Fairness preferences as a cause of inefficient war

M. Christian Lehmann[1]*

Abstract
Social psychologists, behavioral economists and biologists have long documented how fairness concerns motivate behavior, but workhorse bargaining models of war ignore this. I present a simple model in which a nation is challenged by an enemy (e.g., a terrorist attack). The nation's political leader must decide between a peaceful solution (i.e., buying the enemy off) or war (i.e., hunt the enemy down). The leader, who has re-election incentives, knows that a fraction of voters has fairness preferences — they strongly feel that the enemy should be hunted down (i.e., "punished") rather than bought off (i.e., "rewarded"). The model illustrates that fairness preferences can trigger war even if a peaceful solution is ten times cheaper.

JEL Classification: D72; D74; D91

Keywords
bargaining — violence — conflict — justice — vengeance — retaliation

[1] *Department of Economics, University of Brasilia*
***Corresponding author**: clehmann@unb.br

Introduction

On September 11, 2001, Al-Qaeda terrorists attacked the United States, causing the death of thousands and prompting US president George W. Bush to launch the "war on terror", starting with the invasion of Afghanistan and later Iraq. Imagine that, instead of invading Afghanistan, the US president would have announced a peace deal with Al-Qaeda: for an annual cash transfer of one million dollar, the terrorist group agreed to abstain from future attacks against US targets.[1] This would have been a much cheaper solution than a trillion dollar war on terror.[2] However, who would disagree that it would also have been political suicide? Wouldn't voters have cried out with indignation that "President Bush *rewards* Al-Qaeda for attacking the United States"?

The canonical bargaining model of war, however, would predict the peace deal because it is cheaper than war (Blattman and Miguel, 2010). So what is wrong with the model? A peace deal obtained through a transfer from the aggressed nation to the aggressor (the prediction of the canonical model) is incompatible with common sense because it strikes us as an *unfair* solution. However, the canonical bargaining model of war ignores fairness, i.e., it ignores a fundamental motivation of human behavior according to results from lab- oratory experiments (e.g., behavioral games) conducted by social psychologists, behavioral economists and biologists. Consider ultimatum game experiments, for example: the first player proposes a division of a monetary amount and the second player must accept or de- cline the offer, where both players get nothing if the second player declines the offer. The dominant finding is that unfair (small) offers are frequently rejected, when conventional eco- nomic theory predicts that any non-zero offer is accepted (e.g., Fehr & Schmidt, 1999 and Fehr & Fischbacher, 2003).

The ultimatum game is conceptually similar to a war/peace choice. The aggressor (e.g., the terrorist group) offers peace in exchange for concessions and the aggressed player must accept (peace) or decline (war). The ultimatum game literature suggests that the aggressed player would decline an unfair peace offer even if its payoff is positive, i.e., even if the peace deal is cheaper than going to war.[3] Yet no existing model of war allows for this possibility.[4] The word "fairness", for

[1] The point I want to illustrate with this example is that a peace deal obtained through a transfer from the aggressed nation to the aggressor (i.e., the prediction of the canonical bargaining model of war) is at odds with common sense because it is unfair. I do not want to imply that Al-Qaeda could have been bought off with money. The motive of many terrorist attacks is non-monetary concessions or vengeance (in the case of Al-Qaeda perhaps vengeance for US support of Israel, US troops in Saudi Arabia, and US sanctions against Iraq, among others). I examine the vengeance motive in section 2.2.

[2] Total spending for US operations related to the war on terror exceeds one trillion US dollar (Orszag, 2007).

[3] In the context of civil war and terrorism, fairness is a broader concept than in conventional behavioral games. In behavioral games, fairness usually refers to the fair allocation of resources (e.g., ultimatum game). In war and terrorism, fairness can refer to the allocation of resources (e.g., concessions) and other decisions such as the fair amount of retaliation. In this article I use the term fairness mainly to refer to the latter (retaliation), which is related to behavioral games that study preferences for retaliation (e.g., Crockett et al., 2014).

[4] See Blattman & Miguel, 2010 for an excellent review of the literature. In existing bargaining models, war may ensue when there is (i) asymmetric information, such as private information about military strength, and the strategic incentive to misrepresent it to opponents; (ii) commitment problems, especially the inability to commit to peace deals in the absence of a third-party enforcer; and (iii) issue indivisibilities, whereby some issues do not admit compromise.

example, is not mentioned in Blattman & Miguel, 2010's review of the civil war literature and Sandler & Arce, 2007's review of the terrorism literature. Blattman & Miguel, 2010: 21 specifically conclude that "psychological factors (...) have yet to be applied to formal models of civil war."

My objective is to contribute to fill this gap. This article, to my knowledge, is the first attempt to introduce fairness preferences in a formal bargaining model of war. I show that when voters have fairness preferences, war can ensue even if a peace deal is ten times cheaper. The model is deliberately simple to illustrate how fairness preferences influence the decision to go to war and how fairness preference can lead to war even though a peace deal is much cheaper. I see it as a starting point for further theorizing. In the Conclusion I highlight promising avenues for future research.

Model

For ease of exposition, I start by introducing fairness preferences in the simplest bargaining model of war (a one-shot game with no uncertainty). In section 2.2, I introduce uncertainty and repeated interaction, among others. These extensions, however, do not change the conclusion of the simpler model that fairness preferences can be a source of inefficient war.

Baseline model

Consider a nation with an elected incumbent government, populated by a continuum of voters with mass one and national income Y. The sequence of events is as follows:

- t_0: The nation is challenged by an enemy (e.g., a terrorist group). By "challenged" I think of a hostile act in which the enemy signals its strength (e.g., a terrorist attack).
- t_1: Fairness preferences of voters are revealed to the incumbent.
- t_2: The incumbent negotiates a peace deal or declares war.
- t_3: The incumbent runs for re-election.

The incumbent is a rational economist with "standard" preferences, i.e., she believes that her job is to maximize national consumption. But she also has re-election concerns and hence incentives to accommodate voters with "non-standard" (i.e., fairness) preferences.

Going to war in t_2 is the attempt to annihilate the enemy through the use of force (e.g., military, police, etc.), which for the incumbent yields an expected utility of

$$EU_{war} = \alpha[Y - c] \tag{1}$$

where the parameter c reflects the cost of war. National consumption in the case of war is thus given by $Y - c$. The key feature of the model is the parameter $\alpha \in [0, 1]$, which represents the share of voters with fairness preferences. These voters strongly feel that it would be unfair to buy off (reward) the enemy for having attacked the nation and that a fair response is to retaliate. In t_3, they re-elect the incumbent if she goes to war in t_2. Vice- versa, $1 - \alpha$ is the share of voters with "standard" preferences — they prefer the solution that maximizes national consumption: If buying off the enemy is cheaper than war, they re-elect the incumbent in t_3 if she negotiates a peace deal in t_2

I treat α as exogenous, drawn from some cumulative distribution function in t_1. How fairness preferences arise and their distribution in the general population is still an open debate among social psychologists, behavioral economists, and biologists. In ultimatum games, for example, 20 to 30 percent of proposers offer a fair split of the money (i.e., a 50-50 split), which suggests that some but not all individuals are motivated by fairness concerns, i.e., $0 < \alpha < 1$.[5]

Finally, a peace deal in t_2 is achieved through a transfer (concession) to the enemy, denoted τ. National consumption in the case of a peace deal is thus given by $Y - \tau$, which for the incumbent yields an expected utility of

$$EU_{peace} = [1 - \alpha[Y - \tau] \tag{2}$$

It it easy to see that the incumbent is indifferent between war and a peace deal ($EU_{war} = EU_{peace}$) if

$$\alpha^* = \frac{Y - \tau}{2Y - \tau - c} \tag{3}$$

Hence if $\alpha > \alpha^*$, the incumbent opts for war.

Implications. Figure 1 illustrates that fairness preferences can cause war even if a peace deal is substantially cheaper. The figure plots α^* as a function of the ratio of peace deal and war consumption ($[Y - \tau]/[Y - c]$). For example, we see that even if a peace deal is twice as cheap as war (i.e., ratio = 2), it would require only about two-thirds of voters with fairness preferences to prompt the political leader to go to war. That in reality at least two-thirds of voters have fairness preferences is not implausible. In ultimatum games, for example, unfair offers are rejected by more than two-thirds of players. If almost all voters have fairness preferences (e.g., $\alpha = 0.9$) war will ensue even if a peace deal is ten times cheaper.

Extensions

Higher costs of peace. In some contexts, in particular in the case of insurgencies, the enemy's goal is full control over the incumbent's resources (e.g., territory) and the cost of peace can be high. Intuitively, the cost of peace reflect the strength of the enemy: the stronger the enemy the higher the concession

[5] E.g., Güth et al. (1982), Kahneman et al. (1986), Thaler (1988), Güth (1995), and Camerer (1997).

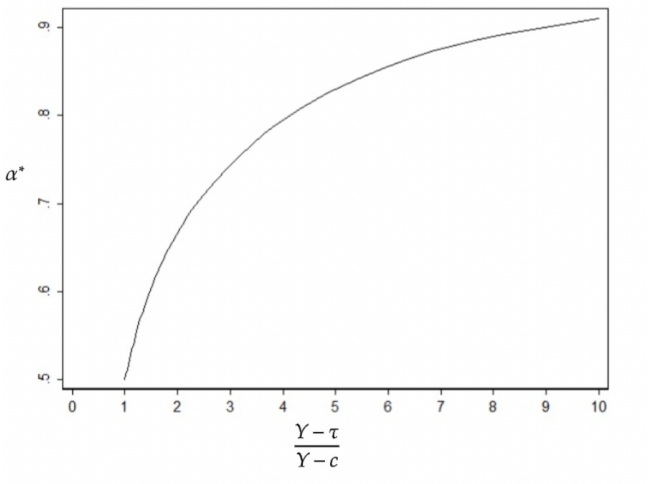

Figure 1. Fairness preferences as a cause of inefficient war.

Notes: The Figure illustrates that fairness preferences can cause war even though a peace deal is substantially cheaper. The figure plots α^* (see equation 3) as a function of the ratio of peace deal and war consumption ($[Y-\tau]/[Y-c]$). For example, we see that even if peace is twice as cheap as war (i.e., ratio=2), it would require only about two-thirds of voters with fairness preferences to prompt the political leader to go to war.

(τ) required to buy off the enemy. A higher cost of peace thus implies a lower expected payoff from fighting because a strong enemy is not easily defeated militarily.

Formally, instead of equation (1), assume that the incumbent's expected utility of going to war is given by

$$EU_{war} = \alpha[\theta Y + (1-\theta)\cdot 0 - c] \tag{4}$$

where θ is a contest success function (Skaperdas, 1996), i.e., the probability of an incumbent military victory. $1-\theta$ reflects the strength of the enemy, i.e., the probability of an enemy military victory. In the baseline model, I implicitly assume that $\theta = 1$, i.e., the incumbent knows with certainty that going to war means winning the war, and that winning the war costs c. Now there is uncertainty about the outcome of war: with probability $1-\theta$ the incumbent loses and gets zero utility.

The minimum concession required to dissuade the enemy from fighting is equal to the enemy's expected payoff of fighting:

$$\tau = (1-\theta)Y - \kappa \tag{5}$$

where κ is the enemy's cost of fighting. The size of τ (the cost of peace) is increasing in the enemy's strength $(1-\theta)$ and the value of the contested resource (Y), and decreasing in the enemy's cost of war (κ). Therefore, the incumbent's expected utility of a peace deal is

$$EU_{peace} = [1-\alpha][Y-\tau] = [1-\alpha] \\ [Y-((1-\theta)Y-\kappa)] = [1-\alpha][\theta Y+\kappa] \tag{6}$$

Now let us compare the incumbent's expected payoff of war (equation (4)) and peace (equation (6)). Without fairness preferences (α) in the model, the expected payoff of a peace deal $(\theta Y+\kappa)$ always exceeds the expected payoff of war $(\theta Y - c)$ if fighting is costly for at least one player (i.e., if $c > 0$ or $\kappa > 0$). With fairness preferences, it is possible that the expected payoff of war (equation (4)) exceeds the expected payoff of peace (equation (6)) if the share of voters with fairness preferences is sufficiently large, i.e., if

$$\alpha > \frac{Y\theta+\kappa}{2Y\theta - c + \kappa} \tag{7}$$

Repeated interaction. A peace deal may not be reached if the enemy is expected to become stronger in the future (Powell, 2006). For example, consider a temporarily weak enemy (e.g., an insurgent without automatic weapons). Today this enemy may be bought off cheaply. But in the future when it is stronger (e.g., with automatic weapons) it has incentives to attack in order to signal its new strength and demand larger concessions. The government will prefer to go to war today if the payoff is greater than the expected payoff of a peace deal sustained by small concessions today but large concessions tomorrow.

Formally, the incumbent's expected utility of a peace deal is

$$EU_{peace} = [1-\alpha][Y - \int \tau(t)dt] \tag{8}$$

where $\int \tau(t)dt$ is the sum of current and future concessions required to sustain a peace deal. The required concession at a given point in time, $\tau(t)$, depends on the enemy's expected payoff of fighting at that point, which in turn depends on the enemy's strength $(1-\theta(t))$ at that point

$$\tau(t) = (1-\theta(t))Y - \kappa \tag{9}$$

The incumbent's expected utility of war is

$$EU_{war} = \alpha[\theta(t)Y - c] \tag{10}$$

The incumbent prefers war if its expected payoff (equation (10)) exceeds the expected payoff of a peace deal (equation (8)), which once more is the case if the share of voters with fairness preferences is sufficiently large, i.e., if

$$\alpha > \frac{Y - \int \tau(t)dt}{[1+\theta(t)]Y - c - \int \tau(t)dt} \tag{11}$$

Enemy with fairness preferences. Thus far I have assumed that only voters have fair- ness preferences. A more realistic assumption is that both voters *and the enemy* have fairness preferences. As we have seen, voters with strong fairness preferences (i.e., a high α) are willing to pay for retaliation (vengeance), i.e., they prefer war even though a peace deal is cheaper. The difference between the cost of war and the cost of a peace deal is what voters are willing to pay for retaliation. Vice-versa, an enemy with fairness preferences is willing to pay for retaliation, i.e., the enemy prefers war even though a peace deal yields a higher material payoff. Formally, suppose that the enemy's expected utility of a peace deal and war is $EU_{peace} = [1-\beta]\tau$ and $EU_{war} = \beta[(1-\theta)Y - \kappa]$, respectively, where $\beta \in [0,1]$ represents the strength of the enemy's fairness preferences. It is easy to see that the enemy's expected utility of a peace deal is decreasing in β. In the extreme case of $\beta = 1$, no transfer τ can dissuade the enemy from war.

Conclusion

In their review of the literature, Blattman & Miguel (2010:21) conclude that "psy- chological factors (...) have yet to be applied to formal models of civil war." This article contributes to fill this gap by introducing fairness preferences in the canonical bargaining model of war. I show that when the share of voters with fairness preferences is sufficiently large, war can ensue even if a peace deal is ten times cheaper.

I see several promising avenues for future research. In this article I focus mainly on the incumbent's incentives and mostly ignore the enemy's incentives. For example, in contexts where the incumbent is clearly stronger from a military point of view, it seems reasonable to assume that the enemy's objective is that the incumbent yields to the enemy's demands and to avoid a military confrontation. In the baseline model of section 2.1, τ can be interpreted as the incumbent's *willingness to yield*, i.e., as the maximum amount of concessions that the incumbent is willing to make. From equation (3) it follows that the incumbent is indifferent between yielding and a military confrontation if $\tau^* = \frac{\alpha[2Y-c]-Y}{\alpha-1}$. It is easy to see that $\frac{\partial \tau^*}{\partial \alpha} < 0$, i.e., fairness preferences reduce the incumbent's willingness to yield. This in turn reduces the enemy's incentives to challenge the incumbent in t_0. Interestingly, the incumbent's willingness to yield is *negative* if the share of voters with fairness preferences is sufficiently large (i.e., $\tau^* < 0$ if $\alpha > \frac{Y}{2Y-c}$). Now the *enemy* would have to have to make concessions in order to dissuade the incumbent from fighting, which eliminates the enemy's incentives to challenge the incumbent in t_0. In this case, interestingly, fairness preferences can foster efficient peace, not inefficient war as in the models in section 2.

The previous example illustrates that there is no categorical answer to the question of whether fairness preferences cause inefficient war. The answer depends on the context. More research is needed to further advance our understanding of how conflicts (e.g., interstate wars, insurgency and counter-insurgency, terrorism and counter-terrorism) are influenced by fairness preferences, and how their influence varies across contexts. In reality, the decision of going to war is more complex than in the simplistic models of this article. Political leaders may be concerned with numerous other factors, such as the incentives that a peace deal creates for other potential enemies. Some of these factors have already been examined by previous research (see Blattman & Miguel, 2010). My model ignores these factors, which creates a simple and transparent "laboratory" setting in which the effect of one specific factor — fairness preferences, which has received little attention thus far — can be studied in isolation from other factors. However, there may be *interactions* between fairness preferences and other factors. Investigating these interactions is another promising avenue of future research.

References

Blattman, C. & Miguel, E. (2010). Civil War. *Journal of Economic Literature, 48*(1), 3–57.

Camerer, C. F. (1997). Progress in behavioral game theory. *Journal of Economic Perspectives, 11*(4), 167–188.

Crockett, M. J., Özdemir, Y. & Fehr, E. (2014). The value of vengeance and the demand for deterrence. *Journal of Experimental Psychology, 143*(6), 2279–2286.

Fehr, E. & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*, 785–791.

Fehr, E. & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics, 114*(3), 817–868.

Güth, W. (1995). On ultimatum bargaining experiments-A personal review. *Journal of Economic Behavior & Organization, 27*(3), 329–344.

Güth, W., Schmittberger, R. & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization, 3*(4), 367–388.

Kahneman, D., Knetsch, J. L. & Thaler, R. H. (1986). Fairness and the assumptions of economics. *Journal of Business, 59*(4), 285–300.

Orszag, P. (2007). Estimated costs of US operations in Iraq and Afghanistan and of other activities related to the war on terrorism. *Congressional Budget Office, Washington DC*.

Powell, R. (2006). War as a commitment problem. *International Organization, 60*(1), 169–203.

Sandler, T. & Arce, D. G. (2007). Terrorism: A game-theoretic approach. In Keith Hartley & Todd Sandler (ed.) *Handbook of Defense Economics*, Volume 2, pp. 775-813. Elsevier.

Skaperdas, S. (1996). Contest success functions. *Economic Theory, 7*(2), 283–290.

Thaler, R. H. (1988). Anomalies: The ultimatum game. *Journal of Economic Perspectives, 2*(4), 195–206.