Ethnic Conflicts, Rumours and an Informed Agent*

Pathikrit Basu[a], Souvik Dutta[b] & Suraj Shekhar[c]

a - The Pennsylvania State University

b - Indian Institute of Management (Bangalore)

c - The Pennsylvania State University

May 24, 2016

Abstract

Rumours often precipitate ethnic conflicts and cause immense damage to life and property. There may exist an agent who knows if the rumour is true or false. We analyze a cheap talk game with multiple audiences (ethnicities) to see how this informed agent (b) may influence the outcome of rumours by sending strategic signals. Since b is biased towards her own ethnicity, she finds it difficult to convince the other ethnicity that she is giving them correct information. We show that even if b is known to be biased towards her own ethnicity, peace is possible in equilibrium. Additionally, we prove that there are only three equilibrium outcomes possible in symmetric strategies. Conflict is inevitable in one. The other outcomes have the following features. One, there may be peace whenever b deems it possible. Two, while b gives more informative signals to her own ethnicity, she may misinform a segment of her own ethnicity in equilibrium. (JEL - D74, D83, P16, D82)

Keywords - Ethnic Conflicts, Strategic Information Disclosure, Rumours, Coordination, Informed Agents

*a - Pathikrit Basu: The Pennsylvania State University. Address - 303, Kern Building, University Park, PA-16802, USA. E-mail - pub143@psu.edu. b- Souvik Dutta: Indian Institute of Management, Bangalore. Address - E001, Indian Institute of Management, Bangalore, India. E-mail - souvik@iimb.ernet.in. c- Suraj Shekhar (Corresponding Author): The Pennsylvania State University. Address - 303, Kern Building, University Park, PA-16802, USA. E-mail - sws5487@psu.edu. We are deeply grateful to Kalyan Chatterjee, Edward Green and Kala Krishna for many helpful discussions. We also want to thank Anirban Mitra, Ryan Fang, Bruno Salcedo, Tetsuya Hoshino, Guillem Roig, Debraj Ray, Vijay Krishna, Syed Nageeb Ali, Ashutosh Varshney, participants at the summer meeting (2014) of the Econometric society at the University of Minnesota, participants at the 10th Annual Conference on Economic Growth and Development at the Indian Statistical Institute, New Delhi, participants at the 33rd Australasian Economic Theory Workshop at Deakin University, seminar participants at IIM Bangalore, participants at the inaugural Society for Economics Research conference at ISB Hyderabad and seminar participants at The Pennsylvania State University for their comments. An early version of this paper was circulated as *Urban Ethnic Conflicts*. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Contents

1 Introduction **4**

2 Literature **7**

3 Model **10**

3.1 Players 10

3.2 Conflict 11

3.3 States of the World 11

3.4 Payoffs 12

3.5 Rumour 13

3.6 Meeting Process/Obtaining Information 13

3.7 Timeline 13

3.8 Beliefs and Information 14

3.8.1 About Rumour 14

3.8.2 On existence of b 14

3.8.3 On Contents of Letter 14

3.9 Game Tree 14

3.10 Assumptions 15

3.11 Equilibrium Concept 15

4 Non-Strategic Informed Agent **16**

4.1 Results 16

5 Strategic Informed Agent **18**

5.1 Strategies 19

5.2 Equilibrium 20

5.2.1 Truth telling 21

5.2.2 Non-truth telling 22

5.3 Outcome Characterization 26

5.4 High probability of b's existence 29

5.5 Other payoff types for b 30

5.5.1 Extremist type 30

6 Discussion **31**

6.1 Correlation of Distributions 31

6.2 Using Letters to Represent Meetings 31

6.3 Non Random Meetings 32

6.4 Uncertainty about b's ethnicity 32

6.5 Limiting k only 32

6.6 Urban Ethnic Conflicts 33

6.7 Cost of participation 33

7 Conclusion **34**

A Appendix **38**

A.1 Type space and prior 38

A.2 Strategies 41

A.3 Results : Non-strategic b 41

A.3.1 Proof of Proposition 3 41

A.3.2 Mixed Strategy Equilibria 43

A.4 Results : Strategic b 45

1 Introduction

Rumours often play a key role in ethnic[1] conflicts. They mobilize people and provide justification to commit violence (Brass, 2011). Many places in the world have suffered greatly due to rumour induced conflicts. The 2014 conflict in New Delhi's Trilokpuri region started when a non-communal fight led to rumours that it was actually a Hindu-Muslim clash. Rumours about the Tutsis shooting down an air-plane carrying Rwandan President Juvenal Habyarimana, who died in the crash led to a massacre of the Tutsi people by Hutu extremists in Rwanda (Bhavnani et al. (2009)). Even developed countries have not escaped the scourge of rumours. The infamous Detroit riots of 1943 in the USA was precipitated by rumours that two white men had thrown a black mother and her child off the Belle Isle Bridge. In a nearby white neighbourhood, another rumour spread that some black men had raped and murdered a white woman on the same bridge. The ensuing riots left over 30 dead and over 75 injured. These rumours were never shown to be true. However, for the people who heard them, there was an apprehension of things to come which led them to commit violent acts. Thus, rumours can precipitate conflicts by a) causing people to have pessimistic beliefs about the state of the world and b) mobilizing people and coordinating their actions via common information.

The disastrous effect of rumours is somewhat puzzling in light of the observation that whenever there is a rumour, there are people who know the truth about the rumour i.e. whether or not it is correct[2]. Why can't these informed agents prevent conflict by revealing their information to all? After all, there have been instances when informed agents have been able to do this. For example, during the Hindu-Muslim riots in Gujarat, India (2002), a small town called Raamrahimnagar in Gujarat remained relatively unaffected because local committees consisting of both Hindus and Muslims were able to prevent large scale riots by squashing rumours with correct information (Berenschot (2011))[3] The problem is, even if we assume that the informed agent wants to have peace (we make this assumption at first and then relax it), the following factors may affect the extent to which informed agents are able to influence the outcome of rumours - a) Reachability - she may be able to share her knowledge with a very small fraction of the society [4] and b)

[1] All conflicts based on ascriptive group identities - race, language, religion, tribe, or caste can be called ethnic (Horowitz (1985)).

[2] For example, in the Trilokpuri incident in New Delhi, natural deaths were being misreported as deaths caused by communal violence ('*Rumour mongers in both sides spew venom'* - Deccan Herald, November 8, 2014). Of course, the families of the deceased knew the truth, and when questioned, confirmed that the deaths were natural.

[3] Similar results are expected from the "Una Hakika" project in Kenya's Tana delta run by the non-profit organization - The Sentinel Project. For details look at - *Using Cell Phones To Stop False Rumors, Before They Lead To Ethnic Violence*- http://www.fastcoexist.com/3029321/using-cell-phones-to-stop-false-rumors-before-they-lead-to-ethnic-violence , April 29, 2014. Upon hearing a rumour, people can text the rumour to this organization and they will check and respond with their findings. The Una Hakika User Satisfaction Survey conducted in April 2015 confirmed that most people believed that the Una Hakika team was providing unbiased information and that the project helped dispel the spread of rumours and made them feel safer (https://thesentinelproject.org/wp-content/uploads/2015/04/Una-Hakika-User-Survey-April-2015.pdf

[4] Generally, the people connected to the rumour know the truth but they can only reveal it to the small population they know. More often than not, they will not have access to a media platform to reveal the truth to all. This would be particularly true if the

Bias - if the informed agent has a known bias towards one ethnicity, she may find it hard to convince the other ethnicity that she is giving them correct information.

The contribution of this paper is to analyze the role of informed agents in influencing the outcomes of rumours in the presence of these issues. We characterize the equilibria possible in such environments. The message of this paper is a hopeful one. As long as the informed agent has high reachability, peace can be an equilibrium outcome. This is true even if the informed agent is known to be biased towards one of the ethnicities. Our paper also contributes to the growing literature on cheap talk game with multiple audiences with the novel addition of payoff externalities (more on this in section 2). Next, we give an intuitive idea of our model and results. We use a simple formulation to describe a good and a bad state of the world and we think of rumours as public signals (news article in the papers or in the media) which push the beliefs of people towards the bad state of the world[5]. A description of the model follows.

There are two ethnicities in a society - E_1 and E_2. The ethnicity of each individual is common knowledge. Additionally, each player could be one of two types - G (good), B (bad). This is private information. Every player has to choose between the actions fight or not fight. A conflict occurs if a large enough fraction of any one ethnicity chooses to fight. The B players are behavioural and always fight. The G players are strategic. There are two states of the world (unobserved by all agents), each characterized by a distribution of G, B types. In the good state of the world, there are sufficiently many G players so that there is no conflict if the G players choose not to fight. In the bad state of the world, conflict is inevitable due to the presence of a large number of B players. A rumour is a public signal which may be true or false. Both the arrival of the rumour and it's veracity (true or false) are correlated with the state of the world[6]. The players cannot distinguish between a true and false rumour. However, they believe that there may exist an informed agent (b) who can[7]. The arrival of the rumour may induce conflict by pushing beliefs towards the bad state of the world. However, before making the decision to fight or not, there is a meeting stage where players may get further information from b about the rumour via private messages. We analyze two cases - in the benchmark case, we consider the simple case[8] where b is non-strategic i.e. she reveals her information truthfully (if the rumour is false, b tells everyone she meets that it is false and if it's true she tells everyone it is true). In the more interesting case, we consider the equilibria when b is strategic and has a commonly known bias towards one ethnicity (b belongs to one ethnicity). b can meet at most k fraction of the population. This

original story was pushed into publication by political heavyweights.

[5]Since we think of rumours as public signals, we are largely concerned with widespread rumours.

[6]Rumours may be more likely to arrive in the bad state of the world. Additionally, depending upon the statement of the rumour, whether the rumour is true or false may also be informative about the state of the world.

[7]We will refer to player b as female and all others as male. Note also that b does not need to be one agent. As in the Una Hakika example, b could also be a group of people or an organization.

[8]Though not unrealistic, as we describe later.

captures the reachability of the informed agent.

To get an idea about the environment described above, consider the example of the Detroit riots. People learn of the existence of such rumours and immediately become pessimistic about the things to come. They don't know if the rumour is true or false. However, just the fact that such rumours are circulating makes them apprehensive about future events. Additionally, everyone realizes that there may have been a person passing by the bridge at the time the incident was rumoured to have occurred. This player would then know if the incident actually happened or not. All players place positive probability on the existence of such a player. We describe our results next.

In the simple benchmark case, if b has high reachability and good information (the veracity of the rumour is closely related to the state of the world), then her superior information can reverse the effect of a bad rumour. This is fairly easy to show. We must point out that the fact that the meeting process is common knowledge, plays an important role in this analyses since it allows players to estimate what other players may know and coordinate their actions[9].

Now, consider the case of a strategic b with a bias and informative signals. In particular, we assume that if b exists she belongs to one of the ethnicities and has a utility function such that if she believes that the state is likely to be good, she wants peace to prevail (we relax this assumption and consider a different payoff type for the informed agent in section 5.5). However, if she thinks that the state is likely to be bad and therefore conflict is inevitable, she wants her own ethnicity to win. The first thing to note here is that truth telling is not an equilibrium strategy. Thus, the benchmark case may be unrealistic if we think of the informed agent as a strategic player. For the strategic informed agent, we show that there are only three equilibrium outcomes possible. In one outcome, everyone fights and conflict occurs with probability one. In the other two equilibrium outcomes, there may be no conflict when b believes that peace is possible (state is likely to be good). Furthermore, in this case, if b's information is imperfect she will give false information to a small fraction of her own ethnicity to hedge against the risk of having incorrect information (her signal leads her to believe the wrong state is the likely one).

Essentially, our results derive from only two conditions. One, the informed agent should not get very low utility from peace. Peace does not even need to be the highest utility outcome for the informed agent. All we need is lower bound to her utility from peace. This provides incentives for the informed player to try and obtain peace when it is possible. Two, every player gets higher expected utility from fighting as more of his own ethnicity chooses to fight and lower expected utility from fighting as more of the opposite

[9]This has been shown to be critical in other environments like global games with noisy information about the payoff structure (Carlsson and Van Damme (1993)).

ethnicity players choose to fight. This sort of complementarity and substitutability helps us establish the mixed strategy equilibrium which leads to peaceful outcomes. The intuition here is that given the strategy of the informed agent and the other ethnicity, if a small fraction of your own ethnicity will fight then the probability of winning is so low that the gains from not fighting and hoping for the peaceful outcome dominates fighting. On the other hand, if a very large fraction of one's own ethnicity is fighting then it pays to fight. Thus, there is a fraction at which the player will be indifferent. We show that the point at which the players mix is below the threshold required for conflict and therefore we can obtain peace as an equilibrium outcome.

The intuitive idea behind the result that the informed agent may provide false information to a section of her own ethnicity is as follows - b estimates that the state is likely to be good so she would like to have the peaceful outcome. However, her information is not perfect so she realizes that her estimation may be wrong and the state may be bad, thereby making conflict inevitable. Thus, she misinforms a fraction of her own ethnicity and persuades them to play fight in order to give her ethnicity a better chance of winning in case her information turns out to be an incorrect signal of the state of the world. This fraction is the largest that can play fight without actually *causing* conflict (remember conflict occurs for sure if for any ethnicity, the fraction of players who play fight is above a threshold). In section 5.5, we consider another natural utility function for the informed agent and show that our results are robust to this alteration.

The paper is organized as follows. Section 2 describes some of the relevant literature. Section 3 describes the main features of the model and the particular equilibrium concept relevant here. Section 4 deals with the case of player b being non-strategic and we discuss the case of a strategic b player in section 5. A discussion of our assumptions, modelling choices and a possible explanation for the high percentage of ethnic conflicts in urban areas (in India) is in section 6. The conclusion is in section 7.

2 Literature

Our paper is related to the literature on cheap talk games with multiple audiences, global games and models of strategic information disclosure.

Two related papers in the cheap talk games with multiple audiences literature are Farrell and Gibbons (1989) and Goltsman and Pavlov (2011). Farrell and Gibbons (1989) considers a standard cheap talk environment with one sender and two receivers. The key differences from our paper are outlined next. While we consider a game where the informed agent (sender) communicates with *both* ethnicities (receivers) si-

multaneously and privately, the games considered in Farrell-Gibbons either include private communication with a single receiver or public communication with both. Secondly, the action chosen by one receiver does not influence the utility of the other receiver i.e there are no payoff externalities. This is in contrast to our set-up where coordination within and across ethnicities is important (payoff externalities). Finally, unlike in Farrell and Gibbons (1989), truth telling cannot be an equilibrium in our model. Goltsman and Pavlov (2011) is closely related to Farrell-Gibbons (1989). They also have one sender and two receivers. Like Farrell-Gibbons, they compare public and private communication. However, they consider private communication with *both* receivers. Another important feature of the paper is the analysis of the game with both public and private communication in which they show that less information will be provided in the private communication stage. The important distinction from our paper is that while this paper considers private communication with both receivers, as with the case of Farrell-Gibbons, there are again no payoff externalities.

In the information disclosure literature, Rauchhaus (2006) shows that a third party mediator can effectively avoid a war if the mediator possesses private information about one or more of the disputant's capabilities. There is a cost of war for both disputants and the mediator receives private information about the defender's cost of fighting (may be high or low). By revealing this information to the challenger the mediator can ensure an offer from the challenger which the defender will accept thereby averting conflict. Our paper differs from this paper in the following way. In Rauchhaus (2006), the mediator can send a signal to only one of the agents whereas in our model the informed agent can send a signal to each and every member of both the groups. If the informed agent could send signals to only one ethnicity, conflict could not be avoided in our model. This is because, unlike Rauchhaus (2006), both groups are uninformed about the state of the world in our model, rumours affect the beliefs of players from both groups and conflict can break out if sufficiently many members of either ethnicity decide to fight. Other papers in this vein consist of - Egorov and Sonin (2014) who show that a dictator would manipulate the beliefs held by the citizen's about his popularity. She is able to do this because she has additional information over the citizens and can use this to deter the citizens from protesting against her. Shadmehr and Bernhardt (2015) also show a similar result where a ruler (an authoritarian state) may censor the information available to citizens to avoid revolution. A lot of the information disclosure literature is in the context of a firm/manager's decision to disclose private information optimally to the investors/buyers which can affect firm's future values and earnings. Dye (1985) and Jung and Kwon (1988), in the context of managers revealing private information to investors, show that there cannot be a policy of full disclosure in equilibrium. This is similar to our result of no-truth telling equilibrium in the case of strategic b.

On the global games front, Lu et al. (2013) assess the impact of circulation of rumours on regime change by studying a coordination game under a global game structure with both public and private signals. In particular, they study the effect of communication (regarding private information about the state of the world) amongst agents. They conclude that under communication, rumours have more impact in mobilizing agents (as they send each other confirmatory messages about their beliefs). Our paper differs from Lu et al. (2013) in the following way. In our paper, post rumour arrival, there may (or may not) exist only one person with more information about the rumour who communicates *strategically* with the masses to align their actions with her own incentives. In contrast, in Lu et al. (2013), everyone receives some *exogenous* private signals about the rumour. Generally, very few people have genuine information about rumours and these players will communicate them strategically. Thus, the assumption that everyone receives signals and that these signals are exogenous seems untenable to us. Another paper on regime change which is related to ours is Tyson and Smith (2014). They study a two-sided coordination problem in a global games environment. There are two groups of citizens (regime adherents and regime opponents). As in our paper, incentives to fight increase (decrease) as more players from the same (opposite) group choose to fight. Also, regime adherents may be better informed than regime opponents. They show that, in spite of differences in private signals, two sided coordination implies that public signals will influence the incentives to act for both groups in the same way. Despite some similarities, certain aspects of our model are not captured by Tyson and Smith (2014). One, these groups have different first best outcomes and only one group may have their preferred outcome in equilibrium. In our model both ethnicities prefer peace over other outcomes and therefore there are coordinating incentives. Two, the 'club' signal which seriously modifies the level of information between groups is a public signal sent to one group. In our model, the strategic agent has the ability to send a different signal to every single player. This plays an important role since it allows the informed agent to send less informative signals to the opposite ethnicity and even misinform some players of her own ethnicity.

There have been a series of papers by Esteban and Ray which offer insights into why ethnic conflicts happen. Esteban and Ray (2008) point out that ethnic conflict may be more likely to occur than class conflict where there is within-group economic inequality. Esteban and Ray (2011) use a theoretical model to show how within-group heterogeneity in radicalism and income help in precipitating ethnic conflicts. These papers highlight income inequality as a source of ethnic conflict. Income inequality could be an explanation for the disproportionate percentage of ethnic conflicts in urban areas as these areas are likely to have more income heterogeneity (see section 6.6 for our explanation).

3 Model

We will only discuss the case of non-strategic informed agent in our base line model and leave the analysis on strategic informed agent for section 5. The analyses of the non strategic informed agent case is fairly simple and serves largely as a benchmark case for the more interesting case of a strategic informed agent. Having said that, we must point out that the case of a non strategic informed agent is neither unrealistic nor unimportant. There could be many interpretations of a non-strategic b player. For example, b could be an unbiased player who has an honest reputation to protect or b could be unaffected by the outcome of conflict (example she does not belong to region where conflict may happen) so she might as well be truthful. During the Gujarat riots of 2002, local committees consisting of both Hindus and Muslims were able to keep Raamrahimnagar relatively peaceful by squashing rumours with the correct information whenever possible. These committees could be interpreted as an example of non-strategic informed agents. The reason they are interpreted to be non-strategic is because they have members from both ethnicities. Therefore, they could not be giving differential information to the two ethnicities.

We describe the components of our model next.

3.1 Players

The environment has a continuum of agents and each player can be one of two ethnicities - $\{E_1, E_2\}$. Ethnicity can be interpreted in many ways ranging from religion to race to nationalities. For each ethnicity, players in it are indexed by $l \in [0,1]$. Hence, the set of all players N can be identified with $[0,1] \times \{E_1, E_2\}$ and an arbitrary player will be denoted as $i \in [0,1] \times \{E_1, E_2\}$. For example, $i = (l, E_1)$ would denote the l-th agent in the ethnic group E_1. For notational convenience, we shall denote as E_1 the set $[0,1] \times \{E_1\}$ and as E_2 the set of agents in $[0,1] \times \{E_2\}$. Hence, E_1 and E_2 will represent the two ethnic groups. We endow N with the natural uniform measure and shall denote it by μ. Hence, $\mu(E_1) = \mu(E_2) = 1$ implying that each ethnicity has the same mass of agents. Note that this is a simplifying assumption which is not crucial to our results. Additionally, a player can be one of two types - *Good* (G) or *Bad* (B). The two types differ in terms of the actions available to them. Players can decide to fight (f) or not (nf). The G type player is strategic. He can choose either action and fights only if it gives him higher payoff. B type players, on the other hand are behavioural and always choose to fight.

After a rumour arrives, there may be one more non-strategic player (outside the population) b who can prove the veracity of the rumour. Note that the assumption that b is outside the population is just for

simplicity of calculation. The players place positive beliefs on the existence of this informed agent. Note that b only knows if the rumour is true or false. She does not know the state of the world.

3.2 Conflict

If a large enough fraction of at least one group chooses to fight then a conflict ensues. If both groups fail to gather enough members to fight, peace prevails. Formally - Let $c \in (0,1)$ be an exogenously given threshold. Given an action profile $a = (a_i)_i$ that is measurable [10], a conflict takes place iff:

$$\max\{n_{E_1}(a), n_{E_2}(a)\} > c$$

where $n_{E_1}(a) = \mu(\{i \in E_1 | a_i = f\})$, $n_{E_2}(a) = \mu(\{i \in E_2 | a_i = f\})$.

Conditional on the conflict happening, probability of winning for any group is given by the following rule:

Given an action profile a, probability of E_1 being the winning group is $\dfrac{n_{E_1}(a)}{n_{E_1}(a) + n_{E_2}(a)}$.

Thus, if there is a conflict then an ethnic group wins with higher probability if more of their members fight than members of the rival group.

3.3 States of the World

At time 0, players are uncertain about the distribution of types in the world. Let ${n^y}_l$ be the fraction of y ethnicity people who are l type. For simplicity, we assume that there are only two kinds of possible type distributions (This is not crucial to our results. A discussion of this assumption is presented in section 6):

Probability ω the type distribution is such that $({n^{E_1}}_G, {n^{E_2}}_G) = (q,q)$.
Probability $(1-\omega)$ the type distribution is such that $({n^{E_1}}_G, {n^{E_2}}_G) = (r,r)$

where $(1-q) < c < (1-r)$.

Thus, if (r,r) is the true distribution of G types, then the number of bad types alone is so high that conflict must happen. On the other hand, if (q,q) is the true distribution of types then conflict may not happen if all the G types choose not to fight. We will call (q,q) the good state of the world and (r,r) the bad

[10] The function induced by the action profile $a : N \to \{f, nf\}$ is measurable

state of the world. We will be interested in the outcome when the true distribution is (q,q). Mathematical details on the type space and the prior distribution on the type can be found in the appendix.

Here on, unless otherwise stated, everything is described for only the *G* type player. This is because the *B* type player is behavioural with fixed actions.

3.4 Payoffs

The payoffs to any player *i* of type *G* depends on his action, whether or not conflict takes place and whether he was part of the winning or losing side if conflict did take place. The payoffs are summarized precisely in table 1. Although the payoff matrix may look contrived, a much more general payoff matrix would also give the same results. We choose this payoff matrix to make the math simpler. Essentially, only two aspects of the payoff table are important. One, peace time payoffs $(\alpha+\delta)$ are higher than the best conflict payoff (α) for the players. Two, the payoffs are such that it always pays to fight when conflict is inevitable.

Why should peace time payoffs be larger than those obtained from victory in war? There are many costs associated with a conflict - like the loss of lives, collateral damage, an atmosphere of uncertainty, apprehension and animosity. Therefore, we feel that that our assumption of any victory being at best a pyrrhic one is not unjustified. That fighting is the best response if conflict is inevitable can arise naturally in a society where players who don't fight for their ethnicities are subsequently ostracized/punished by their own communities. This ex-post social cost may outweigh any private costs to fighting, especially since this cost may have to be suffered by not just the people who did not fight but also by their families (for possibly many generations). Other authors (example Egorov and Sonin (2014)) have justified this sort of assumption by a 'warm glow' effect a player might feel by participating along with his community in a fight against an enemy.

If a player chooses to fight and conflict does not happen then we assume that player's payoff to be negative. This may be interpreted as the cost of getting arrested for unruly behaviour in public.

In the following table, $\alpha,\beta,\gamma,\delta,\varepsilon>0$. ε is small. *CW* means conflict happens and own ethnicity wins, *CL* - conflict happens and own ethnicity loses and *NC* means no conflict occurs.

Table 1: Payoffs

	CW	*CL*	*NC*
f	α	$-\beta+\varepsilon$	$-\gamma$
nf	$-\beta$	$-\beta$	$\alpha+\delta$

3.5 Rumour

A rumour is any piece of news that everyone hears i.e. a public signal. As an example, think of the following news - A news article which declares that E_1 ethnicity people have wantonly killed some E_2 ethnicity people in a neighbouring town. The rumour is either true or false. In the current model, both the arrival of a rumour and its veracity (true or false) will be taken as an informative signal about the underlying state of the world. The arrival of the rumour is modelled as an exogenous event. We abstract away from questions of how a rumour would arise in the first place.

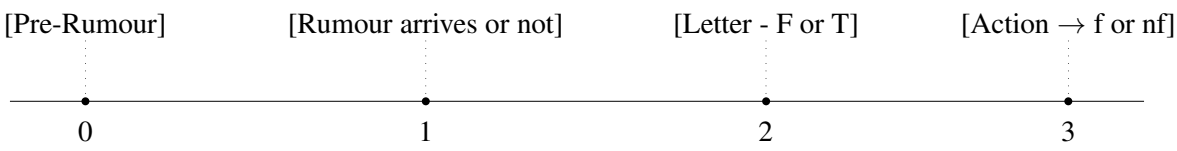
3.6 Meeting Process/Obtaining Information

After the rumour stage, people may get additional information in the following manner: If player b exists, that is - there exists an informed agent who knows if the rumour is true or false, then she randomly picks k fraction of the population and simultaneously sends them letters with one of two signals - True or False. k is a fixed constant to be interpreted as a capacity constraint. The non-strategic b player reveals her signals truthfully to every player she sends letters to. The strategic b (we shall describe this player in more detail in section 5) chooses which signal to send to each and every player. The contents of the letter serve as a signal of the state of the world. This seemingly contrived process of receiving additional information is discussed further in section 6. Our results are not crucially dependent on the above process. We make this modelling choice to make the analysis easier.

3.7 Timeline

The timeline of events is depicted in the figure 1 below. At time 0, players have priors on true distribution of types, whether a rumour arrives or not and if the rumour does arrive then whether there exists a person who will know more about the rumour. They also have priors on the contents of the letter (if it exists). They update these beliefs as events unfold. Action to fight or not will be taken after the letters stage i.e. after the rumour (where some people get the letter from b[11] and others don't).

Figure 1: Timeline



[11] If b exists.

3.8 Beliefs and Information

Any player's ethnicity $\{E_1, E_2\}$, conflict threshold c, the payoff matrix and the meeting process is common knowledge. The type $\{G, B\}$ of a player is private knowledge.

All players have common priors. We describe beliefs about the different components of the model next. These can also be studied in the game tree described in figure 2.

3.8.1 About Rumour

Conditional on the (q, q) being the true distribution, the rumour arrives with probability $1 - \theta_q$. Conditional on (r, r) being the true distribution, the rumour arrives with probability $1 - \theta_r$. Thus the arrival of a rumour is correlated with the distribution of types in the society.

3.8.2 On existence of b

Conditional on the true distribution being (q, q) and the rumour arriving, the probability that there exists one person who knows more about the rumour is given by ζ_q. Similarly define ζ_r.

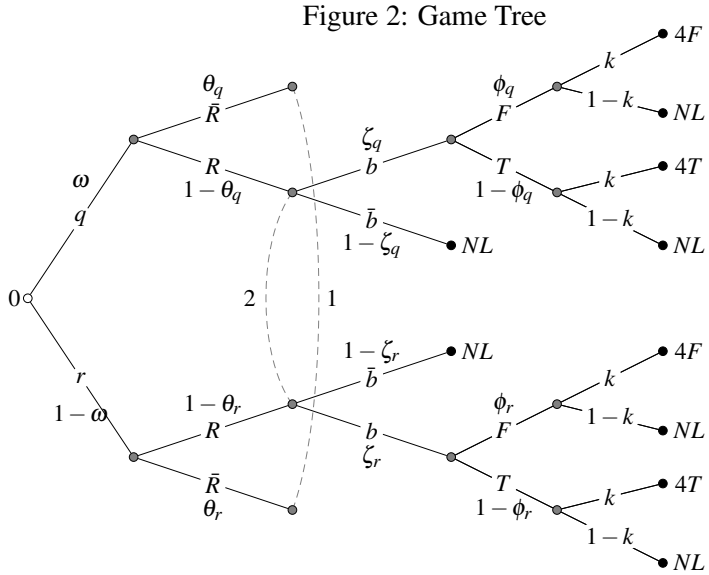
3.8.3 On Contents of Letter

Conditional on the letter arriving, the true distribution of types being (q, q) and b being non-strategic, the probability of receiving the signal F in the letter is given by ϕ_q. Similarly define ϕ_r. Note that before the letter arrived, players were in an information set where they did not know if the rumour was true or false.

3.9 Game Tree

The game as viewed by any G player is described in figure 2. This tree depicts the case of non-strategic b i.e. b always reveals her signal truthfully. $\overline{R}$ and $\overline{b}$ indicate the event - rumour does not arrive and the non-existence of a player who has additional information about the rumour respectively. F and T is when the rumour is true and false respectively. Label k implies that the player is one of the players who is randomly selected into the measure k set of people to whom agent b sends a letter. Conditional probabilities are depicted next to each event. Pre-rumour everyone is at information set 0. Post rumour everyone is at information set 2. Post letters stage, if a player does not get the letter he will be at one of the six nodes in the information set NL. If a player gets the letter from b with the signal F then he is at one of the two nodes

Figure 2: Game Tree



in information set $4F$. If a player gets the letter from b with the signal T then he is at one of the two nodes in information set $4T$.

3.10 Assumptions

1. Players play symmetric (within ethnicity) strategies only.[12]

2. $\zeta_q = \zeta_r = \zeta$. This will make sure that the getting of the letter itself is not informative about the state of the world. We believe this is a reasonable assumption. The fact that there exists someone who knows the truth about the rumour (which is what the arrival of a letter indicates) may be independent of the distribution of types in the world.

3. $\alpha + \varepsilon + 2\gamma < \beta$[13]

3.11 Equilibrium Concept

People are updating beliefs in a Bayesian manner and they choose actions which are optimal given beliefs. Thus, our equilibrium concept is Perfect Bayesian Equilibrium.

[12] Thus, people of the same type and same beliefs play the same strategies.

[13] Needed to guarantee that the no fight equilibrium is Payoff dominant before arrival of rumour.

4 Non-Strategic Informed Agent

As we mentioned before, we will consider the simple case of a non strategic informed agent (i.e. b reveals her signals truthfully to all players that she sends letters to) first. This serves as a benchmark case for the more interesting analysis of a strategic informed agent.

4.1 Results

Suppose first, that there was no informed agent in our model. When would the arrival of a rumour induce conflict? The first two propositions deal with this. Essentially, we will demonstrate conditions under which the following is true: suppose there is no informed agent in the model. Then, peace is an equilibrium outcome if the rumour does not arrive but will no longer be so if the rumour does arrive. These results are fairly simple and only depend upon the correlation of the arrival of the rumour with the bad state of the world. The second question we will answer in this section is that when can this effect of a rumour be negated by the presence of a non strategic informed agent and how is it related to k? Proposition 3 answers this question. Here we use two simple ideas to get a peaceful outcome - the superiority of the informed agent's information over the correlation of the arrival of the rumour with the bad state of the world[14] and the fact that the meeting process is common knowledge.

Note that we will only write strategies for the G type players since the B type players are behavioural and always choose f.

Proposition 1. *Pre rumour, there exists ω^* such that if $\omega > \omega^*$ then there is an equilibrium in which the G type players choose nf, thereby ensuring peace if the state is good. This equilibrium is the highest payoff equilibrium for the G type players.*

Proof. In the appendix. □

The first proposition simply says that - pre-rumour - if the priors on the distribution are such that people place high belief on the distribution with less bad types then there exists an equilibrium in which the good types do not want to fight. Note that there is also an equilibrium in which everyone fights. However, at least peace is possible if the players believe it likely the good state of the world. Note also that the peaceful equilibrium gives the G type players a higher payoff.

Proposition 2 describes the conditions under which the arrival of the rumour make peace impossible.

[14] i.e. the veracity of the rumour (true or false) is a better indication of the state of the world than it's arrival.

Proposition 2. *Post rumour and Pre-letters (information set 2), if* $\frac{1-\theta_r}{1-\theta_q} > \frac{1-\omega^*}{\omega^*}\frac{\omega}{1-\omega}$*, then not fight cannot be supported as an equilibrium. The unique equilibrium is the one in which everyone fights.*

Proof. In appendix. □

The simple idea here is that if the arrival of the rumour pushes people's beliefs towards the bad state of the world (arrival of the rumour is sufficiently correlated with the bad state of the world) then conflict becomes inevitable. This is because conflict is inevitable in the bad state of the world and fighting is best response if conflict is going to happen for sure. So a high enough belief about the state being bad will precipitate conflict.

Corollary 1. *If* $\omega > \omega^*$ *and* $\frac{1-\theta_r}{1-\theta_q} > \frac{1-\omega^*}{\omega^*}\frac{\omega}{1-\omega}$ *but no rumour appears then we can still support the equilibrium in which the good types don't fight.*

The above corollary describes one of the equilibrium outcomes when no rumour arrives. Since the arrival of the rumour is more likely when the true distribution of types is the bad one, the rumour not arriving is more likely when the true distribution of types is the good one. This implies that the posterior on (q,q) if the rumour does not arrive is higher than ω and therefore higher than ω^*.

Proposition 1, proposition 2 and corollary 1 establish conditions under which - before the rumour arrives (or if it does not arrive), G type players may not have chosen to fight but after the rumour arrived everyone chooses to fight and conflict is inevitable. So, if the true distribution of types is (q,q), conflict may not have occurred pre-rumour but it becomes inevitable post rumour. Thus, rumours may *induce* conflict. Proposition 3 talks about when this effect of a rumour may be reversed.

Proposition 3. *Let* $\omega > \omega^*$ *and* $\frac{1-\theta_r}{1-\theta_q} > \frac{1-\omega^*}{\omega^*}\frac{\omega}{1-\omega}$*. Assume* $\frac{\phi_q}{\phi_r} \geq \frac{(1-\omega)(1-\theta_r)\omega*}{\omega(1-\theta_q)(1-\omega^*)} > 1$*. Then, post letters stage, if* $k \approx 1$*, no conflict can be an outcome of an equilibrium. If* $k \approx 0$*, conflict is inevitable.*

Proof. In Appendix. □

The intuition behind proposition 3 is as follows: First note that the condition $\frac{\phi_q}{\phi_r} \geq \frac{(1-\omega)(1-\theta_r)\omega*}{\omega(1-\theta_q)(1-\omega^*)} > 1$ implies that b's information is closely related to the state of the world. That is, whether the rumour is true or false is a very good indicator of the state of the world. The strategy for a player after the letters stage is a triple - (x,y,z) where x gives the action to take if the player does not get a letter, y gives the action for when the player gets a letter and the letter has the signal - 'False' and z describes the action to be taken if the player gets a letter and the signal is - 'True'. Consider the case when the rumour is actually false. This

means that whoever gets the letter - gets the signal -'False' (since the informed agent is non-strategic in this section). If $k \approx 0$, only a very small fraction of the population gets the letter. Hence, a large fraction of the population is at information set 3 and hold the post rumour beliefs as in information set 2 (this is because $\zeta_q = \zeta_r$). Given these beliefs, it is optimal for agents to fight and conflict occurs. The people who get the letter and know that the rumour was false and realize that the state of the world is more likely to be (q,q)[15] but they also realize that too many people have not gotten the letter. Since those people are going to fight, conflict is inevitable. In this case, it is optimal for the player to play f since f is a dominant strategy if conflict is inevitable. Thus, conflict is the only equilibrium outcome if $k \approx 0$. When $k \approx 1$, the players who don't get the letter think that if b had existed they would have received the letter for sure so they conclude b does not exist. This makes them think that everyone has the post-rumour beliefs. This makes it optimal for them to choose to fight (by proposition 2). On the other hand, the people who do get the letter conclude that almost everyone must have gotten the letter which means that everyone must have realized that the rumour was false. This would mean that almost everyone places high probability (higher than ω^* under the conditions described in the proposition) on the state of the world being (q,q). Then, by proposition 1, there exists an equilibrium in which the G types don't fight. Therefore, when $k \approx 1$, if (q,q) is the true distribution of types, then conflict may not take place if the peaceful equilibrium is played. Obviously, our result does not need k to be exactly zero or exactly one. The result will hold if k is small or large enough.

5 Strategic Informed Agent

Until now, we have assumed that b is non-strategic. She knows if the rumour is true or false. She picks/meets a random selection of measure k of players and informs them correctly. In this section we will discuss the case of a strategic b player. First, we need to describe the utility function/payoffs for this player to know how she makes her decisions.

Before describing the utility function for b, we want to point out that we will take as given that a rumour has arrived which has resulted in the inevitability of conflict in the absence of further information. Like all other players, let b have an ethnicity from the set $\{E_1, E_2\}$.[16] In this section, we will assume that if b exists then she has ethnicity E_1. This is common knowledge.[17] Since the b player is not part of the population, she only gets the payoffs from outcomes. Formally, her payoffs (we consider other payoff types

[15] Since the signal F is much more likely in the state (q,q) according to the condition on ϕ_q, ϕ_r.

[16] However, we maintain the assumption that b is outside the population and does not herself fight or not-fight in the conflict. This is just for simplicity of calculations. The same results will go through if b is thought to be a player in the population.

[17] This can be justified by making the assumption that the ethnicity of b is contained in the letter. Letter sending is supposed to represent a meeting process. People know (specially if the player b is one of the players in the population) or can guess the ethnicity of others by observing the name, clothes, look of the person.

for b in section 5.5) are described in table 2.

Table 2: Payoffs for Strategic b

CW	CL	NC
α	$-\beta$	$\alpha+\delta$

Where $\alpha, \beta, \gamma, \delta > 0$. CW means conflict and win, CL - conflict and lose and NC means no conflict occurs. Thus, b gets maximum payoff (we relax this assumption in section 5.5) if conflict does not happen. However, if conflict does happen then she would like her own ethnicity to win.

Next we describe the slightly modified time line of events for the case when b is strategic. After the rumour arrives, if person b exists, then the following happens. b knows whether the rumour is true (T) or false (F). She picks a random selection of measure k of players from the population. For each player i in this selection she sends one of three messages :

- A letter stating that the rumour is false (LF).
- A letter stating that the rumour is true (LT).
- No letter (NL)[18]

From the players point of view, after a rumour arrives, they can be in one of three information sets. 1) They don't get a letter NL, 2) They get a letter with the signal LT and 3) They get a letter with the signal LF.

5.1 Strategies

We focus only on strategies of b that are symmetric within ethnicities. This does not mean that all people of the same ethnicity will receive the same message in equilibrium. This is because we allow b to play mixed strategies. The strategy for b is a function of the ethnicity of the receiving player and the rumour being actually true or false. Her action choices are - No Letter, Letter with signal T, Letter with signal F. This is described formally below:

$$f_b : \{E_1, E_2\} \times \{T, F\} \rightarrow \Delta\{NL, LT, LF\}$$

For simplicity, we will denote letters sent to the opposite ethnicity (opposite from b, i.e. E_2 ethnicity) as NL^d, LT^d, LF^d where superscript 'd' stands for different and same ethnicity as NL^s, LT^s, LF^s (superscript 's' is for same). NL^d refers to a member of the opposite ethnicity receiving no letter, LT^d refers to a

[18]Receiving no letter is an informative signal about the state since it's in the strategy space for b.

member of the opposite ethnicity receiving the signal "rumour is true" from b and LF^d refers to a member of the opposite ethnicity receiving the signal "rumour is false" from b. A similar interpretation follows for NL^s, LT^s, LF^s. Since b has been assumed to be of ethnicity E_1 if she exists, her strategy can be described as a function:

$$f_b : \{E_1\} \times \{T, F\} \to \Delta\{NL^s, LT^s, LF^s\}$$
$$f_b : \{E_2\} \times \{T, F\} \to \Delta\{NL^d, LT^d, LF^d\}$$

Strategy for any player i of the population is a function from his information set to the action set $\Delta\{f, nf\}$.

For ethnicity E_1 :

$$g^{E_1} : \{NL^s, LT^s, LF^s\} \to \Delta\{f, nf\}$$

For ethnicity E_2 :

$$g^{E_2} : \{NL^d, LT^d, LF^d\} \to \Delta\{f, nf\}$$

We will look for equilibria where the strategy of players is symmetric within ethnicity.

5.2 Equilibrium

In this section, we investigate the nature of equilibria in this model. We shall focus on two extreme cases (as before) $k \approx 0$ and $k \approx 1$. Note that any strategy profile where all agents choose the action f for any signal they receive, constitutes an equilibrium. We will call this the *all-fight equilibrium*. We want to investigate if there are other equilibria in which there is a positive probability of conflict being averted.

It is easy to show that when $k \approx 0$, all fight equilibrium is the unique equilibrium. This is because b can influence the beliefs of only a small fraction of players. So even when a player has been informed that the rumour is false, he can reason in the following way: *The rumour is false but k is low. So very few other players will be at the same information set as myself. In fact, most players will not get any further information because of low k. We know that in the absence of further information, players will fight as their beliefs about the state being good is really low after the arrival of the rumour (this result follows*

from proposition 2). Thus, playing fight is the unique best response as conflict is inevitable if such a large fraction of players will fight.

We will henceforth assume that $k = 1$.[19]

The first lemma we prove gives an important intuition. If the probability that an informed agent exists is low, then the informed agent can leverage this to her advantage. The intuition comes from the beliefs held by the players and their response to receiving no signal. If the probability that the informed agent exists is low, then the players will attribute not receiving a signal to this fact rather than assuming that it comes from a strategic move of the informed agent. Lemma 1 shows that the best response to receiving no signal is to play fight. Thus, the informed agent can induce agents to play f by not sending them any signal.

Lemma 1. *There exists $\bar{\zeta}$ such that if $\zeta \leq \bar{\zeta}$ then in any symmetric equilibrium, playing f is the unique best response on receiving the signal NL.*

Proof. In appendix. □

Thus, if the probability that b exists is low enough, in any equilibrium, players have to play f as a response to the signal NL. This lemma will be important because it provides the informed agent with a signal to which the response is guaranteed to be f. If she believes that the state is likely to be bad, then she can use this signal to ensure that all players from her ethnicity fight. We will use this property in the results that follow. Note that the probability of b existing being low is not unusual. Often rumours are about events that happen in far off places. For example, ethnic conflicts in the wake of the Babri Masjid demolition (in Ayodhya, India) led to many rumours about what was truly happening in Ayodhya. Such rumours even pervaded places which were fairly distant from Ayodhya. Obviously, it would be a low probability event that someone in those places actually knows if the rumours were true or false.

5.2.1 Truth telling

In this section, we argue that when b's private signal is highly informative, there cannot exist a symmetric (within ethnicities) strategy profile (apart from playing fight to all signals) where she fully reveals her private information. We express this observation as a proposition below:

Proposition 4. *There exists a $\bar{\sigma}, \bar{\zeta}$ such that if $\frac{\phi_q}{\phi_r} > \bar{\sigma}$ and $\zeta < \bar{\zeta}$, then there cannot be any symmetric*

[19] Equality is for simplicity. The same results will go through if k was close to 1 but not equal to it.

equilibrium (different from all-fight) where b's strategy is truth telling i.e. b's strategy is:

$$f_b(E_1,T) = LT^s$$

$$f_b(E_2,T) = LT^d$$

$$f_b(E_1,F) = LF^s$$

$$f_b(E_2,F) = LF^d$$

Proof. In the appendix. □

The utility function of player b is such that she prefers peace over conflict, but if there is a conflict she wants her ethnicity to win. Player b has an informational advantage over the rest of the players in the game. She knows whether the rumour is actually true or false. The only uncertainty she has, is over the state of the world. Suppose player b's signals are very informative i.e. the state is very likely to be good when the rumour is false and very likely to be bad when the rumour is true. Then, in a truth telling equilibrium, players will respond to the signal F by not fighting. However if the rumour is true, player b would like her ethnicity to win. If the players of the opposite ethnicity respond to the signal F by playing nf then player b would be tempted to lie and send this signal to players of the opposite ethnicity when the rumour is true. This means b would try to use her informational advantage to manipulate the beliefs of the players of the opposite ethnicity so that they do not fight in the conflict. This is a deviation from the truth telling equilibrium. Therefore, there does not exist a truth telling equilibrium.

Thus, the play of the non-strategic informed agent is not equilibrium play when b is strategic.

5.2.2 Non-truth telling

When b's signals are very informative and the probability of existence of b is low, we show that a perfect Bayesian equilibrium exists in the class of strategies given below. This class of strategies has the following desirable property: b can successfully avoid conflict when she is almost sure that the state of the world is good. In this case, b lies to a fraction of her own ethnicity to hedge against the risk of being wrong. If she believes that the state of the world is likely to be bad, she is able to prevent a sufficient mass of the opposite ethnicity from engaging in conflict, thereby providing her own ethnicity with an advantage. Moreover, we show that any equilibrium outcome of the non-truth telling case is in one of the following three. One, all players play f and conflict is inevitable. Two, it is an equilibrium whose outcome can be generated by the class of strategies given below or three, it is an equilibrium where players of opposite ethnicity play pure strategy nf to all signals sent to them along the equilibrium path. This last equilibrium will be described

immediately after the following equilibrium. Consider the following strategy profile.

$$\textbf{b's strategy}: \tag{1}$$

$$f_b(E_1,T) = NL^s$$
$$f_b(E_2,T) = q_b^T LT^d + (1-q_b^T)LF^d$$

$$f_b(E_1,F) = zNL^s + (1-z)(q_b^F LT^s + (1-q_b^F)LF^s)$$
$$f_b(E_2,F) = q_b^F LT^d + (1-q_b^F)LF^d$$

$$where\ zq + (1-q) = c$$

Player's strategies

$$E_1\ ethnicity/Same\ ethnicity$$

$$g^{E_1}(NL^s) = f$$
$$g^{E_1}(LF^s) = nf$$
$$g^{E_1}(LT^s) = nf$$

$$E_2\ ethnicity/Opposite\ ethnicity$$

$$g^{E_2}(Nl^s) = f$$
$$g^{E_2}(LF^d) = p_d f + (1-p_d)nf$$
$$g^{E_2}(LT^d) = p_d f + (1-p_d)nf$$

$$where\ 0 < p_d \leq z,\ q_b^F = q_b^T \in [0,1]$$

Before proceeding to the next proposition which establishes the above strategy profile as an equilibrium for a unique $p_d \in (0,z]$, we make the following parametric restriction on the likelihood of rumour arrival i.e. (θ_q, θ_r). We assume that these probabilities are such that post-rumour, if no conflict happens in the good state and everyone from the other ethnicity fights in the bad state[20] then an agent would:

1. Choose f if a fraction z or more of good types from his own ethnicity choose f.

2. Choose nf if all good types from his own ethnicity choose nf.

The above condition can be interpreted as follows : A fraction z is enough to guarantee a probability of winning high enough to induce f. Given that nobody from the same ethnicity fights, the probability of winning in conflict is low. This makes playing nf optimal. This is because the probability of getting a large peace time payoff is high enough to make playing nf better than playing f and having a small chance of

[20] Shown to hold in equilibrium strategy in proposition 5. This is not surprising, given b's bias.

winning. These conditions can be written explicitly as follows:

$$\omega'(-\gamma)+(1-\omega')[\frac{z+(1-r)}{z+r+2(1-r)}(\alpha)+\frac{r+(1-r)}{z+r+2(1-r)}(-\beta+\varepsilon)] > \omega'(\alpha+\delta)+(1-\omega')(-\beta) \tag{2}$$

$$\omega'(-\gamma)+(1-\omega')[\frac{(1-r)}{r+2(1-r)}(\alpha)+\frac{r+(1-r)}{r+2(1-r)}(-\beta+\varepsilon)] < \omega'(\alpha+\delta)+(1-\omega')(-\beta) \tag{3}$$

Where ω' is the post-rumour belief that the state is good. Note that the first condition immediately implies Proposition 2. These conditions make sure while the arrival of the rumour pushed beliefs towards the bad state of the world, it did not push the beliefs too much. If the people are almost convinced that it is the bad state of the world after the rumour arrives then we may not be able to get our results. We will assume these conditions to hold from here on. We now establish our next result:

Proposition 5. *There exists $\bar{\sigma} > 0, \bar{\zeta} > 0$ such that if $\infty > \frac{\phi_q}{\phi_r} \geq \bar{\sigma}$ and $\zeta \leq \bar{\zeta}$, then there exists a perfect Bayesian equilibrium in the class of strategies described above for a unique p_d.*

Proof. In appendix. □

Corollary 2. *In any equilibrium of the form described above, it is necessarily the case that $q_B^T = q_B^F$. As a result, the informed agent never sends an informative signal to the opposite ethnicity players about the state of the world.*

The intuitive idea behind this equilibrium is the following. Suppose player $b's$ signals are very informative i.e. the state is very likely to be good when the rumour is false and very likely to be bad when the rumour is true. When the rumour is false, player b would like peace to prevail. When the rumour is true, player b thinks that conflict is inevitable. In this case she wants her own ethnicity to win. Suppose, case 1 - there was a signal (LT^d or LF^d) for which, in equilibrium, the opposite ethnicity played nf with a higher probability than when they received the other signal. Then, player b would always send this signal to the opposite ethnicity. This makes this signal uninformative and essentially means that players of the opposite ethnicity will take the same action regardless of whether the rumour is true or false. Case 2 - opposite ethnicity players take the same action to both signals (LT^d or LF^d). Therefore, in any equilibrium which is not the all-fight equilibrium, players of the opposite ethnicity must play the same strategy irrespective of rumour being true or false. In particular if they respond in the same way to both signals, it keeps player b indifferent between the two signals. Moreover, when mixing, the players cannot be playing f with a higher probability than z. This is because in that case, conflict is inevitable which implies the best response for all players is to play f with probability 1 making it the all-fight equilibrium. b sends her own ethnicity informative signals and therefore uses her informational advantage. When the rumour is false, b believes

that the state is extremely likely to be good but hedges against the risk of mistake by sending signals which make a fraction of her own ethnicity fight (just not enough to cause conflict in the good state). When the rumour is true, b believes that conflict is inevitable. She sends signals to ensure everyone from her ethnicity chooses to fight.

There is one other equilibrium outcome possible in the game. In this outcome, players of the opposite ethnicity play pure strategy nf along equilibrium path i.e. they do not fight. The strategies for this equilibrium are described in the proposition below. This equilibrium points out that it is not necessarily the case that b sends completely uninformative signals to members of the opposite ethnicity. If the opposite ethnicity players are choosing the pure strategy nf then the informed agent is indifferent between giving no information and a very small amount of information which will still make make nf optimal for the opposite ethnicity.

Proposition 6. *There exists $\bar{\sigma} > 0, \bar{\zeta} > 0$ such that if $\infty > \frac{\phi_q}{\phi_r} \geq \bar{\sigma}$ and $\zeta \leq \bar{\zeta}$, then the following profile of strategies constitute an equilibrium :*

$$\textbf{\textit{b's strategy}}: \tag{4}$$

$$f_b(E_1, T) = NL^s$$
$$f_b(E_2, T) = q_b^T LT^d + (1 - q_b^T)LF^d$$

$$f_b(E_1, F) = zNL^s + (1-z)(q_b^F LT^s + (1 - q_b^F)LF^s)$$
$$f_b(E_2, F) = q_b^F LT^d + (1 - q_b^F)LF^d$$

$$where\ zq + (1-q) = c$$

Player's strategies

E_1 *ethnicity/Same ethnicity*

$$g^{E_1}(NL^s) = f$$
$$g^{E_1}(LF^s) = nf$$
$$g^{E_1}(LT^s) = nf$$

E_2 *ethnicity/Opposite ethnicity*

$$g^{E_2}(Nl^s) = f$$
$$g^{E_2}(LF^d) = nf$$
$$g^{E_2}(LT^d) = nf$$

$$q_b^F, q_b^T \in [0, 1]$$

Proof. In appendix. □

Corollary 3. *b's messages to the opposite ethnicity in (4) above are barely informative about the state of the world.*

Proof. By barely, we mean that b is indifferent between sending completely uninformative signals and signals which contain so little information that opposite ethnicity players still want to play pure strategy nf. The proof follows from $q_b^F \neq q_b^T$. □

5.3 Outcome Characterization

We will now show that there can only be two kinds of equilibrium outcomes in symmetric strategies (strategies are symmetric within ethnicity) when $b's$ signals are very informative (i.e. the rumour being true or false is very strongly correlated with state being (r,r) or (q,q) respectively). One outcome is due to the equilibrium in which all players choose the action f as a response to all signals. Any other equilibrium outcome can be obtained as an outcome of the equilibrium strategies described in either (1) or (4). To show this result, we will need the help of the following lemmas. Before proceeding, the following definition would be helpful.

Definition (*Equilibrium outcome*) : For any equilibrium E^*, the equilibrium outcome is defined by a pair of tuples - $\{(p_{E_1}^T, p_{E_2}^T), (p_{E_1}^F, p_{E_2}^F)\}$. T, F stands for when the rumour is actually true or false respectively. p_i^t where $t \in \{T,F\}$ and $i \in \{E_1, E_2\}$, describes the fraction of G type i ethnicity players who choose the action f in equilibrium E^* when the rumour is t.

The following lemmas will be useful in characterizing the equilibrium outcomes.

Lemma 2. *In any symmetric equilibrium different from the all fight outcome, there exists $\bar{\zeta}$ such that if $\infty > \frac{\phi_q}{\phi_r} > 0$ and $\zeta < \bar{\zeta}$ then b does not send the signal NL^d to any player of opposite ethnicity.*

Proof. Consider an equilibrium different from one in which players respond to all signals with f. Thus, there exists a signal to which opposite ethnicity players respond with a positive probability of playing nf. This is because if there is no such signal then we will have to be at the *all fight equilibrium*. By the lemma 1, this signal is different from NL^d. Thus b has a choice between sending NL^d and have the opposite ethnicity fight for sure and this signal, for which there is lower probability of fighting.

Regardless of whether the rumour is true or false or how informative her signal is, it is always weakly better for player b if players of the opposite ethnicity do not fight. In fact, if there is positive probability that the state is (r,r) then it is strictly better for player b to not send the signal NL^d to players of opposite

ethnicity and have them play f (lemma 1). The condition $\infty > \frac{\phi_q}{\phi_r} > 0$ guarantees that rumour is true or false is not perfectly informative of state. Thus, b believes that there is always a positive probability that the state is (r,r). Therefore, she will never send the signal NL^d to the opposite ethnicity. □

Lemma 3. *In any symmetric equilibrium where signals are informative ($\frac{\phi_q}{\phi_r}$ is high), all players of the same ethnicity as b will play f when the rumour is true.*

Proof. When signals are very informative, the rumour being true implies that the state is very likely to be (r,r). This implies that conflict is going to happen with a very high probability. In such a situation, player b would like all players of her ethnicity to play f to ensure maximum probability of winning. Since she has the ability to make this happen (send NL^s to all players of her ethnicity), it will happen in any equilibrium. □

We now present our desired result in proposition 7 below. Essentially, the proposition establishes conditions under which any equilibrium outcome different from the all fight one can be obtained from the strategies described in either 1 or 4. The sufficient conditions needed are the same ones we have used before - 1) b's information is sufficiently informative i.e. the state of the world is sufficiently correlated with the veracity of the rumour and 2) The probability that an informed agent exists is low.

Proposition 7. *There exists $\bar{\sigma}, \bar{\zeta} > 0$ such that if $\frac{\phi_q}{\phi_r} > \bar{\sigma}$ and $\zeta < \bar{\zeta}$ then any symmetric equilibrium outcome $O^* = \{(g,h),(j,k)\}$ (distinct from the all fight outcome) can be derived from either (1) or (4).*

Proof. Suppose outcome $O^* = (g,h,j,k)$ is different from all fight i.e. $(g,h,j,k) \neq (1,1,1,1)$. We will show that O^* can be obtained as an outcome of an equilibrium described in (1) or (4) by proving that, in any symmetric equilibrium, the following must hold:

1. $g = 1$
2. $h = k$.
3. $h \leq z$.
4. $j = z$.
5. $h = k = 0$ or $h = k = p_d$

; where z is such that $zq + (1-q) = c$

Proof for 1

$p_h^T = 1$ in any equilibrium outcome. This holds because of lemma 3 and choosing $\frac{\phi_q}{\phi_r}$ high enough.

Proof for 2

Consider any player i of ethnicity E_2. By lemma 2, he will never get the signal NL^d. There are two sub cases: Either he responds in the same way to signals T, F. In this case $h = k$. Or, he responds in a different manner to the signals T, F. In this case, it will be optimal for player b to send him the signal for which the probability of an nf response is higher. If b plays this strategy in equilibrium, then player i gets just one uninformative signal in equilibrium (since he will get the same signal regardless of whether the rumour is true or false). Thus, player i will have only one response on the equilibrium path. This would mean that regardless of whether the rumour is true or false the same fraction of E_2 players play f (since we are looking at symmetric equilibrium). Thus, $h = k$.

Proof for 3

Suppose $h > z$. This implies that conflict will always happen since the fraction of E_2 ethnicity players fighting is above the cutoff when the rumour is true and when the rumour is false. In this equilibrium, it will be optimal for player b to send some signal to all players of her ethnicity (E_1) and have them respond with f. Note that she can make this happen because she can always send them NL^s. However, players of ethnicity E_2 will realize this and play f as best response themselves. This is a contradiction. Hence $h \leq z$.

Proof for 4

Suppose $j > z$. Then, conflict happens regardless of whether the rumour is true or false. This would lead to ethnicity E_2 responding with all players playing f. However, the best response to this is $j = 1$. This implies that we have arrived at the all fight equilibrium. However, we assumed before that we are looking at equilibria different from this one. Thus, $j \leq z$.

Suppose $j < z$. Even when the rumour is false and signals are very informative, the condition $\infty > \frac{\phi_q}{\phi_r}$ guarantees that there is a small probability that the state is (r, r). Player b can be better off by sending the signal NL^s to a slightly higher fraction of E_1 ethnicity players such that the fraction of G type E_1 ethnicity players who play f becomes $j + \varepsilon$ and $j + \varepsilon < z$. b becomes better off since she avoids conflict and gets high peace payoff if state is actually (q, q) and if, despite her very informative signals, the state happens to be (r, r), she has increased the probability of her own side winning. Therefore, b will make this deviation and the fraction of good E_1 who fight in equilibrium cannot be j. This is a contradiction.

Therefore we have that $j = z$.

Proof for 5

Suppose $h \neq 0$. This implies that $h > 0$ i.e. players of the opposite ethnicity are mixing (since $h \leq z$, $h \neq 1$) when they get the signal LT^d, LF^d. However, we derived our unique p_d in the equilibrium described by strategies in (1) by using indifference conditions for players of the opposite ethnicity while players of same ethnicity were playing the same strategies ($p^T_{E_1} = 1, p^F_{E_1} = z$). Therefore, h must be the same as that p_d.

□

Thus, we have characterized all symmetric equilibrium outcomes.

5.4 High probability of b's existence

Up till this point, all results for the strategic b section have been proved under the assumption that the probability of b existing is low. We believe that this assumption is a reasonable one. Rumours are usually started about events that most people know very little about so the existence of an informed agent being low is not unusual. Additionally, notice that even if every event has some informed agents (as we argue in the beginning), these agents may not necessarily be there to provide information in every village. For example, suppose there are rumours in village B that some players of the an ethnicity are being murdered in city A[21]. We think it is reasonable to assume that while there may exist someone in village B who knows if the rumour is true or false, the probability of such an informed agent existing must be low.

Having said that, in this subsection we show that this assumption is not a necessary for our results. Similar results can be obtained under some other parametric conditions. ζ being low guaranteed player b a signal (NL) such that any player who received this signal would choose fight as a best response in any symmetric equilibrium. When b is convinced that the state is very likely to be bad, she uses this signal to make sure that all players of her own ethnicity chose fight and when she thinks that the state is probably good, she used this signal to hedge against the possibility of being wrong. If the probability of existence of b is high then the signal NL may lose it's power. We show that in this case, b can enjoy the same advantages if z (the maximum fraction of good type players from an ethnicity who can play f without inducing conflict.) is low. This is expressed as the following result:

Proposition 8. *Let $\zeta \approx 1$. There exists a $\bar{z}, \bar{\sigma}$ such that if $z < \bar{z}$ and $\infty > \frac{\phi_g}{\phi_r} > \bar{\sigma}$, then there exists an equilibrium which produces the same equilibrium outcome as either* (1) *or* (4).

Proof. Proof can be found in the appendix. □

[21] Rumours about big riots often spread to nearby villages.

The intuition here is very similar to intuition used before. In our analyses before, the players who did not receive a signal chose to fight because they interpreted this to be a strong signal that b did not exist and therefore inferred that no one would have got additional signals which will cause every one else to fight, thereby making fighting optimal. When ζ is low, we propose a strategy in which the players of the same ethnicity always get the signal LT^s if the rumour is true and the state is likely to be bad. So they would like to play fight in response. b also sends this signal to a fraction z of her ethnicity when the rumour is false and the state is likely to be good. This is to induce a fraction z to fight so that her ethnicity has a higher chance of winning in case her information turned out to be an incorrect signal of the state. So why is it that the signal LT^s induces players to fight (just like the signal NL^s before)? Since z is small, the players who receive the signal LT^s infer that this signal is more likely to arrive if the state is bad and therefore choose to fight.

5.5 Other payoff types for b

For the strategic b section we have used a utility specification which describes the informed agent as peace-loving player with a bias towards her own ethnicity. In this subsection we show that this specification is not necessary for our results to go through. In particular, even if player b prefers conflict where her own ethnicity wins to peace, and prefers peace to a conflict where her own ethnicity loses, we can still get the same equilibria as before. The only condition we need is that payoff from peace be above a cut off for player b.

Let $u_b(CW)$ be the payoff to player b if conflict happens and her own ethnicity wins. Similarly define $u_b(CL), u_b(NC)$. Our results up to this point have a used a utility specification where $u_b(NC) > u_b(CW) > u_b(CL)$. Consider the following payoff type for agent b.

5.5.1 Extremist type

This preference type satisfies :

$$u_b(CW) > u_b(NC) > u_b(CL)$$

This payoff type is interpreted as follows : agent b resembles the mindset and payoff specification of an extremist who would prefer conflict to peace but only as long her own ethnicity wins. The following result provides equilibrium possibilities with this payoff specification:

Proposition 9. *Let p_1, p_2 be the probabilities of winning for b when the rumour is true in equilibria* $(1),(4)$ *respectively. Assume conditions under which* $(1),(4)$ *were equilibria before. Then the following hold:*

1. *If $p_2 u_b(CW) + (1-p_2)u_b(CL) < u_b(NC)$, then both* (1) *and* (4) *are equilibria.*
2. *If $p_1 u_b(CW) + (1-p_1)u_b(CL) < u_b(NC) < p_2 u_b(CW) + (1-p_2)u_b(CL)$, then* (1) *is an equilibrium but* (4) *is not.*
3. *If $u_b(NC) < p_1 u_b(CW) + (1-p_1)u_b(CL)$ then all fight is the unique equilibrium.*

Proof. Proof can be found in the appendix. □

Thus, if the informed agent finds little gain from maintaining peace, conflict may never be avoided. Else, there are equilibria where the outcome is peace.

6 Discussion

In this section we discuss some of our assumptions and modelling choices. We show that our claims are robust to some alterations.

6.1 Correlation of Distributions

We assume that only those type distributions are possible which lead to people placing positive weights on (q,q) and (r,r) where $(1-r) > c > (1-q)$. This assumption is not crucial to our results. In particular we could have assumed positive weights on a multitude of distribution states like $(q_1,q_2),...,(q_n,q_{n+1}),(r_1,r_2),...,(r_m,r_{m+1})$ where $max\{1-q_i\}_i < c < min\{1-r_j\}_j$. As long as conflict is inevitable in some states and not in others, our claims will go through. Note that we could allow for beliefs over distributions like (q,r) where $(1-r) > c > (1-q)$ as well but these would be uninteresting (if we allowed for such distributions only) since our definition of conflict makes conflict inevitable if even one ethnicity has enough bad types.

6.2 Using Letters to Represent Meetings

We have used a contrived definition of 'meetings' to say how players find out the truth about the rumour. However, note that our results depend on three things - the fraction of the population who meet b, the meeting process being common knowledge and that only a signal is exchanged in meetings (types are not

revealed). Thus, any meeting process which guarantees that only k fraction will meet b will give the same results. In the case of non-strategic b, we could have the following information dissemination process - b could tell just one person and then that one person may meet others randomly and inform them and then all those people could inform others and so on. If there are finite meeting stages, such that at the end of all meetings only k fraction of players know that the rumour was false, then our claims would go through.

6.3 Non Random Meetings

Realistically, it is more likely that people of the same ethnicity are more likely to meet each other. This spells trouble. Consider an extreme example where player b sends letters to only her own ethnicity. Conflict may be unavoidable now. This is because the other ethnicity does not learn that the rumour is false and most will believe that b does not exist (because ζ is low) which will cause them to fight. This is enough for conflict to occur. This example can be extended to a situation where the two ethnicities seldom meet in the meeting stage. Thus, low levels of inter-ethnic integration/communication may lead to one ethnic group not finding the truth. This would lead to a higher probability of conflict. The inverse relationship between inter-ethnic relationships and ethnic conflicts has been explored in Varshney (2003) and Dutta (2014) among others.

6.4 Uncertainty about b's ethnicity

In the case of strategic b we have assumed that $b's$ ethnicity is common knowledge. Suppose this was not true and that players of each ethnicity had a belief about $b's$ ethnicity. Let belief of players of ethnicity e about player b being of ethnicity E be given by P_E^e. The ideal situation for player b would be the case where $P_e^e \approx 1 \ \forall e \in \{E_1, E_2\}$ i.e. players of both ethnicities believe that player b is of their own ethnicity. In this case, both ethnicity players will believe that b will send them the right signals and will send the wrong signals to players of the other ethnicity. Thus, here, b can guarantee peace when she believes peace is possible and can guarantee success for her own ethnicity when she believes that conflict is inevitable.

6.5 Limiting k only

Throughout the paper we have focussed on two limiting cases of k. Of course, our results are not severely dependent on this. Any result which works when k is 1 will also work when k is high enough but not 1. A similar cut off argument will hold for $k = 0$. The qualitative results will not change. However, if k is nether too high nor too low, then we might need further conditions on the parameters to make any claims.

6.6 Urban Ethnic Conflicts

In this paper we show that regardless of b's preferences (strategic or non strategic), if she can meet only a small fraction of the population (k is small), conflict is inevitable. This is not true if k is large. This points to a possible explanation for the following empirical observation - about 70 percent of all Hindu-Muslim conflicts (and more than 96 percent of deaths in these conflicts) in India between the years 1950 and 1995 have been reported in urban areas - Varshney (2003), Mitra and Ray (2014). This is extremely surprising since over 70 percent of Indians live in rural areas.[22] Presumably, some of this can be explained by the under-reporting of rural conflicts, larger population density in urban areas, some villages having just one ethnicity etc. However, we believe that these may not account for the statistics completely. For example, Varshney (Varshney, 2003) writes that under-reporting of rural deaths would have to be on the scale of $15-20$ times to explain the fact that rural deaths were less than 4 percent of all casualties in ethnic conflicts.

The reason we believe that our results could contribute to explaining such data is that we think of urban areas as areas where the k is small whereas villages are more likely to be areas where $k \approx 1$. This is because rural areas generally have smaller populations where an informed agent may be able to meet a large fraction of the population whereas it is impossible to meet more than a small fraction of a large urban population. Note that it may be easier to transmit information in urban areas because of better technology. However, our results rest on the fact that a *fraction* of population who become informed. This may still be very low in high population urban areas. Moreover, it is likely that the information transmission/meeting process is such that in urban areas most people get to know of the information through second/third hand sources as opposed to rural areas where people hear the new information first hand from b. This may lead to a lower fraction of the population getting good information in urban areas resulting in effects similar to the ones described here.

6.7 Cost of participation

In our model we have assumed that fighting is optimal if conflict is inevitable and we have justified this assumption by pointing out that future social costs may outweigh private fighting costs. Consider now the game where the private fighting cost is larger. Obviously, our results will go through if this cost is really small. If not, people have a trade off - if they fight, then they have a chance of winning and getting α,

[22] Census 2011. Moreover, in the period of $1950-1995$ (for which we have Hindu-Muslim conflict data - Varshney and Wilkinson (2006)), an even larger fraction of the Indian population must have lived in rural areas.

however they have to incur a cost. Suppose conflict is going to happen for sure. If private cost is large then people may not fight. Realistically, this would make playing nf more likely in urban areas since the cost of fighting in a conflict may be larger in urban areas. The cost of fighting may be higher in urban areas because of two reasons - one, it is easier to hurt people you don't know. Two, it may be that people are more violent when they are confident of their anonymity. Now, if nf is more likely in urban areas then this undermines our explanation of higher levels of ethnic conflicts in urban areas. However, note that we have also assumed that gains from winning remains the same always. The more natural assumption would have been that the pie is larger in urban areas. If the increase in gains compensate for the private cost then we will still get the same results.

7 Conclusion

This paper explores the role of informed agents in influencing the impact of rumours on ethnic conflicts. We characterize symmetric equilibria when the informed agent (b) is non-strategic as well as when b is strategic and has a bias towards one ethnicity. We show that there are peaceful equilibria even when b is known to be biased towards one ethnicity.

We present a discussion of our modelling choices and assumptions in section 6 which point to several possibilities which can still be explored. This paper is a step towards understanding the role of informed players in preventing conflicts which are precipitated by rumours. There can be very interesting extensions of this paper. One can look at a repeated environment where a rumour arrives every period and one player may or may not know more about it. It will be useful to understand the dynamics in such an environment. We could also look at an environment where b can choose the portfolio of the people she meets i.e. given her capacity constraint, she can choose exactly what fraction of the players she meets are from either community. In such an environment it will be interesting to look at the optimal portfolio choice and the equilibrium strategies. One of our shortcomings is that we have not made any arguments to say when the peaceful equilibria will be selected over the all fight one. There are many such important and interesting questions which we hope to investigate in the future.

References

Berenschot, W. (2011). The spatial distribution of riots: Patronage and the instigation of communal violence in gujarat, india. *World Development*, 39(2):221–230.

Bhavnani, R., Findley, M. G., and Kuklinski, J. H. (2009). Rumor dynamics in ethnic violence. *The Journal of Politics*, 71(03):876–892.

Brass, P. R. (2011). *The production of Hindu-Muslim violence in contemporary India*. University of Washington Press.

Carlsson, H. and Van Damme, E. (1993). Global games and equilibrium selection. *Econometrica: Journal of the Econometric Society*, pages 989–1018.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451.

Diamond, D. W. (1985). Optimal release of information by firms. *The Journal of Finance*, 40(4):1071–1094.

Dutta, S. (2014). Ethnic conflict and civic engagememt. *PhD Thesis*.

Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of accounting research*, pages 123–145.

Edmond, C. (2013). Information manipulation, coordination, and regime change. *The Review of Economic Studies*, page rdt020.

Egorov, G. and Sonin, K. (2014). Incumbency advantage in non-democracies. Technical report, National Bureau of Economic Research.

Enikolopov, R., Petrova, M., and Zhuravskaya, E. (2011). Media and political persuasion: Evidence from russia. *The American Economic Review*, pages 3253–3285.

Esteban, J. and Ray, D. (2008). On the salience of ethnic conflict. *The American Economic Review*, 98(5):2185–2202.

Esteban, J. and Ray, D. (2011). A model of ethnic conflict. *Journal of the European Economic Association*, 9(3):496–521.

Farrell, J. and Gibbons, R. (1989). Cheap talk with two audiences. *The American Economic Review*, 79(5):1214–1223.

Fearon, J. D. and Laitin, D. D. (1996). Explaining interethnic cooperation. *American political science review*, 90(04):715–735.

Goltsman, M. and Pavlov, G. (2011). How to talk to multiple audiences. *Games and Economic Behavior*, 72(1):100–122.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of law and economics*, pages 461–483.

Horowitz, D. L. (1985). *Ethnic groups in conflict*. Univ of California Press.

Judd, K. L. (1985). The law of large numbers with a continuum of iid random variables. *Journal of Economic theory*, 35(1):19–25.

Jung, W.-O. and Kwon, Y. K. (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research*, pages 146–153.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Kosfeld, M. (2005). Rumours and markets. *Journal of Mathematical Economics*, 41(6):646–664.

Larson, J. M. Interethnic conflict, incendiary rumors, and the networks that help or hurt.

Lu, Y., Suen, W., and Chen, H. (2013). The power of whispers: A theory of rumor, communication and revolution. In *2013 Meeting Papers*, number 411. Society for Economic Dynamics.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391.

Mitra, A. and Ray, D. (2014). Implications of an economic theory of conflict: Hindu muslim violence in india. *Journal of Political Economy*, pages 719–765.

Morris, S. and Shin, H. S. (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review*, pages 587–597.

Morris, S. and Shin, H. S. (2001). Global games: theory and applications.

Morris, S. and Shin, H. S. (2002). Social value of public information. *The American Economic Review*, 92(5):1521–1534.

Petrova, M. (2008). A formal theory of public opinion in conflicts. *Available at SSRN 1149367*.

Rauchhaus, R. W. (2006). Asymmetric information, mediation, and conflict management. *World Politics*, 58(02):207–241.

Shadmehr, M. and Bernhardt, D. (2015). State censorship. *American Economic Journal: Microeconomics*, 7(2):280–307.

Tyson, S. and Smith, A. (2014). Two-sided coordination: Competition in collective action. *Available at SSRN 2159385*.

Vaitla, S. (2011). *Preventing ethnic violence with local capacities: lessons from civil society in India*. PhD thesis, Rutgers University-Graduate School-Newark.

Varshney, A. (2003). *Ethnic conflict and civic life: Hindus and Muslims in India*. Yale University Press.

Varshney, A. and Wilkinson, S. (2006). Varshney-wilkinson dataset on hindu-muslim violence in india, 1950-1995, version 2.

Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the rwandan genocide. *The Quarterly Journal of Economics*, 129(4):1947–1994.

A Appendix

A.1 Type space and prior

Denote as $T = \{G,B\}^N$ the set of all type profiles. Define $T_q = \{t : \mu(\{i \in E_1 : t_i = G\}) = \mu(\{i \in E_2 : t_i = G\}) = q\}$ and similarly define T_r. We endow T with the appropriate sigma algebra such that the sets of the form T_q and T_r are measurable and we assume that the prior $p \in \Delta(T)$ has the following properties:

1. $p(T_q \cup T_r) = 1$
2. For all $i \in N$, $p(T_q|t_i = G) = \omega \geq \omega^*$
3. $p(t_i = G|T_s) = s \; \forall \; i \in N$ and $\forall s \in \{q,r\}$

The construction of such priors has been discussed in Judd (1985). We may do so here by separately performing Judd's construction for T_q and T_r and then naturally extend the measure to the union $T_q \cup T_r$. The first condition says that the type distribution is either (q,q) or (r,r). The second condition says that when an agent learns that he is of type G, his belief about (q,q) is ω. Third, conditional on T_s, the probability of each player being a good type is s.

Proof for Proposition 1

In the pre-rumour stage (information set 0), people have beliefs ω about the good distribution (q,q) being the actual distribution. Strategies are just a function of types and beliefs. Consider the following pure strategy profile:

$S(G,\omega) = nf$

We want to show that if ω is high enough then it will be optimal for the G players to not fight, given that other G players are playing nf. Note that fight or not fight decisions are actually taken after the letters stage. Here, we ask a hypothetical question - If players were asked to make the decision at information set zero, what would they do? This is important because we want to be able to say that the rumour caused conflict i.e. conflict may not have occurred with pre-rumour beliefs but it became inevitable post rumour. Given these strategies, an arbitrary G player will make the following calculations

Payoff from playing $f = \omega(-\gamma) + (1-\omega)(\frac{\alpha-\beta+\varepsilon}{2})$

Payoff from playing $nf = \omega(\alpha+\delta) + (1-\omega)(-\beta)$

Clearly, if $\omega \geq \frac{\alpha+\beta+\varepsilon}{\alpha+\beta+\varepsilon+2(\alpha+\delta+\gamma)}$, then playing nf is best response for G player. So this strategy profile constitutes a Bayesian Nash equilibrium if $\omega \geq \omega^* = \frac{\alpha+\beta+\varepsilon}{\alpha+\beta+\varepsilon+2(\alpha+\delta+\gamma)}$.

To Show - If $\omega > \omega^*$, then this is the Payoff dominant equilibrium for the G players.

Expected payoff from this equilibrium = $\omega(\alpha+\delta)+(1-\omega)(-\beta)$.[23]

There is only one other equilibrium possible in pure strategies - an equilibrium in which both G types and B types play f.

Payoff from this all fight equilibrium = $\frac{\alpha-\beta+\varepsilon}{2}$.

It is easy to see that if $\omega > \omega^*$ and $\alpha+\varepsilon+2\gamma < \beta$ (assumption 3) , then the ex-ante expected payoff from all fight equilibrium is lower than payoff from equilibrium in which G players don't fight.

Let us check to see if there are any mixed strategy equilibria:

First we need this lemma:

Lemma : In any mixed strategy equilibrium where the G types of both ethnicities play the same strategies, the weight on playing f has to be less than or equal to $c-(1-q)$.

Proof : We will prove by contradiction. Suppose the players of any ethnicity play f with a strictly higher weight than $c-(1-q)$. Then the fraction of players playing f for that ethnicity is higher than c in any state of the world. This implies that conflict is inevitable. However, when conflict is inevitable then playing f is strictly dominant strategy. Thus, the ethnicities could not be mixing between f and nf. Contradiction.
□

Consider now the following strategy:

$$\begin{aligned} S(G,\omega) &= f \ ; \ probability \ p \\ &= nf \ ; \ probability \ (1-p) \end{aligned}$$

where $p \leq c-(1-q)$.

For mixing to be optimal, the payoff from f must be equal to the payoff from nf.

Payoff from playing f = $\omega(-\gamma)+(1-\omega)(\frac{\alpha-\beta+\varepsilon}{2})$

[23] We only consider the expected payoffs of the G type when thinking of Payoff dominance. Since the B types are always choosing to fight, clearly they are at least indifferent to the result of their actions.

Payoff from playing $nf = \omega(\alpha+\delta)+(1-\omega)(-\beta)$

If the above payoffs are the same then we have:

$$\omega = \omega^*$$

Payoff from this mixed strategy equilibrium = $\omega^*(\alpha+\delta)+(1-\omega^*)(-\beta)$

Since the ethnicities are symmetric, in any mixed strategy equilibrium, the G players of both ethnicities will play the same strategies. This is obvious from the above proof. Suppose the G players of E_1 ethnicity were playing f with probability p_1 and the G players of the other ethnicity were playing f with probability p_2 where $p_1 \neq p_2$. We can see quite easily from the above proof that a necessary condition for the players of E_1 ethnicity to mix is that $\omega = \omega_1$ and the E_2 ethnicity requires $\omega = \omega_2$ for them to mix in equilibrium where $\omega_1 \neq \omega_2$. Thus, an asymmetric mixed equilibrium is not possible.

Comparing ex ante expected payoffs in the three possible equilibria, it is obvious now that if $\omega > \omega^*$, then the equilibrium in which all G players play nf is the payoff dominant equilibrium (for the G players)[24]

Proof for Proposition 2

We will show this by demonstrating that the posterior belief on the good distribution falls below ω^* under the condition

$\frac{1-\theta_r}{1-\theta_q} > \frac{1-\omega^*}{\omega^*}\frac{\omega}{1-\omega}$

Let $P(q/Rumour)$ be the probability that the true distribution is (q,q) given that the rumour has arrived. Then:

$P(q/Rumour) = \frac{\omega(1-\theta_q)}{\omega(1-\theta_q)+(1-\omega)(1-\theta_r)}$

Then $P(q/Rumour) < \omega^*$

$\Leftrightarrow$

$\frac{\omega(1-\theta_q)}{\omega(1-\theta_q)+(1-\omega)(1-\theta_r)} < \omega^*$

$\Leftrightarrow$

$\frac{1-\theta_r}{1-\theta_q} > \frac{1-\omega^*}{\omega^*}\frac{\omega}{1-\omega}$

Next we look at equilibria possible after the letter receiving stage. The strategies at this information

[24] Since B type players are behavioural, we don't model their payoffs.

set are described first.

A.2 Strategies

The tree represents uncertainty faced by a player of good type. He may be at information set $3, 4T$ or $4F$. A strategy prescribes what action to take at each information set.

Definition : For player i, a *strategy* a function $\sigma_i : \{3, 4T, 4F\} \to \Delta\{f, nf\}$.

The strategy for player b is not described here because we discuss the case of non-strategic b first. We shall focus on symmetric strategy profiles.

A.3 Results : Non-strategic b

A.3.1 Proof of Proposition 3

We now establish Proposition 3. This will imply that if the state is good then there is a high chance of there being no conflict. We will look at pure strategy equilibria first and then we will consider the possibility of mixed strategy equilibria.

Denote as (a, b, c) the strategy $\sigma(3) = a, \sigma(4T) = b$ and $\sigma(4F) = c$ where $\{a, b, c\} \subseteq \{f, nf\}$. We now show conditions under which (f, f, nf) is an equilibrium. Let ϕ_q and ϕ_r be the probability of receiving the message F in the state q and r respectively

Proposition 10. *Suppose the following are true :*

1. $0 < \zeta_q = \zeta_r = \zeta$
2. $k \to 1$
3. $\frac{\phi_q}{\phi_r} > \frac{(1-\omega)(1-\theta_r)\omega*}{\omega(1-\theta_q)(1-\omega^*)} > 1$

then (f, f, nf) can be supported as an equilibrium and no strategy profile with $\sigma(4T) = nf$ can be supported as an equilibrium.

Proof. Since $k \to 1$, any player at information set 3 realizes that the probability that player b exists is close to zero. Also, $\zeta_q = \zeta_r$ implies that this information does not help update the belief about the state of the world. Thus, any player at information set 3 will have post rumour beliefs. Since the belief on the state of the world being (q,q) is strictly less than ω^* (because $\frac{1-\theta_r}{1-\theta_q} > \frac{1-\omega^*}{\omega^*}\frac{\omega}{1-\omega}$) at the post rumour information set (information set 2), it is optimal to play f for any symmetric strategies the other player may be following. It is sufficient to show that this is true when all other G players are playing nf. This has been proved as part of proposition 2.

Now consider the decision at information set $4F$. Under the conditions stated above $Pr(q|4F) > \omega^*$. As the fraction of players getting the letter goes to 1, we know that any player i's belief about the fraction of good types who received the same letter as i goes to 1. Let $k \to 1$. Let $E(f,4F,q)$ and $E(f,4F,r)$ be the expected payoff from playing f at information set $4F$ in state q,r respectively. Let $E(nf,4F,q)$ and $E(nf,4F,r)$ be the corresponding expected payoffs from nf. Then, the expected payoff from playing f at $4F$ is:

$$\begin{aligned}
&E(f,4F,q)p(q/4F) + E(f,4F,r)p(r/4F) \\
&= -\gamma p(q/4F) + \frac{\alpha - \beta + \varepsilon}{2} p(r/4F)
\end{aligned}$$

The expected payoff from playing NP at $4F$ is :

$$\begin{aligned}
&E(nf,4F,q)p(q/4F) + E(nf,4F,r)p(r/4F) \\
&= (\alpha + \delta)p(q/4F) + (-\beta)p(r/4F)
\end{aligned}$$

Since $\frac{\phi_q}{\phi_r} > \frac{(1-\omega)(1-\theta_r)\omega*}{\omega(1-\theta_q)(1-\omega^*)} > 1$, we have that $p(q/4F) > \omega^*$. From proposition 1, this implies that playing nf is optimal for G players.

Under the conditions, $Pr(q|4T) < \omega^*$. Therefore, a similar argument can be used to show that it optimal to play f at information set $4T$. It is also clear that no strategy with $\sigma(4T) = nf$ can be supported as an equilibrium. □

Proposition 11. *Assume the following :*

1. $k \to 0$
2. $\frac{(1-\omega)(1-\theta_r)\omega*}{\omega(1-\theta_q)(1-\omega^*)} > 1$

Then, (f,f,f) *is the unique symmetric equilibrium*

Proof. We first show that no profile of the form (f,l,nf) or (f,nf,b). Since the events - letter has signal T and letter has signal F are mutually exclusive, we put the letter 'l' in place of the non-relevant action choice. Playing f is the only optimal strategy at information set 3 in any symmetric equilibrium. This follows from proposition 2 and the fact that strategies are symmetric across ethnicities in any equilibrium. Now, we have $k \to 0$. Hence, for low values for k, even if the letters are distributed, conflict will inevitably take place since a large fraction of the community (of proportion greater than c) will not have received a letter and will be in information set 3 and will choose to play f. Hence, agents who do receive the letter would know that a conflict will take place irrespective of the state of the world and would choose to play f since it is a dominant strategy under conflict.

Additionally, we know that (f,f,f) can always be supported as an equilibrium and the above argument establishes it as a unique equilibrium. □

A.3.2 Mixed Strategy Equilibria

First we will consider the case of $k \to 1$ and identify necessary and sufficient conditions for there to be a mixed strategy equilibrium.

Proposition 12. *Let* $\hat{\sigma} = \frac{(1-\omega)(1-\theta_r)\omega*}{\omega(1-\theta_q)(1-\omega^*)}$. *If* $\frac{\phi_q}{\phi_r} = \hat{\sigma}$ *or* $\frac{1-\phi_q}{1-\phi_r} = \hat{\sigma}$ *then a mixed strategy equilibrium (with strict randomisation at atleast one message) exists :*

Strategy for each player :

- *NL :* $p_l f + (1-p_l)nf$
- *LT :* $p_t f + (1-p_t)nf$
- *LF :* $p_f f + (1-p_f)nf$

If not, then there are no mixed strategy equilibria.

Proof. Note that since b is giving the same information to players of both ethnicities (since b is non strategic), they will have the same beliefs if they are at the same information set. This, and the fact that both ethnicities have equal populations imply that it is not possible to have an asymmetric equilibrium where players of different ethnicities react differently to a signal. Therefore, we concentrate only on symmetric strategies.

First consider the case where $\frac{\phi_q}{\phi_r} = \hat{\sigma}$. We show that an equilibrium of the above nature exists with $p_l = 1$, $p_t = 1$ and any $p_f \in (0, \frac{c+q-1}{q}]$. Moreover, there is no other mixed strategy equilibrium. Note first that if, for any signal, the response is to fight with probability more than $\frac{c+q-1}{q}$, then the best response is to play f with probability one and the best response to this is to play pure strategy f. Therefore, for a mixed strategy equilibrium to exist, the response to atleast one signal must be to play f with probability less than $\frac{c+q-1}{q}$.

Given a currently held belief $\omega' \in [0,1]$ and fraction $p \in [0, \frac{c+q-1}{q}]$ of good types in each ethnicity playing f , it shall be convenient to define the following values :

- Payoff of an arbitrary player from playing f

$$E_f^{p,\omega'} = \omega'(-\gamma) + (1-\omega')(\tfrac{\alpha-\beta+\varepsilon}{2})$$

- Payoff of an arbitrary player from playing nf :

$$E_{nf}^{p,\omega'} = \omega'(\alpha+\delta) + (1-\omega')(-\beta)$$

From Proposition 1, we know that $E_f^{0,\omega^*} = E_{nf}^{0,\omega^*}$. It is also clear that $E_f^{p,\omega'} = E_f^{p',\omega'}$ and $E_{nf}^{p,\omega'} = E_{nf}^{p',\omega'}$ for all $p, p' \in [0, \frac{c+q-1}{q}]$ and $\omega' \in [0,1]$.

Since, at information set NL a player learns that b does not exist (since $k \to 1$), his posterior equals the post rumour belief (since existence of b is uninformative about the state - assumption 2). Consider the situation of player who receives no letter. He believes that with very high probability, no player received any letter which implies that he is at a situation where with very high probability, everyone received no letters and have to decide whether to fight or not. Thus, the only responses that matter are the equilibrium responses to the signal NL. Suppose the equilibrium response was the following:

$$Ethnicity\ E_1$$

$$NL \to p_{E_1} f + (1-p_{E_1}) nf$$

$$Ethnicity\ E_2$$

$$NL \to p_{E_2} f + (1-p_{E_2}) nf$$

Case 1 - $p_{E_1} > \frac{c+q-1}{q}$

Clearly then conflict happens with probability 1. Therefore, the unique best response is to play f.

Case 2 - $p_{E_1} \in [0, \frac{c+q-1}{q}]$

In this case, the only possible equilibrium corresponds to $p_{E_1} = p_{E_2}$. This is because the ethnicities are completely symmetric so there cannot be an equilibrium where the two ethnicities randomize with different weights. However, if $p_{E_1} = p_{E_2}$, it can be shown as an extension of proposition 2 that given that the other players are playing this way, the unique best response is to play f.

Hence $p_l = 1$ in any equilibrium.

Now, since $\frac{\phi_q}{\phi_r} = \hat{\sigma}$, we have $\Pr(q|LF) = \omega^*$ and since the letters are informative we have $\frac{\phi_q}{\phi_r} = \hat{\sigma} > 1$ which implies $\frac{1-\phi_q}{1-\phi_r} < 1$ and hence $\Pr(q|LT) < \Pr(q|LF) = \omega^*$. Notice that $E(f|LT) = E_f^{p_t,\Pr(q|LT)}$ and $E(f|LF) = E_f^{p_t,\Pr(q|LF)}$ and this holds similarly for nf. So we have $E(f|LT) > E(nf|LT)$ and $E(f|LF) = E_f^{0,\omega^*} = E_{nf}^{0,\omega^*} = E(nf|LF)$. Hence, $p_t = 1$ and any $p_f \in (0, \frac{c+q-1}{q}]$ can be sustained as an equilibrium. By a similar argument, it can be shown that for $\frac{1-\phi_q}{1-\phi_r} = \hat{\sigma}$, an equilibrium can be sustained for $p_t \in (0, \frac{c+q-1}{q}]$ and $p_f = 1$.

Now suppose $\frac{\phi_q}{\phi_r} \neq \hat{\sigma}$ and $\frac{1-\phi_q}{1-\phi_r} \neq \hat{\sigma}$. Then, $\Pr(q|LT)$ and $\Pr(q|LF)$ are both not equal to ω^*. Hence, $E(f|LT) = E_f^{0,\Pr(q|LT)} \neq E_{nf}^{0,\Pr(q|LT)} = E(nf|LT)$ and this similarly holds for LF. Hence, no mixing is possible. □

A.4 Results : Strategic b

In this subsection we give the proofs for the claims made in section 5.

Proof of Lemma 1 :

Consider any symmetric equilibrium E^*. First note that the probability that person b exists given that a player i has received no letter is given by :

$$\frac{\zeta h^*}{\zeta h^* + (1-\zeta)}$$

where h^* is the probability with which b sends no letter to player i in E^*. Clearly, as $\zeta \to 0$, the probability that player b exists given that no letter was received goes to zero. Thus, if we choose a low enough $\bar{\zeta}$, then, when a player doesn't receive a letter he believes that with very high probability, player b does not exist. This implies that with very high probability, no one received any letter.

Beliefs about state for people who received no letters is the same as post rumour beliefs. This is

because existence or non existence of b is not informative about the state (assumption 2). Consider the situation of player who receives no letter. He believes that with very high probability, no player received any letter. Thus the only responses that matter are the equilibrium responses to the signal NL. Suppose the equilibrium response was the following:

$$Ethnicity\ E_1$$

$$NL^s \rightarrow p_{E_1} f + (1 - p_{E_1}) nf$$

$$Ethnicity\ E_2$$

$$NL^s \rightarrow p_{E_2} f + (1 - p_{E_2}) nf$$

Define z ; $zq + (1-q) = c$. z is the maximum fraction of G type players who can play fight and yet not have conflict break out in the good state of the world. If more than z fraction of G types from any one ethnicity decide to fight then conflict is inevitable.

Case 1 - $p_{E_1} > z$

Clearly then conflict happens with probability 1. Therefore, the unique best response is to play f.

Case 2 - $p_{E_1} \in [0, z]$

In this case, the only possible equilibrium corresponds to $p_{E_1} = p_{E_2}$. From proposition 2, we note that when both ethnicities face the same probability of winning ($\frac{1}{2}$), each agent prefers f to nf. If WLOG $p_{E_1} < p_{E_2}$, then every agent in ethnicity E_1 is indifferent between f and nf. However, ethnicity E_1 would now face a probability of winning greater than $\frac{1}{2}$, which would further induce them to play f. But this contradicts the fact that they were indifferent. Hence $p_{E_1} = p_{E_2}$. However, if this indeed were the case, it can be shown as an extension of proposition 2 that given that the other players are playing this way, the unique best response is to play f.

□

Proof of Proposition 4 :

Suppose not i.e. let's suppose that truth telling on $b's$ part can be an equilibrium strategy. Consider an agent of the opposite ethnicity $i \in E_2$. If he receives the message LT^d, then he knows with probability close to

one that the state of the world is (r,r) and hence believes that conflict is inevitable with a high probability. This is because b's signal is extremely informative ($\frac{\phi_q}{\phi_r}$ is high) and she always reveals her signal truthfully. Hence, if strategies are symmetric everyone from the opposite ethnicity chooses to fight when the message LT^d is received. By lemma 1, all players respond to NL with f in any equilibrium. If i receives LF^d, then it cannot be the case that the action nf is played with positive probability. This is because then b, upon receiving signal T, would deviate to the message LF^d to maximize the probability of her ethnicity winning. Hence, the opposite ethnicity always fights making conflict inevitable. The same ethnicity, knowing that conflict cannot be avoided would always choose to fight. This is a contradiction to the assumption that the equilibrium being played was different from the all-fight equilibrium.

□

Proof for Proposition 5

Let $\bar{\sigma} > 0$ and consider a player of the same ethnicity : $i \in E_1$. We claim that an equilibrium can be supported with $p_s = 0$ and $p_d \leq z$. We check optimality for agent i at each information set. Since he is of the same ethnicity as b, receiving the message LT^s or LF^s perfectly reveals to him that $b's$ private signal is F. For high $\frac{\phi_q}{\phi_r}$ (i.e. signals are very informative), the agent knows that with probability close to one the true state of the world is (q,q). In this case, given the strategies of others, he knows that a proportion z of G types from E_1 ethnicity choose to fight and a proportion p_d of G types from E_2 ethnicity choose to fight but this is not enough to start a conflict. Hence, i's optimal strategy is to play nf and get the high peace time payoff. Hence for some σ_1, and any $\frac{\phi_q}{\phi_r} \geq \sigma_1$ the agent's response to LT_s, LF_s is optimal. Now, consider the signal NL^s. We have already shown that the agent will respond with f (lemma 1). We have shown the optimality of strategies of players from the same ethnicity. We concluded that $p_s = 0, p_d \leq z$ can be supported as an equilibrium under $\frac{\phi_q}{\phi_r} \geq \sigma_1$ and $\zeta \leq \bar{\zeta}$ for some $\bar{\zeta}$. Let us now discuss the optimality of b's strategy.

Consider first the case that b receives the private signal T. For high $\sigma_2 > 0$ and $\frac{\phi_q}{\phi_r} \geq \sigma_2$, b believes with high probability that the state is (r,r) i.e conflict cannot be avoided. Her optimal response is then to maximise the probability of her ethnicity winning, which is achieved by persuading all from her own ethnicity to fight (she does this by sending them all the signal NL^s) and dissuading as large a proportion of the opposite ethnicity (in this case $1 - p_d$) from fighting as possible. Given the strategies of the opposite ethnicity, this is achieved by randomly sending each person either LT^d, LF^d.

Now suppose b has received the signal F. For high $\sigma_3 > 0$ and $\frac{\phi_q}{\phi_r} \geq \sigma_3$, b believes with almost full certainty that the state is (q,q). She would, in this case, prefer that conflict be averted (thereby achieving $\alpha + \delta$ as

payoff) and can enforce a no conflict outcome by adhering to the strategy prescriptions (a fraction z fights from own ethnicity and $p_d(\leq z)$ from the opposite ethnicity). Note that b decides to make some players from her own ethnicity fight even when she believes that the state is most likely to be (q,q). This is because as long $\infty > \frac{\phi_q}{\phi_r} > 0$, there is always a positive probability of the state being (r,r). b hedges against this risk by making a fraction z from her own ethnicity fight where z good types fighting is the largest fraction of good types which can fight and not cause conflict in the good state (z; $zq+(1-q)=c$). We have so far shown optimality of the strategies for E_1 ethnicity players and b. We now show optimality for agents of the opposite ethnicity (E_2). In particular, it will be important to find conditions under which the randomization p_d is optimal.

We first notice that in any equilibrium belonging to the above class of strategies, it is necessarily the case that informed agent sends uninformative signals to members of the opposite ethnicity. This is because the opposite ethnicity plays the same randomisation p_d irrespective of the contents of the letter received implying the that belief at LT^d and LF^d must be same (since randomisation is possible only at a unique belief). Hence, $\Pr(q|LT^d) = \Pr(q|LF^d)$. But this implies that $q_b^T = q_b^F$. As a result, receiving a letter takes the opposite ethnicity back to post-rumour beliefs ω'. Now define the function $g : [0,z] \to \mathbb{R}$ as follows :

$$g(p) = \omega'(-\gamma) + (1-\omega')[\tfrac{p+(1-r)}{p+r+2(1-r)}(\alpha) + \tfrac{r+(1-r)}{p+r+2(1-r)}(-\beta+\varepsilon)] - \omega'(\alpha+\delta) + (1-\omega')(-\beta)$$

We know from the conditions (2),(3) that : $g(0) < 0$ and $g(z) > 0$. Also, notice that g is strictly increasing in p. Hence, there exists a unique $p_d \in (0,z)$ such that $g(p_d) = 0$. It is clear that above strategy specification is an equilibrium for this p_d and thresholds by defined by $\bar{\sigma} = \max\{\sigma_1, \sigma_2, \sigma_3\}$ and $\bar{\zeta}$.

□

Proof for Proposition 6

Let $\varepsilon > 0$ such that condition (3) holds for any $\omega'' \in (\omega' - \varepsilon, \omega' + \varepsilon)$. Now define $Pr(q|LF^d; q_b^F, q_b^T)$ and $Pr(q|LT^d; q_b^F, q_b^T)$ be the posteriors of the agents of the opposite ethnicity about the state q conditional on information given by the letters LF^d and LT^d under the signal structure (q_b^F, q_b^T). Now one can find (q_b^F, q_b^T) such that both $Pr(q|LF^d; q_b^F, q_b^T)$,$Pr(q|LT^d; q_b^F, q_b^T) \in (\omega' - \varepsilon, \omega' + \varepsilon)$. We now confirm that this is an equilibrium.

We have already checked that in any symmetric equilibrium the response to the signal NL is to play f. We next show the opposite ethnicities responses to LT^d and LF^d. Under both signals, the posteriors of the agent are in the interval $\omega'' \in (\omega' - \varepsilon, \omega' + \varepsilon)$ and hence satisfy condition (3) when substituted. However, notice that after the substitution the RHS of (3) becomes the payoff from f and the LHS the payoff from nf. Hence playing nf is strictly better at both LT^d and LF^d. Notice that this signal structure is also optimal

for player b. In any state of the world she would want as few of the opposite ethnicity to participate and the suggested strategy achieves that objective. It can be checked that the incentives of the players at other information sets are also optimal. Hence the above specification is an equilibrium.

□

Proof of Proposition 8

Assume $\zeta \approx 1$. Consider the following strategy specification :

$$
\begin{aligned}
&\textbf{b's strategy}: \\
&f_b(E_1,T) = LT^s \\
&f_b(E_2,T) = q_b^T LT^d + (1-q_b^T)LF^d \\
&f_b(E_1,F) = zLT^s + (1-z)(q_b^F LF^s + (1-q_b^F)NL) \\
&f_b(E_2,F) = q_b^F LT^d + (1-q_b^F)LF^d \\
&where\ zq + (1-q) = c \\
&\textbf{Player's strategies} \\
&E_1\ ethnicity/Same\ ethnicity \\
&g^{E_1}(NL^s) = nf \\
&g^{E_1}(LF^s) = nf \\
&g^{E_1}(LT^s) = f \\
&E_2\ ethnicity/Opposite\ ethnicity \\
&g^{E_2}(Nl^s) = f \\
&g^{E_2}(LF^d) = p_d f + (1-p_d)nf \\
&g^{E_2}(LT^d) = p_d f + (1-p_d)nf \\
&where\ 0 < p_d \leq z,\ q_b^F = q_b^T \in [0,1]
\end{aligned}
\tag{5}
$$

This proof will demonstrate how we can obtain the same equilibrium outcome as (1). A similar proof can be written down to get the equilibrium outcome from (4).

Upon receiving the signal LT^s, same ethnicity player thinks that either the rumour is true or the rumour is false but b has sent him LT^s to hedge. When z is low enough, the mass of the belief is largely on the former. Since $b's$ signals are very informative, the rumour being true implies that the state is very likely to be bad. In this case, a same ethnicity player has beliefs below post rumour beliefs. At such beliefs, as we have already shown before, if peace is ensured in the good state, then it is strictly better to fight if the probability of win is above $1/2$ (according to equilibrium strategies). Hence, the same ethnicity strictly prefers to play f at LT^s. Notice that the role of NL in the ζ low case is played here by the signal LT^s. The agent b can use it make his own ethnicity fight. At LF^s or NL, an agent of the same ethnicity learns whether

the rumour is false and hence would want to refrain from fighting.

The optimality of agent b is satisfied since she can induce conflict under T and avert it under F. Notice due to the slight imperfection in her own information, as in proposition 5,6, her hedging incentive appears here as well. Hence at F, she would want to avert conflict but send a maximum proportion z of her own ethnicity to fight.

Notice that the opposite ethnicity does not get any information as in (1). Hence, they always update to post rumour beliefs. At these beliefs, given that conflict is averted in the good state and a fraction $p_d > 0$ of own ethnicity fight and all from the other ethnicity fight in the bad state, the agent is indifferent between f and nf.

□

Proof of Proposition 9 :

1. Consider the strategies outlined in $(1),(4)$. Notice that in both $(1),(4)$ optimality of strategy for players of either ethnicity is satisfied given the behaviour of the agent b. Consider strategies outlined in (4) and consider the incentives of b. If the rumour is true, agent b believes it is the bad state of the world with high probability and hence would want to maximise her probability which is achieved by her strategy. At F, she believes with high probability that it is the good state. Suppose her belief about state being (q,q) if rumour is false is ω_F. She can either avert conflict (which her strategy proposes) or deviate and induce conflict by sending NL^s to her own ethnicity. This would lead to everyone from her ethnicity fighting which gives her a probability of winning equal to $\omega_F \frac{1}{1+1-q} + (1-\omega_F)(\frac{1}{1+1-r}) = p_2$. Then, inducing conflict gives utility $p_2 u_b(CW) + (1-p_2)u_b(CL)$ and averting it gives utility $\omega_F u_b(NC) + (1-\omega_F)[\frac{z+1-r}{(z+1-r)+(1-r)}u_b(CW) + (1-\frac{z+1-r}{(z+1-r)+(1-r)})u_b(CL)]$. Since signals are very informative $\omega_F \to 1$. Now:

$$Payoff\ from\ deviating\ and\ inducing\ conflict = p_2 u_b(CW) + (1-p_2)u_b(CL) < u_b(NC) \leftarrow Payoff\ from\ strategy \tag{6}$$

   Hence, averting conflict is better and her hedging incentive makes her send NL^s to a z proportion of her own ethnicity. Since $p_1 < p_2$, the same argument works for (1) as well.

2. A similar argument as above gives the result.

3. The fact that $(1),(4)$ are not equilibrium strategies any more follows from the arguments made above. We show that the unique equilibrium is *all fight*. Suppose not. Then there is an equilibrium where conflict is averted either when rumour is true or false. Since the rumours are very informative, b may not induce conflict only when the rumour is false. Using arguments similar to those used in proposition 7, it can be established that such an equilibrium has an outcome of the form $((1,s),(z,s)$ where $s = 0$ or $s = p_d$. Suppose $s = p_d$, then when rumour is F, agent b gets $u_b(NC)$. However, she

can deviate and send everyone from her own ethnicity NL^s and secure the payoff $p_1 u_b(CW) + (1 - p_1)u_b(CL) > u_b(NC)$. The deviation is also strictly profitable even when $s = 0$. Hence, *all fight* is the unique equilibrium.

□