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

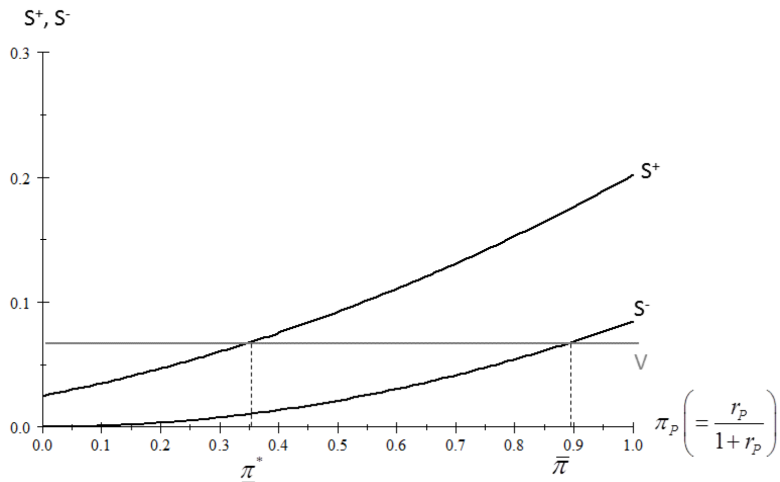

Figure 5: Surplus from trade and war benefits; the case of two traps.

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

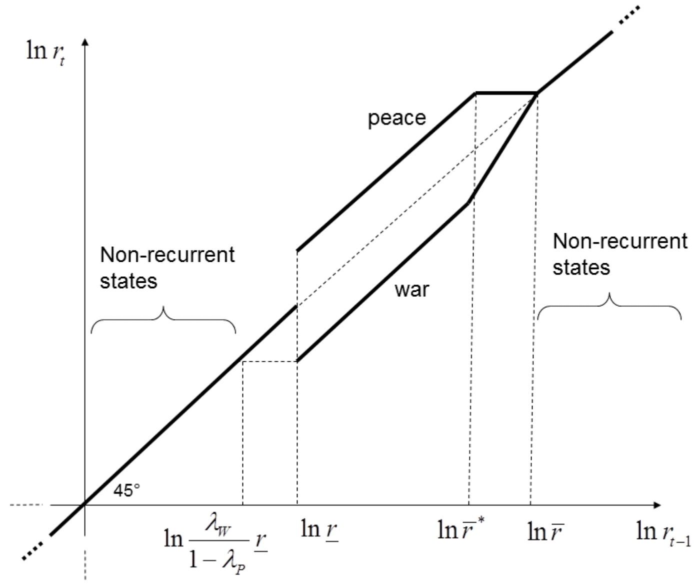

Figure 6: Law of motion of beliefs; the case with two traps.

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:

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

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

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

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

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

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:

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

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

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*

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

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

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:

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

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

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

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 $V - S^-(r_P(r)) > \beta\left(W^-(r_P(r)) - W^-(r_W(r))\right)$, 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:

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

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

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., $\tilde{r}(r_{t-1}) \le \underline{r}$), $\sigma^-(\tilde{r}(r_{t-1})) = \sigma^+(\tilde{r}(r_{t-1})) = 0$. 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*

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

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