacking if they believe the opponent is sufficiently likely to do likewise) since a vulnerable country faces greater losses when its opponent launches a surprise attack.

The following assumptions, illustrated in Figure 2.1, summarize the distribution of cost types in the population.

A1: $0 < \underline{c} < \mu < d^L < c^m < d^H < \overline{c}$

This assumption ensures that the median voter is a co-ordination type if $d_i = d^H$ and a dove if $d_i = d^L$. It implies that, as is expected given the expense, destruction and potential loss of life incurred, a representative citizen is never uncritically in

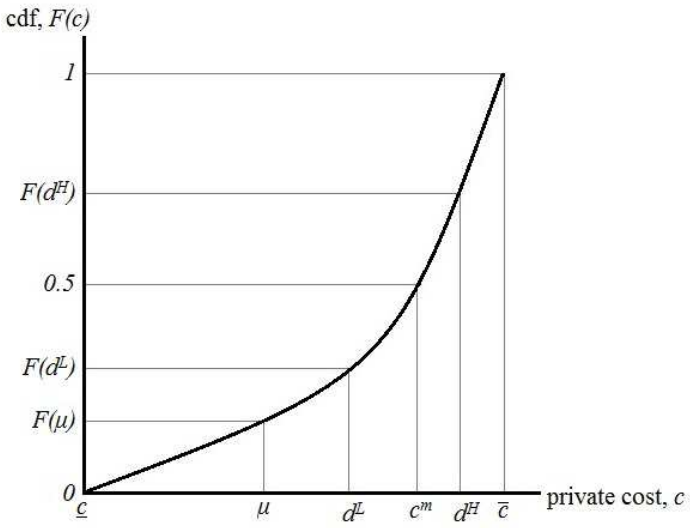


Figure 2.1: Distribution of Private Cost Types

favor of war, but will be willing to support an attack when the costs of attempting a conciliatory strategy are high.[8]

A2: $c^m - \mu > d^H - c^m$

The second assumption is necessary to ensure that N will always be the risk dominant strategy for the median voter. It is relevant when $d_i = d^H$ and the median

[8]The assumption that the median voter is not a pure strategy hawk is standard in conflict theory. See Baliga, Lucca and Sjöström (2011), Jackson and Morelli (2007), Schultz (2005), Tångeras (2009) among many others.

voter is a coordination type (since negotiation is the dominant strategy for doves) and ensures that the gain from choosing to negotiate when the opponent also negotiates, $c^m - \mu$, is greater than the loss from remaining conciliatory when the opponent attacks, $d^H - c^m$.

Leader i is assumed to be chosen from the population of the country, and so has cost type c_i^l independently drawn from F. Thus the leader may be a hawk, a dove or a coordination type. In particular, it means that, while the median voter can only be a dove or a coordination type, the leader may be a hawk, and so unreservedly aggressive.[9] War has frequently been presented as a game where decision makers are dominant strategy hawks.[10] This chapter uses a weaker assumption: it allows for the leader to be more aggressive than the median voter, but does not insist that this is always the case.

A leader's cost type is his private information, although the distribution is common knowledge. It may be argued that, at least in democracies, voters will have some information about the underlying preferences of their leaders before they are elected, based, for example, on election manifestos or previous voting record. Yet it cannot be expected that citizens will have complete information about the preferences of candidates: manifesto promises often remain unfulfilled, and previous political actions may have been constrained by political expediency or party orders, such as whipped votes in legislatures with 'first past the post' electoral systems. Therefore, for the sake of clarity, individual realization of cost type is assumed to be private information.

[9] Equally, it is possible that, when $d_i = d^H$ and the median voter is a coordination type, the leader will be a dove and so less aggressive. The key is that 'attack' will never be a dominant strategy for the median voter, but may be for the leader.

[10] See, for example, Jackson and Morelli (2007) on the selection of biased leaders, or Kirshner's (2007) discussion of leaders' preferences for taking risks.

This chapter compares the two extremes of political regime: autocracy and perfect democracy. It is supposed that an autocrat will remain in power whatever happens, but that a democratic leader requires the support of at least half the population to keep his position. Country i's regime type is given by $T_i \in \{A, D\}$, where A denotes autocracy and D denotes democracy. The regime types of both countries are common knowledge.

All leaders are self interested and want to maintain their power. The utility of doing so is modeled as a reward from staying in office, $R > 0$.[11] The size of R is not restricted, but a larger value magnifies the effect of democracy, since, regardless of regime type, payoffs reduce to those from the basic conflict game as R tends to zero.[12]

If country i is an autocracy the leader will remain in power, and receive reward R, regardless of whether the median voter, or indeed any of the citizens, agree with his choice of action, and thus irrespective of the outcome of the conflict game. Leader i's payoff matrix reduces to that of the basic conflict game given earlier, with his own private cost type, $c = c_i^l$, as R can effectively be dropped.

If country i is a democracy, the leader's decision between attacking and negotiating becomes a principal-agent game, where the citizens act as a principal who can reward the leader (agent) with their support. A citizen will support the leader, which can be interpreted as voting for him in a re-election contest, if the leader has played his preferred action, given the action of the opponent. The decision to support is therefore taken *ex post*: the population observe the actions taken by the leaders of both countries. However, it is made before all payoffs are realized, since it may be

[11] This feature is also used by Bueno de Mesquita *et al* (1999).

[12] Instead of modeling a reward from remaining in office, an alternative approach to the problem would be for the median voter to play the conflict game in a pure democracy.

many years before the final outcome of a war is known. Election of a political leader often takes place in a shorter time frame than the resolution of a conflict - consider two wars in which the United States was a party: there were two Presidential elections during the Iraq War, and five during the Vietnam War.

This is important because the level of vulnerability, d_i, which is not directly observed by the citizens, could be revealed through the payoffs when the conflict is resolved (if i negotiates and j attacks). Instead a citizen's choice of whether or not to support the leader must be based on his belief of the probability that his country is vulnerable, denoted ρ_{ii}. It is assumed that this belief is the same for all citizens, since they should all have access to the same underlying information in the public domain.

The payoff matrix for the democratic leader of country i is given by

i \ j	*Attack*	*Negotiate*
Attack	$-c_i^l + \rho_{ii}R$	$\mu - c_i^l$
Negotiate	$-d_i + (1-\rho_{ii})R$	R

Here ρ_{ii} also represents the probability of the leader's reelection when both countries attack. If the opponent chooses to negotiate the probability of re-election does not depend on the realization of d_i. All except dominant strategy hawks will prefer to negotiate, forming a majority, so a leader who chooses to negotiate when his opponent negotiates will always be re-elected and receive reward R. However, if the opponent attacks, a majority consisting of all except dominant strategy doves will prefer to attack if it is known that the country is vulnerable. If it is not vulnerable, only hawks

and pure coordination types will prefer attacking, and the leader will not have sufficient support to remain in power. Therefore, when choosing whether or not to support the leader those citizens who do not have a dominant strategy will have to consider the probability that the country is vulnerable. They will support attacking if this belief is sufficiently high, and so the probability of re-election, and thus the payoffs to the leader, are a function of the beliefs of the population. (A complete explanation for this support structure is provided in the Second Appendix.)

Information and Communication

The model features both strategic and structural uncertainty. Strategic uncertainty arises from the simultaneous interaction in the conflict game, while structural uncertainty exists because the level of vulnerability of each country is its leader's private information. A leader is involved in international diplomacy and has access to defense and intelligence reports, and so will be better informed than the population about the expected costs of conflict and the relative strengths of the two countries.

Nature randomly draws the state of the world, the vulnerability of each country: $\Omega = \Delta \times \Delta$, where $\Delta = \{d^L, d^H\}$. The vulnerability levels of the two countries are independent, ensuring that their underlying population shares of different private cost types are unrelated.[13]

A country's vulnerability directly affects payoffs, and the preferred conflict strategy of some cost types, so it is clearly of interest to the domestic population.

[13]This is not an unreasonable assumption, since it should not be supposed that because one country is vulnerable to attack, its opponent will not be. Vulnerability is not simply a question of military capability; it also depends on defenses, domestic stability, and economic prosperity, among other factors. This myriad of factors also helps to explain why, despite intelligence sources, which cannot be assumed to reveal full information) a leader does not know his opponent's vulnerability.

Further, when a country is vulnerable it can be considered more aggressive, as there are less dovish types in the population. Vulnerability influences the probability of an attack, the strategic uncertainty in the conflict game, so it is also of interest to the opponent, despite not entering directly into his payoffs. Thus there are dual audiences for any signal sent by leader i about the realization of d_i: the opponent and the domestic population.

This leads to the first stage, the communication game. Leader i sends a signal $m_i \in \{m^L, m^H\}$, indicating whether vulnerability is low or high. For example, if a country engages in military manoeuvres or publicizes its defenses it is signalling that it is not vulnerable, but if the government is shown to be making military cutbacks or claims that the opponent has weapons of mass destruction it is indicating vulnerability. The message is public, and so is observed by both his citizens and the opponent before the conflict game is played. No private communication is possible within a country between the leader and the population. Since such communication is likely to take place through the media, public announcements and debate, it is unlikely that any significant information could be transmitted to the population as a whole but concealed from outside observers, and in any case, it would be practically impossible to conceal broad communication within a country from foreign intelligence agents. The communication game is also a simultaneous move game, so no leader is able to condition his signal on the signal of his opponent.

These messages are 'cheap talk'; the leaders do not pay any cost to send them. In addition, they cannot be verified by either the opponent or the population. Again, this relates to the sources of the leader's private information, ensuring that it would not be possible for citizens to independently verify the leader's claims, even if they invest in research. A general audience cannot have access to sensitive intelligence and

defense sources, and the complexity of amalgamating the the potential costs of being attacked would likely have an unfeasible level of expense for private citizens.

Once the signal is observed all parties will update their prior beliefs according to Bayes rule. Just as ρ_{ii} denotes the citizens of country i's prior probability that their own country is vulnerable, ρ_{ji} denotes leader j's prior probability that country i is vulnerable.[14] All citizens are assumed to have the same prior, but the priors of the domestic population and the opposing country are allowed to differ, since they may well have different underlying information and beliefs - for example, private communication between heads of state.

The posterior will therefore be given by

$$\widehat{\rho}_{ii}(m_i) = \frac{\rho_{ii} Pr(m_i|d^H)}{\rho_{ii} Pr(m_i|d^H) + (1-\rho_{ii}) Pr(m_i|d^L)},$$

where $Pr(m_i|d_i)$ is the conditional probability that signal m_i is observed given that true state for country i is d_i, and likewise for $\widehat{\rho}_{ij}$.

At each stage of the game the leaders maximize their expected utility, conditional on beliefs and the strategy of the opponent.

2.5 Characterization of Equilibria

When choosing whether to attack or negotiate, leader i must consider his own cost type c_i^l, the level of vulnerability d_i, the expected probability of being attacked, and, if a democracy, the domestic population's beliefs. When choosing which signal to send,

[14] For clarity, it is assumed that all priors are common knowledge, but the general results should be unchanged if they are unknown. Assuming ρ_{ii}, $\rho_{ji} \in (0,1)$ ensures that the population and opponent do not know the state of nature with certainty - otherwise the signal choice becomes irrelevant.

he must consider the effect on beliefs and the expected probability of an attack, and therefore the expected payoff in the conflict game.

The probability that country i attacks, denoted q_i, will not be known by the opponent, since d_i is private information. Thus $\widetilde{q}_i$ is used to denote the expected probability that country i will attack.[15]

The equilibrium will be given in cut-off strategies relating to the cost type of the leader. Comparing the payoffs from the conflict game, an autocratic leader of country i will choose to attack if

$$-\widetilde{q}_j c_i^l + (1-\widetilde{q}_j)(\mu - c_i^l) > -\widetilde{q}_j d_i.$$

Thus he attacks if and only if his cost type is sufficiently low, $c_i^l < \mu + \widetilde{q}_j(d_i - \mu)$. The results focus on the underlying probability of an attack; the probability of an autocracy launching conflict is given by

$$q_i = F[\mu + \widetilde{q}_j(d_i - \mu)].$$

Likewise, democratic leader i will choose to attack if,

$$\widetilde{q}_j(-c_i^l + \widehat{\rho}_{ii}(m_i)R) + (1-\widetilde{q}_j)(\mu - c_i^l) > \widetilde{q}_j(-d_i + (1-\widehat{\rho}_{ii}(m_i))R) + (1-\widetilde{q}_j)R,$$

giving the probability of attack by a democracy as

$$q_i = F[\mu + \widetilde{q}_j(d_i - \mu + 2\widehat{\rho}_{ii}(m_i)) - R].$$

The probability of attacking is a function of the expected probability of being attacked, as illustrated in Figure 2.2. It is depicted in the cases of high and low cost

[15] Leader j's belief of the likelihood of an attack by country i is the relevant expected probability. However, equilibrium beliefs must be consistent and so will be common knowledge. Therefore, to simplify notation, only a single subscript is used.

of defensiveness for both an autocracy and a democracy. The underlying probability of attacking is an intermediate value, denoted by

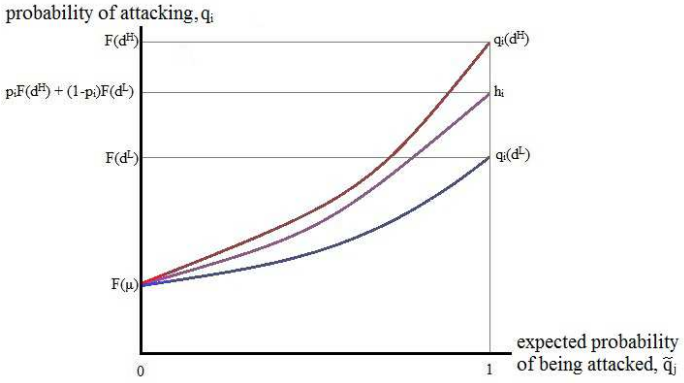
$$h_i = p_i q_i(d^H) + (1 - p_i) q_i(d^L),$$

where p_i is the underlying probability that $d_i = d^H$. Holding beliefs and the true level of vulnerability constant, the probability of a democracy attacking is never higher than the probability of an autocracy attacking when $\widetilde{q}_j < 0.5$, and may not be for higher values of $\widetilde{q}_j$.[16]

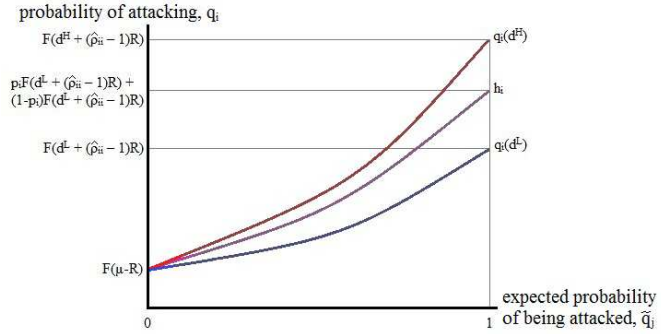
The solution to the game is given by separating Weak Perfect Bayesian equilibria. In addition to the influence of regime type, the solution in cut-off strategies means that leaders choose signals and actions in the conflict game according to their private cost type.

The equilibrium concept specifies that each player's strategy prescribes optimal actions given his beliefs at each decision point and the strategies of all other players. Equilibrium requires consistency of beliefs; that is, all parties will update their beliefs as all others expect them to. Thus all equilibrium beliefs will be common knowledge. The probability of country i attacking is a function of the expected probability that country j will attack, and leader i also knows that the probability that country j attacks will be a function of leader j's expected probability that country i will attack. Therefore, to expand the previous notation, the belief function depends on the regime type and the signal sent, $\widetilde{q}_i^{T_i}(m_i)$:

[16] $q_i^D > q_i^A$ iff $\widetilde{q}_j > \frac{1}{2\widetilde{\rho}_{ii}(m_i)}$.



(a) Autocracy



(b) Democracy

Figure 2.2: Probability of an Attack

$$\widetilde{q}_i^A(m_i) = \widehat{\rho}_{ji}(m_i)F[\mu + \widetilde{q}_j(d^H - \mu)] + (1 - \widehat{\rho}_{ji}(m_i))F[\mu + \widetilde{q}_j(d^L - \mu)], \tag{2.1}$$

$$\widetilde{q}_i^D(m_i) = \widehat{\rho}_{ji}(m_i)F[\mu+\widetilde{q}_j(d^H-\mu+2\widehat{\rho}_{ii}(m_i))-R]+(1-\widehat{\rho}_{ji}(m_i))F[\mu+\widetilde{q}_j(d^L-\mu+2\widehat{\rho}_{ii}(m_i))-R]. \tag{2.2}$$

Consistent equilibrium beliefs will be given by a system of two equations, as illustrated in Figure 2.3. There will be a single point of intersection as, for all regime types, the probability of attacking is strictly increasing in the probability of the opponent attacking. This ensures that equilibrium beliefs about the probability of an attack, a function of the regime types of both countries and the signals sent by both leaders, $\widetilde{q}_1^{TT*}(m_1, m_2)$, $\widetilde{q}_2^{TT*}(m_1, m_2)$, will be uniquely defined.

The true probability of launching an attack is a function of both vulnerability and the expected probability that the opponent will attack, and so of the signals of both leaders: $q_i^{T_1T_2}(m_1, m_2; d_i)$.

The model is analyzed by comparing three pairs of regime types that may be involved in an international dispute: two autocracies (A, A), an autocracy and a democracy (A, D), and two democracies (D, D). Without loss of generality, assume that in an (A, D) dyad $T_1 = A$ and $T_2 = D$. The first step is to characterize the behavior of the leaders in the communication stage.

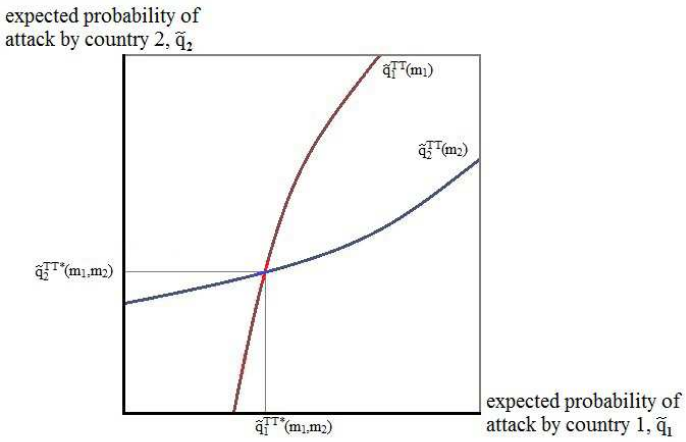


Figure 2.3: Equilibrium Beliefs

Equilibrium Information Transmission

Every leader wants to discourage his opponent from attacking. An autocrat only has a single audience for his signal, the leader of the opposing country, who is aware of the motivation governing signal choice. The autocrat is therefore unable to transmit credible information. In contrast, a democratic leader is constrained by domestic political considerations and his desire for re-election. This can provide an incentive to reveal the country's true level of vulnerability of its citizens, to maximize public support for his choice of action in the conflict game. Knowing this, the citizens and opponent give credibility to his signal, ensuring that some information is transmitted in equilibrium.

Proposition 1 *In all equilibria, an autocrat is never able to transmit credible information. Some information is credibly transmitted by a democratic leader, but there is never full revelation.*

A proof of this proposition is provided in the Second Appendix.

The signal is a cheap talk message and does not directly affect the autocrat's payoff, so when choosing his signal he only considers its effect on his opponent's behavior. Regardless of whether the home country attacks or negotiates, and irrespective of private cost type, all individual payoffs will be higher if the opponent negotiates. If one signal were more likely to dissuade the opponent from attacking, the autocrat would always send this signal, whether or not his country was vulnerable. The opponent is sceptical and will be aware of the autocrat's incentives, so in this case would expect to always observe the same signal. Yet if this is so the signal is not informative and gives the opponent no knowledge about the autocracy's vulnerability. The opponent's posterior belief will remain the same as his prior and his behavior will be unchanged. Thus there cannot be a signal which is more likely to dissuade the opponent from attacking and an autocrat will always choose to randomize between signals. No signal sent by an autocrat can affect the behavior of the opposing leader. This holds regardless of the regime type of the opposing country.

A democratic leader faces two distinct audiences when choosing which signal to send. Like an autocrat, he wants to discourage his opponent from attacking. However, the presence of the domestic audience creates an endogenous cost to sending a false signal, since the population's belief affects the probability of re-election and thus the leader's payoff. The need to publicly coordinate with his citizens provides an incentive to reveal the country's true vulnerability, giving a democratic leader credibility. This

result is a separating equilibrium where the choice of signal is a function of the leader's private cost type.[17]

More specifically, the probability of the opponent attacking is increasing in his belief that the first country will attack, and the probability of starting conflict is greater when vulnerability is high. All else equal, any leader wants to convince his opponent that his country is not vulnerable, and so deter him from launching conflict. In contrast, a democratic leader whose preferred action depends on the realization of d_i will want to reveal this information to his citizens, to ensure that the majority of the population supports his choice of action and he remains in office.

Democratic leaders gain credibility through the need to inform their citizens, but information is not perfectly revealed as there is still an incentive to manipulate both the domestic population and the opponent. Although leaders with intermediate cost types want to reveal the true level of vulnerability, those with particularly high or low private cost of attacking also want to report false information to their citizens.[18] A leader who is a dominant strategy dove, $c_i^l > d^H$, will always prefer to negotiate, and so, to maximize support for this strategy, would always like to inform his citizens that the country is not vulnerable. A sufficiently hawkish leader prefers to attack

[17] This result will be unchanged if the leader is allowed to engage in diplomacy and communicate privately with the leader of the opposing country. As with public signalling by an autocrat, the leader will always choose to send the private message most likely to dissuade the opposing country from attacking; therefore no credible information can be privately communicated between leaders of opposing countries.

[18] Specifically, a leader who could communicate privately with the domestic population would choose to send signal $m_i = m^H$, indicating that the country is vulnerable to attack, for $c_i^l < d_i$. Dominant strategy hawks would always angle for a preemptive strike by signalling that the country is vulnerable, and dominant strategy doves would always indicate that attack can be resisted, to generate support for their non-state dependent preferred strategy.

whenever the opponent does,[19] and thus would like the population to always believe that the country is vulnerable.

The equilibrium is given in cut-off strategies; leaders with a preference for conflict (low private cost of attacking) signal that the country is vulnerable to attack while those who prefer peace (high private cost) indicate the reverse. The individual cost type at which the democratic leader of country 2 is indifferent between signals, denoted $\widetilde{c}_2^{T_1 D}(d_2)$, is a function of the revealed vulnerability level d_2. The leader will send signal $m_2 = m^H$ if $c_2^l < \widetilde{c}_2^{T_1 D}(d_2)$, and signal $m_2 = m^L$ for all other private cost types.[20] The probability of sending signal m^H, $F[\widetilde{c}_2^{T_1 D}(d_2)]$, is higher when the country is vulnerable, since $\widetilde{c}_2^{T_1 D}(d^H) > \widetilde{c}_2^{T_1 D}(d^L)$, thus providing credible, though incomplete, information.

The true state may affect the choice of signal in a democracy, adding credibility and so influencing popular support and potentially the action of the opponent, but the possibility of lying remains. Indeed, a strongly hawkish or dovish leader may lie about the gains from attacking to increase his popular support. His statements will have credibility since, on average, a democratic leader will be telling the truth - thus enabling him to, effectively, get away with lying.

This illuminates the situation leading up to the Iraq War in 2003. The hawkish democratically elected governments in the United States and United Kingdom allowed dubious intelligence, lacking supporting evidence and ignoring contrary information, to be used to exaggerate their countries' vulnerability. British Prime Minister Tony

[19] $c_i^l < d^L$, including coordination types, since the population's belief is only relevant when the opponent attacks.

[20] Note that two other equilibria exist. There is also a babbling equilibrium where the democratic leader randomizes between signals and no credible information is transmitted, and an equilibrium where the signals are 'flipped'. This is identical to the equilibrium described, but the signals are reversed: the democratic leader will send signal m^L when his individual cost type is low and m^H when his cost type is high.

Blair wrote, "[Saddam Hussein's] military planning allows for some of the WMD to be ready within 45 minutes of an order to use them," in the foreword to the notorious September Dossier, while President Bush used the same sources when he stated that, "The British government has learned that Saddam Hussein recently sought significant quantities of uranium from Africa," in his 2003 State of the Union Address."[21] The population, and the legislature, gave credence to their statements, since such democratic governments do not have a history of misinformation, but the claims were proven entirely untrue: the documents from Niger were fakes, and not only were missiles unarmed, but the weapons did not exist at all. Although popular support for the war was always limited, it further declined as the truth about the absence of WMDs became common knowledge.[22]

Truth and Lies

The previous result shows that a democratic leader is able to transmit information about the country's vulnerability, but this credibility also allows him to 'get away' with lying, as the population updates posterior beliefs knowing that democrats tell the truth more often than not. A democratic leader's choice of signal is influenced by the regime type of his opponent as well as his own private cost type, but it is not immediately apparent whether truthful revelation by a democratic leader is more likely when the opposing country is an autocracy or a democracy.

The threshold cost type at which the leader is indifferent between signals depends on the regime type of the opponent but this constraint is a function of equilibrium values as well as parameters. Equilibrium beliefs are determined simultaneously in

[21] Foreign and Commonwealth Office (2002), Bush (2003).

[22] Of course, this example is further clouded by allowing for a degree of costly independent verification, the reports of weapons inspectors.

the signalling game, so the possibility of credible information transmission by the opponent determines the relative significance of the two audience. When the opponent is also a democracy, he too will be constrained by domestic considerations, but an autocrat will be more susceptible to the signal received, as he only needs to consider the direct outcome of the conflict game. When choosing his signal, a democratic leader is more concerned about the effect on the opponent relative to the domestic population when his opponent is an autocrat.

Comparing the cut-off values shows that the probability of truth-telling depends on the true level of vulnerability as well as the opposing regime type. The leader is more likely to truthfully reveal high vulnerability to attack if the opponent is an autocracy, but is more likely to be truthful about having low vulnerability when the opponent is also a democracy.

Proposition 2 *A democratic leader has a higher probability of falsely signalling that the country has high vulnerability ($m_2 = m^H$) when the opposing country is a democracy, but a higher probability of signalling that vulnerability is low ($m_2 = m^L$) when the opposing country is an autocracy. That is,*

$$\widetilde{c}_2^{AD}(d^H) > \widetilde{c}_2^{DD}(d^H) \qquad \textit{and} \qquad \widetilde{c}_2^{AD}(d^L) > \widetilde{c}_2^{DD}(d^L).$$

The parameters of the conflict game enter the constraint in the same way for both dyads, so the effect of the equilibrium values, determined simultaneously, leads to this result. Earlier results show that the expected probability of conflict is always lower when the opponent is a democracy, and this effect dominates, lowering the cut-off values. The range over which signal m^H is chosen is smaller whenever the opponent is a democracy, as $\widetilde{c}_2^{AD}(d_i) > \widetilde{c}_2^{DD}(d_i)$ for both levels of vulnerability. A democratic leader cannot be deemed to be 'more truthful' overall when facing a democracy or an

autocracy. A democratic leader has more incentive to hide vulnerability when facing an autocracy but to hide strength against another democracy.

Echoing the main result, this claim also follows from the fact that an autocrat is only concerned with international interaction in the conflict game, but a democratically elected leader also has the domestic concern of re-election. When country 1 is an autocracy, its leader does not care about his citizens, so will be more influenced by the behavior of country 2. Since he has more influence on an autocratic opponent, and always wants the opponent to think that his country is not vulnerable (and so is himself less likely to attack), he is more likely to lie and falsely state that the country is not vulnerable when the opponent is an autocracy. However, when country 1 is a democracy its leader in constrained by domestic concerns, and so will be relatively less influenced by the signal of country 2. The democratic leader has less influence on his democratic opponent and so is relatively more concerned about influencing his own citizens. Thus a greater range of hawkish leaders will lie to try to increase popular support for attacking, by declaring that the country is vulnerable to attack.

2.6 Conflict in Equilibrium

The previous section characterizes a key difference between autocracies and democracies: only a democratic leader is able to transmit credible information, due to the presence of a second audience, the domestic population. The choice of signal may then influence the behavior of both parties in the conflict game. When choosing the signal in the communication stage, any democratic leader anticipates that he will attack after sending signal m^H and negotiate after sending signal m^L. However, when the opponent is also a democracy the information revealed by the opposing leader may change this intention.

The influence of domestic political concerns on the outcome of the conflict game, both directly and via information revelation, are examined to assess the impact of regime type on the probability of peace. Whatever the country's level of vulnerability, democracy reduces the risk of conflict through the need to appeal to the (never hawkish) median voter, yet when a leader credibly signals that vulnerability is high it increases both countries' probability of attack. It is not obvious which effect will dominate, so it is necessary to compare the equilibria of the game and the likelihood of conflict for different dyads. The first step is to examine equilibrium beliefs about the probability of an attack.

Expected Probability of Conflict

The expected probability of conflict, denoted $\widetilde{w}$ when conditional on the signals observed, equates to the expected probability that at least one country attacks:

$$\widetilde{w}^{T_1T_2}(m_1, m_2) = \widetilde{q}_1^{T_1T_2*}(m_1, m_2) + (1 - \widetilde{q}_1^{T_1T_2*}(m_1, m_2))\widetilde{q}_2^{T_1T_2*}(m_1, m_2).$$

The underlying expected probability of conflict incorporates the likelihood of observing each signal:

$$\widetilde{W}^{T_1T_2} = \sum_{m_1}\sum_{m_2} \widetilde{w}^{T_1T_2}(m_1, m_2)Pr(m_1|T_1T_2)Pr(m_2|T_1T_2).$$

Beliefs must be based on common knowledge of the distribution from which the leader is drawn, as his individual cost of attacking, c_i^l, is his private information. This allows a hawkish democratic leader to use the credibility provided by his election to manipulate for war. However, the equilibrium expected probability of conflict between democracies is lower than that of conflict between autocracies, even if two

democratic leaders both signal that their countries are vulnerable. The direct effect of democratization outweighs the increase in the expected likelihood of conflict which may arise from greater public support for conflict and the opponent's perception of greater vulnerability.

Proposition 3 *The equilibrium expected probability of conflict is lowest for a pair of democracies and highest for a pair of autocracies:*

$$\widetilde{W}^{AA} > \widetilde{W}^{AD} > \widetilde{W}^{DD}.$$

A proof of this proposition is provided in the Second Appendix.

No information is credibly transmitted if both countries are autocracies, so the expected probability of an attack is informed by the prior beliefs of the leaders. Based on common knowledge of the distribution of private cost types, both know that only a minority of potential leaders of the opposing country will be hawks, and indeed if the country is not vulnerable the majority will be doves. Negotiation is the risk dominant strategy for the median citizen - and therefore median potential leader - ensuring that in equilibrium both countries are more likely to negotiate than to attack, $\widetilde{q}_i^{AA*} < 0.5$, as would be expected. The majority of countries in the world remain at peace, even when there are disagreements and tensions between nations.[23]

When country 2 is instead a democracy the best response belief function for the probability of an attack by country 1 remains unchanged, but leader 2 must consider public opinion. This comparison (equations (2.1) and (2.2)) is shown in Figure 2.4. The democratic leader now incorporates popular beliefs when choosing his strategy,

[23]Figures from the Uppsala Conflict Data Program show that, since the end of the Second World War, there has never been more than ten armed conflicts between countries in a single year.

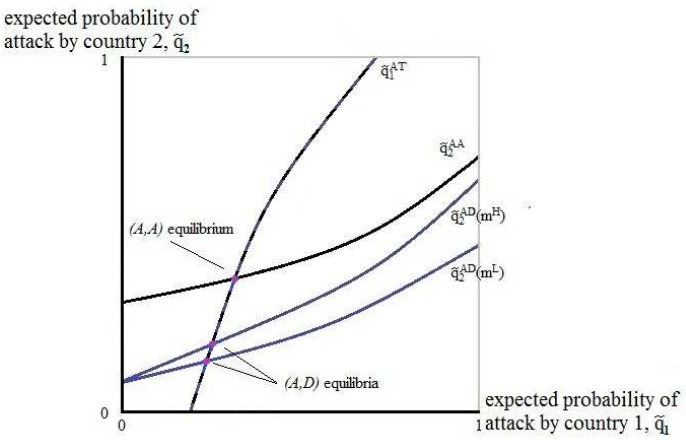


Figure 2.4: Equilibrium Beliefs: (A, A), (A, D) comparison

with the direct effect of reducing the expected probability of an attack; that is, holding beliefs constant, $\widetilde{q}_2^{AD} < \widetilde{q}_2^{AA}$ at $\widetilde{q}_1^{AA*}$.[24] However, beliefs are not constant, since leader 2 is able to credibly transmit information.

If the democratic leader signals that the country is not vulnerable, the posterior will be lower than the prior, $\widehat{\rho}_{i2}(m^L) < \rho_{i2}$, shifting the belief function in the same direction as the direct effect of democracy and further reducing the expected probability of an attack, ensuring that $\widetilde{q}_2^{AD}(m^L) < \widetilde{q}_2^{AA}$ at $\widetilde{q}_1^{AA*}$. If the leader signals instead that the country is vulnerable this effect will be reversed; $\widehat{\rho}_{i2}(m^H) > \rho_{i2}$, implying

[24] Technically, the best response belief function for a democracy may be 'above' the equivalent function for an autocracy, $\widetilde{q}_2^{AD} > \widetilde{q}_2^{AA}$, if the popular belief of the probability of vulnerability is sufficiently large, $\rho_{22} > \frac{1}{2\widetilde{q}_1}$. However, since $\widetilde{q}_1^{AA*} < \frac{1}{2}$ this cannot be the case at the (A, A) equilibrium value.

$\widetilde{q}_2^{AD}(m^H) > \widetilde{q}_2^{AD}(m^L)$ for all $\widetilde{q}_1$. The expected probability of an attack is higher when all believe the country is more likely to be vulnerable, but this effect is mitigated by the fact that there is not full information transmission. Comparing the maximum rise from the increase in posterior beliefs to the minimum fall from the direct effect of democracy shows that the direct effect dominates, and hence $\widetilde{q}_2^{AD}(m^H) < \widetilde{q}_2^{AA}$ at $\widetilde{q}_1^{AA*}$.

The unique intersection of $\widetilde{q}_2^{AD}(m_2)$ and $\widetilde{q}_1^{AD}$ must be lower than the equilibrium with two autocracies, as the belief functions are strictly increasing. Therefore, when a democracy opposes an autocracy the equilibrium expected probability of an attack is lower for both countries than if both were autocracies. Knowing that a democratic leader will have to consider his never-hawkish median voter, his opponent believes that he is less likely to attack than an autocrat, which in turn reduces the opponent's own expected probability of attacking.

The comparison with a pair of democracies is illustrated in Figure 2.5. The democratization of country 1 has a similar effect to the democratization of country 2, further lowering its best response belief function and thus the expected probability that it will attack, $\widetilde{q}_1^{DD*}(m_1) < \widetilde{q}_1^{AD*}$. Democratization of country 1 also effects the level of information transmission by country 2. Proposition 2 shows that the probability of a democracy signalling vulnerability is higher when country 1 has not democratized, and therefore posterior beliefs will reflect this different level of information transmission. When the posterior decreases the decline in the perceived likelihood of vulnerability further reinforces the democratization effect, ensuring that the expected probability of conflict is lower for both countries. Any increase in the posterior will be small, as it is a shift from one level of credible information without perfect revelation to another. Although the possibility of an increase in country 2's probability of

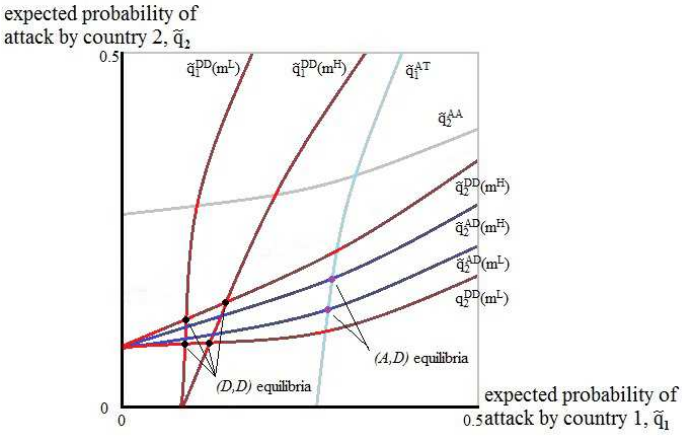


Figure 2.5: Equilibrium Beliefs: (A, D), (D, D) comparison

attacking cannot be ruled out, any such change will be so small that it is dominated by the decline in country 1's probability of attacking, ensuring, the expected probability of conflict is lowest for a pair of democracies, regardless of the information revealed, $\widetilde{w}^{DD}(m_1, m_2) < \widetilde{w}^{AD}(\cdot, m_2)$.[25,26]

[25] The relative levels of information transmission in (A, D) and (D, D) dyads, and therefore a comparison of the posteriors, has not been quantified. It is possible that either $\widehat{\rho}_{i2}^{AD}(m^H) < \widehat{\rho}_{i2}^{DD}(m^H)$ and $\widehat{\rho}_{i2}^{AD}(m^L) > \widehat{\rho}_{i2}^{DD}(m^L)$, or that the inequalities are reversed. The former is shown in Figure 5, but the result holds in both cases.

[26] If instead the stronger, but common, assumption that an autocrat is more likely to be a hawk than a democratic leader (for example, because of the manner in which many come to power, or because autocrats are better at isolating their personal wealth from the fortunes of war) was used this result would only be exacerbated. Technically, if the leader's private cost type is drawn from some distribution G, which is first order stochastically dominated by F, all results will still hold.

The probability of conflict is lowest for a pair of democracies and highest for a pair of autocracies, for all possible signal choices. It follows trivially that the underlying expected probability of conflict has the same ordering.

This reinforces the standard democratic peace result, and shows that the main effect of democratization is to temper the potential aggression of the leader. There will always be both hawks and doves in the population, but the hawks do not form a majority. Hawkish autocrats are able to ignore popular opinion, but a hawkish democrat is constrained by his desire for re-election and a median voter who will be, at best, a coordination type. Signalling enables the democratic leader to sway public opinion in his favor, but he is unable to change the underlying population distribution, and the regime effect dominates the signalling effect. This is supported by evidence that conflict is more often observed among autocracies, even when war is extremely costly, for example when GDP *per capita* is low and investment in war depletes a large share of essential domestic spending, as with conflicts in the Horn of Africa, a region riven by international conflicts as well as internal disputes.[27]

Likelihood of Conflict

The democratic peace hypothesis, the idea that democracy promotes peace, is well established. A comparison of the true probability of conflict for different pairs of regime types confirms this hypothesis and completes the comparison of the equilibria.

Conflict will take place if the leader of at least one country chooses to attack, denoted w when conditional on the signals sent and both countries' level of vulnerability:

$$w^{T_1T_2}(m_1, m_2; d_1, d_2) = q_1^{T_1T_2}(m_1, m_2; d_1) + (1 - q_1^{T_1T_2}(m_1, m_2; d_1))q_2^{T_1T_2}(m_1, m_2; d_2).$$

[27] See, for example, Moaz and Russett (1993).

Similarly to the expected probability of conflict, W denotes the underlying probability of conflict:

$$W^{T_1T_2} = \sum_{d_1}\sum_{d_2}\sum_{m_1}\sum_{m_2} w^{T_1T_2}(m_1, m_2; d_1, d_2)Pr(d_1)Pr(d_2)Pr(m_1|d_1, T_1T_2)Pr(m_2|d_2, T_1T_2).$$

The effect of regime type on the likelihood of conflict reflects its effect on beliefs, since the true probability of an attack is itself an increasing function of the expected probability of an attack by the opposing country. The regime effect dominates the information effect: replacing autocracy with democracy in either country lowers the equilibrium probability of conflict.

Proposition 4 ***(Democratic Peace)*** *The underlying probability of conflict is lowest for a pair of democracies and highest for a pair of autocracies:*

$$W^{AA} > W^{AD} > W^{DD}.$$

A proof of this proposition is provided in the Second Appendix, but the intuition is straightforward. A democratic leader, seeking re-election, must consider the opinions of the median voter, who is never a hawk. If he signals that the country is vulnerable, equilibrium beliefs about the probability of conflict will be higher, as shown previously. As would be anticipated, war is more likely when it is indicated that a country is more vulnerable, resulting from the fear of an unresisted attack by the opponent it creates. However, this is outweighed by the direct effect of democratization. Changing beliefs can only influence the probability that the population will support a coordination strategy: a dominant strategy hawk who attacks when the opponent negotiates will never have popular support, a concern for a democratic leader but ignored by an autocrat.

This does not rule out conflict between two democracies - if a democratic leader is sufficiently hawkish, he will always signal that his country is vulnerable, and proceed

to attack - but this only applies to leaders in the tail of the private cost type distribution, for whom the gains from conflict outweigh the potential loss from failure to remain in office. Democratic leaders who favor conflict but whose private cost type is higher will be willing to compromise their conflict strategy to increase their chances of re-election. This range of potential leaders would, if autocrats, choose to attack; thus the probability of conflict is higher whenever an autocracy is involved.

The presence of strategic complementarities in the conflict game leads to the ordering of the dyads. If the equilibrium expected probability of an attack is reduced for both countries when one autocracy is replaced by a democracy, this in turn reduces the true probability of an attack by both, and so the probability of conflict likewise. Even if the expected probability of an attack by country 2 increases slightly when country 1 democratizes, the ensuing effect on the true probability is small enough to be outweighed by the larger effect for country 1. This ensures that, as with the expected probability of conflict, the true probability also declines whenever one country democratizes.

These complementarities can create a situation of mutual hostility. If the opponent is more likely to attack this increases the probability of the home country attacking, which occurs due to the 'first mover' advantage and the need to defend against anticipated attacks. This can lead to situations of armed hostility and arms race scenarios - as between the United States, USSR and their respective allies during the Cold War; prior to the First World War, when Britain and Germany both invested vast sums in military resources in response to the other country's investment; and, more recently,

the arms race and nuclearization of the tensions between India and Pakistan - and, in the worst case, an attack primarily motivated by the fear of being attacked.[28]

2.7 Conclusion

Politics may be 'the art of the possible' but, as this chapter shows, domestic political concerns can both legitimize and restrain a country's leader. When two countries are in an inherently hostile situation, an autocratic leader cares only about the outcome of the conflict, whereas a democratic leader must also cater to his need for popular support if he is to remain in office. If a democratic leader's favored strategy depends on the country's vulnerability to attack, he has an incentive to reveal this information to the population to gain their support. Telling the truth more often than not, the leader gains credibility, providing him with an opportunity for manipulation which does not exist for an autocrat, whose statements will never be believed. However, since there is not full revelation a democratic leader can, in effect, get away with lying.

A leader involved in conflict must deal with both his opponent and his own country's citizens. This dual audience allows a democratic leader to transmit credible information, in contrast to an autocrat, who always wants to discourage his opponent (a single audience) from attacking and whose signal will therefore be ignored. However, it also provides an incentive to manipulate both parties, and the relative significance of the two audiences can affect the choice of signal. A democratic leader who faces an autocracy is more likely to falsely state that the country is not vulnerable, but if he faces another democracy he is more likely to misleadingly indicate that

[28] The attack driven by the fear of being attacked is exhibited in this simple model, a further dynamic element or repeated interaction should be introduced to fully demonstrate an arms race scenario - see Baliga and Sjöström (2004).

his country is vulnerable. A democratic opponent also has his own domestic political concerns and will thus be less influenced by the signal observed, making the home audience relatively more significant to the democratic leader.

These results reinforce the frequently stated 'democratic peace' hypothesis. In this model the median voter will never be a dominant strategy hawk, but the leader may be. A hawkish democrat is constrained by domestic political concerns, but a hawkish autocrat only considers his personal benefit from attacking, and so the expected probability of conflict is lower the more countries that are democracies, and likewise the true probability of conflict. Although information may be revealed which indicates that the country is vulnerable, and therefore increases public support for, and thus the probability of, an attack, the direct effect of democratization dominates this information effect. The lessening of structural uncertainty about the level of vulnerability is subsumed by the change in strategic uncertainly, the adaption of the best response belief function to incorporate the re-election concerns of a democratic leader.

Democracy, and the political manoeuvering that comes with it, provides a leader with credibility that an autocrat lacks, but this credibility also provides the elected leader with the ability to manipulate. As the population and opponent know that a democratic leader will tell the truth more often than not, he is able to pass on private information, but this also means that he will be believed if he provides false information. Democratization provides an opportunity for manipulation of beliefs, both at home and abroad, yet also puts limits on the ability to do so.

Appendix A

First Appendix

A.1 Proof of Lemma 1

The maximization problems when ethnic alliances have been formed give four first order conditions of the form:

$$\frac{w_j}{n_{ij}y_{ij} - w_j A_{ij}} = \frac{\delta A_{-j}}{(A_T + A_D)^2} E. \tag{A.1}$$

This system of equations gives the conflict investment of each subgroup:

$$\begin{aligned} A_{pT} &= \frac{n_T[(\delta n_T n_D E + 1)n_p - \alpha n_r]}{\gamma(2 + \delta n_T n_D E)}, & A_{pD} &= \frac{n_D[(\delta n_T n_D E + 1)n_p - \alpha n_r]}{\gamma(2 + \delta n_T n_D E)}, \\ A_{rT} &= \frac{n_T[\alpha(\delta n_T n_D E + 1)n_r - n_p]}{\gamma(2 + \delta n_T n_D E)}, & A_{rD} &= \frac{n_D[\alpha(\delta n_T n_D E + 1)n_r - n_p]}{\gamma(2 + \delta n_T n_D E)}. \end{aligned} \tag{A.2}$$

So a rich subgroup contributes more than its partner in an ethnic alliance iff:

$$\begin{aligned} A_{rj} &> A_{pj} \\ \alpha(\delta n_T n_D E + 1)n_r - n_p &> (\delta n_T n_D E + 1)n_p - \alpha n_r \\ \alpha n_r &> n_p \\ \alpha n_r y_{pj} &> n_p y_{pj} \\ Y_{rj} &> Y_{pj}. \end{aligned}$$

A similar procedure gives the result for investment in class alliances: an ethnic minority subgroup contributes more than its class alliance partner if $Y_{iD} > Y_{iT}$. *QED*

A.2 Proof of Proposition 1

Using from the closed form solution for subgroup conflict investment (A.2),

$$A_T = A_{pT}+A_{rT} = \frac{n_T[\delta n_T n_D E(n_p + \alpha n_r)]}{\gamma(2 + \delta n_T n_D E)}, \qquad A_D = A_{pD}+A_{rD} = \frac{n_D[\delta n_T n_D E(n_p + \alpha n_r)]}{\gamma(2 + \delta n_T n_D E)},$$

$$A = A_T + A_D = \frac{\delta n_T n_D E(n_p + \alpha n_r)}{\gamma(2 + \delta n_T n_D E)}. \tag{A.3}$$

Substituting these values into the conflict success function:

$$\sigma_T = \frac{A_T}{A} = \frac{(n_T[\delta n_T n_D E(n_p + \alpha n_r)])/(\gamma(2 + \delta n_T n_D E))}{(\delta n_T n_D E(n_p + \alpha n_r))/(\gamma(2 + \delta n_T n_D E))} = n_T.$$

Likewise, $\sigma_D = \frac{A_D}{A} = n_D$.

As contributions cannot be negative, if the above series of equations gives $A_{ij} < 0$ for some (i, j) there will be a corner solution, so it is necessary to impose $A_{ij} = 0$ when solving the maximization problem. By symmetry, $A_{iT} = 0 \Rightarrow A_{iD} = 0$ for $i = p, r$.

If $A_{pT} = A_{pD} = 0$, then

$$A_T = A_{rT} = \frac{n_T(\alpha n_r - n_p)}{\gamma}, \qquad A_D = A_{rD} = \frac{n_D(\alpha n_r - n_p)}{\gamma}, \qquad A = \frac{\alpha n_r - n_p}{\gamma}. \tag{A.4}$$

If $A_{rT} = A_{rD} = 0$, then

$$A_T = A_{pT} = \frac{n_T(n_p - \alpha n_r)}{\gamma}, \qquad A_D = A_{pD} = \frac{n_D(n_p - \alpha n_r)}{\gamma}, \qquad A = \frac{n_p - \alpha n_r}{\gamma}. \tag{A.5}$$

So $\sigma_j = \frac{A_j}{A} = n_j$ holds in both cases.

A symmetric procedure proves the result for conflict expenditure in class alliances, for both $\beta > 1$ and $\beta < 1$. *QED*

A.3 Proof of Lemma 2

Subgroup (i, j) prefers ethnic alliances to peace if

$$ln\left(y_{ij} - \frac{w_j A_{ij}}{n_{ij}}\right) + s_j^E E > ln(y_{ij}) + s_j^P E. \tag{A.6}$$

The class budget is not included here, as the share obtained will be the same under ethnic alliances and peace since there is no threat of class conflict to influence the decision of the subgroup in government. Under A1 this is approximately equivalent to

$$(s_j^E - s_j^P)E > A_{ij}\frac{w_j}{n_{ij}y_{ij} - w_j A_{ij}}.$$

Using the first order condition (A.1) in (A.6),

$$(s_j^E - s_j^P)E > A_{ij}\frac{A_{-j}}{(A_T + A_D)^2}\delta E.$$

Then substituting in the results from (A.3),

$$s_j^E - s_j^P > \lambda_{ij}\delta n_T n_D,$$

since $\lambda_{ij} = \frac{A_{ij}}{A_j}$.

An equivalent procedure demonstrates that subgroup (i, j) prefers class alliances to peace when $s_i^C - s_i^P > \mu_{ij}\delta n_p n_r$.

Similarly, subgroup (i, j) prefers ethnic alliances to class alliances if

$$ln\left(y_{ij} - \frac{w_j A_{ij}}{n_{ij}}\right) + s_i^E C + s_j^E E > ln\left(y_{ij} - \frac{w_i B_{ij}}{n_{ij}}\right) + s_i^C C + s_i^C E.$$

Under A1 this is approximately equivalent to

$$(s_i^E - s_i^C)C + (s_j^E - s_j^C)E > A_{ij}\frac{w_j}{n_{ij}y_{ij} - w_j A_{ij}} - B_{ij}\frac{w_i}{n_{ij}y_{ij} - w_i B_{ij}},$$

Then using (A.1) and the first order condition for the maximization problem when class alliances have formed, and substituting in (A.3) and the corresponding values for class expenditure,

$$(s_i^E - s_i^C)C + (s_j^E - s_j^C)E > \lambda_{ij}\delta n_T n_D E - \mu_{ij}\delta n_p n_r C.$$

QED

A.4 Proof of Proposition 3

Substituting the closed form solution for expenditure, (A.2) and (A.3), into $\lambda_{ij} = \frac{A_{ij}}{A_j}$,

$$\lambda_{pj} = \frac{(\delta n_T n_D E + 1)n_p - \alpha n_r}{\delta n_T n_D E(\alpha n_r + n_p)}, \qquad \lambda_{rj} = \frac{\alpha(\delta n_T n_D E + 1)n_r - n_p}{\delta n_T n_D E(\alpha n_r + n_p)}.$$

Thus,

$$\frac{\partial \lambda_{pj}}{\partial \alpha} < 0, \qquad \frac{\partial \lambda_{rj}}{\partial \alpha} > 0.$$

The relevant constraint is the preference over ethnic and class alliances of the excluded subgroup sharing neither characteristic with the group in government. This subgroup will receive no share of the class public goods budget if ethnic alliances form and no share of the ethnic public goods budget if class alliances form. When in a generic alliance k it will receive a share s_k of the contested public good: $s_k = \delta\sigma_k = \delta n_k$ by Proposition 1.

Therefore ethnic alliances form if:

$$n_{-J}E - n_{-I}C > \lambda_{-I-J}n_T n_D E - \mu_{-I-J}n_p n_r C. \tag{A.7}$$

λ_{-I-J} is a function of α and the only place this parameter appears in this expression. The right hand side is constant in α. λ_{-I-J}, and therefore the left hand side, is decreasing in α if $-I = p$ and conversely increasing if $-I = r$. Thus the preferences of the decisive subgroup move towards ethnic alliances as vertical inequality increases if a rich subgroup is in government, and towards class alliances if a poor subgroup is in government.

If a corner solution exists for some values of α, λ_{ij} is constant $\forall\, i, j$ throughout this range and therefore preferences are constant in α until such a point that $\lambda_{ij} \in (0, 1)$ and the above result holds.

A.5 Proof of Lemma 3

Using results obtained in the proof of Proposition 1:

- For $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, $A = \frac{\delta n_T n_D E(n_p + \alpha n_r)}{\gamma(2+\delta n_T n_D E)}$ (A.3). Thus, $\frac{\partial A}{\partial \alpha} = \frac{\delta n_T n_D n_r E}{\gamma(2+\delta n_T n_D E)}$, expenditure increasing in α.

- For $\alpha > \overline{\alpha}$, impose $A_{pj} = 0$, so $A = \frac{\alpha n_r - n_p}{\gamma}$ (A.4) and $\frac{\partial A}{\partial \alpha} = \frac{n_r}{\gamma}$, expenditure increasing in α.

- For $\alpha < \underline{\alpha}$, impose $A_{rj} = 0$, so $A = \frac{n_p - \alpha n_r}{\gamma}$ (A.5) and $\frac{\partial A}{\partial \alpha} = -\frac{n_r}{\gamma}$, expenditure decreasing in α.

Combining the results of these three segments gives an ethnic alliance expenditure function which is U-shaped in α. *QED*

A.6 Proof of Proposition 5

Recall that ethnic mobilization takes place if inequality (A.7) is satisfied, and class mobilization otherwise.

First consider variation in α:

- When a rich subgroup is in government,
  - Class alliances exist at $\underline{\alpha}$ ($\lambda_{pj} = 1$) if $n_{-J}E - n_{-I} < n_T n_D E - \mu_{-I-J} n_p n_r C$, which simplifies to $n_{-J}^2 E < (1 - \mu_{pD} n_I) n_{-I} C$.
  - Ethnic alliances exist at $\overline{\alpha}$ ($\lambda_{pj} = 0$) if $n_{-J}E - n_{-I}C > -\mu_{-I-J} n_p n_r C$, that is, $n_{-J}E > (1 - \mu_{-I-J} n_I) n_{-I} C$.

- When a poor subgroup is in government,
  - Ethnic alliances exist at $\underline{\alpha}$ if $n_{-J}E - n_{-I}C > -\mu_{-I-J} n_p n_r C$, that is, $n_{-J}E > (1 - \mu_{-I-J} n_I) n_{-I} C$.
  - Class alliances exist at $\overline{\alpha}$ if $n_{-J}E - n_{-I}C < n_T n_D E - \mu_{-I-J} n_p n_r C$, that is, $n_{-J}^2 E < (1 - \mu_{rD} n_p) n_r C$.

Combining these results shows that in both cases a switch point exists if:

$$n_{-J}E > (1 - \mu_{-I-J}n_I)n_{-I}C > n_{-J}^2E.$$

Now consider variation in β:

- When an ethnic majority subgroup is in government,
  - Class alliances exist at $\underline{\beta}$ ($\mu_{iD} = 0$) if $n_{-J}E - n_{-I}C < \lambda_{-I-J}n_Tn_DE$, that is, $(1 - \lambda_{rD}n_J)n_{-J}E < n_{-I}C$.
  - Ethnic alliances exist at $\overline{\beta}$ ($\mu_{iD} = 1$) if $n_{-J}E - n_{-I}C > \lambda_{-I-J}n_Tn_DE - n_pn_rC$, that is, $(1 - \lambda_{-I-J}n_J)n_{-J}E > n_{-I}^2C$.
- When an ethnic minority subgroup is in government,
  - Ethnic alliances exist at $\underline{\beta}$ if $n_{-I}E - n_{-J}C > \lambda_{-I-J}n_Tn_DE - n_pn_rC$, that is, $(1 - \lambda_{-I-J}n_J)n_{-J}E > n_{-I}^2C$.
  - Class alliances exist at $\overline{\beta}$ if $n_{-J}E - n_{-I}C < \lambda_{-I-J}n_Tn_DE$, that is, $(1 - \lambda_{-I-J}n_J)n_{-J}E < n_{-I}C$.

Combining these results shows that in both cases a switch point exists if:

$$n_{-I}C > (1 - \lambda_{-I-J}n_J)n_{-J}E > n_{-I}^2C.$$

A.7 Proof of Proposition 6

When ethnic-economic subgroup (I, J) is in government $(-I, -J)$ is the decisive excluded subgroup, and ethnic mobilization takes place if:

$$n_{-J}E - n_{-I}C > \lambda_{-I-J}n_T n_D E - \mu_{-I-J}n_p n_r C \tag{A.7}$$

Note that δ is only present on the right hand side (RHS) of this inequality as both λ_{ij} and μ_{ij} are functions of δ.

$$\begin{aligned}
\frac{\partial \lambda_{pj}}{\partial \delta} \geq 0 \text{ iff } \alpha > \frac{n_p}{n_r} &\Leftrightarrow Y_{rj} > Y_{pj}, \\
\frac{\partial \lambda_{rj}}{\partial \delta} \geq 0 \text{ iff } \alpha < \frac{n_p}{n_r} &\Leftrightarrow Y_{pj} > Y_{rj}, \\
\frac{\partial \mu_{iT}}{\partial \delta} \geq 0 \text{ iff } \beta > \frac{n_T}{n_D} &\Leftrightarrow Y_{iD} > Y_{iT}, \\
\frac{\partial \mu_{iD}}{\partial \delta} \geq 0 \text{ iff } \beta < \frac{n_T}{n_D} &\Leftrightarrow Y_{iT} > Y_{iD}.
\end{aligned}$$

The RHS of inequality (A.7) is decreasing if λ_{-I-J} is decreasing and μ_{-I-J} is increasing.

- λ_{-I-J} is decreasing if $-I = p$ and $Y_{pj} > Y_{rj}$, or $-I = r$ and $Y_{rj} > Y_{pj}$, i.e. $Y_{-Ij} > Y_{Ij} \Rightarrow Y_{-I} > Y_I$.

- μ_{-I-J} is increasing if $-J = T$ and $Y_{iD} > Y_{iT}$, or $-J = D$ and $Y_{iT} > Y_{iD}$, i.e. $Y_{i-J} < Y_{iJ} \Rightarrow Y_{-J} < Y_J$.

If the RHS is decreasing, the preferences of decisive subgroup are moving towards ethnic alliances as δ increases: ethnic mobilization at $\delta' \Rightarrow$ ethnic mobilization at all $\delta > \delta'$.

The RHS of inequality (A.7) is increasing if λ_{-I-J} is increasing and μ_{-I-J} is decreasing.

- λ_{-I-J} is increasing if $-I = p$ and $Y_{rj} > Y_{pj}$, or $-I = r$ and $Y_{pj} > Y_{rj}$, i.e. $Y_{-Ij} < Y_{Ij} \Rightarrow Y_{-I} < Y_I$.

- μ_{-I-J} is decreasing if $-J = T$ and $Y_{iT} > Y_{iD}$, or $-J = D$ and $Y_{iD} > Y_{iT}$, i.e. $Y_{i-J} > Y_{iJ} \Rightarrow Y_{-J} > Y_J$.

If the RHS is increasing, the preferences of decisive subgroup are moving towards class alliances as δ increases: ethnic mobilization at $\delta' \Rightarrow$ ethnic mobilization at all $\delta < \delta'$.

The effect of δ on RHS is indeterminate if λ_{-I-J} and μ_{-I-J} are both increasing or both decreasing, i.e. in all other cases results depend on combination of other parameters.

This result holds weakly since for parameter combinations where one subgroup bears the entire cost of an alliance contributions are not a function of δ.

Appendix B

Second Appendix

B.1 Support Structure in Democracies

A democratic leader will remain in power (and receive reward R) if he has the support of at least half the population.

If the opponent negotiates the payoff does not depend on the realization of d_i. All non-hawkish citizens will support negotiation, forming a majority and ensuring that the leader will remain in power if he also chooses to negotiate.

If the opponent attacks, hawks and pure co-ordination types will always support a leader who chooses to attack, and doves will always support a leader who chooses to negotiate. Neither of these groups forms a majority, so the probability of remaining in power depends on the support of those with cost type $d^L < c < d^H$.

A citizen in this cost range will support a leader who attacks when his opponent attacks if he believes that the probability that $d_i = d^H$, ρ_{ii}, is sufficiently high. Formally, suppose that he will vote in support of the leader if ρ_{ii} is greater than some threshold, $\rho_{ii} > \delta$.

δ is drawn from a uniform distribution, $\delta \sim U[0,1]$, where the realization is observed by the citizens but not the leader. δ should be interpreted as a parameter measuring the average (relative) popularity of the leader in the population as a whole. It can be considered to incorporate the performance of the leader on other

policy dimensions and the capabilities of an alternative potential leader, as well as personality issues.

If both countries attack, it follows that all citizens with cost type $d^L < c < d^H$ will vote for the existing leader if $\rho_{ii} > \delta$. The leader will remain in office if $\rho_{ii} > \delta$ and so the expected probability of re-election, and therefore of obtaining reward R, is $Pr(\rho_{ii} > \delta) = \rho_{ii}$.

Similarly the expected probability of a democratic leader remaining in office if he chooses to negotiate when the opponent attacks is $(1 - \rho_{ii})$.

B.2 Proof of Proposition 1

Autocracy

Country 1 is an autocracy. No assumption is made about the regime type of country 2.

Leader 1 has expected payoff $(1-\widetilde{q}_2^{AT*}(m_1, m_2))\mu - c_1^l$ from attacking and expected payoff $-\widetilde{q}_2^{AT*}(m_1, m_2)d_1$ from negotiating.

Comparing expected payoffs, having sent signal m_1, leader 1 will prefer to attack if

$$\mu + \widetilde{q}_2^{AT*}(m_1, m_2)(d_1 - \mu) > c_1^l.$$

The proof is by contradiction. First assume that in equilibrium $\widetilde{q}_2^{AT*}(m^H, m_2) > \widetilde{q}_2^{AT*}(m^L, m_2)$.

Regardless of the signal sent, leader 1 will attack if $\mu + \widetilde{q}_2^{AT*}(m^L, m_2)(d_1 - \mu) > c_1^l$. Comparing payoffs, leader 1 will prefer signal $m_1 = m^L$ for all cost types in this region.

Regardless of the signal sent, leader 1 will negotiate if $c_1^l > \mu + \widetilde{q}_2^{AT*}(m^H, m_2)(d_1 - \mu)$. Comparing payoffs, leader 1 will prefer signal $m_1 = m^L$ for all cost types in this region.

For cost types in the range

$$\mu + \widetilde{q}_2^{AT*}(m^H, m_2)(d_1 - \mu) > c_1^l > \mu + \widetilde{q}_2^{AT*}(m^L, m_2)(d_1 - \mu),$$

leader 1 will attack if he has sent signal $m_1 = m^H$ and negotiate if he has sent signal $m_1 = m^L$. Comparing payoffs, leader 1 will prefer signal $m_1 = m^L$ for all cost types in this region.

Therefore, for both $d_1 = d^H$ and $d_1 = d^L$ leader 1 will send signal $m_1 = m^L$ for all private cost types, c_1^l. Then $Pr(m_1 = m^L | d_1 = d^H) = Pr(m_1 = m^L | d_1 = d^L) = 1$ which implies that there is no information transfer in equilibrium and so $\widehat{\rho}_{21}(m^L) = \rho_{21}$. Since the opponent does not expect to observe the signal $m_1 = m^H$ on the equilibrium path, if he does he must assume that it is a mistake and so $\widehat{\rho}_{21}(m^H) = \rho_{21}$. However, with no information transfer $\widetilde{q}_2^{AT}(m^H) = \widetilde{q}_2^{AT}(m^L)$ for all $\widetilde{q}_1$, so it must be that $\widetilde{q}_2^{AT*}(m^H, m_2) = \widetilde{q}_2^{AT*}(m^L, m_2)$, contradicting the initial assumption.

Now assume instead that $\widetilde{q}_2^{AT*}(m^L, m_2) > \widetilde{q}_2^{AT*}(m^H, m_2)$. A similar procedure shows that for both $d_1 = d^H$ and $d_1 = d^L$ leader 1 will send signal $m_1 = m^H$ for all private cost types. Then $Pr(m_1 = m^H | d_1 = d^H) = Pr(m_1 = m^H | d_1 = d^L) = 1$, so $\widehat{\rho}_{21}(m^H) = \rho_{21}$ and there is no equilibrium transfer in equilibrium. Therefore it must be that $\widetilde{q}_2^{AT*}(m^H, m_2) = \widetilde{q}_2^{At*}(m^L, m_2)$, contradicting the initial assumption.

Finally assume that in equilibrium, $\widetilde{q}_2^{AT*}(m^H, m_2) = \widetilde{q}_2^{AT*}(m^L, m_2) \equiv \widetilde{q}_2^{AT*}$.

Regardless of the signal sent, leader 1 will attack if $\mu + \widetilde{q}_2^{AT*}(d_1 - \mu) > c_1^l$. Comparing payoffs, leader 1 will be indifferent between signals for all cost types in this region.

Regardless of the signal sent, leader 1 will negotiate if $c_1^l > \mu + \widetilde{q}_2^{AT*}(d_1 - \mu)$. Comparing payoffs, leader 1 will be indifferent between signals for all cost types in this region.

For both $d_1 = d^H$ and $d_1 = d^L$, leader 1 will be indifferent between signals for all cost types. He will therefore randomize between signals, so that $Pr(m_1 = m^H | d_1 = d^H) = Pr(m_1 = m^H | d_1 = d^L) = 0.5$, $\widehat{\rho}_{21}(m^H) = \widehat{\rho}_{21}(m^L) = \rho_{21}$, and no information is transmitted in equilibrium. This implies that $\widetilde{q}_2^{AT*}(m^H, m_2) = \widetilde{q}_2^{AT*}(m^L, m_2)$, consistent with the initial assumption.

Therefore when country 1 is an autocracy the only possible equilibrium is one where the leader randomizes between signals and no information is transmitted in equilibrium.

Democracy

Country 2 is a democracy. Leader 2 has expected payoff

$$\mu - c_2^l + \widetilde{q}_1^{T_1 D*}(m_1, m_2)(\widehat{\rho}_{22}^{T_1 D}(m_2)R - \mu)$$

from attacking and expected payoff

$$R - \widetilde{q}_1^{T_1 D*}(m_1, m_2)(d_2 + \widehat{\rho}_{22}^{T_1 D}(m_2)R)$$

from negotiating.

First consider an (A, D) dyad. Comparing expected payoffs, having sent signal m_2, leader 2 will prefer to attack if

$$\mu + \widetilde{q}_1^{AD*}(m_1, m_2)(d_2 - \mu + 2\widehat{\rho}_{22}^{AD}(m_2)R) - R > c_2^l.$$

The proof is by construction. Assume that in equilibrium $Pr(m_2 = m^H | d_2 = d^H) > Pr(m_2 = m^H | d_2 = d^L)$. This implies that $\widehat{\rho}_{i2}^{AD}(m^H) > \rho_{i2} > \widehat{\rho}_{i2}^{AD}(m^L)$ and so $\widetilde{q}_2^{AD}(m^H) > \widetilde{q}_2^{AD}(m^L)$ for all $\widetilde{q}_1$. Thus the equilibrium expected attack probability for both countries is higher after leader 2 has sent signal m^H, $\widetilde{q}_i^{AD*}(\cdot, m^H) > \widetilde{q}_i^{AD*}(\cdot, m^L)$ (independent of the uninformative signal sent by the opposing autocracy).

No credible information is transmitted by country 1 (by Proposition 1), so leader 2 has no uncertainty about the action he will take in the conflict game when choosing which signal to send.

Leader 2 will attack, regardless of the signal he sent, if

$$\mu + \widetilde{q}_1^{AD*}(\cdot, m^L)(d_2 - \mu + 2\widehat{\rho}_{22}^{AD}(m^L)R) - R > c_2^l.$$

Using backwards induction and comparing payoffs, leader 2 will choose signal $m_2 = m^H$ for all cost types in this range.

Leader 2 will negotiate, regardless of the signal he sent, if

$$c_2^l > \mu + \widetilde{q}_1^{AD*}(\cdot, m^H)(d_2 - \mu + 2\widehat{\rho}_{22}^{AD}(m^H)R) - R.$$

Comparing payoffs, leader 2 will choose signal $m_2 = m^L$ for all cost types in this range.

For cost types in the range

$$\mu + \widetilde{q}_1^{AD*}(\cdot, m^H)(d_2 - \mu + 2\widehat{\rho}_{22}^{AD}(m^H)R) - R > c_2^l > \mu + \widetilde{q}_1^{AD*}(\cdot, m^L)(d_2 - \mu + 2\widehat{\rho}_{22}^{AD}(m^L)R) - R,$$

leader 1 will attack if he has sent signal $m_2 = m^H$ and negotiate if he has sent signal $m_2 = m^L$. Comparing payoffs, leader 2 will choose signal $m_2 = m^H$ if he has private cost type

$$c_2^l < \mu + \widetilde{q}_1^{AD*}(\cdot, m^L)(d_2 + \widehat{\rho}_{22}^{AD}(m^L)R) + \widetilde{q}_1^{AD*}(\cdot, m^H)(\widehat{\rho}_{22}^{AD}(m^H)R - \mu) - R \equiv \widetilde{c}_2^{AD}(d_2).$$

Therefore, considering the complete distribution of private cost types, leader 2 will send signal $m_2 = m^H$ if $c_2^l < \widetilde{c}_2^{AD}(d_2)$ and signal $m_2 = m^L$ otherwise. $\widetilde{c}_2^{AD}(d^H) > \widetilde{c}_2^{AD}(d^L)$ so $F[\widetilde{c}_2^{AD}(d^H)] = Pr(m_2 = m^H|d_2 = d^H) > Pr(m_2 = m^H|d_2 = d^L) = F[\widetilde{c}_2^{AD}(d^L)]$, supporting the original assumption.

Since $F[\widetilde{c}_2^{AD}(d^H)] \neq 1$, $F[\widetilde{c}_2^{AD}(d^L)] \neq 0$ there is not complete revelation, but as $F[\widetilde{c}_2^{AD}(d^H)] \neq F[\widetilde{c}_2^{AD}(d^L)]$ credible information is transmitted and $\widehat{\rho}_{i2}^{AD}(m^H) > \rho_{i2} > \widehat{\rho}_{i2}^{AD}(m^L)$.

Now consider instead a (D, D) dyad. The proof takes a similar form, but must account for the possibility that credible information will be transmitted by country 1. When choosing the signal in the communication game leader 2 will be uncertain about his choice of action in the conflict game.

Before any signals are observed, leader 1 expects that, if he sends signal m_2, he will prefer to attack if

$$\mu + Q_1^{DD}(m_2)(d_2 - \mu + 2\widehat{\rho}_{22}^{DD}(m_2)R) - R > c_2^l$$

where

$$\begin{aligned} Q_1^{DD}(m_2) = (p_1 F[\widetilde{c}_1^{DD}(d^H)] \quad &+ \quad (1-p_1)F[\widetilde{c}_1^{DD}(d^L)])\widetilde{q}_1^{DD*}(m^H, m_2) \\ &+ \quad (p_1(1 - F[\widetilde{c}_1^{DD}(d^H)]) - (1-p_1)(1 - F[\widetilde{c}_1^{DD}(d^L)]))\widetilde{q}_1^{DD*}(m^L, m_2) \end{aligned}$$

(i.e. the expected probability that country 1 attacks, given country 2's signal, weighted by the probability of country 1's choice of signal.)

Again assume that in equilibrium, $Pr(m_2 = m^H|d_2 = d^H) > Pr(m_2 = m^H|d_2 = d^L)$, implying that $\widehat{\rho}_{i2}^{DD}(m^H) > \rho_{i2} > \widehat{\rho}_{i2}(m^L)$, and so $Q_1^{DD}(m^H) > Q_1^{DD}(m^L)$.

The same method of comparing expected payoffs over the range of private cost types shows that leader 2 will choose to send signal $m_2 = m^H$ for cost types

$$c_2^l < \mu + Q_1^{DD}(m^L)(d_2 + \widehat{\rho}_{22}^{DD}(m^L)R) + Q_1^{DD}(m^H)(\widehat{\rho}_{22}^{DD}(m^H)R - \mu) - R \equiv \widetilde{c}_2^{DD}(d_2)$$

It follows from $\widetilde{c}_2^{DD}(d^H) > \widetilde{c}_2^{DD}(d^L)$ that $F[\widetilde{c}_2^{DD}(d^H)] = Pr(m_2 = m^H | d_2 = d^H) > Pr(m_2 = m^H | d_2 = d^L) = F[\widetilde{c}_2^{DD}(d^L)]$, supporting the original assumption. As when country 1 is an autocracy, there is not complete revelation, but credible information is transmitted and the posteriors will be updated.

$\widetilde{c}_2^{AD}(d_2) \neq \widetilde{c}_2^{DD}(d_2)$, so the probability of a democratic leader sending signal m^H is influenced by the regime type of his opponent. The result is shown by contradiction.

Assume otherwise, that $\widehat{\rho}_{i2}^{AD}(m^H) = \widehat{\rho}_{i2}^{DD}(m^H)$ and $\widehat{\rho}_{i2}^{AD}(m^L) = \widehat{\rho}_{i2}^{DD}(m^L)$. Then $\widetilde{q}_2^{AD}(m^H) = \widetilde{q}_2^{DD}(m^H)$ and $\widetilde{q}_2^{AD}(m^L) = \widetilde{q}_2^{DD}(m^L)$ for all $\widetilde{q}_1$. However, when country 1 is also a democracy, it too transmits credible information, so $\widetilde{q}_1^{AD} \neq \widetilde{q}_1^{DD}(m^H) \neq \widetilde{q}_1^{DD}(m^L)$. This ensures that $\widetilde{q}_1^{AD*}(m_1, m_2) \neq \widetilde{q}_1^{DD*}(m_1, m_2)$ for all signal combinations, so $\widetilde{q}_1^{AD*}(m_1, m_2) \neq Q_1(m_2)$, $\widetilde{c}_2^{AD}(d_2) \neq \widetilde{c}_2^{DD}(d_2)$ and $\widehat{\rho}_{i2}^{AD}(m_2) \neq \widehat{\rho}_{i2}^{DD}(m_2)$, a contradiction.

Note that regardless of the regime type of country 1, there will also be a 'babbling' equilibrium (the democratic leader randomizes between signals and no credible information is transmitted) and a 'flipped' equilibrium' (the democratic leader sends signal m^L for all private cost types where he would send signal m^H in the equilibrium described above, and *vice versa* - the same level of information transmission is maintained). The proof is omitted as it follows the same method as the equilibrium above, and the equilibrium described is the one of interest for the results.

B.3 Comparison of Equilibria

The proof of Proposition 2 draws on Propositions 3 and 4, evaluating the effect of regime type of the equilibrium (expected) probability of conflict, discussed in section 5. These proofs are therefore presented before the proof of Proposition 2.

Proof of Proposition 3 (Equilibrium Expected Probability of Conflict)

(A, A) Dyad

Equilibrium beliefs are defined by the unique solution to the system of two equations:

$$\widetilde{q}_1^{AA} = \rho_{21}F[\mu + \widetilde{q}_2(d^H - \mu)] + (1 - \rho_{21})F[\mu + \widetilde{q}_2(d^L - \mu)],$$

$$\widetilde{q}_2^{AA} = \rho_{12}F[\mu + \widetilde{q}_1(d^H - \mu)] + (1 - \rho_{12})F[\mu + \widetilde{q}_1(d^L - \mu)].$$

As signals are uninformative, there is a single equilibrium: $\widetilde{q}_1^{AA*}$, $\widetilde{q}_2^{AA*}$.

(A, D) Dyad

Equilibrium beliefs are defined by the unique solution to system of two equations, where $\widetilde{q}_1^{AD} = \widetilde{q}_1^{AA}$ as defined above, and

$$\widetilde{q}_2^{AD}(m_2) = \widehat{\rho}_{12}^{AD}(m_2)F[\mu+\widetilde{q}_1(d^H-\mu+2\widehat{\rho}_{22}^{AD}(m_2)R)-R]+(1-\widehat{\rho}_{12}^{AD}(m_2))F[\mu+\widetilde{q}_1(d^L-\mu+2\widehat{\rho}_{22}^{AD}(m_2)R)-R].$$

There are two possible equilibria, depending on country 2's choice of signal.

Evaluate $\widetilde{q}_2^{AD}(m_2)$ at $\widetilde{q}_1^{AA*}$. When $m_2 = m^L$, the result is trivial: $(2\widehat{\rho}_{22}^{AD}(m^L)\widetilde{q}_1^{AA*} - 1)R < 0$ and $\widehat{\rho}_{12}^{AD}(m^L) < \rho_{12}$, so $\widetilde{q}_2^{AD}(m^L) < \widetilde{q}_2^{AA}$ at $\widetilde{q}_1^{AA*}$ and thus by convexity $\widetilde{q}_1^{AD*}(\cdot, m^L) < \widetilde{q}_1^{AA*}$ and $\widetilde{q}_2^{AD*}(\cdot, m^L) < \widetilde{q}_2^{AA*}$.

When $m_2 = m^H$, $(2\widehat{\rho}_{22}^{AD}(m^H)\widetilde{q}_1^{AA*} - 1)R < 0$ but $\widehat{\rho}_{12}^{AD}(m^H) > \rho_{12}$, so the democratization and information effects are working in opposite directions. When $R = 0$ there is no transmission of information, so $\widetilde{q}_2^{AD}(m^H) = \widetilde{q}_2^{AA}$. Therefore it suffices to show that $\widetilde{q}_2^{AD}(m^H)$, evaluated at $\widetilde{q}_1^{AA*}(m^H)$, decreases as R increases.

$$\begin{aligned}
\frac{d\widetilde{q}_2^{AD}}{dR} &= \frac{\partial\widehat{\rho}_{12}^{AD}(m^H)}{\partial R}\Big(F[\mu + \widetilde{q}_1(d^H - \mu + 2\widehat{\rho}_{22}^{AD}(m^H)R) - R] - F[\mu + \widetilde{q}_1(d^L - \mu + 2\widehat{\rho}_{22}^{AD}(m^L)R) - R]\Big) \\
&+ \Big(\widehat{\rho}_{12}^{AD}(m^H)f[\mu + \widetilde{q}_1(d^H - \mu + 2\widehat{\rho}_{22}^{AD}(m^H)R) - R] + (1 - \widehat{\rho}_{12}^{AD}(m^H))f[\mu + \widetilde{q}_1(d^L - \mu + 2\widehat{\rho}_{22}^{AD}(m^H)R) - R]\Big) \\
&\cdot \left(2\widehat{\rho}_{22}^{AD}(m^H)\widetilde{q}_1 - 1 + \frac{\partial\widehat{\rho}_{22}^{AD}(m^H)}{\partial R}2R\widetilde{q}_1\right)
\end{aligned}$$

The first term is positive and the second term is negative at $\widetilde{q}_1^{AA*}$. Expanding $\frac{\partial\widehat{\rho}_{12}^{AD}(m^H)}{\partial R}$ and $\widehat{\rho}_{12}^{AD}(m^H)$ allows for a comparison of the magnitude of the two terms, and shows that the democratization effect dominates, $\frac{\partial\widetilde{q}_2^{AD}(m^H)}{\partial R} < 0$ at $\widetilde{q}_1^{AA*}$.

Therefore $\widetilde{q}_2^{AD}(m^H) < \widetilde{q}_2^{AA}$ at $\widetilde{q}_1^{AA*}\ \forall\ R$, and so $\widetilde{q}_1^{AD*}(\cdot, m^H) < \widetilde{q}_1^{AA*}$ and $\widetilde{q}_2^{AD*}(\cdot, m^H) < \widetilde{q}_2^{AA*}$.

As $\widetilde{q}_i^{AD}(\cdot, m^L) < \widetilde{q}_i^{AD}(\cdot, m^H) < \widetilde{q}_i^{AA}$ for $i = 1, 2$, it follows trivially that $\widetilde{w}^{AD}(\cdot, m^L) < \widetilde{w}^{AD}(\cdot, m^H) < \widetilde{w}^{AA}$, and so $\widetilde{W}^{AD} < \widetilde{W}^{AA}$.

(D, D) Dyad

Equilibrium beliefs are defined by the unique solution to the system of two equations:

$$\begin{aligned}
\widetilde{q}_1^{DD}(m_1) &= \widehat{\rho}_{21}^{DD}(m_1)F[\mu + \widetilde{q}_2(d^H - \mu + 2\widehat{\rho}_{11}^{DD}(m_1)R) - R] + (1 - \widehat{\rho}_{21}^{DD}(m_1))F[\mu + \widetilde{q}_2(d^L - \mu + 2\widehat{\rho}_{11}^{DD}(m_1)R - R], \\
\widetilde{q}_2^{DD}(m_2) &= \widehat{\rho}_{12}^{DD}(m_2)F[\mu + \widetilde{q}_1(d^H - \mu + 2\widehat{\rho}_{22}^{DD}(m_2)R) - R] + (1 - \widehat{\rho}_{12}^{DD}(m_2))F[\mu + \widetilde{q}_1(d^L - \mu + 2\widehat{\rho}_{22}^{DD}(m_2)R - R].
\end{aligned}$$

There are four potential equilibria, depending on both countries' choice of signal.

The same argument as for the comparison of the (A, D) dyad shows that $\widetilde{q}_1^{DD}(m^L) < \widetilde{q}_1^{DD}(m^H) < \widetilde{q}_1^{AD}$ at $\widetilde{q}_2^{AD*}(\cdot, m^L)$ and $\widetilde{q}_2^{AD*}(\cdot, m^H)$. However, the analysis

is further complicated by the fact that the democratization of country 1 also has an information effect on country 2: either $\widehat{\rho}_{i2}^{AD}(m^H) > \widehat{\rho}_{i2}^{DD}(m^H)$, $\widehat{\rho}_{i2}^{AD}(m^L) < \widehat{\rho}_{i2}^{DD}(m^L)$ or $\widehat{\rho}_{i2}^{AD}(m^H) < \widehat{\rho}_{i2}^{DD}(m^H)$, $\widehat{\rho}_{i2}^{AD}(m^L) > \widehat{\rho}_{i2}^{DD}(m^L)$.

If $\widehat{\rho}_{i2}^{DD}(m_2) < \widehat{\rho}_{i2}^{AD}(m_2)$ both the effect on country 1 and the information effect on country 2 are working in the same direction, so it is clear that $\widetilde{q}_1^{DD*}(m^L, m_2) < \widetilde{q}_1^{DD*}(m^H, m_2) < \widetilde{q}_1^{AD*}(\cdot, m_2)$ and $\widetilde{q}_2^{DD*}(m^L, m_2) < \widetilde{q}_2^{DD*}(m^H, m_2) < \widetilde{q}_2^{AD*}(\cdot, m_2)$, but if $\widehat{\rho}_{i2}^{DD}(m_2) > \widehat{\rho}_{i2}^{AD}(m_2)$ it is necessary to evaluate the comparative statics of the equilibrium values with respect to R_1, while R_2 is held constant, i.e. how the equilibrium expected probabilities of attacking vary together as country 1 democratizes, while the regime type of country 2 does not change.

$$\frac{d\widetilde{q}_1^{DD*}(m_1, m_2)}{dR_1} = \frac{1}{1 - A_1 A_2}(\Delta_1 + A_1\Delta_2) \qquad , \qquad \frac{d\widetilde{q}_2^{DD*}(m_1, m_2)}{dR_1} = \frac{1}{1 - A_1 A_2}(\Delta_2 + A_2\Delta_1)$$

where

$$\begin{aligned} A_i \;&\equiv\; \widehat{\rho}_{-ii}(m_i) f[\mu + \widetilde{q}_{-i}^{*}(m_1, m_2)(d^H - \mu + 2\widehat{\rho}_{ii}(m_i)R_i) - R_i](d^H - \mu + 2\widehat{\rho}_{ii}(m_i)R_i) \\ &+\; (1 - \widehat{\rho}_{-ii}(m_i) f[\mu + \widetilde{q}_{-i}^{*}(m_1, m_2)(d^L - \mu + 2\widehat{\rho}_{ii}(m_i)R_i) - R_i](d^L - \mu + 2\widehat{\rho}_{ii}(m_i)R_i) < 1 \end{aligned}$$

$$\Delta_i = \left.\frac{d\widetilde{q}_i^{DD}(m_i)}{dR_1}\right|_{\widetilde{q}_{-i}^{*}}$$

$\Delta_1 < 0$ and if $\frac{\partial \widehat{\rho}_{i2}(m_2)}{\partial R_1} > 0$ then $\Delta_2 > 0$, which is the non-trivial case of interest.

Assuming symmetry between the two countries, $|\Delta_1| > |\Delta_2|$ so $\frac{\partial \widetilde{q}_1^{*}}{\partial R_1} < 0$: the expected probability that country 1 attacks always decreases when it democratizes.

The expected probability of an attack by country 2 may either increase or decrease, depending on whether the higher belief of the probability of vulnerability or the

lower probability of an attack by the opponent dominates. If $\widetilde{q}_2^{DD*}(m_1, m_2)$ decreases, clearly the expected probability of conflict, $\widetilde{w}^{DD}(m_1, m_2) = \widetilde{q}_1^{DD*}(m_1, m_2) + (1 - \widetilde{q}_1^{DD*}(m_1, m_2))\widetilde{q}_2^{DD*}(m_1, m_2)$, also declines.

If $\widetilde{q}_2^*$ increases the magnitude of this increase is smaller than the decrease in the probability of an attack by country 1:

$$-(A_1 + 1)\Delta_2 > (A_2 + 1)\Delta_1 \qquad \Rightarrow \qquad \left|\frac{d\widetilde{q}_1^{DD*}(m_1, m_2)}{dR_1}\right| > \left|\frac{d\widetilde{q}_2^{DD*}(m_1, m_2)}{dR_1}\right|$$

Thus the change in the expected probability of conflict decreases:

$$\frac{d\widetilde{w}^{DD}(m_1, m_2)}{dR_1} = \frac{d\widetilde{q}_1^{DD*}(m_1, m_2)}{dR_1}(1-\widetilde{q}_2^{DD*}(m_1, m_2))+\frac{d\widetilde{q}_2^{DD*}(m_1, m_2)}{dR_1}(1-\widetilde{q}_1^{DD*}(m_1, m_2)) < 0.$$

Therefore, regardless of the realized levels of vulnerability and the choice of signal by country 2, the expected probability of conflict will be lower when country 1 is also a democracy, $\widetilde{w}^{DD}(m^L, m_2) < \widetilde{w}^{DD}(m^H, m_2) < \widetilde{w}^{AD}(\cdot, m_2)$ for $m_2 = m^L, m^H$, and thus the underlying expected probability of conflict likewise: $\widetilde{W}^{DD} < \widetilde{W}^{AD}$.

Combining these results shows that the expected probability of conflict is highest for a pair of autocracies and lowest for a pair of democracies:

$$\widetilde{W}^{AA} > \widetilde{W}^{AD} > \widetilde{W}^{DD}.$$

QED

PROOF OF PROPOSITION 4 (EQUILIBRIUM PROBABILITY OF CONFLICT)

(A, D) **Dyad**

$$\begin{aligned} q_1^{AA}(\cdot,\cdot;d_1) &= F[\mu + \widetilde{q}_2^{AA*}(d_1 - \mu)] \\ q_2^{AA}(\cdot,\cdot;d_2) &= F[\mu + \widetilde{q}_1^{AA*}(d_2 - \mu)] \end{aligned}$$

$$\begin{aligned} q_1^{AD}(\cdot,m_2;d_1) &= F[\mu + \widetilde{q}_2^{AD*}(\cdot,m_2)(d_1 - \mu)] \\ q_2^{AD}(\cdot,m_2;d_2) &= F[\mu + \widetilde{q}_1^{AD*}(\cdot,m_2)(d_2 - \mu + 2\widehat{\rho}_{22}^{AD}(m_2)R) - R] \end{aligned}$$

As $0 > (2\widetilde{q}_1^{AD*}(\cdot,m^H)\widehat{\rho}_{22}^{AD}(m^H) - 1)R > (2\widetilde{q}_1^{AD*}(\cdot,m^L)\widehat{\rho}_{22}^{AD}(m^L) - 1)R$ and (by Proposition 3) $\widetilde{q}_j^{AA*} > \widetilde{q}_j^{AD*}(\cdot,m^H) > \widetilde{q}_j^{AD*}(\cdot,m^L)$,

$$q_i^{AA}(\cdot,\cdot;d_1) > q_i^{AD}(\cdot,m^H;d_1) > q_i^{AD}(\cdot,m^L;d_1) \qquad \forall\, d_i.$$

It follows trivially that

$$w^{AA}(\cdot,\cdot;d_1,d_2) > w^{AD}(\cdot,m^H;d_1,d_2) > w^{AD}(\cdot,m^L;d_1,d_2) \qquad \forall\, d_1,d_2 \qquad \Rightarrow \qquad W^{AA} > W^{AD}.$$

(D, D) **Dyad**

$$\begin{aligned} q_1^{DD}(m_1,m_2;d_1) &= F[\mu + \widetilde{q}_2^{DD*}(m_1,m_2)(d_1 - \mu + 2\widehat{\rho}_{11}^{DD}(m_1)R) - R] \\ q_2^{DD}(m_1,m_2;d_2) &= F[\mu + \widetilde{q}_1^{DD*}(m_1,m_2)(d_2 - \mu + 2\widehat{\rho}_{22}^{DD}(m_2)R) - R] \end{aligned}$$

If $R_1 = 0$, no information is transmitted by country 1 and thus $q_i^{DD}(m_1,m_2;d_i) = q_i^{AD}(\cdot,m_2;d_i)$. To evaluate the effect of the democratization of country 1, it is again sufficient to consider the comparative statics with respect to R_1.

$$\begin{aligned}
\frac{q_1^{DD}(m_1,m_2;d_1)}{dR_1} &= f[\mu+\widetilde{q}_2^{DD*}(m_1,m_2)(d_1-\mu+2\widehat{\rho}_{11}^{DD}(m_1)R_1)-R_1] \\
&\cdot \left(\frac{d\widetilde{q}_2^{DD*}(m_1,m_2)}{dR_1}(d_1-\mu+2\widehat{\rho}_{11}^{DD}(m_1)R_1)+\frac{\partial\widehat{\rho}_{11}^{DD}(m_1)}{\partial R_1}2\widetilde{q}_2^{DD*}(m_1,m_2)R_1+2\widetilde{q}_2^{DD*}(m_1,m_2)\widehat{\rho}_{11}^{DD}(m_1)-1\right) \\
\frac{q_2^{DD}(m_1,m_2;d_2)}{dR_1} &= f[\mu+\widetilde{q}_1^{DD*}(m_1 1,m_2)(d_1-\mu+2\widehat{\rho}_{22}^{DD}(m_2)R_2)-R_2] \\
&\cdot \left(\frac{d\widetilde{q}_1^{DD*}(m_1,m_2)}{dR_1}(d_2-\mu+2\widehat{\rho}_{22}^{DD}(m_2)R_2)+\frac{\partial\widehat{\rho}_{22}^{DD}(m_2)}{\partial R_2}2\widetilde{q}_1^{DD*}(m_1,m_2)R_2\right)
\end{aligned}$$

Proposition 3 shows that even when $\frac{\partial\widehat{\rho}_{22}^{DD}(m_2)}{\partial R_1}>0$ it may be that $\frac{d\widetilde{q}_2^{DD*}(m_1,m_2)}{dR_1}<0$, in which case the decline in the true probability of an attack follows trivially in all cases. However, it is also possible that $\frac{d\widetilde{q}_2^{DD*}(m_1,m_2)}{dR_1}>0$. To account for this possibility it is necessary instead to consider the change in the probability of conflict when country 1 democratizes:

$$\frac{dw^{DD}(m_1,m_2;d_1,d_2)}{dR_1}=\frac{dq_1^{DD}(m_1,m_2;d_1)}{dR_1}(1-q_2^{DD}(m_1,m_2;d_2))+\frac{dq_2^{DD}(m_1,m_2;d_2)}{dR_1}(1-q_1^{DD}(m_1,m_2;d_1)).$$

For clarity, the result is shown assuming that $\widehat{\rho}_{22}^{DD}(m^H)>\widehat{\rho}_{22}^{AD}(m^H)$. The same procedure is followed if $\widehat{\rho}_{22}^{DD}(m^L)>\widehat{\rho}_{22}^{AD}(m^L)$, and thus is omitted for brevity.

There are two possible signals and two possible levels of vulnerability for each country, giving sixteen total cases for $w^{DD}(m_1,m_2;d_1,d_2)$.

There are eight cases in which $m_2=m^L$. In these cases

$$\frac{d\widetilde{q}_j^{DD*}(m_1,m^L)}{dR_1}<0 \qquad \Rightarrow \qquad \frac{dq_i^{DD}(m_1,m^L)}{dR_1}<0,$$

and thus it follows trivially that $w^{DD}(m_1,m^L;d_1,d_2)<w^{AD}(m_1,m^L;d_1,d_2)$.

If $m_2=m^H$ and $\frac{d\widetilde{q}_2^{DD*}(m^H,m^H)}{dR_1}>0$, then $\left|\frac{d\widetilde{q}_1^{DD*}(m^H,m^H)}{dR_1}\right|>\left|\frac{d\widetilde{q}_2^{DD*}(m^H,m^H)}{dR_1}\right|$.

If $d_1=d_2$ and $m_1=m^H$ then, by symmetry, it is sufficient to show that

$$2\widetilde{q}_2^{DD*}(m^H,m^H)\widehat{\rho}_{11}^{DD}(m^H)-1+\frac{\partial\widehat{\rho}_{11}^{DD}(m^H)}{\partial R_1}2\widetilde{q}_2^{DD*}(m^H,m^H)R_1+\frac{\partial\widehat{\rho}_{22}^{DD}(m^H)}{\partial R_1}2\widetilde{q}^{DD*}(m^H,m^H)R_2<0$$

It is intuitive that

$$\frac{\partial \widehat{\rho}_{22}^{DD}(m^H)}{\partial R_1} < \frac{\partial \widehat{\rho}_{11}^{DD}(m^H)}{\partial R_1} < \widehat{\rho}_{ii}^{DD}(m^H) - \rho_{ii}$$

Setting the above expression to equality and evaluating, it follows that

$$2\widetilde{q}_j^{DD*}(m^H, m^H)\widehat{\rho}_{ii}^{DD}(m^H) - 1 + 2(\widehat{\rho}_{ii}^{DD}(m^H) - \rho_{ii})2\widetilde{q}_j^{DD*}(m^H, m^H) < 0$$

and the result is confirmed. It follows automatically that the same result holds when $m_1 = m^L$, since there is a smaller increase in $\widetilde{q}_2^{DD*}(\cdot, m^H)$ and greater decrease in $\widetilde{q}_1^{DD*}(\cdot, m^H)$.

To complete the result, the final two cases are considered together. Note that, by symmetry, there is the same probability of both cases arising, so this can be dropped in the comparison. First, when $m_2 = m^H$,

$$\begin{aligned}
\frac{dw^{DD}(m^H, m^H; d^L, d^H)}{dR_1} &+ \frac{dw^{DD}(m^H, m^H; d^H, d^L)}{dR_1} \\
&= \frac{dq_1^{DD}(m^H, m^H; d^L)}{dR_1}(1 - q_2^{DD}(m^H, m^H; d^H)) + \frac{dq_2^{DD}(m^H, m^H; d^H)}{dR_1}(1 - q_1^{DD}(m^H, m^H; d^L)) \\
&+ \frac{dq_1^{DD}(m^H, m^H; d^H)}{dR_1}(1 - q_2^{DD}(m^H, m^H; d^L)) + \frac{dq_2^{DD}(m^H, m^H; d^L)}{dR_1}(1 - q_1^{DD}(m^H, m^H; d^H)).
\end{aligned}$$

The probability of conflict cannot be confirmed to decrease in the first case, since the smaller decrease in country 2's probability of attack is given a greater 'weight' than country 1's larger increase. However, any potential increase is countered by the reverse weighting in the second case ($q_1^{DD}(m^H, m^H; d^L) = q_2^{DD}(m^H, m^H; d^L)$ by symmetry), confirming that the average probability of conflict decreases.

Again, it is automatic that the result also holds when $m_1 = m^L$.

Combining these results shows that:

$$
\begin{aligned}
W^{DD} &= \sum_{d_1}\sum_{d_2}\sum_{m_1}\sum_{m_2} w^{DD}(m_1,m_2;d_1,d_2)Pr(d_1)Pr(d_2)Pr(m_1|d_1,DD)Pr(m_2|d_2,DD) \\
&< \sum_{d_1}\sum_{d_2}\sum_{m_2} w^{AD}(\cdot,m_2;d_1,d_2)Pr(d_1)Pr(d_2)Pr(m_2|d_2,AD) = W^{AA}
\end{aligned}
$$

QED

Proof of Proposition 2 (Effect of Opposing Regime Type on Signal Choice)

$$\widetilde{c}_2^{DD}(d_2) = \mu + Q_1^{DD}(m^L)(d_2 + \widehat{\rho}_{22}^{DD}(m^L)R) + Q_1^{DD}(m^H)(\widehat{\rho}_{22}^{DD}(m^H)R - \mu) - R$$

$\widetilde{c}_2^{DD}(d_2) = \widetilde{c}_2^{AD}(d_2)$ when $R_1 = 0$. It is therefore sufficient to show that $\widetilde{c}_2^{DD}(d_2)$ decreases as R_1 increases, i.e. when country 1 democratizes.

$$\widetilde{q}_1^{AD*}(\cdot,m_2) > \widetilde{q}_1^{DD*}(m^H,m_2) > \widetilde{q}_1^{DD*}(m^L,m_2) \qquad \Rightarrow \qquad 0 > \frac{d\widetilde{q}_1^{DD*}(m^H,m_2)}{dR_1} > \frac{dQ_1^{DD}(m_2)}{dR_1}.$$

Therefore $\frac{d\widetilde{c}_2^{DD}(d_2)}{dR_1} < 0$ if

$$\frac{d}{dR_1}\left[\mu + \widetilde{q}_1^{DD*}(m^H,m^L)(d_2 + \widehat{\rho}_{22}^{DD}(m^L)R) + \widetilde{q}_1^{DD*}(m^H,m^H)(\widehat{\rho}_{22}^{DD}(m^H)R - \mu) - R\right] < 0.$$

Denote

$$
\begin{aligned}
C(d_2) &= 2\mu + 2\widetilde{q}_1^{DD*}(m^H,m^L)(d_2 + \widehat{\rho}_{22}^{DD}(m^L)R) + 2\widetilde{q}_1^{DD*}(m^H,m^H)(\widehat{\rho}_{22}^{DD}(m^H)R - \mu) - 2R \\
&= B(d_2) - D(d_2)
\end{aligned}
$$

where

$$
\begin{aligned}
B(d_2) &= \mu + \widetilde{q}_1^{DD*}(m^H, m^H)(d_2 - \mu + 2\widehat{\rho}_{22}^{DD}(m^H)R) - R + \mu + \widetilde{q}_1^{DD*}(m^H, m^L)(d_2 - \mu + 2\widehat{\rho}_{22}^{DD}(m^L)R) - R \\
&= F^{-1}[q_2^{DD}(m^H, m^H; d_2)] + F^{-1}[q_2^{DD}(m^H, m^L; d_2)] \\
D(d_2) &= [\widetilde{q}_1^{DD*}(m^H, m^L) - \widetilde{q}_1^{DD*}(m^H, m^H)](d_2 + \mu)
\end{aligned}
$$

By convexity, $B(d_2)$ is decreasing if $q_2^{DD}(m^H, m^H; d_2) + q_2^{DD}(m^H, m^L; d_2)$ is decreasing as country 1 democratizes. This result follows directly from the proof of Proposition 4, since any possible increase in one term is outweighed by the decrease in the other.

$$
\frac{d\widetilde{q}_1^{DD*}(m^H, m^L)}{dR_1} - \frac{d\widetilde{q}_1^{DD*}(m^H, m^H)}{dR_1} > 0 \qquad \Rightarrow \qquad \frac{dD(d_2)}{dR_1} > 0
$$

This result follows directly from the comparison of the change in equilibrium expected probability in the proof of Proposition 3, as the decrease is lower when it is indicated that the country is likely to be vulnerable.

$$
\frac{dC(d_2)}{dR_1} < 0 \qquad \Rightarrow \qquad \frac{d\widetilde{c}_2^{DD}(d_2)}{dR_1} < 0.
$$

This implies that

$$
\widetilde{c}_2^{AD}(d_2) > \widetilde{c}_2^{DD} \qquad \text{for } d_2 = d^H, d^L.
$$

QED

Bibliography

[1] Acemoglu, Daron, Davide Ticchi and Andrea Vindigni (2010), "A Theory of Military Dictatorships," *American Economic Journal: Macroeconomics* **2**(1), 1-42.

[2] Akerlof, George A. and William T. Dickens (1982), "The Economic Consequences of Cognitive Dissonance," *American Economic Review* **72**(3), 307-319.

[3] Alesina, Alberto, Reza Baqir and William Easterly (1999), "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics* **114**(4), 1243-1284.

[4] Anderlini, Luca, Dino Gerardi and Roger Lagunoff (2010), "Social Memory, Evidence and Conflict," *Review of Economic Dynamics* **13**(3), 559-574.

[5] Aumann, Robert J. (1976), "Agreeing to Disagree," *The Annals of Statistics* **4**(6), 1236-1239.

[6] Baliga, Sandeep, David O. Lucca and Tomas Sjöström (2011), "Domestic Political Survival and International Conflict: Is Democracy Good for Peace?," *Review of Economic Studies* **78**(2), 458-486.

[7] Baliga, Sandeep and Tomas Sjöström (2004), "Arms Races and Negotiations," *Review of Economic Studies* **71**(2), 351-369.

[8] Baliga, Sandeep and Tomas Sjöström (2011), "The Strategy of Manipulating Conflict", *American Economic Review*, forthcoming.

[9] Bandiera, Oriana and Gilat Levy (2008), "Diversity and the Power of the Elites in Democratic Societies," mimeo LSE.

[10] Bénabou, Roland (2008), "Groupthink: Collective Delusions in Organizations and Markets," mimeo Princeton.

[11] Bester, Helmut and Karl Wärneryd (2006), "Conflict and the Social Contract," *Scandinavian Journal of Economics* **108**(2), 231-249.

[12] Bueno de Mesquita, Bruno, James D. Morrow, Randolph M. Siverson and Alastair Smith (1999), "An Institutional Explanation of the Democratic Peace," *The American Political Science Review* **93**(4), 791-807.

[13] Bush, George W. (2003), "State of the Union," available at http://georgewbush-whitehouse.archives.gov.

[14] Easterly, William and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics* **112**(4), 1203-1250.

[15] Esteban, Joan, Laura Mayoral and Debraj Ray (2010), "Ethnicity and Conflict: An Empirical Study," mimeo NYU.

[16] Esteban, Joan and Debraj Ray (2001), "Collective Action and the Group Size Paradox", *The American Political Science Review* **95**(3), 663-672.

[17] Esteban, Joan and Debraj Ray (2010), "A Model of Ethnic Conflict," *Journal of the European Economic Association*, forthcoming.

[18] Esteban, Joan and Debraj Ray (2008), "On the Salience of Ethnic Conflict," *American Economic Review* **98**(5), 2185-2202.

[19] Esteban, Joan and Debraj Ray (2009), "Linking Conflict to Inequality and Polarization," *American Economic Review*, forthcoming.

[20] Farrell, Joseph and Robert Gibbons (1989), "Cheap Talk with Two Audiences," *American Economic Review*, **79**(5), 1214-1223.

[21] Fearon, James D. and David D. Laitin (2000), "Violence and the Social Construction of Ethnic Identity," *International Organization* **54**(4), 845-877.

[22] Foreign and Commonwealth Office (2002), "Iraq's Weapons of Mass Destruction: The Assessment of the British Government," available at www.nationalarchives.gov.uk.

[23] Garfinkel, Michelle and Stergios Skaperdas (2007), "Economics of Conflict: An Overview," in T. Sandler and K. Hartley (eds.), *Handbook of Defense Economics*, Vol.II, North Holland, 649-709.

[24] Genicot, Garance and Stergios Skaperdas (2002), "Investing in Conflict Management," *Journal of Conflict Resolution* **46**, 154-170.

[25] Glewwe, Paul (1986), "The Distribution of Income in Sri Lanka in 1969-70 and 1980-81: A Decomposition Analysis," *Journal of Development Economics* **24**(2), 255-274.

[26] Grossman, Herschel I. and Minseong Kim (1995), "Swords or Plowshares? A Theory of the Security of Claims to Property," *The Journal of Political Economy* **103**(6), 1275-1288.

[27] Hirshleifer, Jack (1995), "Anarchy and its Breakdown," *The Journal of Political Economy* **103**(1), 26-52.

[28] Horowitz, Donald (2001), "Ethnic Groups in Conflict" (second edition), UC Press.

[29] Jackson, Matthew O. and Massimo Morelli (2007), "Political Bias and War," *American Economic Review* **97**(4),1353-1373.

[30] Kirshner, Jonathan (2000), "Rationalist Explanations for War?" *Security Studies* **10**(1), 143-150.

[31] Köszegi, Botond (2006), "Emotional Agency," *Quarterly Journal of Economics*, **121**(1), 121-156.

[32] Latham, Oliver (2011), "Political Divisions and Foreign Intervention," mimeo University of Cambridge.

[33] Le Breton, Michel (1989), "A Note on Balancedness and Nonemptiness of the Core in Voting Games," *International Journal of Game Theory* **18**(1), 111-117.

[34] Leeds, Brett Ashley and David R. Davis (1999), "Beneath the Surface: Regime Type and International Interaction, 1953-1978" *Journal of Peace Research* **63**(1), 5-21.

[35] Levy, Gilat and Ronny Razin (2004), "It Take Two: An Explanation for the Democratic Peace," *Journal of the European Economic Association* **2**(1), 1-29.

[36] Maoz, Zeev and Bruce Russett (1993), "Normative and Structural Causes of Democratic Peace, 1946-1986," *The American Political Science Review* **87**(3), 624-638.

[37] McKenzie, David, John Gibson and Steven Stillman (2007), "A Land of Milk and Honey with Streets Paved with Gold: Do emigrants have over-optimistic expectations about incomes abroad?", CReAM Discussion Paper.

[38] Østby, Gudrun (2008), "Inequalities, the Political Environment and Civil Conflict: Evidence from 55 Developing Countries," in Frances Stewart (ed.), *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*, Basingstoke: Palgrave Macmillan, 136-157.

[39] Ray, Debraj (2009), "Remarks on the Initiation of Costly Conflict," mimeo NYU.

[40] Reiter, Dan and Allan C. Stam III (1998), "Democracy, War Initiation and Victory," *American Political Science Review* **92**(2), 377-389.

[41] Riboni, Alessandro (2011), "Doubts and Dogmatism in Conflict Behavior," mimeo Université de Montréal.

[42] Rousseau, David L., Christopher Gelpi, Dan Reiter and Paul K. Huth (1996), "Assessing the Dyadic Nature of the Democratic Peace", *The American Political Science Review* **90**(3), 512-533.

[43] Samii, Cyrus (2010), "Do Quotas Exacerbate or Reduce Ethnic Conflict? Evidence from Burundi's Military," mimeo NYU.

[44] Schultz, Kenneth A. (2005), "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch", *International Organization* **59**(1), 1-38.

[45] Schwarz, Michael and Konstantin Sonin (2007), "A Theory of Brinkmanship, Conflicts, and Commitments", *The Journal of Law, Economics and Organization* **24**(1), 163-183.

[46] Skaperdas, Stergios (1991), "Conflict and Attitudes Toward Risk," *American Economic Review*, **81**(2), 116-120.

[47] Skaperdas, Stergios (1996), "Contest Success Functions," *Economic Theory* **7**, 283-290.

[48] Sriskadarajah, Dhananjayan, "Development, Inequality and Ethnic Accomodation: Clues from Malaysia, Mauritius and Trinidad and Tobago," *Oxford Development Studies* **33**(1), 63-79.

[49] Sriskandarajah, Dhananjayan (2003), "Inequality and Conflict in Fiji: From Purgatory to Hell?" *Asia Pacific Viewpoint* **44**(3), 305-324.

[50] Sriskandarajah, Dhananjayan (2005), "Socio-Economic Inequality and Ethno-Political Conflict: Some Observations from Sri Lanka," *Contemporary South Asia* **14**(3), 341-356.

[51] Stewart, Frances, "Horizontal Inequalities: A Neglected Dimension of Development", CRISE Working Paper.

[52] Tångeras, Thomas P. (2009), "Democracy, Autocracy and the Likelihood of International Conflict," *Economics of Governance* **10**(2), 99-117.

[53] Uppsala Conflict Data Program (Date of retrieval: 2012/03/22) UCDP Conflict Encyclopedia: www.ucdp.uu.se/database, Uppsala University.

[54] Wimmer, Andreas, Lars-Erik Cederman and Brian Min (2009), "Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Dataset," *American Sociological Review* **74**(2), 316-338.

[55] YouGov (2010), "Who Do You Trust?," available at www.YouGov.com.