War signals: a theory of trade, trust and conflict

Rohner, Dominic ; Thoenig, Mathias ; Zilibotti, Fabrizio

Abstract: We construct a theory of persistent civil conflicts, where persistence is driven by the endogenous dynamics of inter-ethnic trust and trade. In times of peace, agents belonging to two groups are randomly matched to trade bilaterally. Trade hinges on trust and cooperation. The onset of conflict signals that the aggressor has a low propensity to cooperate, harming future trust and trade. Agents observe the history of conflicts and update their beliefs over time. The theory predicts that civil wars are persistent. Moreover, even accidental conflicts that do not reflect economic fundamentals erode trust, and can plunge a society into a vicious cycle of recurrent conflicts (a war trap). The incidence of conflict can be reduced by policies abating cultural barriers, fostering inter-ethnic trade and human capital, and shifting beliefs. Coercive peace policies, such as peacekeeping forces or externally imposed regime changes, have no enduring effects.

DOI: https://doi.org/10.1093/restud/rdt003

Originally published at:
Rohner, Dominic; Thoenig, Mathias; Zilibotti, Fabrizio (2013). War signals: a theory of trade, trust and conflict. Review of Economic Studies, 80(3):1114-1147.
DOI: https://doi.org/10.1093/restud/rdt003

War Signals: A Theory of Trade, Trust and Conflict*

Dominic Rohner[†], Mathias Thoenig[‡], Fabrizio Zilibotti[§]

This version: December 2012
First version: July 2011

Abstract

We construct a theory of persistent civil conflicts, where persistence is driven by the endogenous dynamics of inter-ethnic trust and trade. In times of peace, agents belonging to two groups are randomly matched to trade bilaterally. Trade hinges on trust and cooperation. The onset of conflict signals that the aggressor has a low propensity to cooperate, harming future trust and trade. Agents observe the history of conflicts and update their beliefs over time. The theory predicts that civil wars are persistent. Moreover, even accidental conflicts that do not reflect economic fundamentals erode trust, and can plunge a society into a vicious cycle of recurrent conflicts (a *war trap*). The incidence of conflict can be reduced by policies abating cultural barriers, fostering inter-ethnic trade and human capital, and shifting beliefs. Coercive peace policies, such as peacekeeping forces or externally imposed regime changes, have no enduring effects.

JEL classification: D74, D83, O15, Q34.

Keywords: beliefs, civil war, conflict, cooperation, cultural transmission, ethnic fractionalization, human capital, learning, matching, peacekeeping, stochastic war, strategic complementarity, stag hunt game, trade.

*For comments on this paper, we are grateful to Philipp Kircher, the editor, three anonymous referees, and Daron Acemoglu, Philippe Aghion, Alberto Alesina, Roland Benabou, Joan Esteban, Julien Hugonnier, Massimo Morelli, Jody Ono, Gerard Padro i Miquel, Maria Saez-Marti, Gilles Saint-Paul, Pierre Yared, as well as conference participants at the CEPR conference on Political Economy of Development, NBER Summer Institute, Nemmers Prize conference, IAE concentration on conflict, SED annual meeting, and seminar participants at CREi, Paris School of Economics, SciencesPo Paris, Tilburg University, Toulouse School of Economics, University of Geneva, University of Lausanne, University of Oslo, and University of Zurich. Dominic Rohner acknowledges financial support from the Swiss National Science Foundation (grant no. 100014-122636). Mathias Thoenig acknowledges financial support from the ERC Starting Grant GRIEVANCES-313327. Fabrizio Zilibotti acknowledges financial support from the ERC Advanced Grant IPCDP-229883.

[†]Department of Economics, University of Lausanne. Email: dominic.rohner@unil.ch.

[‡]Department of Economics, University of Lausanne. Email: mathias.thoenig@unil.ch.

[§]Department of Economics, University of Zurich. Email: fabrizio.zilibotti@econ.uzh.ch.

1 Introduction

More that 16 million people are estimated to have died because of civil conflicts in the second half of the 20th century (Fearon and Laitin, 2003). Civil conflicts are persistent: 68 percent of all outbreaks took place in countries where multiple conflicts were recorded (Collier and Hoeffler, 2004). DeRouen and Bercovitch (2008) document that more than three quarters of all civil wars stem from enduring rivalries among ethnic groups that enter repeatedly into conflicts with each other. Weak institutions likely are part of the explanation, but are not the sole cause. Democracy, for instance, appears to have no systematic effect on the risk of civil war after controlling for other factors such as ethnic diversity, GDP per capita and natural resource abundance.[1] Moreover, several developing countries with relatively solid institutions experience recurrent conflicts, whereas some other countries with weak institutions and deep ethnic cleavages never see civil conflicts.[2]

In this paper, we propose a theory arguing that trust is a main determinant of civil conflict, and that inter-ethnic trade is the channel linking the dynamics of trust and conflict. On the one hand, conflict disrupts business relationships among the groups involved. A thriving inter-ethnic trade, therefore, deters war by raising the opportunity cost of war. On the other hand, trade hinges on trust, since inter-ethnic partnerships (e.g., seller-buyer, employer-employee, supplier-producer, lender-borrower relationships) typically go beyond spot transactions. By fostering trade, trust deters civil conflict.

We formalize our ideas through a dynamic model in which agents belonging to two ethnic groups are randomly matched to engage in bilateral partnerships (*trade*), which we model as a variant of the classic *stag hunt game* augmented with individual heterogeneity in the propensities to cooperate. There are strategic complementarities: the proportion of cooperators in each group increases in the perceived trustworthiness of the other group. Over time, beliefs get updated based on public signals (and, in an extension, on private information acquired by traders) and transmitted across generations. Finally, one group can wage war against the other, at the cost of destroying trade in the current period. Conflict undermines future trust by signaling to the victimized group that the aggressor has a low propensity to trade cooperatively. Thus, a war today carries the seed of distrust and future conflict.

The theory yields two main predictions that are borne out in the data. First, civil wars are persistent: each outbreak of conflict increases the probability that a country will fall again into civil war in the future. Imperfect and, possibly, incorrect learning is the source of endogenous persistence. Second, trust is negatively correlated with civil conflict. The causation runs both ways: war causes

[1] See, e.g., Fearon and Laitin (2003), Collier an Hoeffler (2004), Montalvo and Reynal-Querol (2005), Collier and Rohner (2008), Collier *et al.* (2009), and Esteban *et al.* (2012).

[2] Columbia, India, Turkey, Sri Lanka and the Philippines fare relatively well in terms of democracy and other institutional indicators, conditional on their stage of development. Yet, they are prone to civil conflicts. Interestingly, these countries have a lower average level of trust than the average non-OECD country (0.16 vs. 0.22). On the opposite front, Bhutan, Cameroon, Gabon, Kazahstan, Togo, China and Vietnam have low scores on democracy and high ethnic fractionalization, but no recent history of civil war. Data on trust are available only for China and Vietnam among these countries. Their average trust is 0.51, even larger than in the average OECD country.

trust and trade to plummet; conversely, low trust and scant inter-ethnic trade increase the probability of future wars. "Accidental wars", e.g., aggressions initiated by a belligerent minority of a group, or conflicts triggered by exogenous factors (such as a lower threat of international sanctions against aggressors) may lead to the permanent breakdown of peace (a *war trap*). War traps are characterized by endemic conflict and low inter-ethnic cooperation even during peace spells.

In the benchmark model, we introduce a number of simplifications to achieve a sharp characterization of the dynamic equilibrium. In section 6 we relax some of the assumptions. First, we consider shocks affecting over time the groups' propensity to cooperate. Second, we extend the analysis to a richer and more realistic information setting where traders acquire private information throughout their business experience about the other group's type. These and other extensions show that war traps is a robust result. In addition war cycles, i.e., periods of endemic recurrent wars followed by more peaceful periods, can arise.

The analysis yields a number of policy implications. First, policies increasing the profitability of inter-ethnic trade reduce the incidence of conflict. Examples include policies abating barriers, such as educational policies promoting the knowledge of several national languages, or subsidies for human-capital investments. Second, policies directly targeting people's held beliefs may be useful. These include educational campaigns promoting civic values and cross-group empathy, as well as repressive interventions outlawing the diffusion of hateful messages demonizing other groups. Likewise, interventions designed to nurture and foster cohesive values (e.g., national over ethnic identity) can be important. Credible campaigns documenting and publicizing success stories of inter-ethnic business relationships, joint ventures, and so on, are other relevant examples. On the contrary, attempts to *impose* peace through coercion – e.g., peacekeeping forces or externally-imposed regime changes – ultimately have no persistent effects, especially if they fail to restore trade links. Forcing the separation of groups may even be harmful, since such measures would stifle any potential for trade cooperation which otherwise may emerge during peace spells, and which over time may slowly restore confidence. These predictions are consistent with empirical studies in the conflict literature, which we discuss in more detail below.

1.1 Motivating evidence

We start by documenting that conflicts are highly persistent. We construct an indicator of civil war incidence taking on the unit value in each five-year interval during which a country experiences a civil conflict causing at least twenty-five casualties in a single year. The source data (originally at the annual level) are from the "UCDP/PRIO Armed Conflict Dataset" (UCDP, 2012).[3] We use a panel of 174 countries in the sample period 1949-2008. We run an autoregressive pooled logit regression of civil war incidence on its lagged value. Table 1 reports the marginal effects of the main variables of interest. The webpage Appendix A provides the complete set of regression results.

We find that a country experiencing war in the five-year period t has a 36 percentage points higher

[3] This dataset has been used, among others, by Besley and Persson (2011) and Esteban *et al.* (2012).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
War (t-1)	0.36***	0.22***	0.30***	0.17***	0.32***	0.17***	0.24***	0.10	0.24***	0.05
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.04)	(0.07)	(0.02)	(0.04)
Trust (t-1)							-0.37*	-0.56***	-0.48***	-0.46***
							(0.21)	(0.20)	(0.08)	(0.17)
Conflicts coded as war	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.
Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	WVS	WVS	WVS	WVS	WVS	WVS
Observations	1426	1426	1026	939	409	378	101	101	564	439
Pseudo R-squared	0.304	0.322	0.363	0.358	0.460	0.392	0.575	0.572	0.695	0.597

Dependent variable: Civil war incidence (five-year intervals). The dependent variable is coded as 1 if a conflict causing at least 25 (1000) fatalities is recorded in at least one of the five years. Sample period: 1949-2008. Number of countries for which observations are available: 174. The set of controls include: lagged democracy, lagged GDP per capita, oil exporter, lagged population, ethnic fractionalization, mountainous terrain, noncontiguous state, region fixed effects and time dummies. Columns 5-6 restrict the sample to only the 61 countries for which at least one trust observation is available. Columns 9-10 have as dependent variable civil war incidence at the annual level (details in the text). The table reports the marginal effects of logit regressions with robust standard errors, clustered at the country level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 1: Persistence of civil conflicts and correlation between conflict and lagged trust (frequency: five-years).

probability of experiencing war in the five-year period t than a country that did not experience war at period $t-1$ (column 1). The autoregressive coefficient is highly significant (>1%). We see similar results if we code as civil wars only conflicts causing at least a thousand fatalities in a single year (column 2). Since the persistence in conflict could be driven by persistent differences in institutional or political factors, we add controls for a standard set of explanatory variables used in the conflict literature; see, e.g., Fearon and Laitin (2003), Collier and Hoeffler (2004), Montalvo and Reynal-Querol (2005), Cederman and Girardin (2007), Collier and Rohner (2008), and Esteban *et al.* (2012). In particular, we control for an index of democracy, natural resources (oil), population, an index of ethnic fractionalization, geography (i.e., the proportion of mountainous terrain and a dummy for non-contiguous states), and the lagged GDP per capita. In addition, we control for time dummies and regional fixed effects. The marginal effects are affected only slightly by the control variables, and remain highly statistically significant (columns 3 and 4). The results are robust also to country fixed effects (see Table 2, columns 11-12 in Appendix A), which absorb the effect of any time invariant heterogeneity, consistent with the results of Martin *et al.* (2008b). Therefore, the result is not driven by persistent differences in institutional factors that make some countries more prone to war.

Next, we document that civil war incidence is negatively correlated with a lagged measure of trust from the World Values Survey (2011). Trust is measured by the proportion of respondents answering "Most people can be trusted" to the question: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" (A165).[4] This is a coarse measure from our perspective, since it does not focus specifically on the inter-ethnic dimension. However, we expect it to be positively correlated with inter-ethnic trust. Lagged trust is coded as the average trust level across all World Values Survey (WVS) observations available in the five-year

[4] In our sample, an average 27.5% of all respondents in a given country and five year period declare that they generally trust others (with a s.d. of 15.0%). Moreover, 23.2% of country five year period observations experience either a minor or a major conflict, whereas only 9.7% of all country five year periods experience a major war.

interval preceding the interval in which the dependent variable is measured. First, since adding trust to the regressions causes a major reduction in the sample size, we document that the results of columns 3-4 are robust to restricting the sample to the 61 countries in the WVS for which at least one trust observation is available (columns 5-6). Second, columns 7-8 show that the effect of lagged trust is negative and statistically significant (note that the sample falls further to 101 observations, since even in the restricted sample of 61 countries there is an average of only 1.7 observations per country). The effect is quantitatively large: a one standard deviation increase in lagged trust is associated with a fall in the probability of conflicts of 5.2 percentage points for all wars (column 7), and of 7.9 percentage points for big wars (column 8). It is not possible to run this specification with country fixed effects: there are only six countries – totaling sixteen observations – that have both multiple observations of lagged trust and variation in the war variable for the periods in which lagged trust is available. The results are robust to exploiting the variation of conflict at the annual frequency. In columns 9 and 10, we repeat the regressions in columns 7 and 8 using annual variables, with two qualifications. First, lagged trust continues to be the average across all observations available in the five years preceding the observation of the dependent variable.[5] Second, we continue to measure the lagged war dummy also over a five-year interval, since it is plausible that civil wars have a persistent effect on the probability of future conflict that exceeds the one-year horizon.[6] The results are robust: war is persistent, and lagged trust is highly significant.[7]

Most control variables have the expected sign (see Tables 2 and 3 in Appendix A), although they are significant only in some of the specifications. Consistent with the literature, we find that countries that are oil exporters, and have large ethnic fractionalization, a large population, a high proportion of mountainous terrain, and non-contingous territory tend to experience more conflict. Democracy has no robust effect on the risk of conflict. Its sign changes across specifications and is often insignificant.

Although we use *lagged* trust as a right-hand side variable, we do not claim to identify causal effects of trust on civil wars. To do so would require credible exclusion restrictions that are difficult to find. Moreover, our theory predicts that the causality runs both ways: whilst distrust increases the probability of civil conflict, war erodes cross-community trust. Evidence of the latter effect is documented by a number of empirical case studies. For instance, in a companion empirical paper (Rohner *et al.* 2012a), we study the effect of civil conflict on trust and other measures of social capital in Uganda. We find that an exogenous outburst of ethnic conflicts in 2002-05 driven by an international shock (the US administration's change of policy against insurgent movements of the

[5] This choice is motivated mainly by data limitations: The WVS does not provide yearly surveys. We also considered alternative measures of lagged trust where (i) we take the average over the past ten years, and (ii) we use only the most recent available survey. In both cases, the results are robust.

[6] The results are robust to a number of alternative specifications including (i) replacing the lagged war dummy with a variable counting the number of wars over the past five years; (ii) augmenting the specification discussed in the text with a dummy switching on if, in the year before the observation of the dependent variable, there was a civil conflict; (iii) lagged war is measured over a ten-year instead of a five-year window.

[7] If we use a measure of ethnic polarization instead of fractionalization, the coefficients of lagged war and lagged trust remain statistically significant in every column and of the same magnitude. Analogously, all results are robust to controlling for diamond production.

after the terrorist attack of 9/11), had a large negative effect on trust. On the one hand, districts and counties that were subject to more intense fighting experienced a decline in general trust relative to areas where fighting was lighter. On the other hand, conditional on the extent of regional violence, trust fell in relative terms for ethnic groups that were directly involved in fighting. Similar findings are documented by Cassar *et al.* (2011), who conduct experiments and surveys in Tajikistan after the end of the 1992-1997 civil war. They find that exposure to conflict reduced significantly trust within local communities.

In our theory, the channel linking war and trust is inter-ethnic trade. While the effect of international trade on both civil and international wars is well documented (see, e.g., Martin *et al.* 2008a and 2008b), it is more difficult to find systematic data on trade within-country between different communities. However, the evidence from a number of case studies conforms well with the predictions of our theory. We return to the discussion of the case-study evidence in section 7.

1.2 Related literature

Our paper connects to various streams of economic literature. The link between trust, social norms, specific investments and business relationships is emphasized by a large body of literature on contractual incompleteness. The salience of this issue in the context of cross-community trade is emphasized by Dixit (2003). In Hauk and Saez-Marti (2002) and Tabellini (2008) pro-social norms preventing opportunistic behavior are key to sustaining efficient trade.

Learning traps are related to the literature on herding, social learning, and informational cascades. This includes Banerjee (1992); Bikhchandani *et al.* (1992); Ely and Valimaki (2003); Fernandez (2012) and Piketty (1995). The connection with these papers is discussed in more detail in section 5.2. The theory is related also to the theoretical literature on supermodular games with strategic complementarities (Baliga and Sjostrom, 2004; Chamley, 1999 and 2004; Chassang and Padro-i-Miquel, 2010 and Cooper and John, 1988). While most of these papers emphasize the possibility of static multiplicity, in our paper we constrain parameters to yield a unique equilibrium under perfect information. The dynamic nature of the model of conflict is related to Yared (2010). The importance of luck and the persistent effects of negative shocks link our contribution to Acemoglu and Zilibotti (1997). Also related to our research are the recent papers by Aghion *et al.* (2010; 2011) focusing on the relation between public policy, on the one hand, and beliefs and norms of cooperation in the labor market, on the other hand.

In a recent paper, Acemoglu and Wolitzky (2012) propose a theory of conflict snowballs in which mistaken signals can trigger conflict between two groups. Their main focus is on war cycles, and specifically on how conflicts that are not driven by fundamental reasons may come to an end. In their model, as a bilateral conflict escalates, aggressive actions become uninformative, eventually inducing a group to experiment with cooperation. As long as the other group is not inherently bad, this move can bring the conflict to an end. Their model differs from ours in many respects - in particular, there is no explicit link between war and trade, which is the focal point of our analysis.

Our paper is related more generally to the economic literature on conflict. A set of prominent papers focus on institutional failures, such as weak state capacity and weak institutions (Besley and Persson, 2010, 2011; Fearon, 2005). In Besley and Persson (2011) the lack of checks and balances implies that rent-sharing strongly depends on who is in power, thereby strengthening incentives to fight. Acemoglu and Robinson (2012) discuss the importance of inclusive political and economic institutions to avoid civil wars. Poverty and natural resource abundance also have been found to fuel conflict, as the former reduces the opportunity cost of fighting, while the latter results in a larger "pie" that can be appropriated (Torvik, 2002; Collier and Hoeffler, 2004). The importance of ethnic polarization is emphasized, among others, by Esteban and Ray (2008, 2011), Caselli and Coleman (2012) and Rohner (2011). Some theories try to explain, as we do, why civil conflicts recur. Collier and Hoeffler (2004) argue that wars increase the availability of conflict-specific capital, such as cheap military equipment, that can be used for further conflicts. Acemoglu *et al.* (2010) argue that in weak states civilian governments select small and weak armies in order to prevent coups. However, this makes these states more vulnerable to insurgency and rebellion.

The plan of the paper is the following. Section 2 presents the benchmark model of inter-ethnic trade and conflict. Section 3 characterizes the perfect information equilibrium, while section 4 presents the equilibrium for asymmetric information. Section 5 extends the framework to a dynamic environment where beliefs are transmitted across generations, and derives the main results. Section 6 presents a number of extensions of the benchmark theory. Section 7 discusses some empirical evidence consistent with the predictions of the theory. Section 8 discusses some policy implications. Section 9 concludes. The proofs of all Lemmas, Propositions and Corollaries in Sections 3–5 are in Appendix I. The webpage appendixes A, B and C contain proofs not in the text, and additional technical material.

2 Model Environment

The model economy is populated by a continuum of risk-neutral individuals belonging to two *ethnic* groups (A and B), each of unit mass. The interactions between the two groups are described by a two-stage game. First, group A decides whether to wage war against group B. Next, cross-ethnic economic interactions (*trade*) occur. No economic decisions are made under the shadow of war. In the case of peace, each agent in group A is randomly matched to trade with an agent in group B.

Trade is modelled as a classic stag hunt game. If both trade partners in a match cooperate, a high economic surplus is generated, and each trader receives the payoff c. If both sides defect, each receives $d < c$. If only one side cooperates, the cooperator suffers a loss, and receives $d - l$, while the defector receives d. The matrix of material payoffs is:

	C	D
C	c, c	$d - l, d$
D	$d, d - l$	d, d

Absent other elements, the payoff matrix above describes a coordination game with multiple equilibria:

C-C is a Pareto-dominant Nash equilibrium, D-D is also a Nash equilibrium. Note that defection is a "safer" strategy: If one player perceives it unlikely that the partner cooperates, she will go for the safe payoff d from defection.

The matrix above describes only part of the surplus arising from inter-ethnic partnerships. The total utility includes, in addition, a psychological component related to the compliance with existing social norms of cooperation. The salience of such norms is assumed to be heterogeneous across individuals and groups. More formally, we denote by $\mathcal{P} \in \mathbb{R}$ the psychological benefit enjoyed by an agent from playing cooperatively, irrespective of his opponent's action. Then, the matrix of total payoffs is given by

$$\begin{array}{c|cc} & C & D \\ \hline C & c+\mathcal{P}_i, c+\mathcal{P}_j & d-l+\mathcal{P}_i, d \\ D & d, d-l+\mathcal{P}_j & d, d \end{array} \tag{1}$$

The heterogeneity reflects the fact that individuals enjoy cooperation to varying extents. Agents with high $\mathcal{P}$ exhibit strong civic norms. When $\mathcal{P} > l$, the norm is so strong that the player would even cooperate with a partner known to defect. However, $\mathcal{P}$ can be low and even negative, indicating *hatred,* i.e., pleasure from inflicting losses to a member of the other group. The distribution from which $\mathcal{P}$ is drawn is assumed to be group specific: for example, a more clannish group may be on average less inclined to cooperate with strangers.

More formally, $\mathcal{P}$ is assumed to be a continuous random variable, *i.i.d.* across agents in each group, and drawn from a probability density function (p.d.f.), $h^J : \mathbb{R} \to \mathbb{R}^+$, where $J \in \{A, B\}$, which is assumed to have no mass points. We denote by $H^J : \mathbb{R} \to [0,1]$ the corresponding cumulative distribution function (c.d.f.). Group A can be of two types: $h^A \in \{h^+, h^-\}$, and accordingly $H^A \in \{H^+, H^-\}$, where H^+ first-order stochastically dominates H^-. Since $\mathcal{P}$ reflects the propensity to cooperate with the other group, we say that group A is trustworthy (or *civic*) when $H^A = H^+$, and not trustworthy (or *uncivic*) when $H^A = H^-$. Instead, we assume that H^B has a unique realization. This is for tractability, as it avoids complications arising from a multi-dimensional learning process. We denote by $k \in \{-, +\}$ the type of group A. Note that throughout the paper we refer to a group's type as a particular cross-sectional distribution of attitudes towards cooperation, rather than as the propensities of specific individuals (which vary within each group). Therefore, the assumption that H^B has a unique realization means that group B's type is common knowledge, whereas group B ignores whether $H^A = H^+$ or $H^A = H^-$. It is useful to rescale some variables in order to simplify computations.

Notation 1 *Let (i)* $z \equiv c-(d-l)$; *(ii)* $\mathcal{L}_i \equiv l-\mathcal{P}_i$; *(iii)* $f^J(\mathcal{L}) \equiv h^J(l-\mathcal{L})$ *and* $F^J(\mathcal{L}) \equiv 1-H^J(l-\mathcal{L})$, *with* $J \in \{+, -, B\}$.

Intuitively, z is the payoff difference for a cooperator when its opponent switches from defecting to cooperating. $\mathcal{L}_i$ is the payoff difference (including psychological benefits) between cooperating and defecting, when the opponent turns out to be a defector. The group-specific p.d.f. and c.d.f. of $\mathcal{L}$ are simple transformations of the respective p.d.f. and c.d.f. of $\mathcal{P}$.

Finally, we introduce a technical assumption that is maintained throughout the rest of the paper.

Assumption 1 *There exists $\varepsilon > 0$ such that the p.d.f. $f^B(\mathcal{L}), f^+(\mathcal{L})$ and $f^-(\mathcal{L})$ are non-decreasing in the subrange $\mathcal{L} \in [-\varepsilon, z+\varepsilon[$.*

Assumption 1 requires that, at least in the interval $\mathcal{P} \in [-(c-d), l]$, there are no more people with high than with low civic norms. As explained below, this is a sufficient condition to guarantee existence and uniqueness of the equilibrium in the trade game.

2.1 Interpretation

The trade game with social norms, (1), is related to Hauk and Saez-Marti (2002) and Tabellini (2008). We assume, as they do, that individuals drawn from different populations are randomly matched to form business partnerships, and that there is individual heterogeneity in norms and taste for cooperation. However, in their models social norms of good conduct are endogenous and evolve according to parental cultural transmission. In our model, we take norms as exogenous and emphasize, instead, the effects of asymmetric information about a group's propensity to cooperate, or its *trustworthiness* in the other group's eyes. Cooperation is inherently risky, and defection can be motivated by the drive to play it safe against a distrusted partner. For instance, an efficient partnership may require one trader (the *producer*) to deliver goods to a partner (the *middleman*) in exchange for a deferred payment (e.g., the middleman can be credit constrained and unable to pay the producer before selling the good). However, if the middleman's group is reputed to be prone to defection, the producer may refrain from entering into the arrangement. Due to the strategic complementarity in the trade game, the public reputation of group A affects the behavior of group B: if group A is perceived to be little trustworthy, for instance, its members will expect that many agents in group B will defect, and will themselves defect.

Finally, we note that the normal-form game, (1), is susceptible to an alternative interpretation that abstracts from norms and psychological payoffs. Suppose that the success of an inter-ethnic business opportunity requires that both partners make a costly *pre-trade investment* (in physical or human capital) such as familiarizing oneself with the other group's language and customs, or building an inter-ethnic trade network. The investment cost is heterogeneous across agents and equal to $\mathcal{L}_i \equiv l - \mathcal{P}_i$. When both partners invest, they can trade and earn each a payoff c. Otherwise, there is no trade and agents receive a default payoff d. In this context, cooperation and defection correspond to investing and not investing, respectively, before random matching. The payoff matrix is observationally equivalent to (1).

While in our discussion we emphasize the notion of trade, which we believe to be especially important, the gist of our argument extends to a broader set of inter-ethnic social interactions within countries. These can include inter-ethnic marriages, public good provision in villages and communities, and more general situations involving bilateral cooperation.

2.2 Stochastic wars

We denote by $\mathcal{V}$ the net benefit of war accruing to group A, comprising both the value of the resources seized in war and any military or psychological costs associated with war. While, for simplicity, we do not model explicitly the conflict dynamics, $\mathcal{V}$ can be interpreted as the expected payoff of a war with an uncertain outcome. We make the important natural assumption that $\mathcal{V}$ is a stochastic variable.[8] Shocks to $\mathcal{V}$ reflect shifts in a variety of factors affecting the cost and the risk of military operations. These include the amount of expropriable resources (or, to the opposite, the losses in case of defeat), the internal cohesion of the group, the state of organization and morale of the army, and international factors such as the willingness of the international community to impose sanctions on the aggressor. An important assumption is that the realization of $\mathcal{V}$ cannot be (perfectly) observed by group B. For instance, the state of efficiency of the Soviet army after the 1936 purges was uncertain to its enemies. Its weaknesses became patent (and even exaggerated) after the poor performance of the Red Army in the Winter War against Finland.

The decision to wage war is driven by the comparison between the realization of $\mathcal{V}$ and the aggregate *trade surplus* accruing to group A in case of peace, denoted by S^k with $k \in \{+, -\}$. Note that S^k corresponds to the sum of all the payoffs from trade obtained by the whole population of group A, given the payoff matrix 1, the distribution of psychological benefits $\mathcal{P}$ from cooperation, and the proportions in both groups of agents cooperating. S^k is endogenous and will be determined in equilibrium. We will show below that S^k is bounded, $S^k \in [S^{\min}, S^{\max}]$. Although different members of group A might disagree about the desirability of war, depending on the individual realization of $\mathcal{P}$, we assume that group A has access to internal transfer mechanisms that compensate losers. Thus, the decision to go to war weighs the expected benefit $\mathcal{V}$ against the total opportunity cost of war to group A.[9]

Finally, to ease tractability, we discretize the support of the distribution of $\mathcal{V}$, allowing for three possible realizations, $\mathcal{V} \in \{V_L, V, V_H\}$.[10] We make the following assumption:

Assumption 2 $V_L < S^{\min} < V < S^{\max} < V_H$.

Under the realization V_H, the net benefit of war is so high that waging war is optimal, for any feasible trade surplus.[11] Conversely, V_L corresponds to a situation in which the cost of waging war is prohibitive, due, for example, to a strong international pressure or a failure in the collective action

[8] Our assumptions echo the recent literature that views the onset of war as "stochastic" (Gartzke, 1999; Caselli *et al.*, 2012), in particular due to stochastic shocks to coordination costs of rebellion (Collier and Hoeffler, 1998), or to rebel capability (Gates, 2002; Buhaug *et al.*, 2009).

[9] A voting process *à la* Lindbeck and Weibull (1987) would yield the same outcome.

[10] The main results of the paper can be extended to a model where $\mathcal{V}$ is drawn from a continuous distribution. However, the analysis is more involved. See Rohner *et al.* (2012b).

[11] $\mathcal{V} = V_H$ can reflect also a temporary explosion of hatred (Gurr, 1970) or the capture of the political process by a biased political elite (Jackson and Morelli, 2007).

process. We refer to the realizations V_H and V_L as a *war shock* and a *peace shock*, respectively, with associated probabilities λ_W and λ_P such that $\lambda_P + \lambda_W < 1$. The intermediate realization, V, will be referred to as *business as usual* (BAU). This state captures the situation in which economic considerations, and in particular the state of the trade surplus, make group A swing for either war or peace.

3 Perfect information equilibrium

To establish a benchmark, we first consider the case in which group A's type is public knowledge. In this case, war spoils trade but conveys no information. We solve the game backwards, starting from the trade game. Let n_J denote the probability that a random member of group $J \in \{A, B\}$ plays cooperatively. The expected payoff of cooperation for agent i in group A is $\mathcal{P}_i + c \times n_B + (d-l) \times (1-n_B)$, whereas the safe payoff of defection is d. Using the definitions of z and $\mathcal{L}_i$ provided above, agent i chooses cooperation if $\mathcal{L}_i \leq n_B \times z$. Conversely, agent j in group B cooperates if $\mathcal{L}_j \leq n_A \times z$. Hence, the Nash equilibrium conditional on group A's type, $k \in \{+, -\}$, is given by the fixed point

$$\left\{n_A^k, n_B^k\right\} = \left\{F^k(zn_B^k), F^B(zn_A^k)\right\} \tag{2}$$

Due to the strategic complementarity, the trade game can feature multiple Nash equilibria. Since the mechanism underlying coordination failures in this environment is well understood (see Cooper and John 1988), we abstract from it in this paper. In particular, Assumption 1 is sufficient to ensure that a unique Nash equilibrium exists under perfect information. This restriction allows us to focus more sharply on the dynamic interaction between belief formation and warfare.

The trade surplus accruing to group A under peace is given by

$$\begin{aligned} S^k &\equiv n_A^k \times \left[n_B^k \times c + \left(1-n_B^k\right) \times (d-l)\right] + \int_{-\infty}^{zn_B^k} (l-\mathcal{L})\, dF^k(\mathcal{L}) + \left(1-n_A^k\right) \times d \\ &= d + z \times \left(n_A^k \times n_B^k\right) - \int_{-\infty}^{zn_B^k} \mathcal{L} dF^k(\mathcal{L}). \end{aligned} \tag{3}$$

This expression has a simple interpretation: (i) d is the basic surplus from defection; (ii) $z \times \left(n_A^k \times n_B^k\right)$ is the extra surplus generated by mutual cooperation, where $n_A^k \times n_B^k$ is the measure of successful trade relationships; (iii) $\int_{-\infty}^{zn_B^k} \mathcal{L} dF^k(\mathcal{L})$ is the total loss (net of aggregate psychological benefit from following the cooperative norm) suffered by cooperators who are matched with defectors.

Proposition 1 *Under Assumption 1 and perfect information, the Nash Equilibrium of the trade game conditional on peace exists and is unique. Denoting by $k \in \{+, -\}$ group A's type, the equilibrium proportions of cooperators are given by $\{n_A^k, n_B^k\}$ as given by equation (2) where, $n_A^- \leq n_A^+$ and $n_B^- \leq n_B^+$. The equilibrium trade surplus accruing to group A is $S^{\min} = d + zn_A^- n_B^- - \int_{-\infty}^{zn_B^-} \mathcal{L} dF^-(\mathcal{L})$ for the uncivic type and by $S^{\max} = d + zn_A^+ n_B^+ - \int_{-\infty}^{zn_B^+} \mathcal{L} dF^+(\mathcal{L})$ for the civic type, where $S^{\min} \leq S^{\max}$. The probability of war is larger for the uncivic type $(1-\lambda_P)$ than for the civic type (λ_W).*

Since $S^{\min} < V < S^{\max}$, under BAU the uncivic type wages war while the civic type retains peace. Hence, when group A is civic (uncivic) war occurs with the constant probability λ_W $(1 - \lambda_P)$.

3.1 Welfare implications

The Nash equilibrium is generically suboptimal because of two interacting sources of inefficiency. First, conditional on peace, trade cooperation is inefficiently underprovided. This result echoes the analysis of supermodular games with random matching and externalities in Cooper and John (1988). In our model, an increase in the proportion of cooperators in group A (B) has a direct positive externality on group B (A). Moreover, by inducing more agents in group B (A) to cooperate, it has a beneficial feedback effect on group A (B). A "social planner" maximizing the sum of the welfare of the two groups would attain a more efficient outcome by inducing more cooperation. This result holds true irrespective of group A's type.[12] Second, war imposes welfare losses on group B (in the form of physical destruction, human lives, etc.) that are not internalized by the unilateral war decision of group A.[13] If we denote such costs by V_B, then war is always globally inefficient whenever $\mathcal{V} < (S^k + S^B + V_B)$, where S^B is group B trade surplus. Perhaps more interestingly, the Nash equilibrium may feature too many wars even from the self-regarding standpoint of group A, due to the inefficiency in the trade game. Namely, more cooperation could lead group A to retain peace in situations where the Nash equilibrium does prescribe war. In particular, it is easy to construct examples in which the efficient level of cooperation in the trade game would induce both the civic and the uncivic type to preserve peace under BAU, and yet the Nash equilibrium features war when A is uncivic. In other words, the underprovision of cooperation induces a second layer of inefficiency in the form of too frequent wars.

4 Asymmetric Information

From now on in the paper, we assume that group B can observe neither group A's type nor the realization of $\mathcal{V}$. In this environment, war and peace become public signals of group A's type. Beliefs are common knowledge, and are updated using Bayes' rule after the observation of war or peace.

We denote by $\pi_{-1} \in [0, 1]$ the prior belief held by group B that group A's type is civic (or, less formally, the extent to which group B *trusts* group A) and by π_W and π_P, the posterior beliefs conditional on war and peace, respectively. It is often convenient to take the likelihood ratios: $r_{-1} \equiv \pi_{-1}/(1 - \pi_{-1})$ and $r_s \equiv \pi_s/(1 - \pi_s)$, for $s \in \{W, P\}$. Finally (σ^-, σ^+) denote the probability that peace is chosen under BAU by the uncivic and civic type, respectively.

The timing of the game is the following.

[12] More formally, the efficient solution prescribes that $zn_B^{k*}\left(1 + f^B\left(\bar{\mathcal{L}}_A^k\right)\right) = \bar{\mathcal{L}}_A^k$ and $zn_A^{k*}\left(1 + f^k\left(\bar{\mathcal{L}}_B^k\right)\right) = \bar{\mathcal{L}}_B^k$, where both n_A^{k*} and n_B^{k*} are larger than in the decentralized equilibrium. In the case of uniform distributions, this yields $n_A^{k*} = F^k\left(zn_B^{k*}\left(1 + \bar{f}^B\right)\right) > F^k\left(zn_B^k\right)$ and $n_B^{k*} = F^B\left(zn_B^{k*}\left(1 + \bar{f}^k\right)\right) > F^B\left(zn_A^k\right)$, where $\bar{f}^k$ and $\bar{f}^B$ are the constant density functions. See Appendix C for details.

[13] We do not allow for cross-group transfers that could reduce the scope of wars. As usual, such transfers entail commitment problems (cf. Fearon, 1995; Powell, 2006). In the case of *ex-ante* transfers, group A could cash in, and still wage war. In the case of *ex-post* transfers, group B could refuse to pay once the risk of war is gone.

1. The *war stage*: agents in group B receive the common prior belief r_{-1}, agents in group A observe the state $\mathcal{V}$. Conditional on its own type, $k \in \{-,+\}$, group A decides to wage war with probability $1-\sigma^k$.

2. The *trade stage*: agents in group B update their beliefs. If there is peace, agents are randomly matched, gains from trade are realized, and consumption occurs.

The equilibrium concept is Perfect Bayesian Equilibria (PBE).

Definition 1 *A strategy for an agent in population A specifies for each possible types, $k \in \{+,-\}$ and for each state $\mathcal{V} \in \{V_L, V, V_H\}$, conditional on public beliefs, a war action ("wage war" or "keep peace"), and, for each possible realization of the idiosyncratic preference shock, $\mathcal{P}$, a trade action ("cooperate" or "defect"). A strategy for an agent in population B specifies a trade action ("cooperate" or "defect") for each of the possible realizations of the idiosyncratic preference shock, $\mathcal{P}$, conditional on public beliefs. A PBE is a strategy profile, a belief system and a triplet $\left(n_A^-, n_A^+, n_B\right) \in [0,1]^3$ such that: (i) in the trade continuation game all agents choose their action so as to maximize the expected pay-off conditional on the posterior likelihood ratio of beliefs after peace, r_P, and the realization of the preference shock $\mathcal{P}$; $\left(n_A^-, n_A^+, n_B\right)$ yields the associated measure of agents who optimally cooperate in group A for each type, $k \in \{+,-\}$, and for group B, respectively. (ii) group A chooses the probability $1-\sigma^k$ of waging war on group B so as to maximize the total expected utility of its members, given group A's type (k), the shock $\mathcal{V} \in \{V_L, V, V_H\}$ and the prior likelihood ratio of beliefs, r_{-1}; (iii) beliefs are updated using Bayes' rule.*

4.1 Trade game

We solve the PBE backwards, starting from the Nash equilibrium of the trade game under peace. The reaction function of group A, which is fully informed, continues to be given by $n_A^k = F^k(zn_B)$, with $k \in \{+,-\}$. Since n_A depends on type k which group B does not observe, group B's reaction function becomes $F^B(z\mathbb{E}_B(n_A \mid \pi_P)) = F^B\left(z\left[\pi_P n_A^+ + (1-\pi_P)\, n_A^-\right]\right)$. Rescaling the belief π_P in term of its likelihood ratio r_P, we obtain that the equilibrium proportions of cooperators in the two groups satisfy the following fixed-point equation:

$$\left\{n_A^-, n_A^+, n_B\right\} = \left\{F^-(zn_B), F^+(zn_B), F^B\left(\frac{r_P}{1+r_P} zF^+(zn_B) + \frac{1}{1+r_P} zF^-(zn_B)\right)\right\} \tag{4}$$

Proposition 2 characterizes the set of Nash equilibria of the trade game.

Proposition 2 *Under Assumption 1 and given a likelihood ratio of posterior beliefs $r_P \in [0,+\infty)$, the Nash Equilibrium of the trade game conditional on group A's type $k \in \{+,-\}$ exists and is unique. The equilibrium proportions of cooperators are given by $\{n_A^-(r_P), n_A^+(r_P), n_B(r_P)\}$, implicitly defined by equation (4). $n_A^-(r_P)$, $n_A^+(r_P)$ and $n_B(r_P)$ are continuous, weakly increasing and lie strictly within*

the unit interval. Moreover, $n_A^-(r_P) \leq n_A^+(r_P)$.
The equilibrium trade surplus accruing to group A, $S^k(r_P)$, *is given by*

$$S^k(r_P) \equiv d + zn_A^k(r_P)\, n_B(r_P) - \int_{-\infty}^{zn_B(r_P)} \mathcal{L}\; dF^k = d + \int_{-\infty}^{zn_B(r_P)} F^k(\mathcal{L})\; d\mathcal{L}. \tag{5}$$

$S^k(r_P)$ *is weakly increasing in* r_P. *Moreover,* $S^-(r_P) \leq S^+(r_P)$.

Proposition 2 shows that trust has a social value: optimistic prior beliefs about group A's type (large r_P) induce more cooperative behavior from both groups (large n_A and n_B), and a larger trade surplus, due to the strategic complementarity: the trust in group A makes group B more cooperative; this in turn makes group B more trustworthy in group A's eyes, further enhancing the cooperation of group A, and so on. To the opposite, a vicious circle of low trust and trustworthiness makes cooperation unravel, reducing the trade surplus. Figure 1 plots the equilibrium trade surplus (conditional on group A's actual type) against the posterior belief, $\pi_P = r_P/(1+r_P)$, in the case of a uniform distribution of $\mathcal{P}$.[14]

4.2 War decision and PBE

In this subsection, we analyze the group A's war decision, based on the comparison between the trade surplus, $S^k(r_P)$, and the realization of the stochastic benefit of war, $\mathcal{V}$. Since the trade surplus depends on posterior beliefs, we must characterize the mapping from prior to posterior. Bayes' rule yields

$$\ln r_P(r_{-1}) = \ln r_{-1} + \ln \frac{\lambda_P + (1-\lambda_W-\lambda_P)\,\sigma^+(r_{-1})}{\lambda_P + (1-\lambda_W-\lambda_P)\,\sigma^-(r_{-1})}, \tag{6}$$

$$\ln r_W(r_{-1}) = \ln r_{-1} - \ln \frac{1-\lambda_P - (1-\lambda_W-\lambda_P)\,\sigma^-(r_{-1})}{1-\lambda_P - (1-\lambda_W-\lambda_P)\,\sigma^+(r_{-1})}, \tag{7}$$

where the optimal choice of maintaining peace is given by

$$\sigma^k(r_{-1}) = \begin{cases} 0 & \text{if} \quad S^k(r_P(r_{-1})) < V \\ \in [0,1] & \text{if} \quad S^k(r_P(r_{-1})) = V \\ 1 & \text{if} \quad S^k(r_P(r_{-1})) > V \end{cases}. \tag{8}$$

Note that, since S^k depends on r_P, which is endogenous to the war/peace decision, the characterization of the optimal choice of σ^k involves a fixed-point problem that can yield multiple solutions.

Proposition 3 establishes the existence of the PBE. The proof follows immediately from Proposition 2 and is omitted.

Proposition 3 *A PBE exists and is fully characterized by the set of equations (4), (5), (6), (7), (8), given a likelihood ratio of prior beliefs* $r_{-1} \in [0,+\infty)$.

[14] We set $z = 0.9$, $d = 0$, and assume $F^B \sim [0,1]$, $F^+ \sim [-0.25, 1]$, $F^- \sim [0, 1.25]$. The geometric representation entails some slight abuse of notation, as S^k was defined to be a function of the likelihood ratio r_P rather than of π_P. We prefer this representation since $\pi_P \in [0,1]$, and we can display the range of variation of the surplus for any belief.

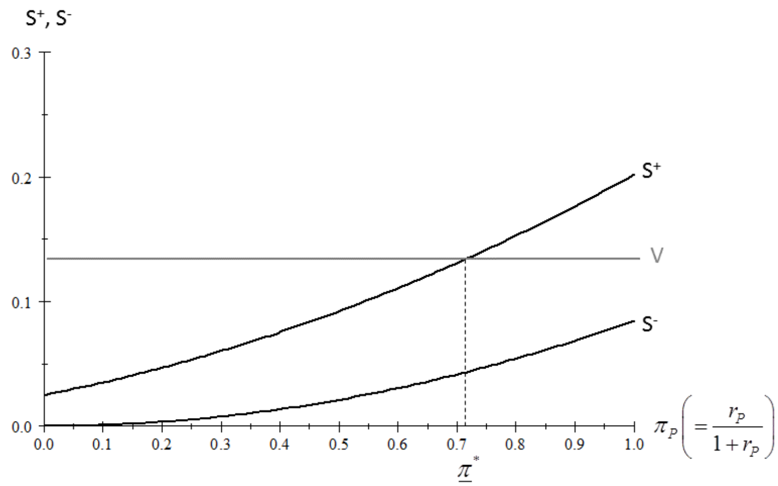


Figure 1: Surplus from trade as function of the posterior belief, and war benefit under BAU.

The posterior belief is independent of the observation of war or peace whenever $\sigma^+ = \sigma^-$. This may happen when priors are either very optimistic or, to the opposite, very pessimistic. In particular, if $V < S^-(r_P)$, then both types retain peace under BAU ($\sigma^+(r_{-1}) = \sigma^-(r_{-1}) = 1$), whereas if $V > S^+(r_P)$, then both types wage war under BAU ($\sigma^+(r_{-1}) = \sigma^-(r_{-1}) = 0$). On the contrary, the observation of war or peace is informative whenever $S^-(r_P) \leq V \leq S^+(r_P)$ – with one inequality being strict. In this case, the uncivic type would wage war whereas the civic type would preserve peace ($\sigma^+(r_{-1}) = 1$ and $\sigma^-(r_{-1}) = 0$). Thus, peace strengthens the trust of group B towards group A, while war undermines it. More formally, $S^-(r_P) \leq V \leq S^+(r_P) \Leftrightarrow r_P > r_{-1} > r_W$. We refer to this situation as an informative PBE.

Definition 2 *Given r_{-1} a PBE is "informative" iff $\sigma^+(r_{-1}) > \sigma^-(r_{-1})$, or identically iff $r_P(r_{-1}) > r_{-1} > r_W(r_{-1})$. A PBE is "uninformative" iff $\sigma^+(r_{-1}) = \sigma^-(r_{-1})$, or identically iff $r_P(r_{-1}) = r_{-1} = r_W(r_{-1})$.*

Throughout the rest of the paper, we maintain that $V > S^+(0)$. This restriction sharpens the focus on the more interesting case, in which pessimistic beliefs can induce even the civic type to wage war. For expositional purposes, it is convenient to consider two cases separately: high value of war ($S^-(\infty) < V < S^+(\infty)$) and low value of war ($V \leq S^-(\infty) < S^+(\infty)$).[15] High value of war will be our main case, since it delivers the main economic insights in a simpler way. In this case, the uncivic type always wages war under BAU, irrespective of public beliefs. The case of low value of war, where the uncivic type may find it optimal to mimic the civic type by maintaining peace under BAU, is discussed in Appendix B. Figure 1 represents the case of high value of war: $S^- < V$ for all π_P, and there exists $\underline{\pi}^*$ such that $S^+ > V$ if $\pi_P > \underline{\pi}^*$ and $S^+ < V$ if $\pi_P < \underline{\pi}^*$.

[15] In the case of low value of war, we assume also that $S^+(0) < S^-(\infty)$. The set of parameters for which this inequality holds is non-empty.

We now discuss the equilibrium mapping from prior to posterior.

Lemma 1 *Assume $V > S^{+}(0)$ and $S^{-}(\infty) < V < S^{+}(\infty)$. Let $\underline{r}^{*}$ be such that $V = S^{+}(\underline{r}^{*})$, and let $\underline{r} \equiv \lambda_P \underline{r}^{*}/(1-\lambda_W)$. For $r_{-1} \leq \underline{r}$ the PBE is unique and uninformative. For $r_{-1} \in [\underline{r}, \underline{r}^{*}]$ there are multiple PBE. For $r_{-1} > \underline{r}^{*}$ the PBE is unique and informative.*

$\underline{r}^{*} \equiv \underline{\pi}^{*}/(1-\underline{\pi}^{*})$ is the threshold posterior likelihood ratio below which both types wage war under BAU. Uninformative PBE are associated with pessimistic priors. Intuitively, when trust is low, trade opportunities are scant and both the civic and the uncivic type wage war under BAU. When $r_{-1} \in [\underline{r}, \underline{r}^{*}]$, the mapping (8) yields multiple PBE: one uninformative and two informative ones (one of which involves mixed strategies). While none of our results depends on a specific selection criterion, we make the following convenient assumption.

Assumption 3 *In the range of prior beliefs such that multiple PBE exist, the most informative equilibrium is selected.*

Since our analysis emphasizes the possibility that economies fall into uninformative equilibria, this is the most conservative selection criterion.

In summary, this section has established the existence and characterization of the set of PBE (Proposition 3). It has, in addition, introduced a formal distinction between informative and uninformative PBE and established that for a pessimistic prior the PBE is unique and uninformative, while for an optimistic prior the PBE is unique and informative, and for a range of intermediate beliefs there are multiple equilibria (Lemma 1). For this latter case, we provide a simple selection criterion that simplifies the analysis of the dynamic model.

5 The dynamic model

In this section, we extend the analysis to a dynamic economy populated by non-altruistic overlapping generations of two-period lived agents. The purpose here is to characterize the endogenous intertemporal link between stochastic wars and beliefs, which is the main contribution of our paper. In the first period of their lives agents make no economic choice, and acquire a common prior belief, based on the public history regarding warfare.[16] In the second period agents make all economic decisions. After group A decides whether or not to wage war, agents in group B update their beliefs. In the case of peace, agents are randomly matched to trade. We assume that the information set of young agents comprises *only* the history of warfare. In particular, we rule out that young agents can observe the success of inter-ethnic trade during peace.[17]

[16] Since young agents are passive and earn no pay-off, one could interpret the model alternatively in terms of a sequence of one-period lived agents.

[17] If agents in group B could observe the average success of current inter-ethnic trade, they could attain a perfect inference of group A's type. Information about the success of inter-community trade in reality is sparse, and difficult to collect or to distinguish from intra-ethnic trade. In the benchmark model discussed in this section, we also rule out that

The dynamics of beliefs drives the stochastic process of war, peace and trade. We denote by r_0 the common prior of the adult generation at time zero.

Definition 3 *A Dynamic Stochastic Equilibrium (DSE) is a sequence of PBE with an associated sequence of beliefs such that, given an initial likelihood ratio r_0, the posterior likelihood ratio at $t-1$ is the prior likelihood ratio at t, for all $t \geq 0$.*

Combining equations (6)-(8), Lemma 1, and Assumption 3, the equilibrium law of motion of the likelihood ratio of beliefs is given by the following stochastic difference equation:

$$\ln r_t = \begin{cases} \ln r_{t-1} & \text{if} \quad r_{t-1} \in [0, \underline{r}] \\ \ln r_{t-1} + (1 - \mathbb{W}_t) \ln\left(\frac{1-\lambda_W}{\lambda_P}\right) - \mathbb{W}_t \ln\left(\frac{1-\lambda_P}{\lambda_W}\right) & \text{if} \quad r_{t-1} > \underline{r} \end{cases} \tag{9}$$

where $\mathbb{W}_t \in \{0, 1\}$ is a random variable taking on the unit (zero) value if there is war (peace) at date t. Also, recall that $1 > \lambda_W + \lambda_P$. In the low-trust region, $r_{t-1} \in [0, \underline{r}]$, the probability of peace is low ($\mathbb{P}(\mathbb{W}_t = 0) = \lambda_P$) and beliefs are stationary irrespective of group A's type. The observation of a peace event is attributed rationally to a peace shock. In contrast, in the high-trust region, $r_{t-1} > \underline{r}$, the probability of peace is low ($\mathbb{P}(\mathbb{W}_t = 0) = \lambda_P$) if A is uncivic, and high ($\mathbb{P}(\mathbb{W}_t = 0) = 1 - \lambda_W$) if A is civic. In addition, the equilibrium is informative and peace leads – through Bayesian updating – to an increase in the belief that group A is civic. The stochastic process (9) is represented in Figure 2, which is zoomed around the threshold $\underline{r}$.

Suppose that group A is civic, implying that peace and high cooperation would prevail under perfect information (Proposition 1). Under imperfect information, this low-conflict outcome is an equilibrium only if trust is sufficiently high, $r_t > \underline{r}$. Moreover, an unlucky sequence of war shocks can spoil trust forever. As soon as r_t crosses (from above) the threshold $\underline{r}$, group A starts waging war under BAU even though it is civic, and the probability of war jumps from λ_W to $1 - \lambda_P$. Group B rationally stops updating its beliefs, and no peace spell – no matter how long – can restore trust. The following Proposition summarizes the formal characterization of the dynamic equilibrium (proof in the text).

Proposition 4 *Assume that $V > S^+(0)$ and $S^-(\infty) < V < S^+(\infty)$. Then, the DSE is characterized by Proposition 3 and by the law of motion (9). If group A is uncivic, the probability of war is $\mathbb{P}(\mathbb{W}_t = 1) = 1 - \lambda_P$. If group A is civic, the probability of war is $\mathbb{P}(\mathbb{W}_t = 1) = 1 - \lambda_P$ if $r_{t-1} < \underline{r}$, and $\mathbb{P}(\mathbb{W}_t = 1) = \lambda_W$ if $r_{t-1} \geq \underline{r}$.*

5.1 War traps and long-run distribution

The next step is to characterize the long-run equilibrium outcome. We start by defining a *war trap.*

young agents can observe private signals (e.g., a finite sample of trade outcomes). Otherwise, agents would enter the second period of their lives with heterogenous beliefs, and the dynamic model would become intractable. We relax this assumption in section 6.2, where we assume that cohorts are connected through dynastic links, and allow information acquired by traders to be transmitted within families.

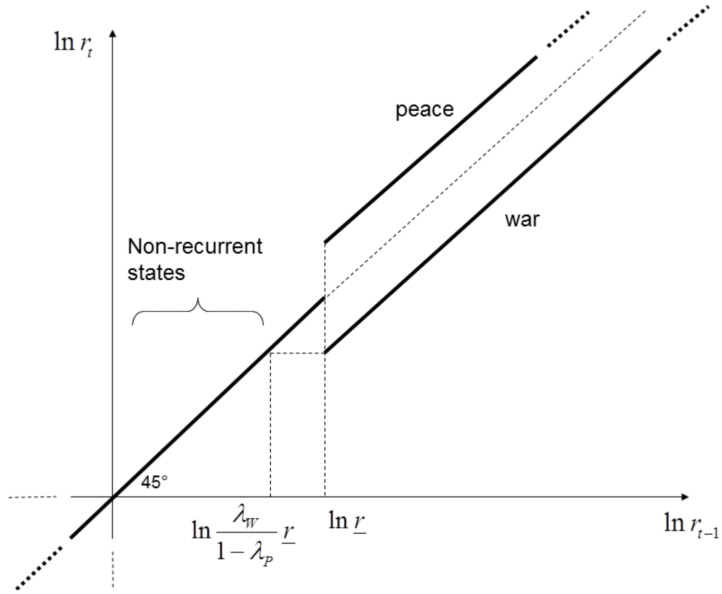


Figure 2: Stochastic law of motion of beliefs.

Definition 4 *A "war trap" is a set of states,* $\Omega_{TRAP} \subset \mathbb{R}^+$*, such that, for all* $r_t \in \Omega_{TRAP}$*, group A strictly prefers war under BAU irrespective of its type, and beliefs are stationary:* $\forall s \geq t, r_s = r_t$*.*

It follows immediately from (9) that $\Omega_{TRAP} = [0, \underline{r})$. Whatever the state of beliefs, $r_t < \infty$, the economy falls into the war trap with a positive probability, since a finite number of war shocks can plunge trust below the threshold $\underline{r}$. Once the economy is in the trap, beliefs are stationary, trade is scant and war is frequent. Note that both war and peace shocks are essential for a war trap to exist. If only war shocks existed, peace would become a perfectly revealing signal since, contrary to the civic type, the uncivic type would never retain peace under BAU. Thus, the civic type could reveal his type by not waging war under BAU, and the equilibrium would be identical to the case of perfect information.

Since the dynamics have no rest point in the informative equilibrium region, one might wonder whether the economy is deemed to fall into the war trap in the long run. The answer to this question depends on whether group A is civic or uncivic.

Proposition 5 *Assume that* $V > S^+(0)$, $S^-(\infty) < V < S^+(\infty)$, *and* $r_0 > \underline{r}$*.*

(i) If group A is uncivic, then the DSE enters the war trap *in finite time with probability one.*

(ii) If group A is civic, then the DSE enters the war trap *in finite time with probability* $\mathbb{P}_{TRAP} > 0$*, and stays out of the war trap forever with probability* $1 - \mathbb{P}_{TRAP} > 0$*. If the economy stays out of the trap, the DSE converges to perfect learning, i.e.,* $r_t \to \infty$*, and war incidence stays permanently low,* $\mathbb{P}(\mathbb{W}_t = 1) \longrightarrow \lambda_W$*.*

(iii) The probability $\mathbb{P}_{TRAP}$ *has the following bounds:*

$$0 < \frac{\lambda_W}{1 - \lambda_P} \frac{\underline{r}}{r_0} < \mathbb{P}_{TRAP} \leq \frac{\underline{r}}{r_0} < 1. \tag{10}$$

If group A is civic, the economy averts the trap in the long run with a positive probability. However, if group A is uncivic, the war trap is attained almost surely. The proof is based on the Martingale Convergence Theorem, ensuring that the stochastic belief π_t converges almost surely to a limit. This limit – as the proof shows – cannot lie in the interior of the informative region, thus either π_t enters the trap or $\lim_{t\to\infty} \pi_t = 1$. In order for $\lim_{t\to\infty} \pi_t = 1$, the economy must remain forever in the informative region. This possibility can be ruled out if group A is uncivic, since the Strong Law of Large Numbers would then imply that group B could observe an infinite sample of realizations of the war/peace process, and eventually learn the truth, i.e., that group A is uncivic. This would cause a contradiction. On the contrary, when group A is civic, group B can learn asymptotically the truth (i.e., $\lim_{t\to\infty} \pi_t = 1$) with positive probability. However, this is not the sole possible outcome. Alternatively, the economy can fall into the war trap with positive probability. So, the long-run fate of the economy hinges on luck, or the realization of the stochastic process of peace/war (similarly to Acemoglu and Zilibotti, 1997).[18] Proposition 5 establishes a lower and an upper bound to the probability that the economy falls into the trap. However, in a particular case, we can obtain an exact characterization of $\mathbb{P}_{TRAP}$:

Corollary 1 *Assume that group A is civic,* $V > S^+(0)$, $S^-(\infty) < V < S^+(\infty)$, $\lambda_W = \lambda_P = \lambda$ *and* $r_0 > \underline{r}$. *Then, the probability that the economy falls into the war trap is* $\mathbb{P}_{TRAP} = \left(\frac{\lambda}{1-\lambda}\right)^{\Delta(r_0)}$, *where* $\Delta(r_0) \equiv \lceil (\ln r_0 - \ln \underline{r}) / (\ln(1-\lambda) - \ln \lambda) \rceil$.

Note that $\Delta(r_0)$ counts the number of war episodes minus that of peace episodes which is necessary in order for r to cross the threshold $\underline{r}$, given the initial belief r_0. Intuitively, $\mathbb{P}_{TRAP}$ decreases with r_0. Thus, the probability of an economy falling into a war trap increases (falls) after each war (peace) episode, since this reduces (increases) r. Our theory predicts, then, that war is endogenously persistent: each conflict increases the probability of future conflicts.

When group A is civic, war traps are inefficient outcomes, even relative to the second-best implemented under perfect information (where, recall, peace is retained under BAU). There are too many wars and, in addition, there is less cooperation than under perfect information. When the economy converges to perfect learning, the equilibrium is asymptotically identical to the perfect information outcome, although there is inefficiency along transition. Conversely, when A is uncivic, war traps may yield higher welfare than the perfect information equilibrium. In the trap, group B fails to learn perfectly about the uncivicness of group A. Consequently, peace spells are characterized by more economic cooperation than under perfect information. Since cooperation tends to be suboptimally low (see section 3.1), imperfect learning in this case can improve the welfare of both groups.

[18] The war decision entails an intergenerational spillover, as war depletes trust and harms future generations. In Appendix B, we show that an equilibrium isomorphic to that of Propositions 4-5 can be sustained as the Markov Perfect Equilibrium of an extended model where the decision to stage war incorporates an altruistic concern towards the next generation.

5.2 War traps and related literature

In our model, the key for the existence of war traps is that, in some range of priors, the equilibrium has a separating nature (i.e., different types take different actions – waging war or retaining peace), whereas for another range of priors the equilibrium has a pooling nature. In a pooling equilibrium there is no informative public signal. The dynamics can push the equilibrium towards the pooling low-trust equilibrium, even though group A is civic. The reason is that, although beliefs are updated rationally, the updating process is subject to disturbances over time. Signal jamming due to war and peace shocks makes it possible that agents update beliefs "in the wrong direction". While a long enough further sampling eventually would correct temporary mistakes and lead to correct learning, sampling *de facto* ends (in the sense that agents stop receiving informative signals) as soon as the economy enters the pooling equilibrium region.

While learning traps are not per se a novel finding, their source and mechanism here are different from those described in the existing literature, to the best of our knowledge. In models of informational cascades (see, e.g., Bikhchandani *et al.* 1992), agents make decisions sequentially, having access to private information and a public signal. They may end up in an informational cascade where each agent rationally ignores her private information and conforms to the behavior of the majority. In contrast, in our model there is no private information. In Piketty (1995), learning traps are driven by costly experimentation and imperfect common knowledge. In his paper, people with identical preferences, but different sampling histories, may end up having persistently heterogeneous beliefs about the cost of redistribution. The mechanism is again very different from ours, where all information is public. Experimentation is key also in Aghion *et al.* (2011). Chamley (1999) studies coordination and social learning in a dynamic setting with uncertainty and strategic complementarities. The environment is different from ours insofar as the game features multiple equilibria under perfect information, whereas in our stag hunt game the equilibrium is always unique. The main difference in the social learning process is that in his model the dynamics are driven by exogenous changes in unobservable fundamentals, while in our model the dynamics depend on the individual decisions of players that determine the surplus from trade. While in our setting learning can break down fully, this is never the case in Chamley's model. In the extension of section 6.1, we move closer to Chamley's setting by studying the possibility that group A's type changes stochastically.

5.3 War persistence

Our analysis is motivated by the evidence that civil wars feature hysteresis, namely, a conflict today increases the probability of future conflicts (see Section 1.1). In the model presented above, each war episode increases smoothly the long-term frequency of conflicts.[19] Similarly, one can compute the

[19] If we define the long run frequency of conflicts (when group A is civic) as $\mathbb{F}(r_t) \equiv \lim_{T\to+\infty} \mathbb{E}\left[\frac{1}{T}\sum_{s=t+1}^{T} \mathbb{W}_s \mid k=+, r_t\right]$, then Proposition 5 (replacing r_0 by r_t) implies that $\mathbb{F}(r_t) = \mathbb{P}_{TRAP}(r_t) \times (1-\lambda_P) + [1-\mathbb{P}_{TRAP}(r_t)] \times \lambda_W$. Clearly, $\mathbb{F}(r_t)$ is decreasing in r_t, and each war (peace) episode, by decreasing (increasing) r_t, increases (reduces) the frequency of future conflict.

effect of a war at t on the probability of conflict at $t+T$, for $\infty > T \geq 1$ (as long as group A is civic). For each lead T, there exists a threshold $\underline{r}_T \equiv \underline{r} \times \left(\frac{1-\lambda_P}{\lambda_W}\right)^T$ such that, for all $r_t \in]\underline{r}, \underline{r}_T[$, a war at t increases the probability of war at $t+T$. The nature of the persistence becomes somewhat stark if one looks at the short-run dynamics ($T=1$): then, the probability of war at $t+1$ increases discretely after a war at t only if r_t is in a right-hand neighborhood of $\underline{r}$ such that the war immediately pushes the economy into the trap. This is not a robust feature of the theory. Rather, it hinges on the simplifying assumption that the stochastic process $\mathcal{V}$ has a discrete support. If one assumes that $\mathcal{V}$ is drawn from a continuous distribution, the theory predicts a smoother form of short-run persistence.

In this section, we sketch the argument for the case of a particular continuous distribution of $\mathcal{V}$.[20] Suppose $\mathcal{V}$ is distributed uniformly in the interval $[V_L, V_H]$ where $V_L < S^-(0) < S^+(\infty) < V_H$. Note that the range of realizations $\mathcal{V} > S^+(\infty)$ and $\mathcal{V} < S^-(0)$ can be interpreted as war and peace shocks, respectively. Bayes' rule implies that

$$r_P(r_t) = \frac{\mathbb{P}[\mathcal{V} < S^+(r_P(r_t))]}{\mathbb{P}[\mathcal{V} < S^-(r_P(r_t))]} \times r_t = \frac{S^+(r_P(r_t)) - V_L}{S^-(r_P(r_t)) - V_L} \times r_t > r_t, \tag{11}$$

$$r_W(r_t) = \frac{\mathbb{P}[\mathcal{V} > S^+(r_P(r_t))]}{\mathbb{P}[\mathcal{V} > S^-(r_P(r_t))]} \times r_t = \frac{V_H - S^+(r_P(r_t))}{V_H - S^-(r_P(r_t))} \times r_t < r_t. \tag{12}$$

As in the benchmark model, war (peace) causes a fall (increase) in r_t. In Rohner *et al.* (2012b) we prove that (i) for any r_t, there exists a unique PBE; (ii) the gap between S^+ and S^- increases with r_t, implying that the war/peace signal becomes more informative as r_t grows. The result that the speed of learning increases smoothly with the level of trust is a generalization of the result in the benchmark model above, that the speed of learning stays constant in the informative region ($r_t \geq \underline{r}$) and falls discretely to zero in the learning trap ($r_t < \underline{r}$). The endogenous probability of war becomes now a *continuous* decreasing function of r_t, rather than a decreasing step function. More formally,

$$\mathbb{P}[\mathbb{W}_{t+1} = 1 \mid k, r_t] = 1 - \mathbb{P}\left[\mathcal{V} < S^k(r_P(r_t))\right] = \frac{V_H - S^k(r_P(r_t))}{V_H - V_L}. \tag{13}$$

This is a (continuous) generalization of the second part of Proposition 4. A corollary of (13) is that, for any r_t, a civil conflict today – by reducing r_t – increases strictly the probability of a conflict next period. In conclusion, war persistence is a hard prediction of our theory. The particular form that the hysteresis takes in the benchmark model is instead driven by the simplifying assumption that $\mathcal{V}$ has a discrete support.

6 Extensions

In this section, we discuss some important extensions of the benchmark model.

[20] See Rohner *et al.* (2012b, available upon request) for a more detailed and formal analysis of the continuous case. There, we also show that war traps carry over to a version of the model with a generic continuous distribution of $\mathcal{V}$.

6.1 Stochastic types

So far, group A's type was assumed to be a permanent characteristic. In this section, we generalize the analysis to an environment in which the group type is subject to stochastic shocks, driven, for instance, by cultural shifts. For instance, ancient Vikings were an aggressive population prone to war and looting, whereas their current Scandinavian descendants are regarded as peaceful and cooperative people. Cultural shifts in types can be related to low-frequency changes in political regimes, institutions, social structures or population mixture (see Tabellini 2008).

We assume that group A's type follows a two-state first-order stochastic Markov process with the transition matrix

$$\begin{array}{c|cc} t-1 \backslash t & + & - \\ \hline + & 1-\psi & \psi \\ - & \phi & 1-\phi \end{array}. \tag{14}$$

We make the realistic assumption that $\phi \leq 1/2$ and $\psi \leq 1/2$, implying a positive serial autocorrelation. The ergodic distribution is summarized by the unconditional likelihood ratio that A is civic $\hat{r} = \phi/\psi$.

Without loss of generality, we assume that the *type shock* is realized between the end of period $t-1$ and the beginning of period t. This implies that the posterior belief at $t-1$ is no longer identical to the prior belief at t. In particular, the mapping from posterior at $t-1$ to prior at t is given by:

$$\tilde{r}\left(r_{t-1}\right) = \frac{(1-\psi)r_{t-1}+\phi}{\psi r_{t-1}+1-\phi}, \tag{15}$$

where $\tilde{r}'\left(r_{t-1}\right) > 0$ and $\tilde{r}''\left(r_{t-1}\right) < 0$. The equilibrium law of motion of beliefs, (9), can be generalized according to the following stochastic difference equation:

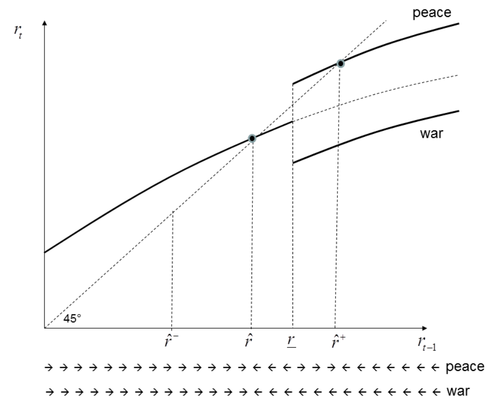
$$\ln r_t = \begin{cases} \ln \tilde{r}\left(r_{t-1}\right) & \text{if} \quad \tilde{r}\left(r_{t-1}\right) \in [0, \underline{r}] \qquad \text{ODE} \\ \ln \tilde{r}\left(r_{t-1}\right) + (1-\mathbb{W}_t) \ln\left(\frac{1-\lambda_W}{\lambda_P}\right) - \mathbb{W}_t \ln\left(\frac{1-\lambda_P}{\lambda_W}\right) & \text{if} \quad \tilde{r}\left(r_{t-1}\right) > \underline{r} \qquad \text{StoDE} \end{cases} \tag{16}$$

In the uninformative region, $\tilde{r}\left(r_{t-1}\right) \in [0, \underline{r}]$, group A wages war under BAU irrespective of its type, and thus the observation of war/peace conveys no information. The dynamics of beliefs is governed by an ordinary difference equation (ODE) such that r tends to the unconditional log-likelihood ratio, $\ln \hat{r}$. In the informative region, $\tilde{r}\left(r_{t-1}\right) > \underline{r}$, group A wages war when it is uncivic and retains peace when it is civic, under BAU. Thus, the dynamics is governed by a stochastic difference equation (StoDE) comprising both stages of the Bayesian updating: the type shock and war/peace.

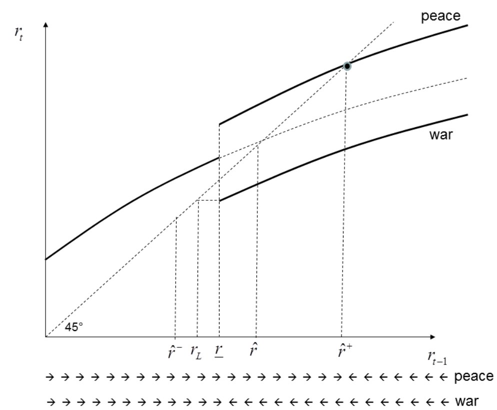
A qualitative description of the belief dynamics, (16), is provided by Figure 3.[21] It is useful to define two constants:

$$\hat{r}^+ = \frac{1-\lambda_W}{\lambda_P} \times \tilde{r}\left(\hat{r}^+\right) \text{ and } \hat{r}^- = \frac{\lambda_W}{1-\lambda_P} \times \tilde{r}\left(\hat{r}^-\right),$$

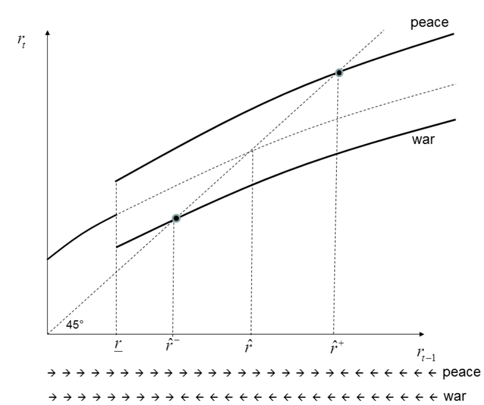
[21] Note that the figure plots likelihood ratios rather than their logarithms, since the graphical representation is more effective in levels than in logs.



Panel a: $\underline{r} > \hat{r}$**.** The economy converges in the long run to $\hat{r} < \underline{r}$ (war trap).



Panel b : $\underline{r} \in \left[\hat{r}^{-}, \hat{r}\right]$**.** The economy cycles in the long run in the region $\left[r_L, \hat{r}^{+}\right]$, including both informative and uninformative PBE.



Panel c : $\underline{r} < \hat{r}^{-}$**.** The economy cycles in the long run in the region $\left[\hat{r}^{-}, \hat{r}^{+}\right]$, where the equilibrium is informative (no war trap).

Figure 3: Dynamics of beliefs with stochastic types.

denoting the limits to which the dynamics of beliefs (16) would converge after an infinite sequence of war and peace episodes, respectively, if the realization of war and peace were informative everywhere (i.e., if the dynamics were governed by the StoDE for all r_{t-1}). Note that $0 < \hat{r}^- < \hat{r} = \phi/\psi < \hat{r}^+$. Three cases are possible (see Appendix C for analytical details and proofs):

1. If $\underline{r} > \hat{r}$ (panel a of Figure 3), beliefs converge with probability one to the war trap. Conditional on trust being initially high, the dynamics is governed by the StoDE part of (16) with frequent spells of peace. However, cooperation is deemed to collapse. As soon as r enters the uninformative region, $[0, \underline{r}]$, the ODE of (16) drives the belief process to a monotonic convergence to $\hat{r}$. Since $\hat{r} < \underline{r}$, $\hat{r}$ is an absorbing state. Intuitively, the long-run frequency of the civic type is very low, making peace and cooperation fragile.

2. If $\underline{r} \in [\hat{r}^-, \hat{r}]$ (panel b of Figure 3), the economy cycles between periods of low trust with frequent wars (uninformative PBE) and periods of peace with thriving trade (informative PBE). Since the dynamics has no rest point – in particular, $\hat{r}$ is not a steady state, as the figure shows – the economy wanders indefinitely in the ergodic set, $[\hat{r}^-, \hat{r}^+]$, which comprises portions of both the informative and uninformative region. Suppose that $r_0 < \underline{r}$. Then, the PBE starts uninformative, and the dynamics is governed initially by the ODE part of (16): r_t grows slowly until it crosses the threshold $\underline{r}$.[22] Thereafter, the dynamics is governed by the StoDE part of (16); the PBE remains informative until the stochastic sequence brings r back below the threshold $\underline{r}$. Although the economy does not get stuck into the uninformative equilibrium region, now and then it falls into times of raging conflict.

3. If $\underline{r} < \hat{r}^-$ (panel c of Figure 3), the uninformative region is never visited in the long run. The ergodic set is again the interval $[\hat{r}^-, \hat{r}^+]$, but this lies entirely within the informative region. As soon as the informative region is reached, it is never left, nor do beliefs ever settle to a rest point.

In summary, allowing for shifts in group A's type yields richer dynamics without altering the main insights of our analysis. If the unconditional probability that A is civic is low (panel a of Figure 3) the economy necessarily falls into a war trap. For an intermediate range of parameters (panel b of Figure 3), cycles emerge in which periods of recurrent conflicts and periods of peace and trade alternate. Spells of recurrent wars are preceded by decreases in reciprocal trust. Finally, if the unconditional probability that group A is civic is sufficiently high (panel c of Figure 3), recurrent wars are not observed in the long run. This case may represent well-functioning developed societies for which conflicts are rare.

Structural factors such as weak institutions, low human capital, or abundance of natural resources can determine which of the three regimes prevails. However, these factors alone would not fully explain

[22] When the uninformative equilibrium region is visited, trust may recover only very slowly, especially if types are very persistent. In particular, as $\psi \to 1$ and $\phi \to 1$, the dynamics of (15) which drive the exit from the trap become arbitarily slow.

the dynamics of conflict. In particular, in the intermediate regime described by panel b of Figure 3, societies that are fundamentally identical may behave differently for prolonged periods.

6.2 Learning from trade

So far in our analysis, the information set of group B has been limited to the history of warfare. The inference about group A's type could be improved if agents were able to directly observe part of the trade history. For instance, if public records of the outcome of inter-ethnic trade existed, group B could infer group A's type perfectly. Clearly, this is not a realistic scenario, since in reality cross-community trade and business links are decentralized and hardly distinguishable from intra-community trade. In this section, we extend the model by assuming that agents are born into (non-altruistic) dynastic *families* where information can be transmitted from parents to children.

We expand the information set of group B by allowing agents to acquire (stochastically) some information through their trade experience. In particular we assume that in each peaceful period where trade is active, each agent can observe, with a positive probability τ (i.i.d. across agents), group A's type. Such "hard information" can be transmitted to the agent's offspring.[23] In this environment, without additional assumptions, all dynasties would learn perfectly group A's type asymptotically. To prevent the informational friction from vanishing in the long run, we make the realistic assumption that the inter-generational transmission of hard information is subject to frictions: with the exogenous probability θ, an informed parent's child fails to receive the information. In this model, θ is an inverse measure of the efficiency of learning from trade history, and $1/\theta$ is the average number of generations to which the information is transmitted. Note that the model of this section nests the benchmark model in the particular case in which $\theta = 1$ (or $\tau = 0$).

In every period there is both a hard information inflow (uninformed traders learning about group A's type) and an exogenous outflow (some informed families forget). In wartime, noone trades and the net inflow necessarily is negative. In peacetime, the net inflow can be positive. Intuitively, information depreciates: if trade was intense in the far past, but then waned due to frequent conflicts, the information gathered through the past trade fades away.

We define by ι the share of informed agents in group B. The state space is now $(r, \iota) \in \mathbb{R}^+ \times [0, 1]$, specifying the public belief of uninformed agents and the proportion of informed agents. The PBE of the trade game is characterized formally in Appendix C, Proposition 9. There, we establish that it is the unique 4-tuple $\{n_A^-, n_A^+, n_B^-, n_B^+\} \in [0, 1]^4$ such that, for $k \in \{-, +\}$, $n_A^k (r_P, \iota) = F^k \left(zn_B^k (r_P, \iota)\right)$ and

$$n_B^k (r_P, \iota) = \iota F^B \left[zF^k \left(zn_B^k (r_P, \iota)\right)\right] + (1 - \iota)F^B \left[\frac{r_P}{1 + r_P} zF^+(zn_B^+) + \frac{1}{1 + r_P} zF^-(zn_B^-)\right]. \quad (17)$$

Note that uninformed agents know how many players are informed, but ignore the content of that information. Different from Proposition 2, n_B depends now on the state k, since a proportion ι of

[23] Although we abstract, for simplicity, from horizontal transmission of information within the group, this could be added in principle , as long as one retains some frictions in the transmission process.

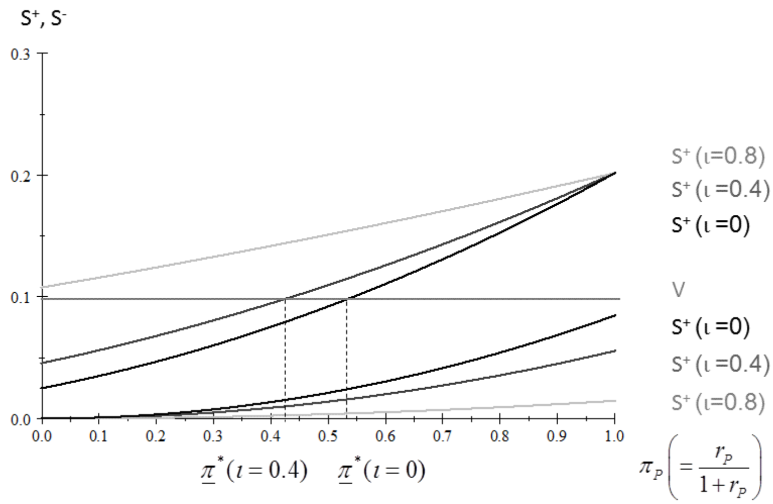


Figure 4: Trade surplus with different proportion of informed players, ι.

group B players can condition its behavior on k. Note that, in order to set the strategy that maximizes the expected payoff, uninformed players must infer the behavior of group A conditional on either state of nature, $k \in \{+, -\}$. Since n_A^k depends on n_B^k, the characterization of equilibrium involves keeping track of both n_B^+ and n_B^-, i.e., the proportion of cooperators conditional on the "true" as well as on the "counterfactual" type.

To progress further, we specialize the analysis to a uniform distribution of psychological payoffs $\mathcal{P}$ in both populations. Let $S^k(r_P, \iota)$ denote the trade surplus of group A. This depends now on A's type, public beliefs, and the proportion of informed agents in B. Corollary 2 in Appendix C proves that $\partial S^-/\partial\iota \leq 0$, whereas $\partial S^+/\partial\iota \geq 0$, implying that $\partial(S^+ - S^-)/\partial\iota \geq 0$. Intuitively, as the share of informed agents increases, the gap in the extent of cooperation across the two states of nature (i.e., n_A^+ vs. n_A^-, and n_B^+ vs. n_B^-) becomes wider, and is largest as we approach the perfect information equilibrium, $\iota \to 1$. Such a divergence between the two trade surpluses makes the observation of war/peace more and more informative for any given r_P, destabilizing the war trap. The comparative statics is shown in Figure 4, drawn for the same parameter values as Figure 1.

When $\iota = 0$ the schedules S^+ and S^- are identical to those in Figure 1, and a war trap exists. Increasing ι to 0.4 reduces the range of posterior beliefs consistent with the existence of the traps. Eventually, as ι is increased further, the *war trap* vanishes, as shown by the gray lines in Figure 4, representing the case of $\iota = 0.8$. In summary, the larger the share of informed agents, the harder it is to sustain the war trap.

In the rest of the section, we show that the war trap is robust to this environment, when ι evolves endogenously. Consider the law of motion of ι. The set of informed agents at date $t+1$ comprises children of either uninformed traders who acquired (and did not lose) information at t, as long as there was peace, or informed agents at t who did not experience any memory loss. More formally, the law of motion is:

$$\iota_{t+1} = (1-\theta)[\iota_t + (1-\mathbb{W}_t) \times \tau \times (1-\iota_t)]. \tag{18}$$

Note that, if $\theta = 0$, the share of informed agents would converge to unity since, due to peace shocks, some trade occurs even in the war trap and there is no memory loss. In terms of Figure 4, the schedules S^+ and S^- would shift progressively outwards every time there is peace, and would never shift inwards. On the contrary, when $\theta > 0$ the share of informed agents can either increase or decrease over time, implying that the schedules S^+ and S^- can shift either inwards or outwards. Moreover, the share of informed agents is bounded away from unity. In particular, equation (18) implies that, as long as $\iota_0 < \iota_\infty(\theta) \equiv \left(1 + \frac{\theta}{\tau(1-\theta)}\right)^{-1}$, then $\iota_t < \iota_\infty(\theta)$ for any $t > 0$ and for any realization, $\mathbb{W}_t$, of the war stochastic process. Note that $\iota_\infty(\theta)$ corresponds to the limit to which the economy converges after an infinite sequence of peace shocks ($\mathbb{W}_t = 0$, for all t).[24]

Looking back at Figure 4, a war trap exists if $S^+(0;\iota_\infty(\theta)) < V$. In this case, the S^+ schedule crosses the horizontal V line for all feasible values of ι. In particular, there exists a low range of public beliefs ($0 < r < \underline{r}(\theta)$) such that both the civic and the uncivic type wage war even though the share of informed agents is at its upper bound. In Appendix C, we provide a complete closed form characterization of the war trap. We first show that there exists a constant $\theta_W < 1$ such that a war trap exists if and only if $\theta \geq \theta_W$. Then, given θ, we characterize the threshold $\underline{r}(\theta)$ such that, for any $r < \underline{r}(\theta)$ the economy is in a war trap.

When agents learn through trade, it is also possible to observe war cycles. Suppose that group A is civic and that the economy initially is in the informative region. Suppose, in addition, that $S^+(0;\iota_\infty(\theta)) > V$, so that there is no war trap. Yet, a sequence of war shocks can push the economy into a region where both groups wage war under BAU, and thus the probability of war is high $(1-\lambda_W)$. Moreover, each war spell brings about the collapse of trade, causing a further reduction in the share of informed agents (ι) who trust group A. Even though, for any state of the economy, there exist sequences of peace shocks that can bring the economy out of the uninformative region, each war episode makes the economy sink deeper into the vicious cycle of low trust and low trade. Conversely, peace episodes become very important. Although peace is per se no informative public signal, during each peace spell agents trade and informed traders build up and transmit to future generations trust towards the other group. This result is related to Acemoglu and Wolitzky (2012) where trust cycles also hinge on a (different) form of limited memory.

6.3 Other extensions

In order to achieve tractability, we have introduced a number of important simplifications. For instance, we have assumed that the net benefits from war are exogenous, and orthogonal to observable and unobservable characteristics of the aggressor, including its propensity to cooperate while trading.

[24]The region $\iota > \iota_\infty(\theta)$ corresponds to a set of non-recurrent states. If the economy starts in such a set, it abandons it in finite time with probability one.

Another *caveat* is that so far the analysis has been restricted to two groups. While a thorough generalization along these and other dimensions is left to future research, in this section we sketch two first-step extensions.

In the first extension, we introduce correlation between group A's propensity to cooperate and the stochastic process $\mathcal{V}$. In particular, we assume that the uncivic type has weaker psychological barriers against attacking the other group. We consider two cases. First, we assume that the benefits of war under BAU are type-dependent: $\mathcal{V}^k \in \{V_L, V^k, V_H\}$, where $k \in \{+, -\}$ and $V^- > V^+$ (the uncivic enjoys war more than does the civic under BAU). In this case, the qualitative results of the benchmark model remain unchanged. A war trap continues to exist, as long as $V^+ > S^+(0)$. Second, we assume that $V^- = V^+ = V$, but the probability of war (peace) shocks is higher when group A is uncivic (civic): $\lambda_W^- > \lambda_W^+$ and $\lambda_P^- < \lambda_P^+$. The probabilities of BAU stay unchanged ($\lambda_W^- + \lambda_P^- = \lambda_W^+ + \lambda_P^+$), for simplicity. The main new result is that the observation of war and peace is now always informative, irrespective of beliefs. Even in the region where both types wage war under BAU, the probability of a peace (war) shock is now higher when A is civic (uncivic), and thus $r_P(r_{-1}) > r_{-1} > r_W(r_{-1})$.[25] Strictly speaking, this extension features no war trap. However, wars become highly frequent and persistent whenever $r < \underline{r}$, since both types wage war under BAU. In that region, every war spell sinks the economy deeper and deeper into the vicious cycle of distrust and low cooperation.

In the second extension (analyzed more formally in Appendix C), we assume that the economy is inhabited also by a third group, C, which can trade with A. Group A can only wage war against B. During wartime, there is no trade between A and B, but A can trade with C. Trade between A and C is subject to no informational friction (for simplicity), but is less productive than trade between A and B, reflecting the lower skill of group C. The intensity of the link between A and C is parameterized by the payoff of their bilateral trade.

This analysis highlights a new interesting implication. If from group A's standpoint trade with C is a substitute of trade with B, then opening a trade link with a third group increases the range of beliefs such that A wages war on B. Intuitively, the opportunity cost of war is lower if during wartime A can replace the destroyed trade with B with newly created trade with C (even though this trade is less productive).[26] On the contrary, if trade with C is a complement of trade with B, a stronger trade link between A and C reduces the probability of an aggression against B.

For example, if entrepreneurs in group A can hire workers in either group B or C, and these workers

[25] More formally, the log-likelihood ratios evolve now as follows:

$$\begin{aligned}
\ln r_P(r_{-1}) &= \ln r_{-1} + \ln \frac{\lambda_P^+ + (1 - \lambda_W^+ - \lambda_P^+)\,\sigma^+(r_{-1})}{\lambda_P^- + (1 - \lambda_W^- - \lambda_P^-)\,\sigma^-(r_{-1})}, \\
\ln r_W(r_{-1}) &= \ln r_{-1} - \ln \frac{1 - \lambda_P^- - (1 - \lambda_W^- - \lambda_P^-)\,\sigma^-(r_{-1})}{1 - \lambda_P^+ - (1 - \lambda_W^+ - \lambda_P^+)\,\sigma^+(r_{-1})}.
\end{aligned}$$

[26] Similar implications obtain in a model where, in peacetime, C trades with B, regarded to be a better partner than A. A successful attack of A on B destroys any trade involving B, eventually forcing C to trade with A. In this case, the competition motive strengthens the drive to attack B.

have substitutable skills, the availability of workers from C reduces the economic losses suffered by A during an inter-ethnic war with B. However, if members of group B – alongside trading directly with A – act as middlemen between A and C, then the trade link between A and C turns into a war deterrent. Intuitively, not only does war disrupt the trade between A and B, but it also causes collateral damage to the trade between A and C. In the next section, we show that these predictions are consistent with the empirical evidence that the probability of inter-ethnic conflict is indeed higher when groups hold substitutable skills rather than complementary skills.

7 Trade and civil conflict: evidence from case studies

In this section, we document case-study evidence of the link connecting trust, inter-ethnic trade (or, more generally, social interaction), and war.

The surplus of inter-ethnic trade clearly depends on how much different communities are specialized in complementary activities, as we discuss in section 6.3. In line with this prediction, Horowitz (2000) shows that strong economic inter-group complementarities contribute to inter-ethnic peace. Examples include Indonesia, Myanmar, Malaysia and India where middleman minorities often have been shielded from political violence, as they provide valuable services to the local ethnic majority. A similar conclusion is reached by Jha (2008) who studies Hindu-Muslim interactions using town-level data for India. He finds that during medieval times in India's trade ports, Hindus and Muslims provided complementary services to each other, and argues that this business interaction led to religious tolerance and a lower level of political violence in these trade ports than in other Indian towns. Interestingly, this kind of situation persists today. In a similar vein, Varshney (2001, 2002) argues that the existence of inter-ethnic business and civic associations can stunt the potential for riots in India. According to Varshney, for the prevention of ethnic conflict "trust based on *interethnic*, not *intraethnic*, networks is critical" (2001: 392). Bardhan (1997) discusses how waning inter-ethnic business links, due to exogenous factors lowering the opportunity costs of conflict, resulted in the outbreak of riots.[27] Moving to Africa, Olsson (2010) shows that an exogenous change in climatic conditions (i.e. a severe drought) has brought to a collapse of the inter-ethnic trade between farmers and herders in Sudan's Darfur region, and that this breakdown has been followed by the outbreak of conflict.[28] Similarly, Porter *et al.* (2010) carry out in-depth interviews with market traders in areas of Nigeria characterized by inter-ethnic tensions, and find that the existence of inter-ethnic trade links

[27] "On the Moradabad riots of 1980: The higher wages in the brass industry and entrepreneurship brought about not only greater prosperity among the Muslims, it also began to lessen the importance of the middlemen, often Hindu, in business transactions. Some of the Muslim entrepreneurs even managed to get direct orders from West Asian countries. The Hindu middlemen thus edged out began to rally round the Jan Sangh (now BJP) which has its base among petty businessmen" (Bardhan, 1997: 1397).

[28] Also UNICEF (2003) finds that exogenous collapses in inter-ethnic markets have resulted in conflicts in Darfur: "The groups confronting each other in the current conflicts have a long history of guarded cooperation and relative peaceful coexistence. In the past, they exchanged goods and services; indeed some of the herds that the Arab nomads reared belonged to wealthy Fur who did not opt to become nomads themselves. The Fur sold most of their herds on the onset of the drought in 1982/83. This was considered a severance of economic relations, which strained the relations between the Fur and the Arabs" (UNICEF, 2003: 53).

prevents the outbreak of full-blown riots.

Another prediction of our theory is that, whenever war depletes trust, inter-group trade tends to collapse. In line with this prediction, Guiso *et al.* (2009) provide evidence of a causal negative effect of the long-run intensity of bilateral warfare (over the 1000-1970 period) on the current level of bilateral trust and trade in a sample of European countries (see also Glick and Taylor, 2010). Looking at a more recent episode, in Rohner *et al.* (2012a) we find that post-conflict economic recovery in Uganda was especially slow in counties that both had been subject to intensive fighting *and* were more ethnically fractionalized (an interaction effect), likely because of the collapse of inter-ethnic business cooperation. Similarly, Cassar *et al.* (2011) find that exposure to ethnic conflict in Tajikistan undermined the former victims' willingness to participate in market activity involving trade with people with whom they do not have a personal connection.

Trust, trade and conflict appear to have been intertwined in Rwanda. Case studies from Ingelaere (2007) and Pinchotti and Verwimp (2007) document that throughout the 1980s inter-ethnic trust was relatively high and sustained symbiotic business relationships, cooperation in agricultural production associations and mixed rotating savings groups involving both Hutus and Tutsis. Survey data indicate that trust plummeted as of 1990, after localized ethnic fighting erupted in northern Rwanda (Ingelaere, 2007). The collapse of trust was followed by fading trade and business links between the communities, until inter-ethnic cooperation ceased altogether at the onset of the 1994 genocide.[29] Even several years after the conflict, the average inter-ethnic trust levels are significantly lower than in the 1980's (Ingelaere, 2007), and also inter-ethnic trade is persistently low (Colletta and Cullen, 2000).

Similarly, UNICEF (2003) documents that in several of Darfur's conflicts inter-group trust and trade broke down in the aftermath of fighting. For example, the civil war has resulted in the disintegration of the traditional economic arrangements between nomads and farmers regulating the use of pastureland and access to water in the Upper Nile region. This collapse of economic cooperation spurred kidnapping and other forms of inter-ethnic violence, triggering more local conflicts. The same pattern is also found elsewhere in Africa. Dercon and Gutierrez-Romero (2012) study the 2007 Kenyan electoral violence. Their survey data indicate that violence decreased trust between ethnic groups (while increasing trust within ethnic groups). Furthermore, after episodes of violence, people indicated that they tend to do less business with people from other ethnic groups and that they find violence more justifiable.

There are also case studies documenting the detrimental effect of war on trust and trade in Europe. Blagojevic (2009) finds that after the Bosnian war inter-ethnic trust collapsed and that the economic cooperation between the Serbian population on the one hand, and the Bosnian and Croatian population on the other declined sharply, being often replaced by intense inter-ethnic competition. Similarly, according to Kaufman (1996) the war in Moldova ushered in a climate of distrust between

[29] Colletta and Cullen, (2000:45) find that while vertical (within-group) social capital remained intact, "conflict deeply penetrated such forms of horizontal social capital as exchange, mutual assistance, collective action, trust and the protection of the vulnerable. [...] The use of credit in exchanges was common in preconflict Rwanda. This practice has diminished over time, in part due to decreased levels of trust as a consequence of warfare".

the Moldavans and the Russian-speaking minority, and ultimately a substantial decline in inter-ethnic business cooperation.

8 Policy implications

Our model implies that policies intended to increase the return from inter-ethnic trade – e.g., by reducing cross-community trade frictions, or facilitating the economic integration of minorities – increase the opportunity cost of war, narrowing the range of beliefs that sustain war traps.[30] Formally, scaling up the economic pay-off of the stag hunt game causes an upward shift of both S^+ and S^- in Figure 1, shifting $\underline{\pi}$ to the left, and possibly even ruling out the trap altogether. This prediction lines up with the empirical results of Horowitz (2000) on affirmative action and ethnic conflict. He finds that preferential programs aiming to improve the integration of less advanced ethnic groups in the national economy have reduced the potential for conflict in various countries such as India, Indonesia, Malaysia and Nigeria.[31] Since trade typically thrives in fast-growing economies, our theory is also broadly consistent with the evidence that high economic growth reduces the risk of war recurrence (Sambanis, 2008; Walter, 2004). The theory also suggests a role for policy incentivizing inter-ethnic cooperation (in our model, increasing z). For instance, improving contract enforcement and punishing fraudulent behavior in business relationships would reduce the payoff of cheating and opportunistic behavior, thereby increasing the trade surplus. In the alternative interpretation of the trade game provided in section 2.1, z could be increased by policies supporting human capital formation, such as promoting the knowledge of multiple national languages or of other cultural aspects that affect inter-ethnic barriers. This finding is in line with the evidence that higher education expenditures and enrollment rates decrease the risk of civil wars (Thyne, 2006).

To the opposite, larger windfall gains from war (i.e., larger V) expand the potential for war traps. For instance, the discovery of oilfields or diamond mines can drive an economy in which trust was initially low into a war trap, while the same discovers would have no detrimental effects in economies where trust and cooperation were high in the first place. Here the predictions of our theory conform well with the evidence that more abundant natural resources fuel war recurrence and hinder post-war recovery in fragile states dominated by low trust (see, e.g., Doyle and Sambanis, 2000; Fortna, 2004; Sambanis, 2008), while typically they have no harmful consequences in high-trust strong states like Norway. International sanctions, such as embargoes on arms or on natural resource imports from regimes that have seized power through ethnic wars could reduce V.

[30] While we have assumed, for simplicity, that the matching process is frictionless, one can easily extend the model to incorporate search frictions. In this environment, only a fraction of the agents seeking partnerships would be able to trade. The stronger the frictions, the lower the trade surplus. Active integration policies reducing such frictions would ultimately shrink the range of war traps.

[31] Horowitz (2000) and Whah (2010) show that the programs of state-induced inter-ethnic joint venture companies in Malaysia dating from the 1970's have in many instances enhanced trust between the Malay and the Chinese population reducing social tensions. Similarly, Augenbraun et al. (1999) find that microenterprise lending by donors in Bosnia for inter-ethnic joint ventures has worked well, not only in a purely economic sense, but also calming tensions between groups.

Our theory has ambiguous implications about the effectiveness of *international peacekeeping*. The benchmark model predicts that international peacekeeping interventions limited to "stopping the shooting" will have only short-lasting effects on political stability. External military-backed peacekeeping may even be detrimental, if local groups attribute any peace episode to the mere presence of foreign troops, and then stop updating beliefs. However, according to the extension of section 6.2, even an externally imposed truce can be useful if it restores some inter-ethnic trade that induces some learning and trust rebuilding. The nature of the peace-keeping intervention becomes therefore crucial: If the intervention simply keeps the fighting groups apart, it will not achieve permanent effects. Yet, if peacekeeping is first complemented and then replaced by trade- and trust-enhancing measures, it can be effective. These ambiguous predictions are in line with the gloomy appraisal of a variety of studies on the survival of peace. According to Sambanis (2008: 30): "UN missions have a robust positive effect on peacebuilding outcomes, particularly participatory peace, but the effects occur mainly in the short run and are stronger when peacekeepers remain." Luttwak (1999: 37) goes further, and argues that mere peacekeeping – unless accompanied by trade-promoting or trust-restoring measures – does not lead to lasting peace; it simply interrupts hostilities that will recur once the UN troops leave: "(Peacekeeping), perversely, can systematically prevent the transformation of war into peace. The Dayton accords are typical of the genre: they have condemned Bosnia to remain divided into three rival armed camps, with combat suspended momentarily but a state of hostility prolonged indefinitely... Because no path to peace is even visible, the dominant priority is to prepare for future war rather than to reconstruct devastated economies and ravaged societies."

Our theory also suggests that policies directly targeting beliefs may be important when distrust hinges on no fundamental reasons (in the model, $k = +$), yet cooperation is hindered by bad beliefs. In such cases, creating pro-trade public information, for instance by documenting and publicizing successful episodes of inter-ethnic business cooperation, can help end the vicious cycle. There is indeed empirical evidence that inter-group prejudices can be weakened by targeted media exposure (Paluck, 2009; Paluck and Green, 2009). According to Paluck's (2009) findings the listeners exposed to the "social reconciliation" radio soap opera in Rwanda were significantly more likely to find it "not naive to trust" and to feel empathy for other Rwandans than the control group exposed to a "health" radio soap opera. Similarly, Bardhan (1997) shows that in India direct targeting of Muslims' and Hindus' beliefs by spreading success stories of cooperation has reduced distrust and the potential for conflicts: "Public information on what actually happened, on how a disturbance started, on who tried to take advantage of it, on instances of intercommunity cooperation in the face of tremendous odds, etc., if effectively transmitted in the early stages, can stop some of the vicious rumors that fuel communal riots and calm group anxieties" (Bardhan, 1997: 1395).

Finally, when the fundamentals are bad (in the model, $k = -$), public campaigns aimed to shift collective preferences may be effective. Such campaigns may aim either to foster pro-cooperative norms, or to eradicate prejudice and dislike towards specific communities (both cases can be modeled as shifts in the distribution of $\mathcal{P}$) . For instance, Hauk and Saez-Marti (2002) document that a strong press and education campaign was key to slashing corruption and opportunistic practices in Hong Kong during the 1970's. Denazification after World War II is an example in which the campaign targeted people's preferences pointedly, aiming to eradicate racial hatred against minorities. In terms of our model, these campaigns may be viewed as measures to promote cultural shifts, such as those happening exogenously in section 6.1. While the statement that campaigns against racial hatred can reduce conflict may sound obvious, our theory warns that such campaigns may be of limited help if they do not target, simultaneously, the beliefs (trust) of the victims. For instance, our model of section 6.1 shows that if group B perceives a change of group A's type to be very unlikely (case 1), even a campaign that successfully turns group A from uncivic to civic may prove ineffective. This issue is salient in countries prone to civil conflicts where the ethnic group in majority replaces a radical sectarian government with a more inclusive one, but fails to restore the confidence of minorities.

While the theory suggests a number of useful policy implications, a *caveat* is that these are valid under a set of specific assumptions that may not capture salient aspects of particular conflicts. We focus on a broad category of inter-group social interactions taking the form of a stag hunt game. However, there may be other forms of inter-group interactions with different payoff structures and incentives. For instance, recent empirical literature shows that ethnic diversity may have different effects depending on whether cross-group interactions are about rival (private) or non-rival (public) goods (Esteban et al., 2012; Spolaore and Wacziarg, 2012).

9 Conclusion

This paper provides a theory in which asymmetric information and cultural transmission of beliefs can explain why societies enter into recurrent civil conflicts even when there is no fundamental reason for conflicts to occur. We emphasize the link between trade and war, which has been highlighted in the recent literature on international conflicts. We believe that this connection can be even more salient in inter-group conflicts within countries where business relationships are decentralized and do not need the mediation of institutions that can aggregate and diffuse information. Our discussion has emphasized the role of inter-ethnic trade, which we hold to be especially important. However, our theory can be extended to a broader set of social interactions within countries, including inter-ethnic marriages, public good provision in villages and communities, and more general situations involving bilateral cooperation across ethnic cleavages.

While our current study presents a rational-agent theory, integrating psychological aspects may reveal more about these issues. For instance, exposure to war during childhood may be especially important, since many beliefs and values are formed at tender ages. In our model, heterogeneous preferences are exogenous, while in reality these may be shaped by parental transmission, as in Hauk and Saez-Marti (2002). Our analysis also abstracts from general equilibrium effects that may affect the opportunity costs of wars and the returns to violent resource appropriation (as emphasized by Dal Bo and Dal Bo, 2011). Finally, we have abstracted from institutions and state capacity (Besley and Persson, 2010). As emphasized by Aghion *et al.* (2011), institutions and beliefs are not independent factors: on the one hand, institutions can affect the trust-building process, while on the other trust can influence institutional developments that can deter conflict. Reforming institutions in a more inclusive direction (in the sense of Acemoglu and Robinson, 2012) can help build inter-ethnic trust and trade. Studying these interconnections is left to future research.

REFERENCES

Acemoglu, Daron, and James Robinson, 2012, *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, New York: Crown Publishing.

Acemoglu, Daron, Davide Ticchi, and Andrea Vindigni, 2010, "Persistence of Civil Wars", *Journal of the European Economic Association* 8: 664–676.

Acemoglu, Daron, and Fabrizio Zilibotti, 1997, "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth", *Journal of Political Economy* 105: 709-751.

Acemoglu and Wolitzky, 2012, "Cycles of Distrust: An Economic Model", mimeo, MIT.

Aghion, Philippe, Yann Algan, and Pierre Cahuc, 2011, "Civil society and the state: The interplay between cooperation and minimum wage regulation", *Journal of the European Economic Association* 9: 3–42.

Aghion, Philippe, Yann Algan, Pierre Cahuc, and Andrei Shleifer, 2010, "Regulation and Distrust", *Quarterly Journal of Economics* 125: 1015-1049.

Augenbraun, Eliene, Karl Feld, Iain Guest, and Susan Merrill, 1999, "Reconciling Ethnic Conflicts: A Case Study of Bosnia-Herzegovina and Croatia", USAID Evaluation Special Study 80.

Baliga, Sandeep, and Tomas Sjostrom, 2004, "Arms Races and Negotiations", *Review of Economic Studies* 71: 351-369.

Banerjee, Abhijit, 1992, "A simple model of herd behavior", *Quarterly Journal of Economics* 107: 797-817.

Bardhan, Pranab, 1997, "Method in the Madness? A Political-Economy Analysis of the Ethnic Conflicts in Less Developed Countries", *World Development* 25: 1381-98.

Besley, Timothy, and Torsten Persson, 2011, "The Logic of Political Violence", *Quarterly Journal of Economics* 126: 1411-1445.

Besley, Timothy, and Torsten Persson, 2010, "State Capacity, Conflict and Development", *Econometrica* 78: 1-34.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, 1992, "A theory of fads, fashion, custom, and cultural change as information cascades", *Journal of Political Economy* 100: 992-1026.

Blagojevic, Bojana, 2009, "Causes of Ethnic Conflict: A Conceptual Framework", *Journal of Global Change and Governance* 3: 1-25.

Blattman, Chris, and Edward Miguel, 2010, "Civil War", *Journal of Economic Literature* 48: 3-57.

Buhaug, Halvard, Scott Gates, and Päivi Lujala, 2009, "Geography, Rebel Capability, and the Duration of Civil Conflict", *Journal of Conflict Resolution* 53: 544-69.

Caselli, Francesco, and Wilbur John Coleman II, 2012, "On the Theory of Ethnic Conflict", forthcoming in the *Journal of the European Economic Association.*

Caselli, Francesco, Massimo Morelli, and Dominic Rohner, 2012, "The Geography of Inter-State Resource Wars", mimeo, LSE.

Cassar, Alessandra, Pauline Grosjean, and Sam Whitt, 2011, "Civil War, Social Capital and Market Development: Experimental and Survey Evidence on the Negative Consequences of Violence", mimeo, University of San Francisco.

Cederman, Lars Erik, and Luc Girardin, 2007, "Beyond Fractionalization: Mapping Ethnicity onto Nationalist Insurgencies", *American Political Science Review* 101: 173-85.

Chamley, Christophe, 1999, "Coordinating Regime Switches", *Quarterly Journal of Economics* 114: 869-905.

Chamley, Christophe, 2004, *Rational Herds: Economic Models of Social Learning*, Cambridge: Cambridge University Press.

Chassang, Sylvain and Gerard Padro i Miquel, 2010, "Conflict and Deterrence under Strategic Risk", *Quarterly Journal of Economics* 125:1821-1858.

Collier, Paul and Anke Hoeffler, 1998, "On economic causes of civil war", *Oxford Economic Papers* 50: 563-73.
Collier, Paul and Anke Hoeffler, 2004, "Greed and grievance in civil war", *Oxford Economic Papers* 56: 563-95.
Collier, Paul, Anke Hoeffler, and Dominic Rohner, 2009, "Beyond Greed and Grievance: Feasibility and Civil War", *Oxford Economic Papers* 61: 1-27.
Collier, Paul and Dominic Rohner, 2008, "Democracy, Development, and Conflict", *Journal of the European Economic Association* 6: 531-40.
Colletta, Nat and Michelle Cullen, 2000, *Violent Conflict and the Transformation of Social Capital: Lessons from Cambodia, Rwanda, Guatemala, and Somalia*, Washington: World Bank.
Cooper, Russell, and Andrew John, 1988, "Coordinating Coordination Failures in Keynesian Models", *Quarterly Journal of Economics* 103: 441-463.
Dal Bó, Ernesto, and Pedro Dal Bó, 2011, "Workers, Warriors, and Criminals: Social Conflict in General Equilibrium", *Journal of the European Economic Association* 9: 646–677.
Dercon, Stefan, and Roxana Gutierrez-Romero, 2012, "Triggers and Characteristics of the 2007 Kenyan Electoral Violence", *World Development* 40: 731-44.
DeRouen, Karl, and Jacob Bercovitch, 2008, "Enduring Internal Rivalries: A New Framework for the Study of Civil War", *Journal of Peace Research* 45: 55-74.
Dixit, Avinash Kamalakar, 2003, "Trade Expansion and Contract Enforcement", *Journal of Political Economy* 111: 1293-1317.
Downes, Alexander, 2006, "More Borders, Less Conflicts? Partition as a Solution to Ethnic Civil Wars", *SAIS Review* XXVI: 49-61.
Doyle, Michael, and Nicholas Sambanis, 2000, "International Peacebuilding: A Theoretical and Quantitive Analysis", *American Political Science Review* 94: 779-801.
Ely, Jeffrey, and Juuso Valimaki, 2003, "Bad Reputation", *Quarterly Journal of Economics* 118: 785-813.
Esteban, Joan, Laura Mayoral, and Debraj Ray, 2012, "Ethnicity and Conflict: An Empirical Investigation", *American Economic Review* 102: 1310-1342.
Esteban, Joan and Debraj Ray, 2008, "On the Salience of Ethnic Conflict", *American Economic Review* 98: 2185-202.
Esteban, Joan, and Debraj Ray, 2011, "Linking Conflict to Inequality and Polarization", *American Economic Review* 101: 1345-374.
Fearon, James, 1995, "Rationalist Explanations for War", *International Organization* 49: 379-414.
Fearon, James, 2005, "Primary Commodity Exports and Civil War", *Journal of Conflict Resolution* 49: 483-507.
Fearon, James, and David Laitin, 2003, "Ethnicity, Insurgency, and Civil War", *American Political Science Review* 97: 75–90.
Fernandez, Raquel, 2012, "Culture as Learning: The Evolution of Female Labor Force Participation over a Century", *American Economic Review,* forthcoming.
Fortna, Virginia, 2004, "Does Peacekeeping Keep Peace? International Intervention and the Duration of Peace After Civil War", *International Studies Quarterly* 48: 269-292.
Gartzke, Erik, 1999, "War Is in the Error Term", *International Organization* 53: 567-87.
Gates, Scott, 2002, "Recruitment and Allegiance: The Microfoundations of Rebellion", *Journal of Conflict Resolution* 46: 111-30.
Glick, Reuven, and Alan Taylor, 2010, "Collateral Damage: Trade Disruption and the Economic Impact of War", *Review of Economics and Statistics* 92: 102-27.
Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2009, "Cultural biases in Economic Exchanges?", *Quarterly Journal of Economics* 124: 1095-1131.
Hauk, Esther and Maria Saez-Marti, 2002, "On the Cultural Transmission of Corruption", *Journal of Economic Theory* 107: 311-15.
Horowitz, Donald, 2000, *Ethnic Groups in Conflict*, Berkeley: University of California Press (2nd edition).
Ingelaere, Bert, 2007, "Living the Transition: A Bottom-up Perspective on Rwanda's Political Transition", Discussion Paper, University of Antwerp.
Jackson, Matthew and Massimo Morelli, 2007, "Political Bias and War", *American Economic Review* 97: 1353-1373.

Jha, Saumitra, 2008, "Trade, institutions and religious tolerance: evidence from India", mimeo, Stanford University.
Kaufman, Stuart, 1996, "Spiraling to Ethnic War", *International Security* 21: 108-138.
Licklider, Roy, 1995, "The consequences of negotiated settlements in civil wars, 1945-1993", *American Political Science Review* 89: 681-690.
Lindbeck, Assar, and Jörgen Weibull, 1987, "Balanced-budget redistribution as the outcome of political competition", *Public Choice* 52: 273-97.
Luttwak, Edward, 1999, "Give War a Chance", *Foreign Affairs* July/August: 36-44.
Martin, Philippe, Thierry Mayer, and Mathias Thoenig, 2008a, "Make Trade not War?", *Review of Economic Studies*, 75: 865-900.
Martin, Philippe, Thierry Mayer, and Mathias Thoenig, 2008b, "Civil Wars and International Trade", *Journal of European Economic Association* 6: 541–550.
Montalvo, José, and Marta Reynal-Querol, 2005, "Ethnic Polarization, Potential Conflict, and Civil Wars", *American Economic Review* 95: 796-816.
Olsson, Ola, 2010, "Market Collapse in Vulnerable Environments: A Model Applied to Darfur", mimeo, University of Gothenburg.
Paluck, Elizabeth Levy, 2009, "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda", *Journal of Personality and Social Psychology* 96: 574-87.
Paluck, Elizabeth Levy, and Donald Green, 2009, "Prejudice Reduction: What Works? A Review and Assessment of Research and Practice", *Annual Review of Psychology* 60: 339-67.
Piketty, Thomas, 1995, "Social Mobility and Redistributive Politics", *Quarterly Journal of Economics* 110: 551-584.
Pinchotti, Shanley and Philip Verwimp, 2007, "Social Capital and the Rwandan Genocide: A Micro-Level Analysis", HiCN working paper 30.
Porter, Gina, Fergus Lyon, Fatima Adamu, and Lanre Obafemi, 2010, "Conflict and Cooperation in Market Spaces: Learning from the Operation of Local Networks of Civic Engagement in African Market Trade", *Human Organization* 69: 31-42.
Powell, Robert, 2006, "War as a Commitment Problem", *International Organization* 60: 169-203.
Rohner, Dominic, 2011, "Reputation, Group Structure and Social Tensions", *Journal of Development Economics* 96: 188-199.
Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti, 2012a, "Seeds of Distrust? Conflict in Uganda", mimeo, University of Zurich.
Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti, 2012b, "A note on the "War Signal" model with a continuous stochastic process of war", mimeo in progress, University of Zurich.
Sambanis, Nicholas, 2001, "Do Ethnic and Nonethnic Civil Wars Have the Same Causes?", *Journal of Conflict Resolution* 45: 259-282.
Sambanis, Nicholas, 2008, "Short- and Long-Term Effects of United Nations Peace Operations", *World Bank Economic Review* 22: 9-32.
Spolaore, Enrico, and Romain Wacziarg, 2012, "War and Relatedness", mimeo, Tufts University and UCLA.
Tabellini, Guido, 2008, "The Scope of Cooperation: Values and Incentives", *Quarterly Journal of Economics* 123: 905-950.
Thyne, Clayton, 2006, "ABC's, 123's, and the Golden Rule: The Pacifying Effect of Education on Civil War, 1980-1999", *International Studies Quarterly* 50: 733-54.
Torvik, Ragnar, 2002, "Natural resources, rent seeking and welfare", *Journal of Development Economics* 67: 455-70.
UCDP, 2012, "UCDP/PRIO Armed Conflict Dataset", dataset, www.pcr.uu.se/research/ucdp.
UNICEF, 2003, "Analysis of Nine Conflicts in Sudan", working paper, UNICEF.
Varshney, Ashutosh, 2001, "Ethnic Conflict and Civil Society", *World Politics* 53: 362-98.
Varshney, Ashutosh, 2002, *Ethnic Conflict and Civic Life. Hindus and Muslims in India*, Yale University Press, New Haven.

Walter, Barbara, 2004, "Does Conflict Beget Conflict? Explaining Recurring Civil War", *Journal of Peace Research* 41: 371-88.
Whah, Chin Yee, 2010, "Towards Inter-Ethnic Business Development and National Unity in Malaysia", CRISE Working Paper 73, University of Oxford.
World Values Survey, 2011, WVS Five Wave Aggregated File v.20090901, dataset, www.worldvaluessurvey.org/
Yared, Pierre, 2010, "A Dynamic Theory of War and Peace", *Journal of Economic Theory* 145: 1921-1950.

Appendix I

Proof of Propositions 1 and 2. The proof of Proposition 2 encompasses the proof of Proposition 1, since perfect information is a particular case of imperfect information, where either $\pi_P = 0$ (if $k = -$) or $\pi_P = 1$ (if $k = +$).

We start by proving that a PBE exists. Given the definition $\pi_P \equiv \frac{r_P}{1+r_P}$, we can write the third element of equation (4) as

$$n_B = \tilde{F}^B(n_B, \pi_P) \equiv F^B\left(\pi_P z F^+(z n_B) + (1 - \pi_P) z F^-(z n_B)\right). \tag{19}$$

We claim three properties of $\tilde{F}^B$: (i) $\tilde{F}^B : [0,1]^2 \to [0,1]$ is a continuous function, with $0 < \tilde{F}^B(n_B, \pi_P) < 1$; (ii) $\tilde{F}^B(n_B, \pi_P)$ is increasing in π_P, for any n_B; (iii) $\tilde{F}^B(n_B, \pi_P)$ is non-decreasing and convex in n_B, for any π_B.

Note, first, that Assumption 1 implies that F^B, F^+ and F^- are continuous non-decreasing convex functions. Then, (i) follows from the fact that $\tilde{F}^B$ is a convex combination of two c.d.f.'s, F^+ and F^-, which are continuous functions. Assumption 1 guarantees that $0 < \tilde{F}^B(n_B, \pi_P) < 1$; (ii) follows from the fact that $\tilde{F}^B$ is a convex combination of two c.d.f.'s F^+ and F^-, with weights π_P and $1-\pi_P$, respectively, where $F^-(z n_B) < F^+(z n_B)$ which in turn follows from the assumption that H^+ first-order stochastically dominates (FOSD) H^-; (iii) follows from the fact that $\tilde{F}^B$ is a non-decreasing convex transformation of the convex combination of two non-decreasing convex functions of n_B.

Given (i), the existence, for any given $\pi_P \in [0,1]$, of $n_B = n_B^*(\pi_P)$ satisfying (19) follows from the Brouwer fixed-point theorem. Since $0 < \tilde{F}^B(n_B; \pi_P) < 1$, then the fixed-point of (19) must be strictly within the unit interval. (ii)-(iii) guarantee that the mapping $n_B^*(\pi_P)$ is unique and is monotonically weakly increasing. We prove uniqueness by contradiction. Assume that, for some π_P, there exists a second fixed point $\hat{n}_B = \tilde{F}^B(\hat{n}_B; \pi_P)$. Without loss of generality, let $n_B^*(\pi_P) < \hat{n}_B$. The second fixed point, $\hat{n}_B \in [n_B, 1]$, can be expressed as the convex combination $\hat{n}_B = \frac{1-\hat{n}_B}{1-n_B^*(\pi_P)} \times n_B^*(\pi_P) + \frac{\hat{n}_B - n_B^*(\pi_P)}{1-n_B^*(\pi_P)} \times 1$. Then, applying to $\hat{n}_B$ the convexity criterion of $\tilde{F}^B$ yields

$$\tilde{F}^B(\hat{n}_B; \pi_P) \leq \frac{1-\hat{n}_B}{1-n_B^*(\pi_P)} \tilde{F}^B(n_B^*(\pi_P); \pi_P) + \frac{\hat{n}_B - n_B^*(\pi_P)}{1-n_B^*(\pi_P)} \tilde{F}^B(1; \pi_P)$$

and this leads to

$$\hat{n}_B \leq \frac{1-\hat{n}_B}{1-n_B^*(\pi_P)} n_B^*(\pi_P) + \frac{\hat{n}_B - n_B^*(\pi_P)}{1-n_B^*(\pi_P)} \tilde{F}^B(1; \pi_P)$$

which in turn implies (since $n_B^*(\pi_P) < \hat{n}_B$) that $\tilde{F}^B(1; \pi_P) > 1$, contradicting property (i). This establishes the uniqueness of the equilibrium mapping $n_B^*(\pi_P)$. That $n_B^*(\pi_P)$ is monotonically weakly increasing follows immediately from claim (ii).

Next, define $n_B^*(\pi_P) = n_B^*\left(\frac{r_P}{1+r_P}\right) \equiv n_B(r_P)$. The analysis above establishes that the function $n_B(r_P)$ exists, is unique, is weakly increasing and lies strictly within the unit interval. That the same properties carry over to $n_A^-(r_P) = F^-(z n_B(r_P))$ and $n_A^+(r_P) = F^+(z n_B(r_P))$ is immediate. Thus, equation (4) has a unique fixed point. Since F^- FOSD F^+, then $n_A^-(r_P) \leq n_A^+(r_P)$. This proves the first part of the proposition.

The formula of the trade surplus is derived from integrating by parts $\int_{-\infty}^{z n_B(r_P)} \mathcal{L}\, dF^k = -\left[\mathcal{L}\ F^k\right]_{-\infty}^{z n_B(r_P)} + \int_{-\infty}^{z n_B(r_P)} F^k(\mathcal{L})\, d\mathcal{L}$, where $\left[\mathcal{L}\ F^k\right]_{-\infty}^{z n_B(r_P)} = z n_A^k(r_P)\, n_B(r_P)$, and simplifying terms. Since F^k is non negative and $n_B(r_P)$ is non-decreasing, then $S^k(r_P)$ must be non-decreasing. That $S^-(\pi_P) \leq S^+(\pi_P)$, finally, follows again from the fact that F^- FOSD F^+. **QED**

Proof of Lemma 1. We proceed in two steps: (i) we prove that an uninformative PBE exists if and only if $r_{-1} \leq \underline{r}^*$; (ii) we prove that informative PBE exist if and only if $r_{-1} \geq \underline{r}$. (i) and (ii) together prove the Lemma.

(i) First, we prove (by guess-and-verify) that an uninformative PBE exists if $r_{-1} \leq \underline{r}^*$. Guess that an uninformative PBE exists. Then, $r_{-1} \leq \underline{r}^* \Rightarrow r_P(r_{-1}) = r_{-1} \leq \underline{r}^*$, which in turn implies that $V \geq S^+(r_P(r_{-1})) > S^-(r_P(r_{-1}))$. Thus, $\sigma^+(r_{-1}) = \sigma^-(r_{-1}) = 0$, verifying the guess. Second, we prove (by contradiction) that an uninformative PBE exists only if $r_{-1} \leq \underline{r}^*$. Suppose that an uninformative PBE exists in the range $r_{-1} > \underline{r}^*$. Then, by the definition of uninformative equilibrium, $r_P(r_{-1}) = r_{-1} > \underline{r}^*$, which in turn implies that $S^+(r_P(r_{-1})) > V$ and $S^-(r_P(r_{-1})) < V$. Thus, $\sigma^+(r_{-1}) = 1$ and $\sigma^-(r_{-1}) = 0$, contradicting that the PBE is uninformative.

(ii) First, we prove (by guess-and-verify) that informative PBE exist if $r_{-1} \geq \underline{r}$. Guess that an informative PBE exists such that $\sigma^+(r_{-1}) = 1$ and $\sigma^-(r_{-1}) = 0$ (note: it can be established that a mixed strategy equilibrium such that $0 < \sigma^+(r_{-1}) < 1$ and $\sigma^-(r_{-1}) = 0$ also exists in this range, but this is not essential to prove the Lemma, so the proof is omitted). Then, $r_{-1} \geq \underline{r} \Rightarrow \infty > r_P(r_{-1}) \geq \underline{r}^*$, which in turn implies that $S^-(r_P(r_{-1})) < V \leq S^+(r_P(r_{-1}))$. Thus, $\sigma^+(r_{-1}) = 1$ and $\sigma^-(r_{-1}) = 0$, verifying the guess. Second, we prove (by contradiction) that an informative PBE exists only if $r_{-1} \geq \underline{r}$. Suppose that an informative PBE exists in the range $r_{-1} < \underline{r}$. Then, $r_{-1} < \underline{r} \Rightarrow r_P(r_{-1}) < \underline{r}^*$, which in turn implies that $V > S^+(r_P(r_{-1})) > S^-(r_P(r_{-1}))$. Thus, $\sigma^+(r_{-1}) = \sigma^-(r_{-1}) = 0$, contradicting that the equilibrium is informative. **QED**

Proof of Proposition 5. The process π_t is a bounded non-negative martingale. Hence, the Martingale Convergence Theorem implies that, almost surely, the limit $\pi_\infty = \lim_{t\to\infty} \pi_t$ exists and is unique. Let Γ_∞^k denote the support of π_∞ when group A's type is $k \in \{-,+\}$. Let $\tilde{\Omega}_{TRAP} = \left]\underline{r}/(\frac{1-\lambda_P}{\lambda_W} + \underline{r}), \underline{r}/(1+\underline{r})\right]$ denote the war trap (in term of the state space of beliefs, π_t) without the range of non-recurrent states, $\left[0, \underline{r}/(\frac{1-\lambda_P}{\lambda_W} + \underline{r})\right]$. Also, let the open set $\Gamma_I =]\frac{\underline{r}}{1+\underline{r}}, 1[$ denote the interior of the informative region.

We start by proving two key results: (1) for any $k \in \{-,+\}$, $\Gamma_I \cap \Gamma_\infty^k = \varnothing$; (2) $1 \notin \Gamma_\infty^-$ (i.e., when $k = -$, π_t cannot converge to the "wrong" perfect learning). Both claims are proven by contradiction.

1. Suppose that $\pi_\infty \in \Gamma_I$. Then, there exists $T < \infty$ such that, for all $t \geq T$, $\pi_t \in \Gamma_I$. Next, recall that within the informative region the dynamics of (the likelihood ratio of) beliefs is governed by (9). Taking limits as $t \to +\infty$ and replacing r by its definition yields: $\lim_{t\to\infty}(\ln r_t - \ln r_{t-1}) = \ln \frac{\pi_\infty}{1+\pi_\infty} - \ln \frac{\pi_\infty}{1+\pi_\infty} = 0 = \lim_{t\to\infty}\left((1-\mathbb{W}_t)\ln\left(\frac{1-\lambda_W}{\lambda_P}\right) - \mathbb{W}_t \ln\left(\frac{1-\lambda_P}{\lambda_W}\right)\right)$. This is a contradiction, since the last term does not converge.

2. Suppose that $k = -$ and $\pi_\infty = 1$. Then, there exists $T < \infty$ such that, for all $t \geq T$, π_t must remain forever in the informative region ($\pi > \frac{\underline{r}}{1+\underline{r}}$). But then, the Strong Law of Large Numbers implies that, as $t \to \infty$, the empirical frequencies of war/peace converge to the true underlying probabilities. Then, we should have that $\pi_\infty = 0$, leading to a contradiction.

(1) and (2) jointly establish that, (i) if $k = -$, then $\Gamma_\infty^- \subset \tilde{\Omega}_{TRAP}$ (the DSE enters the war trap almost surely); (ii) if $k = +$, then $\Gamma_\infty^+ \subset \tilde{\Omega}_{TRAP} \cup \{1\}$ (the DSE either enters the war trap or converges to truthful learning).

Let $\mathbb{P}_{TRAP} \equiv \mathbb{P}[\pi_\infty \in \tilde{\Omega}_{TRAP} \mid k = +]$. Since π_t is a martingale, then, $\pi_0 = \mathbb{E}[\pi_t], \forall t$. Moreover,

$$\begin{aligned} \pi_0 &= \lim_{t\to\infty} \mathbb{E}_0[\pi_t] \\ &= \pi_0 \times \mathbb{E}[\pi_\infty \mid k = +] + (1-\pi_0) \times \mathbb{E}[\pi_\infty \mid k = -], \end{aligned} \tag{20}$$

where

$$\mathbb{E}[\pi_\infty \mid k = +] = \mathbb{P}_{TRAP} \times \mathbb{E}\left[\pi_\infty \mid k = +, \pi_\infty \in \tilde{\Omega}_{TRAP}\right] + (1-\mathbb{P}_{TRAP}) \times 1. \tag{21}$$

Combining (20) and (21) we obtain, after standard algebra,

$$\mathbb{P}_{TRAP} = \frac{1-\pi_0}{\pi_0} \frac{\mathbb{E}[\pi_\infty \mid k = -]}{1 - \mathbb{E}\left[\pi_\infty \mid k = +, \pi_\infty \in \tilde{\Omega}_{TRAP}\right]}. \tag{22}$$

From the definition of $\tilde{\Omega}_{TRAP}$ we immediately obtain $\frac{\underline{r}}{\frac{1-\lambda_P}{\lambda_W}+\underline{r}} < \mathbb{E}\left[\pi_\infty \mid k=+, \pi_\infty \in \tilde{\Omega}_{TRAP}\right] \leq \frac{\underline{r}}{1+\underline{r}}$. Similarly, since $\Gamma_\infty^- \subset \tilde{\Omega}_{TRAP}$, then $\frac{\underline{r}}{\frac{1-\lambda_P}{\lambda_W}+\underline{r}} < \mathbb{E}\left[\pi_\infty \mid k=-\right] \leq \frac{\underline{r}}{1+\underline{r}}$. Combining these two inequalities with (22), and using the definition $r_0 \equiv \frac{\pi_0}{1-\pi_0}$ yields, after standard algebra, (10). This completes the proof of Proposition 5. **QED**

Proof of Corollary 1. Consider now the case of Corollary 1, $\lambda_W = \lambda_P = \lambda$. In this case, the state space of the stochastic process (9) is isomorphic to $\mathbb{Z}$: at any time horizon, the termination value of the process is fully characterized by the initial condition and by the total number of wars minus the total number of peace episodes (e.g., the termination value is the same after the sequence war-war-peace as after the sequence peace-war-war). This implies that, given π_0, the set of possible realizations belonging to the war trap, $\pi_\infty \in \tilde{\Omega}_{TRAP}$, is reduced to a singleton $\underline{\pi}(\pi_0)$ characterized by $\frac{\underline{\pi}(\pi_0)}{1+\underline{\pi}(\pi_0)} \equiv \frac{\pi_0}{1+\pi_0}\left(\frac{\lambda}{1-\lambda}\right)^{\Delta(r_0)}$ and $\Delta(r_0) \equiv \lceil (\ln r_0 - \ln \underline{r}) / (\ln(1-\lambda) - \ln\lambda) \rceil$. Hence, $\Gamma_\infty^+ = \{\underline{\pi}(\pi_0)\} \cup \{1\}$ and $\Gamma_\infty^- = \{\underline{\pi}(\pi_0)\}$. Consequently, $\mathbb{E}\left[\pi_\infty \mid k=-\right] = \mathbb{E}\left[\pi_\infty \mid k=+, \pi_\infty \in \tilde{\Omega}_{TRAP}\right] = \underline{\pi}(\pi_0)$. Thus, equation (22) simplifies to $\mathbb{P}_{TRAP} = \frac{1-\pi_0}{\pi_0}\frac{\underline{\pi}(\pi_0)}{1-\underline{\pi}(\pi_0)} = \left(\frac{\lambda}{1-\lambda}\right)^{\Delta(r_0)}$. This concludes the proof of Corollary 1. **QED**

APPENDIXES A, B and C (webpage)

This document comprises three appendixes. Appendix A provides details about the data and empirical analysis of section 1.1. Appendix B presents two extensions not discussed in the text. Finally, Appendix C provides technical details of the analysis in sections 3 and 6 in the text.

A Empirical analysis

In this appendix we will describe in more detail the data and specification used for the empirical analysis of Table 1. In Table 2 we provide a more detailed and complete version on the columns 1-8 of Table 1 (where the unit of observation is a country in a given five-year-period). Similarly, Table 3 contains a more detailed and complete version of columns 9-10 of Table 1 (where the unit of observation is a country in a given year).

Data and specification

Our dependent variable is civil war incidence, taken from the "UCDP/PRIO Armed Conflict Dataset" (UCDP, 2012), which is the most commonly used standard data source for civil wars at the country level.[32] While this data does not contain information about the number of fatalities in a given conflict and year, it codes two different intensity levels, "minor armed conflicts" (between 25 and 999 battle-related deaths in a given year) and "wars" (at least 1000 battle-related deaths in a given year). Our dependent variable in Table 2 (resp., in Table 3) is a dummy taking a value of 1 if in a given country and at any point in a given five-year period (resp., in a given year) a civil conflict took place, and 0 otherwise. While in all "odd" columns of the regression analysis we code as 1 all civil conflict observations, i.e. minor armed conflicts as well as wars, in all "even" columns we only code the dependent variable as 1 when a major war with at least 1000 battle-related deaths took place.

The first main explanatory variable is in both tables a dummy taking a value of 1 when there has been a war at any point during the last five years, where the war definition used is obviously the same as for the dependent variable (i.e. all conflicts for all odd columns, and only big wars for all even columns). The second main explanatory variable in both Tables, lagged average trust in a given country and year, is taken from the World Values Survey (2011). The World Values Survey trust surveys are only conducted every few years since 1981, and the newest available data —which is the one we are using— covers maximum five waves of surveys per country. The trust measure we use for a given country and five-year-period is the average proportion of respondents in the survey wave(s) taking place during this period and location who answer "Most people can be trusted" to the question "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" (A165).

We use a standard battery of control variables, which results in a specification that is extremely close to the core specifications run by Fearon and Laitin (2003), Collier and Hoeffler (2004), Montalvo and Reynal-Querol (2005), Cederman and Girardin (2007), Collier and Rohner (2008), and Esteban *et al.* (2012). Like these papers, we control for democracy, GDP per capita, natural resources (oil exporter), population size, ethnic fractionalization, and geography (mountainous terrain and noncontiguous states). These variables are described in more detail at the end of this appendix.

Results

Table 2 displays exactly the same regression results in the first eight columns as in columns 1-8 of Table 1 in the main text. The only difference is that we display –to avoid duplication– the point estimates,

[32] Recent papers that also focus on civil war incidents using the same war data like us include for example Besley and Persson (2011) and Esteban *et al.* (2012).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
War (t-1)	**3.13******	**3.67******	**2.65******	**2.92******	**3.12******	**2.69******	**3.28*****	**2.08**	**3.44******	**2.30**	**0.96******	**0.84*****
	(0.21)	**(0.26)**	**(0.24)**	**(0.28)**	**(0.46)**	**(0.47)**	**(1.29)**	**(1.50)**	**(1.19)**	**(1.64)**	**(0.23)**	**(0.34)**
Trust (t-1)							**-5.13*****	**-11.11******				
							(2.33)	**(3.37)**				
Democracy (t-1)			-0.00	-0.02	0.01	0.02	-0.04	-0.29*	-0.01	-0.18	0.01	-0.08*
			(0.02)	(0.03)	(0.02)	(0.04)	(0.13)	(0.18)	(0.11)	(0.17)	(0.03)	(0.05)
ln GDP p.c.(t-1)			-0.22	-0.34**	-0.10	-0.43*	0.63	0.02	0.32	-0.94	0.20	0.36
			(0.15)	(0.15)	(0.19)	(0.23)	(0.62)	(1.30)	(0.58)	(1.16)	(0.45)	(0.64)
Oil exporter (t-1)			0.30	0.40	0.63	0.62	1.15	-2.38	1.16	-1.68	0.18	0.85
			(0.31)	(0.43)	(0.41)	(0.75)	(0.79)	(2.29)	(0.81)	(2.51)	(0.89)	(1.42)
ln Popul.(t-1)			0.24***	0.34***	-0.13	0.09	0.72	2.38***	0.57	1.58***	0.57	-2.02
			(0.08)	(0.10)	(0.16)	(0.25)	(0.50)	(0.44)	(0.48)	(0.61)	(1.22)	(2.69)
Ethnic fractionaliz.			1.04**	0.82*	1.84***	1.86**	4.04***	1.66	3.59***	1.28		
			(0.44)	(0.49)	(0.64)	(0.83)	(1.29)	(3.00)	(1.31)	(2.46)		
Mountainous Terrain			0.01*	0.01**	0.01*	0.04**	0.02	0.12***	0.02	0.13***		
			(0.00)	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)	(0.02)	(0.04)		
Noncontiguous state			0.76*	0.51	1.52***	1.59***	0.64	4.31***	0.47	3.93**		
			(0.41)	(0.44)	(0.49)	(0.53)	(1.19)	(1.33)	(1.24)	(1.72)		
Conflicts coded as war	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.
Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	WVS	WVS	WVS	WVS	WVS	WVS	All	All
Observations	1426	1426	1026	939	409	378	101	101	101	101	530	265
Pseudo R-squared	0.304	0.322	0.363	0.358	0.460	0.392	0.575	0.572	0.565	0.544	0.106	0.182

Dependent variable: civil war incidence (five-year intervals). The dependent variable is coded as 1 if a conflict causing at least 25 (1000) fatalities is recorded in at least one of the five years. Sample period: 1949-2008. Number of countries for which observations are available: 174. The set of controls include the variables listed as well as region fixed effects and time dummies. Columns 1 to 10 contain logit regressions with robust standard errors, clustered at the country level. Columns 11-12 contain country fixed effects logit regressions. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 2: Persistence of civil conflicts and correlation between conflict and lagged trust (frequency: five-years).

while in Table 1 we displayed the marginal effects. Further, Table 2 also displays the coefficients of the control variables. In addition to these eight first columns, Table 2 contains four additional columns not included in Table 1. Columns 9-10 display the same specification without lagged trust as in columns 5-6, but restricting the sample to the same 101 observations as in columns 7-8 where lagged trust is included. This allows us to see that the change in the magnitude of coefficients of the lagged war variable between 5-6 and 7-8 is mostly due to the drop in sample size rather than to the change in specification.

Columns 11-12 display the same specification as in columns 5-6, but including country fixed effects (which leads to a drop of the time invariant controls, ethnic fractionalization, mountainous terrain, and noncontiguous state).[33] We find that there is still statistically significant persistence of war, even when controlling for country fixed effects.

Table 3 is the mirror image of Table 2, but displays the point estimates of the variables of interest and of all controls for the specification where the unit of observation is a country in a given year. After displaying in columns 1-6 the results on persistence when the lagged trust measure is not included, the full specification with all controls and lagged trust in columns 7-8 of Table 3 displays the point estimates corresponding to columns 9-10 of Table 1. In columns 9-10 of Table 3 we display again the results when running the specification of columns 5-6 on the restricted sample of observations included in columns 7-8. While, again, most of the change in the lagged war coefficients from 5-6 to 7-8 comes from the drop in sample size, the inclusion of lagged trust accounts for a substantial additional drop of the lagged war coefficient in column 8.

Columns 11-12 show that the persistence of war holds up to the inclusion of country fixed effects.

[33] As discussed in the main text, we are not able to include lagged trust in the presence of country fixed effects, as there are only very few countries that have both multiple observations of lagged trust and variation in the war variable for the periods in which lagged trust is available.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
War (t-1)	**4.17*** **	**4.41*** **	**3.93*** **	**4.43*** **	**4.22*** **	**4.10*** **	**5.38*** **	**1.57**	**4.98*** **	**2.38*** **	**2.66*** **	**2.68*** **
	(0.18)	**(0.24)**	**(0.26)**	**(0.31)**	**(0.38)**	**(0.48)**	**(0.85)**	**(1.18)**	**(0.79)**	**(0.69)**	**(0.15)**	**(0.25)**
Trust (t-1)							**-10.85*** **	**-15.44** **				
							(2.40)	**(6.82)**				
Democracy (t-1)			0.04*	0.02	0.06**	0.05	0.02	0.08	0.07	0.18	0.01	-0.02
			(0.02)	(0.03)	(0.03)	(0.05)	(0.07)	(0.13)	(0.08)	(0.15)	(0.02)	(0.03)
ln GDP p.c.(t-1)			0.06	-0.24*	0.05	-0.24	0.55	-0.15	-0.19	-1.76***	0.63**	1.30**
			(0.16)	(0.12)	(0.19)	(0.19)	(0.36)	(1.49)	(0.44)	(0.59)	(0.31)	(0.64)
Oil exporter (t-1)			0.14	0.30	0.14	0.78	1.28**	0.25	0.37	0.99	0.91**	2.01**
			(0.31)	(0.33)	(0.42)	(0.48)	(0.65)	(2.23)	(0.83)	(1.74)	(0.44)	(0.83)
ln Popul.(t-1)			0.29***	0.35***	0.17	0.39***	0.80***	2.96**	0.36	1.11***	1.31	-0.21
			(0.10)	(0.12)	(0.17)	(0.15)	(0.28)	(1.37)	(0.28)	(0.42)	(0.80)	(1.59)
Ethnic fractionaliz.			0.27	-0.22	0.07	0.34	1.75*	2.11	0.17	-0.57		
			(0.46)	(0.44)	(0.60)	(0.80)	(1.01)	(5.58)	(1.39)	(1.89)		
Mountainous Terrain			1.00**	0.47	1.63**	2.61***	1.32	15.97	3.22	15.47**		
			(0.49)	(0.80)	(0.74)	(0.79)	(1.78)	(10.29)	(2.38)	(6.75)		
Noncontiguous state			0.42	0.18	0.55	0.71*	1.29	4.68	1.37	4.48*		
			(0.48)	(0.51)	(0.57)	(0.37)	(0.79)	(2.93)	(0.85)	(2.42)		
Conflicts coded as war	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.	>25 Fat.	>1000 Fat.
Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	WVS	WVS	WVS	WVS	WVS	WVS	All	All
Observations	7613	7613	5222	4808	2707	2248	564	439	564	439	2955	1269
Pseudo R-squared	0.443	0.430	0.502	0.523	0.539	0.523	0.695	0.597	0.652	0.579	0.259	0.337

Dependent variable: civil war incidence (annual observations). The dependent variable is coded as 1 if a conflict causing at least 25 (1000) fatalities is recorded in the year of observation. Sample period: 1946-2008. Number of countries for which observations are available: 175. The set of controls include the variables listed as well as region fixed effects and time dummies. Columns 1 to 10 contain logit regressions with robust standard errors, clustered at the country level. Columns 11-12 contain country fixed effects logit regressions. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Persistence of civil conflicts and correlation between conflict and lagged trust (frequency: annual).

Description of variables used

The dependent variable, civil war incidence, and the main independent variables, lagged war and lagged trust, have been described above. In what follows we describe the control variables.

Democracy: Polity scores ranging from -10 (strongly autocratic) to +10 (strongly democratic). From Polity IV (2012).

GDP per capita: PPP adjusted GDP per capita at constant prices. From the Penn World Tables (Heston *et al.*, 2011).

Oil exporter: Dummy variable taking a value of 1 if in a given country and year the fuel exports (in % of merchandise exports) is above 33%. Variable from Fearon and Laitin (2003), but updated with recent data of the variable "fuel exports (in % of merchandise exports)" from World Bank (2012).

Population: Total population. From World Bank (2012).

Ethnic Fractionalization: Index of ethnic fractionalization. From Fearon and Laitin (2003).

Mountainous Terrain: Percentage of territory covered by mountains. From Collier *et al.* (2009).

Noncontiguous State: Dummy taking a value of 1 if a state has noncontiguous territory. From Fearon and Laitin (2003).

B Additional extensions

This appendix discuss two extensions.

B.1 Low value of war

In the analysis in the text, we have restricted attention to parameters such that the uncivic type always wages war under BAU. In this appendix we assume, instead, that $V \leq S^{-}(\infty) < S^{+}(\infty)$, implying that the uncivic type chooses peace for a region of high beliefs. The new insight is the existence of two

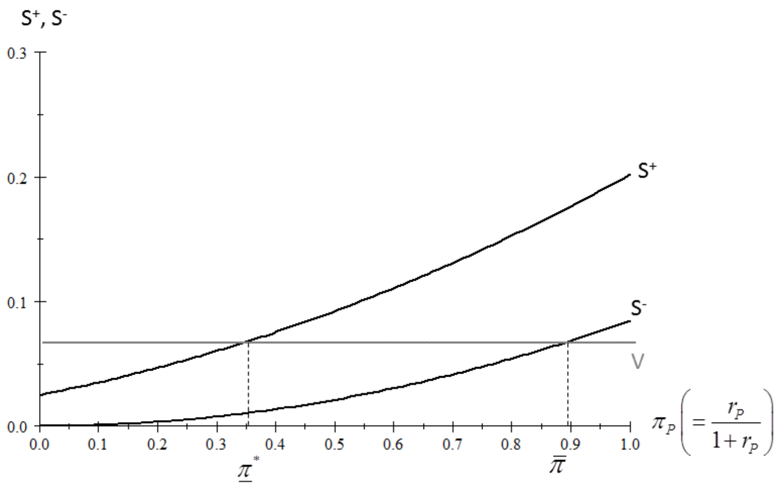


Figure 5: Surplus from trade and war benefits; the case of two traps.

learning traps, one with frequent and one with rare wars. Consider Figure 5. The difference relative to Figure 1 is that there exists a high range of posterior beliefs, $\pi_P \geq \bar{\pi}$, such that neither types find it optimal to wage war. In such a range, the equilibrium is uninformative and peace prevails even though group A is uncivic.

In this section, we first outline the results in an intuitive fashion. Then, we present technical details in section B.1.1.

The equilibrium dynamics continues to be characterized by equations (6),(7) and (8). However, for a range of prior r_t in the left-hand neighborhood of $\bar{r}$, the uncivic type is now indifferent between waging war and keeping peace. Then, the (unique) PBE prescribes that an uncivic group A chooses a strictly mixed strategy under BAU, $\sigma^- \in (0, 1)$ (see Lemma 2 in section B.1.1). In such a range, the equilibrium is informative (since $\sigma^+ = 1 > \sigma^-$), but σ^- increases with r_{t-1}, and war/peace becomes less informative as r_{t-1} grows. Finally, as $r \geq \bar{r}$, both groups stick to peace and the equilibrium turns uninformative.

The equilibrium dynamics is represented in Figure 6 (see Proposition 6 in section B.1.1). For $r < \bar{r}^*$, it is isomorphic to Figure 2. In the interval $[\bar{r}^*, \bar{r}]$, group A (if uncivic) randomizes between war and peace. If peace is the outcome of the randomization at $t-1$, beliefs get stuck to $r_{t+s} = \bar{r}$ for all $s \geq 0$. $\bar{r}$ is an absorbing state: if the prior beliefs is $\bar{r}$, both types retain peace under BAU, and the posterior belief is also $\bar{r}$. The set of priors $r > \bar{r}$ would also give rise to stationary beliefs, but it is never reached in equilibrium unless the economy starts in that region. If we define a peace trap (Π_{TRAP}) to be the mirror image of a war trap, then $\Pi_{TRAP} = [\bar{r}, \infty)$.

If the initial prior lies in the informative region $[\underline{r}, \bar{r}]$, the belief process follows initially the stochastic dynamics given by (6)-(7), converging eventually to either the war or the peace trap. The process cannot stay forever in the informative region, or, otherwise, agents could observe an infinitely large sample of realizations of the war/peace process, and thus learn the true type of group A, by virtue of the strong law of large numbers. However, perfect learning would be inconsistent with the beliefs staying in the informative region, $[\underline{r}, \bar{r}]$.

In section B.1.1 (Proposition 7), we provide a formal characterization of the long-run probability distribution. Intuitively, the economy can get stuck in the "wrong" beliefs with a positive probability. Namely, it is possible that group A is civic, and yet the economy falls into a war trap after a sequence of war shocks. Conversely, it is possible that group A is uncivic, and yet the economy falls into a peace trap after a sequence of peace shocks.

Peace traps dominate in welfare terms the perfect information equilibrium outcome when A is

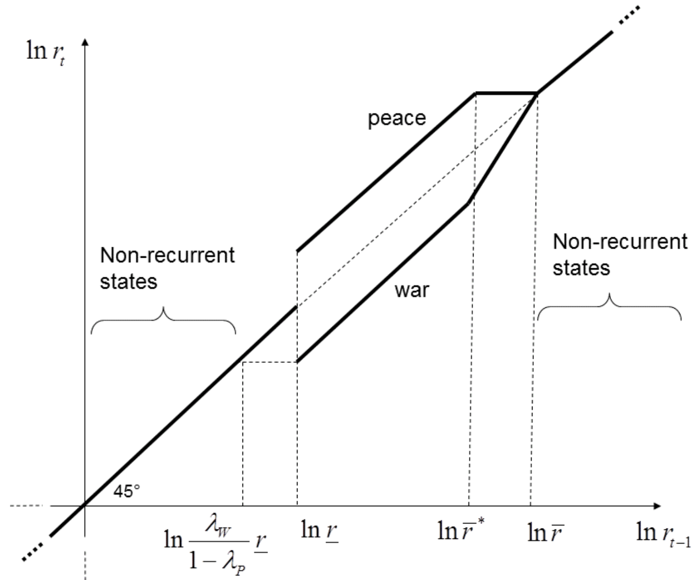


Figure 6: Law of motion of beliefs; the case with two traps.

uncivic. They entail fewer wars and more cooperation.[34] Interestingly, when A is civic, the best long run outcome is less efficient than the best long run outcome in the benchmark of a high value of war. To see why, recall that when the value of war is high, group B can learn (almost) perfectly that group A is civic, and the allocation converges to the perfect information equilibrium. In contrast, in the peace trap there is some persistent signal jamming, implying that the likelihood ratio never exceeds $\bar{r}$. Thus, the peace trap features the same (low) probability of war as the perfect information equilibrium but delivers less trust and cooperation.

B.1.1 Technical analysis

Notation 2 *Let $\bar{r}$ be such that $V = S^{-}(\bar{r})$ and let $\bar{r}^{*} \equiv \frac{\lambda_P}{1-\lambda_W}\bar{r}$.*

Intuitively, $\bar{r}$ is the threshold posterior belief such that both types retain peace under BAU if $r_P \geq \bar{r}$. As long as $r_{-1} \geq \frac{\lambda_P}{1-\lambda_W}\bar{r}$, the posterior can be larger or equal to $\bar{r}$.

Remark 1 $\underline{r} < \underline{r}^{*} < \bar{r}^{*} < \bar{r}$.

Given these definitions, the following Lemma can be established.

Lemma 2 *Assume $V > S^{+}(0)$ and $V \leq S^{-}(\infty) < S^{+}(\infty)$. For $r_{-1} \leq \underline{r}$ the PBE is unique and uninformative. For $r_{-1} \in [\underline{r}, \underline{r}^{*}]$ there are multiple PBE. For $r_{-1} \in [\underline{r}^{*}, \bar{r}^{*}[$ the PBE is unique and informative. For $r_{-1} \in [\bar{r}^{*}, \bar{r}[$ the PBE is unique and informative but involves mixed strategy: $\sigma^{+}(r_{-1}) = 1$ and $\sigma^{-}(r_{-1}) = \frac{(1-\lambda_W)\frac{r_{-1}}{\bar{r}} - \lambda_P}{1-\lambda_W-\lambda_P}$. For $r_{-1} \geq \bar{r}$ the PBE is unique and uninformative.*

Proof. The analysis of the range $r_{-1} < \bar{r}^{*}$ is identical to the proof of Lemma 1. Therefore, we only focus here on the range $r_{-1} \geq \bar{r}^{*}$.

[34] Note, though, that group B may suffer losses in the trade game due to an excessive optimism, which induces its members to overcooperate vis-a-vis an uncivic group with a high propensity to defect.

We start by proving that, if we restrict attention to the range $r_{-1} \geq \bar{r}^*$, an uninformative PBE exists if and only if $r_{-1} \geq \bar{r}$. To this aim, we first prove the "if" part. Consider a prior $r_{-1} \geq \bar{r}$. The posterior $r_P(r_{-1})$ cannot be lower than r_{-1}. Hence $r_P(r_{-1}) \geq \bar{r}$, and this implies that $V \leq S^-(r_P(r_{-1})) < S^+(r_P(r_{-1}))$. Thus, $\sigma^+(r_{-1}) = \sigma^-(r_{-1}) = 1$, and this in turn means that the PBE is uninformative with $r_P(r_{-1}) = \bar{r}$. Second, we prove (by contradiction) the "only if" part. Suppose that an uninformative PBE exists in such a range. Then, by the definition of uninformative equilibrium, $r_P(r_{-1}) = r_{-1} \in [\bar{r}^*, \bar{r}[$, which in turn implies that $S^+(r_P(r_{-1})) > V$ and $S^-(r_P(r_{-1})) < V$. Thus, $\sigma^+(r_{-1}) = 1$ and $\sigma^-(r_{-1}) = 0$, contradicting that the PBE is uninformative.

Next, we prove that a unique informative equilibrium exists in the range $r_{-1} \in [\bar{r}^*, \bar{r}]$. We start by proving that an informative pure-strategy PBE does not exist. Suppose, to derive a contradiction, that such a PBE exists. The PBE would then feature $\sigma^+(r_{-1}) = 1$ and $\sigma^-(r_{-1}) = 0$. But, then, Bayes' rule implies that $r_P(r_{-1}) = \frac{1-\lambda_W}{\lambda_P} r_{-1} > \bar{r}$. This would imply $V \leq S^-(r_P(r_{-1})) < S^+(r_P(r_{-1}))$ and, thus, $\sigma^+(r_{-1}) = \sigma^-(r_{-1}) = 1$. This would imply a contradiction. As a consequence the PBE must be a mixed-strategy equilibrium. We guess that the equilibrium has the following form: $\sigma^+(r_{-1}) = 1$ (the civic type chooses peace with probability one) and $\sigma^-(r_{-1}) = \frac{(1-\lambda_W)\frac{r_{-1}}{\bar{r}} - \lambda_P}{1-\lambda_W-\lambda_P} \in (0,1)$ (the uncivic type randomizes, choosing peace with probability larger than zero and smaller than one). Bayes' rule implies then that $r_P(r_{-1}) = \bar{r}$. Since $V = S^-(\bar{r}) < S^+(\bar{r})$, we have verified the guess by showing that indeed the civic type strictly prefers peace, whereas the uncivic type is indifferent between war and peace. ■

The next proposition follows immediately from the selection criterion 3 and from Lemma 2.

Proposition 6 *Assume that $V > S^+(0)$ and $V \leq S^-(\infty) < S^+(\infty)$. The DSE is characterized as follows:*
The PBE at time t is unique and characterized by Proposition 3 and by the following law of motion:

$$\ln r_t = \begin{cases} \ln r_{t-1} & \text{if} \quad r_{t-1} \in [0, \underline{r}) \cup]\bar{r}, \infty) \\ \ln r_{t-1} + (1 - \mathbb{W}_t) \ln\left(\frac{1-\lambda_W}{\lambda_P}\right) - \mathbb{W}_t \ln\left(\frac{1-\lambda_P}{\lambda_W}\right) & \text{if} \quad r_{t-1} \in [\underline{r}, \bar{r}^*] \\ (1 - \mathbb{W}_t) \ln \bar{r} + \mathbb{W}_t \left[\ln r_{t-1} - \ln \frac{1 - r_{t-1}(1-\lambda_W)/\bar{r}}{\lambda_W}\right] & \text{if} \quad r_{t-1} \in [\bar{r}^*, \bar{r}] \end{cases} . \tag{23}$$

If group A is civic ($k = +$), the probability of war is

$$\mathbb{P}(\mathbb{W}_t = 1) = \begin{cases} 1 - \lambda_P & \text{if} \quad r_{t-1} \in [0, \underline{r}[\\ \lambda_W & \text{if} \quad r_{t-1} \geq \underline{r} \end{cases} .$$

If group A is uncivic ($k = -$), the probability of war is

$$\mathbb{P}(\mathbb{W}_t = 1) = \begin{cases} 1 - \lambda_P & \text{if} \quad r_{t-1} \leq \bar{r}^* \\ 1 - (1 - \lambda_W)\frac{r_{-1}}{\bar{r}} & \text{if} \quad r_{t-1} \in [\bar{r}^*, \bar{r}] \\ \lambda_W & \text{if} \quad r_{t-1} \in]\bar{r}, \infty] \end{cases} ,$$

We now characterize the asymptotic dynamics of beliefs in the following proposition:

Proposition 7 *$V > S^+(0)$ and $V \leq S^-(\infty) < S^+(\infty)$, and $r_0 \in [\underline{r}, \bar{r}]$. Then, both when $k = +$ and when $k = -$, the DSE exits the informative equilibrium regime almost surely, and learning comes to* *a halt in finite time. The final belief is such that with probability $\mathbb{P}_{TRAP} > 0$ the economy is stuck in a war trap and with probability $1 - \mathbb{P}_{TRAP} > 0$ it is stuck in a peace trap. The probability has the following bounds*

$$\frac{\frac{\bar{r}}{1+\bar{r}} - \frac{r_0}{1+r_0}}{\frac{\bar{r}}{1+\bar{r}} - \frac{\underline{r}}{\frac{1-\lambda_P}{\lambda_W}+\underline{r}}} < \mathbb{P}_{TRAP} \leq \frac{\frac{\bar{r}}{1+\bar{r}} - \frac{r_0}{1+r_0}}{\frac{\bar{r}}{1+\bar{r}} - \frac{\underline{r}}{1+\underline{r}}}$$

Proof. We apply the same type of argument as in the proof of Proposition 5. Belief π_t being a bounded martingale taking values in $[0, 1]$, the Martingale Convergence Theorem implies that π_t converges almost surely to a random variable π_∞ with a support Γ_∞. Clearly, $\underline{r}$ being an absorbing state of the dynamics (23), we have: $\Gamma_\infty = \tilde{\Omega}_{TRAP} \cup \{\frac{\bar{r}}{1+\bar{r}}\}$ with $\tilde{\Omega}_{TRAP} \equiv]\frac{\underline{r}}{\frac{1-\lambda_P}{\lambda_W}+\underline{r}}, \frac{\underline{r}}{1+\underline{r}}]$. Let us characterize now $\mathbb{P}_{TRAP} = \mathbb{P}[\pi_\infty \in \tilde{\Omega}_{TRAP}] = 1 - \mathbb{P}[\pi_\infty = \frac{\bar{r}}{1+\bar{r}}]$. Since the belief π_t is a Martingale, we have $\forall t, \pi_0 = \mathbb{E}[\pi_t]$. Taking the limit as $t \to +\infty$ this leads to

$$\pi_0 = \mathbb{P}_{TRAP} \times \mathbb{E}\left[\pi_\infty \mid \pi_\infty \in \tilde{\Omega}_{TRAP}\right] + (1 - \mathbb{P}_{TRAP}) \times \frac{\bar{r}}{1+\bar{r}} \tag{24}$$

This yields

$$\mathbb{P}_{TRAP} = \frac{\frac{\bar{r}}{1+\bar{r}} - \pi_0}{\frac{\bar{r}}{1+\bar{r}} - \mathbb{E}\left[\pi_\infty \mid \pi_\infty \in \tilde{\Omega}_{TRAP}\right]} \tag{25}$$

We now aim to bound $\mathbb{P}_{TRAP}$ in the previous equation. Given that $\tilde{\Omega}_{TRAP} =]\frac{\underline{r}}{\frac{1-\lambda_P}{\lambda_W}+\underline{r}}, \frac{\underline{r}}{1+\underline{r}}]$ we have

$$\frac{\underline{r}}{\frac{1-\lambda_P}{\lambda_W}+\underline{r}} < \mathbb{E}\left[\pi_\infty \mid \pi_\infty \in \tilde{\Omega}_{TRAP}\right] \leq \frac{\underline{r}}{1+\underline{r}} \tag{26}$$

Combining (25) and (26) and noting that $\pi_0 = \frac{r_0}{1+r_0}$ we obtain

$$\frac{\frac{\bar{r}}{1+\bar{r}} - \frac{r_0}{1+r_0}}{\frac{\bar{r}}{1+\bar{r}} - \frac{\underline{r}}{\frac{1-\lambda_P}{\lambda_W}+\underline{r}}} < \mathbb{P}_{TRAP} \leq \frac{\frac{\bar{r}}{1+\bar{r}} - \frac{r_0}{1+r_0}}{\frac{\bar{r}}{1+\bar{r}} - \frac{\underline{r}}{1+\underline{r}}}$$

■

We refer the interested reader to the working paper version (Rohner, Dominic, Mathias Thoenig and Fabrizio Zilibotti, 2011. "War Signals: A Theory of Trade, Trust and Conflict," CEPR Discussion Papers 8352) where we use analytical tools from stopping time theory. These methods are more involved but allow us to provide (more accurate) type-dependent bounds $\mathbb{P}^-_{TRAP}$ and $\mathbb{P}^+_{TRAP}$.

B.2 Altruism and dynamic game

The war decision entails an intergenerational spillover. War depletes trust and harms future generations in both groups. However, in the analysis in the text, agents have no concern for future generations, and ignore such a spillover. In this appendix, we consider an extension of the basic model where the decision to wage war incorporates an altruistic concern towards the next generation.[35]

Since within-group inefficiencies are ruled out by intra-group transfers, each war can be viewed as the decision of a single agent (which we label group A *planner*) in each cohort. Thus, war becomes a dynamic game between subsequent group A planners. We follow the recent politico-economic literature (see, e.g., Hassler *et al.,* 2003, and Song *et al.*, 2012), and focus on Markov Perfect Equilibria (MPE),

[35] The goal of this extension is to establish that the results of the benchmark model are robust to intergenerational altruism, namely, an equilibrium isomorphic to the DSE of section 5 can be sustained when agents are forward-looking. A characterization of the whole set of equilibria in the new environment is beyond the scope of the analysis.

in which strategies are conditioned on a vector of payoff relevant state variables. In our case, the only state variable is the state of beliefs.

Consider the following recursive representation of group A's value function:

$$W^k(r_{t-1}) = \max_{\sigma^k \in [0,1]} \tilde{W}^k(\sigma^k; r_{t-1}), \tag{27}$$

$$\begin{aligned} \tilde{W}^k(\sigma^k; r_{t-1}) \quad \equiv \quad & \Big[(1-\lambda_W-\lambda_P)\sigma^k + \lambda_P\Big] \times \Big[S^k(r_P(r_{t-1})) + \beta W^k(r_P(r_{t-1}))\Big] \\ & +(1-\lambda_W-\lambda_P)(1-\sigma^k) \times \Big[V + \beta W^k(r_W(r_{t-1}))\Big] + \lambda_W \times \Big[V_H + \beta W^k(r_W(r_{t-1}))\Big] \end{aligned} \tag{28}$$

where $k \in \{+,-\}$, $\beta \in (0,1)$ is an intergenerational discount factor, and S^k is determined as in section 4.

Definition 5 *A Markov Perfect Political Equilibrium (MPE) is a pair of functions, $\langle \Sigma^+, \Sigma^- \rangle$, where $\Sigma^k : [0,\infty] \to [0,1]$ is a "conflict rule" such that $\Sigma^k(r) = \sigma^k = \arg\max_{\{\sigma^k \in [0,1]\}} \tilde{W}^k(\sigma^k; r)$ where $\tilde{W}^k(\sigma^k; r)$ is defined as in (27)-(28) and where the dynamics of posterior beliefs are given by*

$$\ln r_P(r) \quad = \quad \ln r + \ln \frac{\lambda_P + (1-\lambda_W-\lambda_P)\Sigma^+(r)}{\lambda_P + (1-\lambda_W-\lambda_P)\Sigma^-(r)}, \tag{29}$$

$$\ln r_W(r) \quad = \quad \ln r - \ln \frac{1-\lambda_P - (1-\lambda_W-\lambda_P)\Sigma^-(r)}{1-\lambda_P - (1-\lambda_W-\lambda_P)\Sigma^+(r)}, \tag{30}$$

In words, group A makes the current war/peace decision conditional on current beliefs, under the rational expectation that future decisions to wage war will follow the equilibrium conflict rules, $\langle \Sigma^+, \Sigma^- \rangle$. Furthermore, the vector of policy functions determined by the optimal choice is a fixed point of the system of functional equations resulting from the constrained maximization in (27). Note that cooperation in the trade game continues to be determined by a sequence of static decisions, since individual traders act atomistically.

Proposition 8 *Suppose that $\lambda_P \frac{\beta}{1-\beta} \leq \frac{V-S^-(\infty)}{S^-(\infty)-S^-(0)}$ (sufficient condition). Then, there exists a MPE such that, conditional on parameters, the probability of war, the extent of cooperation under peace, and the equilibrium dynamics of beliefs are identical to those in Proposition 4.*

Proof. We proceed by guessing the equilibrium policy (war/peace) function, $\langle \Sigma^+, \Sigma^- \rangle$, and then verifying that the guesses are consistent with the equilibrium. We guess:

$$\Sigma^-(r) = 0 \text{ and } \Sigma^+(r) = \begin{cases} 0 & \text{if} \quad r < \underline{r} \\ 1 & \text{if} \quad r \geq \underline{r} \end{cases}. \tag{31}$$

We prove that neither group has any incentive to deviate from (31), if future generations follow the equilibrium policy guessed, (31).

Equation (27) implies that

$$\frac{d\tilde{W}^k(\sigma^k; r_{t-1})}{d\sigma^k} = (1-\lambda_W-\lambda_P) \times \Big[S^k(r_P(r_{t-1})) - V + \beta \times \Big(W^k(r_P(r_{t-1})) - W^k(r_W(r_{t-1}))\Big)\Big] \tag{32}$$

Consider, first, the range $r < \underline{r}$, where $\Sigma^k(r) = 0$ for both types. Substituting the guess, (31), into (29)-(30) yields $r_P(r_{t-1}) = r_W(r_{t-1}) = r_{t-1}$. Thus, the sign of $d\tilde{W}^k(\sigma^-; r_{t-1})/d\sigma^k$ is determined by the sign of $S^k(r_{t-1}) - V$, which is negative for both $k \in \{+,-\}$. This implies that the optimal choice is $\sigma^k(r_{t-1}) = 0 = \Sigma^k(r)$ for both k. This proves that neither type faces a profitable deviation from the guessed policy rule, (31), in the range $r < \underline{r}$.

Next, we move to the range $r \geq \underline{r}$. Consider the civic type: we claim that $d\tilde{W}^+(\sigma^-; r_{t-1})/d\sigma^+ > 0$, since $S^+(r_P(r_{t-1}) \geq V$, and $W^+(r_P(r_{t-1})) > W^+(r_W(r_{t-1}))$. Thus, the optimal choice is $\sigma^+(r_{t-1}) = 1 = \Sigma^+(r)$. This proves that there is no profitable deviation from (31) for the civic type in the range $r \geq \underline{r}$. Consider, next, the uncivic type: the sign of $d\tilde{W}^-(\sigma^-; r_{t-1})/d\sigma^-$ is in general ambiguous, since $S^-(r_P(r_{t-1}) < V$, and $W^-(r_P(r_{t-1})) > W^-(r_W(r_{t-1}))$. We now prove that under the parameter restriction of Proposition 8, the first term dominates, and $d\tilde{W}^-(\sigma^-; r_{t-1})/d\sigma^- < 0$. From (32), this is the case if and only if

$$\beta \times \left(W^-(r_P(r_{t-1})) - W^-(r_W(r_{t-1}))\right) < V - S^-(r_P(r_{t-1})) \tag{33}$$

First, note that $V - S^-(\infty)$ is a lower bound to the right hand side of (33). Second, note that $\beta \times (W^-(\infty) - W^-(0))$ is an upper bound to the left hand side of (33), since W is increasing in r. Therefore, the following condition is sufficient to ensure that $d\tilde{W}^-(\sigma^-; r_{t-1})/d\sigma^- < 0$:

$$\beta \times \left(W^-(\infty) - W^-(0)\right) < V - S^-(\infty) \tag{34}$$

Given that, under the proposed policy rule ($\Sigma^- = 0$), $W^-(\infty) = \frac{\lambda_P}{1-\beta} S^-(\infty) + \frac{1-\lambda_P-\lambda_W}{1-\beta} V + \frac{\lambda_W}{1-\beta} V_H$ and $W^-(0) = \frac{\lambda_P}{1-\beta} S^-(0) + \frac{1-\lambda_P-\lambda_W}{1-\beta} V + \frac{\lambda_W}{1-\beta} V_H$, the sufficient condition (34) can be rewritten as

$$\frac{\beta\lambda_P}{1-\beta} \leq \frac{V - S^-(\infty)}{S^-(\infty) - S^-(0)},$$

which is the condition given in the statement of the Proposition. Under this sufficient condition, $\sigma^-(r_{t-1}) = 0 = \Sigma^-(r)$, so the uncivic type faces no profitable deviation from the guessed policy rule Σ^-. This concludes the proof of the Proposition. ■

The intuition of the proof is the following. When the economy is in the war trap there is no learning, hence, no intergenerational spillover. Thus, the optimal war/peace decision is not affected by altruism. In the informative region, however, there is an intergenerational spillover, and the altruistic concern can induce the uncivic to mimic the civic type and keep peace under BAU in order to increase trust and cooperation during future peace spells. The sufficient condition of Proposition 8 rules out that such a mimicking deviation is profitable. In particular, it guarantees that $V - S^-(r_P(r)) > \beta(W^-(r_P(r)) - W^-(r_W(r)))$, i.e., the static gain from waging war over retaining peace exceeds the discounted difference between the continuation values after peace and war, respectively. Intuitively, a low β reduces the scope of a one-period deviation since the benefits of the deviation accrue to future generations. The probability of peace shocks, λ_P, also matters, since along the equilibrium path (after the deviation), an uncivic group A will only retain peace when peace shocks arise. It is therefore only under peace shocks that trade and beliefs matter for future generations. Hence, frequent peace shocks make it harder, *ceteris paribus*, to sustain the equilibrium of Proposition 8.

C Technical details of the analysis in the main text

C.1 Analysis of Section 3.1 (welfare analysis)

Consider two planners who are entrusted to choose which agents cooperate and which defect. Each planner maximizes the sum of the trade surplus of its own group. Note that the pay-off includes the taste for cooperation, or its opposite, i.e., the dislike of some individuals for cooperation with the other group. We consider two cases. In the first, the two planners play a Nash game, i.e., they decide taking the behavior of the other group as given. In the second, they cooperate so as to maximizes the sum of the welfare of the two groups (economy-wide level efficiency).

In all cases, the planners ask more cooperative players (high $\mathcal{P}_i$) to cooperate, and, possibly, some less cooperative players (low $\mathcal{P}_i$) to defect. Thus, the planners will adopt a threshold rule. We shall denote by $\bar{\mathcal{P}}_A^k$ and $\bar{\mathcal{P}}_B^k$, respectively, the threshold agent in group A and group B conditional on the group A's type being $k \in \{+,-\}$.

It is useful to recall here the notation introduced in the text:

Notation 3 *Let (i)* $z \equiv c-(d-l)$; *(ii)* $\mathcal{L}_i \equiv l-\mathcal{P}_i$; *(iii)* $f^J(\mathcal{L}) \equiv h^J(l-\mathcal{L})$ *and* $F^J(\mathcal{L}) \equiv 1-H^J(l-\mathcal{L})$, *with* $J \in \{+,-,B\}$.

C.1.1 Nash equilibrium between group planners

The Nash planner of group A chooses $\bar{\mathcal{P}}_A^k$ in order to maximize S^k, taking as given the proportion $n_B^k \equiv F^B\left(\bar{\mathcal{L}}_B^k\right)$ of cooperators in group B. S^k is defined as

$$\begin{aligned} S^k &\equiv \int_{-\infty}^{\bar{\mathcal{P}}_A^k} d\, dH^k(\mathcal{P}) + \int_{\bar{\mathcal{P}}_A^k}^{\infty} \left(n_B^k(c+\mathcal{P}_i) + \left(1-n_B^k\right)(d-l+\mathcal{P}_i)\right) dH^k(\mathcal{P}) \\ &= d\int_{-\infty}^{\bar{\mathcal{P}}_A^k} dH^k(\mathcal{P}) + \int_{\bar{\mathcal{P}}_A^k}^{\infty} \left(n_B^k(c-(d-l)) + (d-l)+\mathcal{P}_i\right) dH^k(\mathcal{P}) \\ &= d + \int_{\bar{\mathcal{P}}_A^k}^{\infty} \left(n_B^k z + (\mathcal{P}_i - l)\right) dH^k(\mathcal{P}) \end{aligned}$$

The maximization with respect to $\bar{\mathcal{P}}_A^k$ yields

$$\max_{\bar{\mathcal{P}}^k} \int_{\bar{\mathcal{P}}_A^k}^{\infty} \left(n_B^k z + (\mathcal{P}_i - l)\right) dH^k(\mathcal{P}) \Rightarrow z n_B^k = l - \bar{\mathcal{P}}_A^k$$

Using the definition $n_B^k = F^B\left(\bar{\mathcal{L}}_B^k\right)$ we get

$$\bar{\mathcal{L}}_A^k = zF^B(\bar{\mathcal{L}}_B^k) \tag{35}$$

The same argument leads one to conclude that the planner of group B chooses

$$\bar{\mathcal{L}}_B^k = zF^k(\bar{\mathcal{L}}_A^k) \tag{36}$$

Using again the definitions of n_A^k and n_B^k the first order conditions (35) and (36) give

$$\begin{aligned} n_A^k &= F^k(n_B^k) \\ n_B^k &= F^B(n_A^k) \end{aligned}$$

We conclude that the allocation chosen by the Nash planners is identical to the decentralized equilibrium.

C.1.2 Cooperative solution (economy-wide first best)

The efficient solution maximizes

$$\begin{aligned} S^k + S^B &\equiv \int_{-\infty}^{\bar{\mathcal{P}}_A^k} d\, dH^k(\mathcal{P}) + \int_{\bar{\mathcal{P}}_A^k}^{\infty} \left(n_B^k(c+\mathcal{P}_i) + \left(1-n_B^k\right)(d-l+\mathcal{P}_i)\right) dH^k(\mathcal{P}) + \\ &\quad \int_{-\infty}^{\bar{\mathcal{P}}_B^k} d\, dH^k(\mathcal{P}) + \int_{\bar{\mathcal{P}}_B^k}^{\infty} \left(n_A^k(c+\mathcal{P}_i) + \left(1-n_A^k\right)(d-l+\mathcal{P}_i)\right) dH^k(\mathcal{P}) + \\ &= 2d + \int_{\bar{\mathcal{P}}_A^k}^{\infty} \left(n_B^k z + (\mathcal{P}_i - l)\right) dH^k(\mathcal{P}) + \int_{\bar{\mathcal{P}}_B^k}^{\infty} \left(n_A^k z + (\mathcal{P}_i - l)\right) dH^B(\mathcal{P}), \end{aligned}$$

where $n_A^k = H^k\left(\bar{\mathcal{P}}_A^k\right)$ and $n_B^k = H^B\left(\bar{\mathcal{P}}_B^k\right)$. Plugging in the expressions of n_A^k and n_B^k yields

$$S^k + S^B \equiv 2d + \int_{\bar{\mathcal{P}}_A^k}^{\infty}\left(zH^B\left(\bar{\mathcal{P}}_B^k\right) + (\mathcal{P}_i - l)\right) dH^k(\mathcal{P}) + \int_{\bar{\mathcal{P}}_B^k}^{\infty}\left(zH^k\left(\bar{\mathcal{P}}_A^k\right) + (\mathcal{P}_i - l)\right) dH^B(\mathcal{P})$$

Taking FOCs yields

$$\begin{aligned}
\frac{d\left(S^k + S^B\right)}{d\bar{\mathcal{P}}_A^k} &= 0 \Rightarrow \\
\left(zH^B\left(\bar{\mathcal{P}}_B^k\right) + \left(\bar{\mathcal{P}}_A^k - l\right)\right) + zh^k\left(\bar{\mathcal{P}}_A^k\right)\int_{\bar{\mathcal{P}}_B^k}^{\infty} dH^B(\mathcal{P}) &= 0 \Rightarrow \\
\left(zH^B\left(\bar{\mathcal{P}}_B^k\right) + \left(\bar{\mathcal{P}}_A^k - l\right)\right) + zh^k\left(\bar{\mathcal{P}}_A^k\right) H^B\left(\bar{\mathcal{P}}_B^k\right) &= 0 \Rightarrow \\
zn_B^k\left(1 + h^k\left(\bar{\mathcal{P}}_A^k\right)\right) &= l - \bar{\mathcal{P}}_A^k
\end{aligned}$$

$$\begin{aligned}
\frac{d\left(S^k + S^B\right)}{d\bar{\mathcal{P}}_B^k} &= 0 \Rightarrow \\
\left(zH^B\left(\bar{\mathcal{P}}_A^k\right) + \left(\bar{\mathcal{P}}_B^k - l\right)\right) + zh^B\left(\bar{\mathcal{P}}_B^k\right)\int_{\bar{\mathcal{P}}_A^k}^{\infty} dH^k(\mathcal{P}) &= 0 \Rightarrow \\
\left(zH^B\left(\bar{\mathcal{P}}_A^k\right) + \left(\bar{\mathcal{P}}_B^k - l\right)\right) + zh^B\left(\bar{\mathcal{P}}_B^k\right) H^k\left(\bar{\mathcal{P}}_A^k\right) &= 0 \Rightarrow \\
zn_A^k\left(1 + h^B\left(\bar{\mathcal{P}}_B^k\right)\right) &= l - \bar{\mathcal{P}}_B^k
\end{aligned}$$

Using the notation conventions above, this yields:

$$\bar{\mathcal{L}}_A^k = zF^B(\bar{\mathcal{L}}_B^k) \times \left(1 + f^B\left(\bar{\mathcal{L}}_A^k\right)\right) \quad (37)$$

$$\bar{\mathcal{L}}_B^k = zF^k(\bar{\mathcal{L}}_A^k) \times \left(1 + f^k\left(\bar{\mathcal{L}}_B^k\right)\right) \quad (38)$$

Comparing (35)-(36) to (37)-(38) we note the presence of two new terms relative to the Nash equilibrium above. These terms reflect the cross-group spillover. In general, both groups gain from additional investment relative to the laissez-faire equilibrium.

A particular transparent case is one in which $\mathcal{P}$ is drawn from uniform (type- and group-specific) distributions. Then (37)-(38) lead to

$$\begin{aligned}
zn_B^k\left(1 + \bar{f}^B\right) &= \bar{\mathcal{L}}_A^k \Rightarrow n_A^{k*} = F^k\left(zn_B^{k*}\left(1 + \bar{f}^B\right)\right), \\
zn_A^k\left(1 + \bar{f}^k\right) &= \bar{\mathcal{L}}_B^k \Rightarrow n_B^{k*} = F^B\left(zn_B^{k*}\left(1 + \bar{f}^k\right)\right).
\end{aligned}$$

Then,

$$\begin{aligned}
S^{k*} &= d + z \times \left(n_A^{k*} \times n_B^{k*}\right) - \int_{-\infty}^{zn_B^{k*}} \mathcal{L} dF^k(\mathcal{L}) > S^k, \\
S^{Bk*} &= d + z \times \left(n_A^{k*} \times n_B^{k*}\right) - \int_{-\infty}^{zn_A^{k*}} \mathcal{L} dF^k(\mathcal{L}) > S^{Bk}.
\end{aligned}$$

We have therefore shown that the trade game underprovides cooperation relative to the first best.

C.2 Analysis of Section 6.1 (stochastic types)

In this section, we provide the details of the analysis in Section 6.1. The *type shock* is realized at the beginning of each period, before group A decides whether to go to war. We continue to denote by r_{t-1} the posterior belief (likelihood ratio) that group A is of the good type after the realization of war/peace in time $t-1$. However, this is now different from the prior belief at t, which drives war and trade decisions, due to the mean reversion induced by (14). We denoted by $\tilde{\pi}$ such a prior, and by $\tilde{r}$ the corresponding likelihood ratio. Bayes rule yields $\tilde{\pi}(\pi_{t-1}) = (1-\psi)\pi_{t-1} + \phi(1-\pi_{t-1})$, hence,

$$\tilde{r}(r_{t-1}) = \frac{(1-\psi)r_{t-1}+\phi}{\psi r_{t-1}+1-\phi}. \tag{39}$$

Thus, the posterior likelihood ratio after war and peace are, respectively,

$$\begin{aligned} \ln r_P(r_{t-1}) &= \ln \tilde{r}(r_{t-1}) + \ln \frac{\lambda_P + (1-\lambda_W-\lambda_P)\,\sigma^+(\tilde{r}(r_{t-1}))}{\lambda_P + (1-\lambda_W-\lambda_P)\,\sigma^-(\tilde{r}(r_{t-1}))}, \\ \ln r_W(r_{t-1}) &= \ln \tilde{r}(r_{t-1}) - \ln \frac{1-\lambda_P - (1-\lambda_W-\lambda_P)\,\sigma^-(\tilde{r}(r_{t-1}))}{1-\lambda_P - (1-\lambda_W-\lambda_P)\,\sigma^+(\tilde{r}(r_{t-1}))}. \end{aligned} \tag{40}$$

The Bayesian updating process is described by the system (39)–(40).

In the region of uninformative PBE (i.e., $\tilde{r}(r_{t-1}) \leq \underline{r}$), $\sigma^-(\tilde{r}(r_{t-1})) = \sigma^+(\tilde{r}(r_{t-1})) = 0$. Thus, $r_P = r_W = \tilde{r}(r_{t-1})$, and the dynamics are governed by the following ordinary difference equation (see ODE in (16)),

$$r_t = \tilde{r}(r_{t-1}) = \frac{(1-\psi)r_{t-1}+\phi}{\psi r_{t-1}+1-\phi}. \tag{41}$$

In the uninformative region, group A always wages war under BAU. The unconditional likelihood ratio that A is civic, $\hat{r} \equiv \phi/\psi$, is the unique rest point of (41): $\hat{r} \equiv \tilde{r}(\hat{r})$, with $\hat{r} > 0$.

In the region of informative PBE (i.e., $\tilde{r}(r_{t-1}) > \underline{r}$), $\sigma^-(\tilde{r}(r_{t-1})) = 0$ and $\sigma^+(\tilde{r}(r_{t-1})) = 1$, and the dynamics are governed by the stochastic difference equation (see StoDE in (16))

$$r_t = \begin{cases} \tilde{r}(r_{t-1}) \times \frac{1-\lambda_W}{\lambda_P} & \text{if} \quad \mathbb{W}_t = 0 \\ \tilde{r}(r_{t-1}) \times \frac{\lambda_W}{1-\lambda_P} & \text{if} \quad \mathbb{W}_t = 1 \end{cases} \tag{42}$$

In this region, group A wages war under BAU when it is of the uncivic type and retains peace under BAU when it is of the civic type. We define $\hat{r}^+ > 0$ and $\hat{r}^- > 0$ to be, respectively, the upper and lower bound of the ergodic set induced by the stochastic equation (42): $\hat{r}^+ = \tilde{r}(\hat{r}^+) \times \frac{1-\lambda_W}{\lambda_P}$ and $\hat{r}^- = \tilde{r}(\hat{r}^-) \times \frac{\lambda_W}{1-\lambda_P}$. Intuitively, $\hat{r}^+$ ($\hat{r}^-$) corresponds to the "quasi-steady state" to which beliefs would converge after an infinite sequence of peace (war) observations. The equations $\hat{r}^+ = \tilde{r}(\hat{r}^+) \times \frac{1-\lambda_W}{\lambda_P}$ and $\hat{r}^- = \tilde{r}(\hat{r}^-) \times \frac{\lambda_W}{1-\lambda_P}$ are polynomials of the second degree. They admit the following roots:

$$\begin{aligned} \hat{r}^+ &= \left(\frac{1-\lambda_W}{\lambda_P}(1-\psi) - (1-\phi)\right)/2\psi \\ &\pm \frac{1}{2\psi}\sqrt{\left(\frac{1-\lambda_W}{\lambda_P}(1-\psi) - (1-\phi)\right)^2 + 4\frac{1-\lambda_W}{\lambda_P}\phi\psi} \\ \hat{r}^- &= \left(\frac{\lambda_W}{1-\lambda_P}(1-\psi) - (1-\phi)\right)/2\psi \\ &\pm \frac{1}{2\psi}\sqrt{\left(\frac{\lambda_W}{1-\lambda_P}(1-\psi) - (1-\phi)\right)^2 + 4\frac{\lambda_W}{1-\lambda_P}\phi\psi} \end{aligned}$$

for $k \in \{+,-\}$. The smaller roots are always negative and can therefore be discarded. As a consequence $\hat{r}^-$ and $\hat{r}^+$ are uniquely defined, and $\hat{r}^- < \hat{r} < \hat{r}^+$.

Note that neither $\hat{r}^+$ nor $\hat{r}^-$ nor $\hat{r}$ depend on V. Since $\underline{r} \equiv \lambda_P \left(S^+\right)^{-1}(V)/(1-\lambda_W)$, it is possible to choose V (or, alternatively, ϕ/ψ) consistent with each of the three cases analyzed in text, $\underline{r} > \hat{r}$, $\underline{r} \in [\hat{r}^-, \hat{r}]$ and $\underline{r} < \hat{r}^-$.

C.3 Analysis of Section 6.2 (learning from trade)

We start by a general characterization of the PBE in the environment of Section 6.2.

Proposition 9 *For any $(r_P, \iota) \in [0,+\infty) \times [0,1]$, the Perfect Bayesian Equilibrium of the trade game exists and is unique. It is characterized by the 4-tuple $\{n_A^-, n_A^+, n_B^-, n_B^+\} \in [0,1]^4$ such that, for $k \in \{+,-\}$,*

$$n_A^k = F^k(zn_B^k) \text{ and } n_B^k = G^k\left(zn_B^+, zn_B^-\right), \tag{43}$$

where $G^k\left(zn_B^+, zn_B^-\right) \equiv \iota F^B(zF^k(zn_B^k)) + (1-\iota)F^B\left[\frac{r_P}{1+r_P}zF^+(zn_B^+) + \frac{1}{1+r_P}zF^-(zn_B^-)\right]$. The equilibrium trade surplus accruing to group A, $S^k(r_P,\iota)$, is given by

$$S^k(r_P,\iota) = d + \int_{-\infty}^{zn_B(r_P,\iota)} F^k(\mathcal{L})\ d\mathcal{L}.$$

Proof. First, we derive (43). Suppose $k = +$. Then, all informed agents in B such that $\mathcal{P}^i \geq zn_A^+$ will cooperate. Likewise, all agents in A such that $\mathcal{P}^i \geq zn_B^+$ will cooperate. However, some agents in B are uninformed, and agents in A know it. An uninformed player in B will cooperate as long as $\mathcal{P}^i \geq \pi_P \times zF^+(zn_A^+) + (1-\pi_P) \times zF^-(zn_A^-)$. As this inequality shows, in order to determine the behavior of the uninformed players, we must solve for the counterfactual distribution n_A^-, which in turn requires that we solve for n_B^-. More formally, if $k = -$, all informed agents in B such that $\mathcal{P}^i \geq zn_A^-$ would cooperate, and all agents in A such that $\mathcal{P}^i \geq zn_B^-$ will cooperate. And so on.

Thus, the complete system yields

$$\begin{aligned}
n_A^+ &= F^+(zn_B^+), \\
n_A^- &= F^-(zn_B^-), \\
n_B^+ &= \iota F^B(zF^k(zn_B^+)) + (1-\iota)F^B\left[\frac{r_P}{1+r_P}zn_A^+ + \frac{1}{1+r_P}zn_A^-\right], \\
n_B^- &= \iota F^B(zF^k(zn_B^-)) + (1-\iota)F^B\left[\frac{r_P}{1+r_P}zn_A^+ + \frac{1}{1+r_P}zn_A^-\right],
\end{aligned}$$

which is equivalent, after substituting the expressions of n_A^+ and n_A^- into the third and fourth equality, to (43).

Given $(r_P,\iota) \in [0,+\infty) \times [0,1]$, the system of equations (43) defines a continuous mapping $\mathbf{G}:[0,1]^4 \to [0,1]^4$ such that $\left(n_A^-, n_A^+, n_B^-, n_B^+\right) = \mathbf{G}\left(n_A^-, n_A^+, n_B^-, n_B^+\right)$. Brouwer's fixed point theorem implies that $\mathbf{G}$ has at least one fixed point.

Let $\hat{\mathbf{G}} \equiv \{G^+, G^-\} : [0,1]^2 \to [0,1]^2$ denote the third and fourth equation of (43), $\left(n_B^-, n_B^+\right) = G\left(n_B^-, n_B^+\right)$. Note that this sub-fixed-point problem can be solved without reference to n_A^k. Brouwer's fixed point theorem implies that $\hat{\mathbf{G}}$ has also at least one fixed point. Moreover, the fixed point is unique, since $\hat{\mathbf{G}}\left(n_B^-, n_B^+\right)$ is a continuous, (weakly) monotonically increasing, convex mapping. This follows, in turn, from F^B, F^+ and F^- being continuous non-decreasing convex functions. From the uniqueness of $\left(n_B^-(r_P,\iota), n_B^+(r_P,\iota)\right)$ it follows immediately that $\left(n_A^-(r_P,\iota), n_A^+(r_P,\iota)\right)$ is also unique, establishing that the fixed point $\left(n_A^-, n_A^+, n_B^-, n_B^+\right) = \mathbf{G}\left(n_A^-, n_A^+ n_B^-, n_B^+\right)$ is unique.

The derivation of the expression for the trade surplus is as in the proof of Proposition 2. **QED** ■

In the rest of this section, we specialize the analysis to uniform distributions of psychological costs and benefits of cooperation, as discussed in the text.

Assumption 4 *$F^+ \sim [-x, 1]$, $F^- \sim [0, 1+x]$, $F^B \sim [0, 1]$ with $x \geq 0$.*

Remark 2 *Assumption 4 is consistent with Assumption 1 if and only if $z \leq 1$.*

In the rest of the section, we also focus on the particular case in which $z = 1$, and normalize the payoff matrix so that $d = 0$. These assumptions entail no loss of generality and are only aimed at obtaining simple algebraic expressions. The generalization to $z \leq 1$ and $d \neq 0$ is straightforward, if more cumbersome. Note that, under perfect information, $z = 1$ implies that the Nash equilibrium features $\left(n_A^-, n_B^-\right) = (0,0)$ and $\left(n_A^+, n_B^+\right) = (1,1)$ [note that this is no corner solution, i.e., for any $z < 1$ the solution is strictly in the interior of $[0,1]^2$].

Under these distributional and parametric restrictions, we can provide a closed form solution of the PBE in Proposition 9.

Corollary 2 *Under Assumption 4, and the (algebra-simplifying) assumptions that $z = 1$ and $d = 0$, the PBE has the following characterization*

$$
\begin{aligned}
n_B^-(r_P, \iota) &= \frac{r_P}{1+r_P}\left(1 - \frac{x\iota}{1-\iota+x}\right), \\
n_B^+(r_P, \iota) &= \frac{r_P}{1+r_P}\left(1 + \frac{x\iota}{1-\iota+x}\right) + \frac{x}{\frac{1+x}{\iota} - 1}, \\
n_A^-(r_P, \iota) &= \frac{n_B^-(r_P, \iota)}{1+x}, \\
n_A^+(r_P, \iota) &= \frac{n_B^+(r_P, \iota) + x}{1+x}, \\
S^+(r_P, \iota) &= \frac{[n_B^+(r_P, \iota) + x]^2}{2(1+x)}, \\
S^-(r_P, \iota) &= \frac{[n_B^-(r_P, \iota)]^2}{2(1+x)}.
\end{aligned}
$$

Note that (i) all expressions are increasing in r_P; (ii) $n_B^+(r_P, \iota)$ and $n_A^+(r_P, \iota)$ (respectively, $n_B^-(r_P, \iota)$ and $n_A^-(r_P, \iota)$) are increasing (respectively, decreasing) in ι; (iii) $S^+(r_P, \iota)$ is increasing in ι and $S^-(r_P, \iota)$ is decreasing in ι.

Proof. The Corollary follows from Proposition 9, after standard algebra. **QED** ■

Consider next the dynamics of ι. We establish an upper bound to the proportion of informed agents.

Lemma 3 *Let $\iota_\infty(\theta) \equiv \left(1 + \frac{\theta}{\tau(1-\theta)}\right)^{-1}$. Assume $\iota_0 < \iota_\infty(\theta)$. Then, for any $t \in [0, \infty)$ and any realization of the war/peace process, $\iota_t < \iota_\infty(\theta) = \left(1 + \frac{\theta}{\tau(1-\theta)}\right)^{-1}$.*

Proof. The lemma follows from (18), after setting $\mathbb{W}_t = 0$ for all t. **QED** ■

The upper bound $\iota_\infty(\theta)$ corresponds to the proportion of informed agents accumulated after an infinite sequence of peace shocks. Note that $\iota_\infty(\theta)$ is decreasing in θ and increasing in τ. Moreover, $\iota_\infty(0) = 1$ and $\iota_\infty(1) = 0$. The model of this section nests the benchmark model in the particular case in which $\theta = 1$ (or $\tau = 0$).

Next, we turn to the definition of war traps.

Definition 6 *A war trap is a set of states,* $\Omega_{TRAP} \subset \mathbb{R}^+ \times [0,1]$, *such that if* $(r_t, \iota_t) \in \Omega_{TRAP}$ *then* $\forall s \geq t, r_s = r_t$ *for all continuation paths* $[r_s, \iota_s]_{s=t}^{\infty}$.

Note that we do not require the stationarity of ι_t for an economy to be in a war trap. The test for the existence of a trap is that, for a non-empty set of beliefs, both types follow the same strategy (i.e., either wage war or retain peace) under BAU, when the number of informed agents is at its upper bound. Moreover, this must remain true for any subsequent sequence of war and peace shocks.

We continue to focus on the region of the parameter space such that, absent learning from trade (e.g., when $\theta = 1$), $V > S^+(0)$ and $S^-(\infty) < V < S^+(\infty)$, which are the conditions of Propositions 4 and 5. Combined with the Corollary 2 this translates into

$$V > \frac{x^2}{2(1+x)} \text{ and } x > 1 \tag{44}$$

The following Lemma and Proposition establish that (i) θ must be sufficiently large for a war trap to be sustained – i.e., a sufficiently large friction in the information transmission is crucial for the war trap to be robust; (ii) the size for the war trap depends on θ.

Lemma 4 *Suppose condition (44) holds. Then, under the conditions of Corollary 2,* $\Omega_{TRAP} \neq \varnothing$ *if and only if* $1 > \theta > \theta_W \equiv [[1 + \frac{1+x}{\tau x}((\sqrt{2(1+x)V} - x)^{-1} - 1]^{-1}]^{-1}$.

Proof. Using the expressions in Corollary 2, we obtain that $S^+(0, \iota_\infty(\theta)) = \frac{1}{2}\frac{(1+x)x^2}{(1+x-\iota_\infty(\theta))^2}$, where S^+ is increasing in ι_∞, and ι_∞ is decreasing in θ. In particular, $S^+(0, \iota_\infty(\theta_W)) = V$. (i) Suppose $\theta < \theta_W$. Then, $S^+(0, \iota_\infty(\theta)) > S^+(0, \iota_\infty(\theta_W)) = V$. Hence, $\theta < \theta_W$ implies that $\Omega_{TRAP} = \varnothing$. (ii) Suppose that $\theta > \theta_W$. Then, the same argument implies that $S^+(0, \iota_\infty(\theta)) < S^+(0, \iota_\infty(\theta_W)) = V$. But, then, for any such θ, there exists $\hat{r} > 0$, such that $S^+(\hat{r}, \iota_\infty(\theta)) < V$. Hence, $\theta > \theta_W$ implies that $\Omega_{TRAP} \neq \varnothing$. Finally, the case of $\theta = \theta_W$ is degenerate, as in this case a trap exists only as long as $r = 0$. ■

Proposition 10 *Suppose condition (44) holds and* $\theta \geq \theta_W$. *Then, under the conditions of Corollary 2, an economy is in a war trap if and only* $r < \underline{r}(\theta) \equiv \lambda_P \underline{r}^*(\theta)/(1 - \lambda_W)$, *where*

$$\underline{r}^*(\theta) \equiv \frac{\sqrt{2(1+x)V} - x - \frac{1}{1+(1+x)\theta/\tau x(1-\theta)}}{1 - \sqrt{2(1+x)V} + x}, \tag{45}$$

and $\underline{r}(\theta)$ *is an increasing function of* θ.

Proof. For a given θ, the upper bound of Ω_{TRAP} is denoted $r_P = \underline{r}^*(\theta)$, characterized by

$$S^+(\underline{r}^*(\theta), \iota_\infty(\theta)) = V.$$

Using the expression of $S^+(r, \iota)$ given by Corollary 2 and the expression of $\iota_\infty(\theta)$ given by Lemma 3, and simplifying terms, yields (45). The assumption that $1 > \theta \geq \theta_W$ ensures that $\underline{r}^*(\theta) > 0$. Standard algebra establishes that $\underline{r}^*(\theta)$, and, hence, $\underline{r}(\theta)$, is an increasing function of θ. **QED** ■

C.4 Analysis of Section 6.3 (three groups)

In this section, we extend the analysis to an environment in which the economy is inhabited by three groups: A, B and C. We focus on the trade links between A and B, and between A and C, and on how the presence of a third group (C) affects the probability that A attacks B.

Let $\mathbb{E}[y_i^{kB}(r_P)] \geq d$ denote the expected payoff to agent i in group A who is randomly matched with an agent belonging to group B, and who plays the strategy (either cooperate or defect) that maximizes the expected payoff in the trade game described by the payoff matrix 1. Note that, irrespective of beliefs $\mathbb{E}[y_i^{kB}(r_P)] \geq d$, since an agent can always choose to defect and earn the safe payoff d. In the benchmark model, the trade surplus accruing to A under peace can be expressed as $S^k(r_P) = \int_i \mathbb{E}[y_i^{kB}(r_P)]$.

Each agent in A can choose how much time to spend trading with a partner in B and how much trading with a partner in C. Let τ be the fraction of time endowment trader i in group A spends trading with the randomly matched partner in group B. War implies that $\tau_i = 0$ for all i's.

Let us focus on two alternative polar opposite cases (substitution and complementarity):

Case 1 (substitution) $y_i = \tau_i \times y_i^B + (1 - \tau_i) \times y^C$

Case 2 (complementarity) $y_i = \min\{\tau_i, 1 - \tau_i\} \times y_i^B + \min\{1 - \tau_i, \tau_i\} \times y^C$

In the case of substitution, trade with B and C are completely independent activities. In this case y^C simply denotes the productivity of (full-time) trade with any agent of group C. In the case of complementarity, there are spillovers across trade activities. The pay-off from trading with B requires time spent with C, and the pay-off from trading with C requires time spent with B. In this case, the maximum trade off from trade with group C is $y^C/2$.

Assumption 5 $y^C < d$.

C.4.1 The case of substitution

Assumption 5 implies that, in this case, trading with B is strictly preferred to trading with C. Thus, all agents in A choose $\tau = 1$ in peacetime. So $S^k(r_P) = \int_i \mathbb{E}[y_i^{kB}(r_P)]$. In wartime, all agents must trade full time with C, i.e., $\tau = 1$. Thus, $S^{WAR} = y^C$, where S^{WAR} denotes the trade surplus during wartime. The difference between $S^k(r_P)$ and S^{WAR} is the opportunity cost of war, and determines the size of the trap. War is chosen under BAU whenever

$$V + y^C > S^k(r_P).$$

Proposition 11 *Given $k \in \{+, -\}$, in the case of substitution, the range of beliefs such that group A chooses war under BAU is an increasing function of the productivity of the trade link, y^C.*

Note that $y^C = 0$ is equivalent to the model with only two groups, which yields the lowest probability of war.

C.4.2 The case of complementarity

In this, under peace, each agent chooses $\tau = 1/2$. This yields $S^k(r_P) = \frac{1}{2}\left(\int_i \mathbb{E}[y_i^{kB}(r_P)] + y^C\right)$. Under war, $\tau = 1$ and thus $S^{WAR} = 0$. Note that group B provides services that have both some intrinsic value, and increase the productivity of trade with C. Destroying trade with B also destroys the surplus from trading with C. War is chosen under BAU whenever

$$V > \frac{1}{2}\left(\int_i \mathbb{E}[y_i^{kB}(r_P)] + y^C\right).$$

In this case, war becomes less likely as y^C increases.

Proposition 12 *Given $k \in \{+, -\}$, in the case of complementarity, increasing y^C reduces the size of the war trap.*

ADDITIONAL REFERENCES

Hassler, John, José Rodríguez Mora, Kjetil Storesletten, and Fabrizio Zilibotti, 2003, "The Survival of the Welfare State", *American Economic Review* 93: 87-112.

Heston, Alan, Robert Summers, and Bettina Aten, 2011, "Penn World Table Version 7.0", dataset, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, http://pwt.econ.upenn.edu/.

Polity IV, 2012, "Political Regime Characteristics and Transitions", dataset, www.systemicpeace.org/polity4.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, 2012, "Rotten parents and Disciplined Children", *Econometrica* 80: 2785–2803.

World Bank, 2012, "World Development Indicators", Dataset, http://data.worldbank.org/indicator.