Market anticipations of conflict onsets

Thomas Chadefaux

Trinity College Dublin
Department of Political Science

Abstract

Does the recurrence of wars suggest that we fail to recognize dangerous situations for what they are, and are doomed to repeat the errors of the past? Or rather that policy-makers correctly anticipate the consequences of their actions but knowingly choose conflict? Unfortunately, little is known about how well wars are anticipated. Do conflicts tend to come as a surprise? We estimated the risk of war as perceived by contemporaries of all inter- and intra-state conflicts between 1816 and 2007. Using historical financial data of government bond yields, we find that market participants tend to underestimate the risk of war prior to its onset, and to react with surprise immediately thereafter. This result illustrates how conflict forecasts can be self-fulfilling or self-defeating. Present predictions may affect future behavior, such that wars may be less likely to occur when they are predicted, but more likely when they are not. We also show that the forecasting record has not improved over the past 200 years, and that wars involving democracies lead to greater market shocks. These findings also have implications for the way decision-makers respond to new information, and how audiences perceive the risk of war and hence their leaders' actions.
Keywords. forecasting; war; conflict; financial markets; government bonds
Corresponding author: thomas.chadefaux@tcd.ie

The recurrence of wars despite their tremendous economic, social and institutional costs, may suggest that we are doomed to repeat the errors of the past. Time after time, policy-makers seem to mispredict the consequences of their actions and fail to recognize dangerous situations for what they are. Can the risks of war be correctly estimated, or do we really only learn from history that we do not learn from it?

Unfortunately, little is known about how well wars are anticipated. Do conflicts indeed tend to come as a surprise to their contemporaries? Or are they correctly anticipated, but decision-makers choose to engage in them anyway? Using financial data, we examine the reaction of market participants to the onset of all civil and interstate conflicts from 1816 to 2007. If wars are correctly predicted, then those who have a stake in them should not be surprised by their onset. Yet we find the opposite: investors have historically underestimated the probability of war prior to its outbreak and the onset typically led to a large correction. Market participants, in particular, could often have obtained better returns had they correctly estimated the risk of war.

Whether observers correctly estimate the risk of war matters for several reasons. First, understanding how past observers have fared is a first step in identifying possible ways to improve future forecasts. Second, are there types of war or attributes of the warring countries that increase the predictability of conflict? And are forecasts improving? Third, the findings are relevant to the large literature on the public's reaction to their leaders' foreign policy choices. One important assumption in that literature is that the leaders' choices are clearly and unambiguously understood by those who decide their fate. Audiences may, for example, punish leaders for reckless actions that escalate the risk of war. Yet if observers misestimate the risks of war, then there are important implications for our understanding of audience costs and costly signals, for example. Can leaders really tie their hands or more generally signal their intentions if the associated risk of war is misestimated?

Finally, the (in)ability of contemporaries to predict might in fact not be an indictment of their predictive ability, but actually be a sign of the reactivity of policy. Indeed, if policy-makers adjust their policies by incorporating predictions and reacting to them—perhaps by trying to avoid the war or instead by precipitating it—then wars would not happen when they are expected, and hence would appear to be difficult to predict. Far from implying that we do not learn from history, then, it may in fact suggest that wars are difficult to anticipate precisely because decision-makers incorporate current predictions into their assessment and react accordingly.

Our results also relate to the evaluation of applied research on forecasting. To assess the quality of our predictions, we must acknowledge their possible effect on policy. This endogeneity means that the difficulty lies not only in forecasting war, but also in evaluating our performance doing so. Forecasts that are based on static variables are unlikely to perform well, and our results therefore call for more dynamic estimation of risk.

The article proceeds in three steps. We first discuss the relevant literature and present hypotheses relating regime type, war type, and predictability of the onset of war. We then review the data used to test this conjecture, including data on government bond yields and control variables. Finally, we show three main results: contemporaries tend to underestimate the risks of war; our ability to estimate this risk has not improved over the past 200 years; and conflicts involving democracies lead to greater shifts in market prices than others.

Markets' estimation of geopolitical risk

Conflict forecasting has received increasing attention in political science (Beck, King & Zeng, 2000, 2004; De Marchi, Gelpi & Grynaviski, 2004; Gleditsch & Ward, 2013; Hegre et al., 2013). The availability of increasingly fine-grained spatio-temporal data in particular has allowed more refined predictions (Brandt, Freeman & Schrodt, 2011). Data ranges from stock market prices (Schneider, Hadar & Bosler, 2017), to news reports (Chadefaux, 2014), urban violence (Bhavnani et al., 2014), or climate data (Witmer et al., 2017).

However, we know little about how well wars are predicted by their contemporaries. The existing literature focuses instead on the detection of early warning signals. Yet showing that, say, market fluctuations can help improve forecasts (Schneider, Hadar & Bosler, 2017; Chadefaux, 2015a) does not mean that the market's forecasts were accurate. For example, a small but systematic change in the price of an asset before the onset of war may be sufficient to improve the researcher's forecasts, but the large shift in price following war would still suggest that the market had misestimated the probability of war.

Guidolin & La Ferrara (2010) do study the reaction of markets to conflict onsets, but are concerned with what can be inferred about their economic costs rather than the adequacy of the market's forecasts itself. Their goal is to estimate the cost of the conflict by using market reactions as a metric, and not to study the quality of the market's forecast itself. More importantly, the data they use is not country-specific, with the exception of the USA, the UK, France, and Japan. As a result, these data are not fine-grained enough to infer the market's reaction to a particular war, except for those involving these four countries. Finally, the limited time-span of their data (1974–2004) precludes the analysis conducted here on the evolution and possible improvement over time of forecasts. Other studies that focus on the reaction of financial markets to the onset of conflict are limited to case studies (Rigobon & Sack, 2005; Leigh, Wolfers & Zitzewitz, 2003; Amihud & Wohl, 2004; Hall, 2004; Chen & Siems, 2004; Schneider & Troeger, 2006; Schneider, Hadar & Bosler, 2017; Brune et al., 2015). Tetlock (2005) is more specifically focused on the quality of forecasts, but has no data on conflict and also a more limited time-frame.

What we need is an estimate of contemporaries' beliefs around the time of the onset. Several measures are possible. Reading newspapers, for example, might give us a sense of the perceived probability of war (Ramey, 2011; Chadefaux, 2014). News, however, suffer from a major drawback. They are likely to respond to novelty more than to reflect true underlying concerns. Thus, the number of articles about the war after its onset is likely to increase sharply, but that need not indicate surprise—simply that its onset has put it to the fore. That interest may wane once the novelty wears out (Fig. A4).

What we need instead is the perception of those who have an incentive to reveal their true perception of the risks of war. Financial markets are particularly well suited for that purpose, because they aggregate the opinion of a large number of participants who have a stake in correctly estimating risk. Through prices, then, market participants reveal their true beliefs about geopolitical risk.

Government bond yields, in particular, are an ideal source of information about the market's perception of a country's probability of war. Government bonds (or 'sovereign' bonds) are the standard way by which national governments borrow from the market. They are typically issued in exchange for regular interest payments and the promise to repay the principal once the bond reaches its maturity. The price of the bond (and hence its yield) depends on the perceived sovereign risk. A high yield will be demanded when the perceived risk is high, whereas 'safe' countries will be able to borrow at low interest rates. If the yield is too low in relation to the perceived risk, investors will prefer other financial assets such as equities, commodities (e.g. gold), or even cash.

A simple model of bond pricing will be useful to understand the effect of the onset of war on bond yields. The price at time t of a government bond with periodic interest payment C (coupons), N payments (e.g. 40 payments for a 10-year bond with quarterly payments), market interest rate r_t (typically the central bank's rate), and value at maturity M (typically 100) can be evaluated as the time-discounted sum of coupon payments plus the discounted value of the repayment at maturity:

$$\begin{aligned} P_t &= \frac{C}{(1+r_t)} + \frac{C}{(1+r_t)^2} + \cdots + \frac{C}{(1+r_t)^N} + \frac{M}{(1+r_t)^N} \\ &= \left[\frac{1-\frac{1}{(1+r_t)^N}}{r_t}\right] C + \frac{M}{(1+r_t)^N}, \end{aligned} \tag{1}$$

The current yield is then simply the nominal value of the coupon C as a percentage of the current bond price P_t, i.e., $\text{Yield}_t = C/P_t$.[1]

Wars, in turn, generate three main kinds of sovereign risks for investors. First, the government may fail to pay its debt back, for two main reasons: (a) it may incur so much debt to finance the war that it is unable to repay the principal fully once the bond matures; and (b) the economy may contract so much as a result of the war that the government's fiscal receipts will plummet and the burden of repaying the debt will become too high. In the notation above, this implies that the expected value of M—the value at maturity—decreases or vanishes entirely, thereby driving down the prices of bonds (and hence increasing their yields, as bond prices and yields move in opposite direction). A second type of sovereign risk is that periodic interest payments might be reduced or cut entirely (i.e., the number or value of C above may decrease). Finally, even if the government honors the terms of repayment without any 'haircut', a third risk is the inflation in the currency of the bond that is likely to be associated

[1]Since bonds are priced at all times, and not just once every quarter, we have two indicators of time. Time t is a continuous variable corresponding to calendar time (e.g. 18 May 2017), whereas $n \in N$ refers to discrete (e.g. quarterly) payment events (e.g. Q2 of 2017).

with a costly conflict. In (1) above, the market interest rate i, mostly determined by the central bank and inflation, may increase. This inflation reduces the investor's *real* return, and hence a higher *nominal* yield will be demanded today to compensate for this risk. Note that this last scenario implies that central bank rates may mediate the effect of the onset of war on bond yields—a possibility we explore using proper controls and by estimating mediation models (see appendix A.3).

Together, these risks imply that a bondholder aware of a forthcoming war should demand a higher yield today. Investors calculate the expected (and discounted) return from a given bond, and all information available is immediately incorporated into the price (Fama, 1991). They trade to reconcile residual differences in their beliefs, and shocks in the yield of bonds therefore signal the emergence of new information that was not expected by market participants. Shocks (or 'jumps') in bond prices—changes in prices over a short period of time—therefore mean that new information is at odds with the market's prior belief. Just as well-anticipated central bank announcements have no effect on asset prices (Poole, Rasche & Thornton, 2002), wars should also not cause any unusual variation if correctly anticipated. A shock then implies either a surprise at the onset of war, or at least that markets believed until the end that war was avoidable. Either way, it means that they misestimated the risk of war. An illustrative example can be found in appendix A.1.

Regardless of what drives the jump, a simple way to consider the problem is to think about a 'no-regret' clause. Savvy investors who had anticipated the war (or the central bank rate hike associated with it) should have no regret over their investment decision once the war has started. A jump in prices, however, necessarily implies regrets, as many will wish they had sold before the price drop (remember that yields move in opposite direction from prices) and stored their wealth in cash. Even if inflation is a concern, and hence cash is a risky choice, gold, commodities or other assets classes would have been preferable options.

Hypotheses

Our first hypothesis relates to observers' estimates of the risks of war. We conjecture here that wars will be poorly predicted on average, and that investors will tend to underestimate their probability. That is because the estimated probability of war at time t may affect the actions taken by leaders, and hence change the actual probability of war at time $t+1$. Leaders assess the future and base their choices on what they have learned from history and their rational expectations given available information. Those who recognize the risks might adjust their behavior and strategy. For example, aggressive states may tone down their rhetoric, demands may be softened, troops withdrawn from the border, or rising states may make concessions to alleviate the fears their growth generates (Chadefaux, 2011). Alternatively, forecasts of a distant war may prompt countries to attack now, perhaps before a power shift, so that the initial predictions are again invalidated. On the contrary, states who underestimate the risk of war may behave more recklessly or demand larger concessions in negotiations.

This endogeneity makes it particularly difficult to predict wars with any certainty, as predictions are based on available information, but that information also affects behavior and hence is likely to invalidate the initial prediction—a point related to Lucas' critique of macroeconomic forecasting based on parameters that are not policy-invariant (Lucas, 1976). We conjecture in particular that predictions of a coming war increase the probability that decision-makers alter their plans, and hence reduce the probability that war will actually happen then. As a result, wars will be less likely to occur when they are predicted, but more likely when they are not. As a result, we expect the probability of war to be underestimated on average.[2]

Hypothesis 1 (Shock) *The onset of a war involving country i leads on average to a sudden increase in the yield of its government bond.*

Second, market participants buy or sell assets such as bonds based on the expected revenue stream and price. Expectations of a costly war should therefore lead to a larger impact on the price of the asset. We therefore anticipate that the expected cost of war will negatively affect the yield.

Hypothesis 2 (Costly wars) *The onset of a war with large initial fatalities leads, on average, to a larger increase in yields than for wars with low initial fatalities.*

Third, if indeed political decision-makers constantly incorporate newly available information into their policy choices, then counterintuitively we expect to observe that wars will be very difficult to forecast. Indeed, state leaders informed of a looming war are likely to take steps that will affect its onset. They may strive to prevent it altogether, delay its onset, or on the contrary rush its preparation. These steps will affect the path leading to the onset, and hence possibly invalidate the initial prediction. Because of this feed-forward effect, the wars that are left are those that may not have been predicted, perhaps because they are particularly hard to forecast. In other words, because information and forecasts affect behavior itself, wars may always be 'in the error term', and no matter how much our prediction ability improves, the wars that do occur would always come as a surprise. If they had not, they might have been prevented, postponed, or rushed. If that is the case, i.e., if wars are indeed in the error term because of this feed-forward mechanism, then the lessons from history may help prevent wars, but they will not avoid our surprise at those wars that do occur—the wars that we failed to forecast. An implication of this argument is that the wars that we do observe should be as surprising today as they were at the

[2] Note that this issue of endogeneity is related to, but distinct from Gartzke's (1999) idea that incomplete information also creates a limit to our ability to forecast. Gartzke argues that, because war is caused by incomplete information (Fearon, 1995), its onset itself must logically be uncertain. Rational actors update their beliefs using public information and adjust their bargaining strategy accordingly. Additional information in favour of one party will simply lead her to demand more, again pushing the negotiation to the point where both parties are indifferent between war and peace—i.e., where the onset of war is 'in the error term.' Whereas Gartzke's work applies to crisis bargaining, the focus here is on how leaders incorporate the (potentially distant) probability of war into their decisions. In other words, forecasts affect behavior, which in turn affects forecasts. Gartzke's argument, on the other hand, is about offers and counteroffers in the context of crisis bargaining, and the idea that concessions by one country will lead the other to push for further concessions, up to the point where war is again 'in the error term.'

beginning of our sample in 1816, and no significant pattern should emerge over time.

Hypothesis 3 (Constant predictability) *The average magnitude of the shock associated with the onset of war is constant over time.*

Our next hypothesis relates to regime type. The effect of regime type on forecast is difficult to assess a priori. On the one hand, democracies tend to be more transparent, and hence their policies and decisions are more easily and reliably observable, both to other states and to domestic audiences. This would intuitively lead to easier predictability of their actions. Yet transparency implies that their policies are also more likely to be challenged domestically or to receive unwanted attention from the media. This has two effects. First, policy will tend to nimbler, and hence more reactive to updates in the perceived probability of war. Just as liquid financial markets are less predictable than illiquid ones, decision-makers who incorporate new information or parameters rapidly push the current policy to the point where it is no longer easily predictable. The public's reaction to the expectation of war, for example, may lead to adjustment in the government's policies, making prediction more difficult (e.g. the Fashoda crisis, see Schultz 2001, pp. 175–96). A second effect is that this attention and the potential challenges from the opposition and the media may lead democratic leaders to be particularly discreet about their plans, so that their opponents may not discover them, and wars are therefore more likely to come as a surprise (Chadefaux, 2015b). As a corollary, autocracies and their leaders may be more predictable, and their preparation for war more obvious. Counterintuitively, then, the transparency that characterizes democracies may lead to a lower predictability of their foreign policy choices.

Hypothesis 4 (Regime type) *The onset of conflict in democracies is associated with a larger average shock than in autocracies.*

We also expect civil wars to be more predictable than interstate wars. First, actors in civil wars are less clearly defined than in interstate wars. For rebels to even identify themselves may be risky, and their forces may need to build over a significant period of time before they reach a size sufficient to challenge the central government. These buildups will therefore be more visible and predictable than the sudden mobilization that characterizes interstate wars. In addition, low-level skirmishes, which do not reach the level of conflict per se, may be more frequent than in interstate wars, thereby signaling the rising level of tensions to market observers. Bargaining tends to be stickier. Moreover, civil wars often rely on deep animosities and built-up tensions. These may be harder to reverse than in the case of interstate wars, where a clear chain of command will help prevent escalation and accidents. Civil wars, then, are expected to be easier to predict because their dynamic may be harder to reverse and their buildup slower and more visible.

Hypothesis 5 (Conflict type) *The onset of interstate conflicts is associated with a larger average shock than the onset of intrastate conflicts.*

Do wars really come as a surprise?

Demonstrating surprise is difficult. In the absence of an explicit estimate from market participants of their beliefs about the probability of conflict, we must infer them from observed valuations and fluctuations. We adopt a threefold strategy. First, we estimate a regression of changes in bond yields on the onset of conflict using the entire sample. Second, we examine the evolution of yields in the time surrounding the onset of war. Finally, we address the possibility that the corrections we observe simply reflect the stochasticity of the war process, and would hence not be indicative of any market underestimation of the risk of war. Just like weather forecasters may be correct in predicting a 90% chance of rain when in fact it ended up not raining, markets may correctly estimate the pre-onset probability of war, and react with an upward correction once the event is certain. If that were the case, then markets' forecast would be correct *on average*. Yet we find that their forecasts are systematically lower than the actual probability of war, and hence conclude that markets do, in fact, underestimate the risk of war.

Effect of the onset on yields

Data on government bond yields from 1816 to 2007 were collected from *Global Financial Data*, a leading provider of financial data. The country-level time series are either weekly or monthly depending on the country and period. The 10-year bond was used to the extent it existed, and instruments with shorter maturities were used otherwise.[3] Because many countries never issued them, or only recently started to, data are limited to 45 countries and an average of 3,788 observations by country (see Table A1 for details). Even though this sample may be biased towards countries with well-established financial systems (typically advanced democracies), it still exhibits significant variation in terms of GDP, polity and historical background.[4]

We first estimate the effect of the onset of war on yields using the full sample of bond yields from 1816 to 2007. This allows us to compare the effect of onsets to other types of events. The dependent variable ΔYield_{it} is the change in country i's yield in week t (i.e., $\Delta\text{Yield}_{it} = \ln\left[\text{Yield}_{it}/\text{Yield}_{it-1}\right]$) and the main independent variable is the onset of conflict (WAR ONSET_{it}), which is coded as 1 for the week of the onset and the following one, 0 otherwise).[5] Data on conflict onsets was obtained from

[3] This should have little impact on the results, since we are comparing bonds with their previous value. The only drawback is that bonds with shorter maturities may respond differently to distant events, but the direction of the change should not be affected. The inclusion of a dummy variable for bonds with shorter maturities had no significant effect on our results.

[4] One concern could be that a state with a more developed financial market may face higher costs than a country with a less mature financial infrastructure. Indeed, countries with less developed financial markets may be able to finance the war through other means, such as coercion, natural resources, budget surpluses. As a result, the reaction of bond markets may be reduced in those countries. However, this concern is addressed (a) by the variation in our data, which encompasses countries with various levels of development and (b) by controling for the advancement of the financial markets using proxies such as GDP per capita and inflation rates.

[5] Similar results apply if we code that variable 1 only for the very week of the onset, but we think

the Correlates of War (CoW) and the Militarized Interstate Disputes (MIDs) data (Sarkees & Wayman, 2010). This includes all 2,516 inter- and intra-state country-conflicts with a starting date of 1816 to 2007 for which bond yield data is available (see Fig. A7).

Control variables include: the lag of the dependent variable ($\Delta \text{Yield}_{it-1}$); the average change in yield in the world ($\Delta \text{Yield}_{world,t} = \sum_{j \neq i} \Delta \text{Yield}_{jt}/n$, where j denotes countries other than the country of interest) to control for possible shocks that may affect all n countries, independently of the onset of war; the change in central bank rate (weekly) and inflation (quarterly), $\Delta \text{CBRATE}_{it}$ and ΔCPI_{it} (defined in the same way as ΔYield_{it}) from Reinhart & Rogoff (2009); GDP per capita (GDPPC) and Government debt levels (Govt Debt) were also gathered from the same source; and polity scores were obtained from the Polity IV data (Marshall, Jaggers & Gurr, 2002).

Bond yields of a country's relevant network will also have an effect on that country's own yields, for two main reasons. First, a change in bond prices in its network may indicate that those countries are likely to engage in conflict. Because allies are more likely to be dragged into a war, such changes in bond yields abroad may suggest an upcoming conflict and hence may push the yields up at home. A second mechanism works in the opposite direction: looming conflict abroad will push investors to search for safer alternatives. 'Neighboring' countries are a likely outlet for these investors, who are therefore likely to push down the yields of members of the warring country's network. To include these effects, we therefore need to include in our model the yields of the members of the network. To that end, we adopt a simple spatial regression framework. First, we obtain a measure of foreign policy similarity ('fps') from Häge (2011), from which we infer a connectivity matrix of each country's network, W_i, an $n \times n$ matrix whose i, j element is $\frac{\text{fps}_{i,j}}{\sum_{k=1}^{N} \text{fps}_{i,k}}$. $W_i \text{Yield}_j$ thus gives us an average of the yields in country i's network, weighted by foreign policy similarity.

Finally, more severe conflicts tend to be costlier, and hence will likely lead to a larger shift. We would therefore like to include a measure of severity, but information about those variables remains undefined at the time of the onset—they will only be known at the end of the war. What we need then is information about severity that was available at the time of the onset.[6] Unfortunately there is no data that documents the breakdown of casualties over the course of the conflict. We address this difficulty in two ways. First, we limit our attention to conflicts of a very short duration. Indeed, if the conflict lasted less than, say, a week, then the market's correction will reflect information about severity that was available to the market at

that a two-week period is better to capture some of the uncertainty around the dating of the onset.

[6]To be sure, markets form expectations not only about the immediate cost of the war, but also about its long-term costs, in particular as they may impact the ability of the state to repay its debt. How markets estimate this expected cost is difficult to say, and is likely to be a function, among other, of the relative strength of the participants, the expected duration of the conflict, and the expected outcome. In particular, one party may be expected to bear a disproportionate share of the cost and therefore suffer a far worse adjustment to its yield. Some of this asymmetry is included here, since the casualties are country-specific. In that sense, we expect those who suffer larger casualties to suffer a large cost—a result that is supported by our analysis below.

the onset. We therefore focus on those short conflicts only and include the (logged) number of casualties in that week (Fatalities, dispute) and alternatively a dummy variable coded as 1 if this short dispute generated any casualties, and zero otherwise. While this approach deprives us of another important dimension of the cost of conflict, its duration, it allows us to isolate the effect of the fatalities dimension of the cost. Our second approach is to use incident-level data from the Correlates of War, which documents the number of casualties associated with a specific incident (a dispute may include several incidents). While these data are only available post-1993, they allow us to examine the effect of the onset of all wars (not only the short ones) on yields while controlling for the severity of the initial event. Summary statistics of all variables are reported in Table A4. Finally, we note that similar tests using the entire sample and measures of duration and fatalities for the entire conflict yield very similar results.

We estimated a simple OLS with standard errors clustered by country (clustering by year as well made no difference), with country-level fixed effects (adding yearly effects also had little effect). The results are reported in Table 1.

We note two things. First, overall R-squared is low (it ranges from 0.3% to 2.3%), but not surprisingly so. We are explaining fluctuations in bond yields—i.e., financial market returns—which typically are stochastic. Campbell & Thompson (2008), for example, review variables listed in the financial literature to account for variations in monthly stock market returns, and show that these variables lead to in-sample R-squared values ranging from 0.02% to a maximum of 2.6%.

Second, we find that the onset of a war in a given week has a positive and strongly significant effect on that country's yields. The effect is small, but keep in mind that we are dealing with the change in yield over a single week.[7] This shows that even when looking at the universe of all changes in yield, which are caused by countless factors—economic crises, exchange rate regime changes or regime changes—war onset has a positive and significant impact on yields. This result strongly supports hypothesis 1.

However, the setup also has major disadvantages. First, it compares the change in yield associated with the onset of war to all changes in history. Many of these shifts have causes that are of no interest here, including financial crises or changes in exchange rate regime (still, we note that controlling for the gold standard or the type of exchange rate regime had no substantive effect on our results). That is, the absolute size of the shift tells us little about the extent of the investors' surprise. While it is remarkable that we find a positive and significant shift even when compared with all shifts in the time-span covered, our ultimate goal is not to compare the shock of war with other shocks, and hence the model's interpretation is limited. Second, this model is inflexible. As it is, the setup focuses only on the week of the onset, but we would like to understand more about the path leading to war and what happens after the onset—not just in that very week. We therefore now evaluate our hypotheses using yields around the onset of conflict.

[7] To ensure that our results are not artificial flukes of the data, we generated a synthetic data of 'conflicts' with the same characteristics as the actual conflicts, but randomly assigned to another country-date. The results show that there are no breaks for wars that did not occur.

Table 1. Evidence of a level shift after the onset of war. The dependent variable is the change ($\ln(yield_t/yield_{t-1})$, multiplied by 100) in the 10-year government bond yield for country i and time t. OLS run with country fixed effects (yearly FE make no difference) and standard errors clustered by country.

	(1) DV: ΔYield$_{it}$ (log)	(2) DV: ΔYield$_{it}$ (log)
War onset$_{it}$	0.228**	0.265**
	(0.045)	(0.086)
ΔYield$_{it-1}$ (log)	0.553	-11.207^{**}
	(4.660)	(3.828)
ΔYield$_{world,t}$ (log)	1.176**	0.653
	(0.413)	(0.453)
W Yield$_j$	-0.0017	0.0029
	(0.0014)	(0.0021)
ΔCBRATE$_{it}$ (log)		4.506
		(2.581)
ΔCPI$_{it}$ (log)		10.082
		(5.150)
GDPPC (log)		-0.011
		(0.019)
Govt debt (%, log)		-0.0010^{**}
		(0.0004)
Polity		0.0032
		(0.0042)
N	241,379	130,115
Overall R^2	0.003	0.022

$^{**}p < .01$; $^{*}p < .05$

Standard errors clustered by country in parentheses

Fluctuations around the onset

Here we study the evolution of yields shortly before and after war to assess the market's reaction to the onset. For each of the 2,516 conflicts in our data, our dependent variable in this section is the yield of that country's sovereign bond three months before and after the onset of conflict. We standardize these time series as z-scores based on the pre-war distribution.[8]

Figure 1 illustrates the data by plotting the evolution of the 10-year government bond yield around the onset of the two World Wars for three different countries. The time series are standardized over the interval for the purpose of comparability. We note that WWI, for example, led to a large jump in bond yields, whereas the slow escalation of the 1930s and Hitler's clear intentions left no one incredulous in 1939. For illustration purposes, we also aggregated these standardized time series for all 176

[8] We chose a two-year pre-war interval as the basis for the standardization to allow for a sufficient distance from the onset. However, using the entire pre- and post-war period makes little difference.

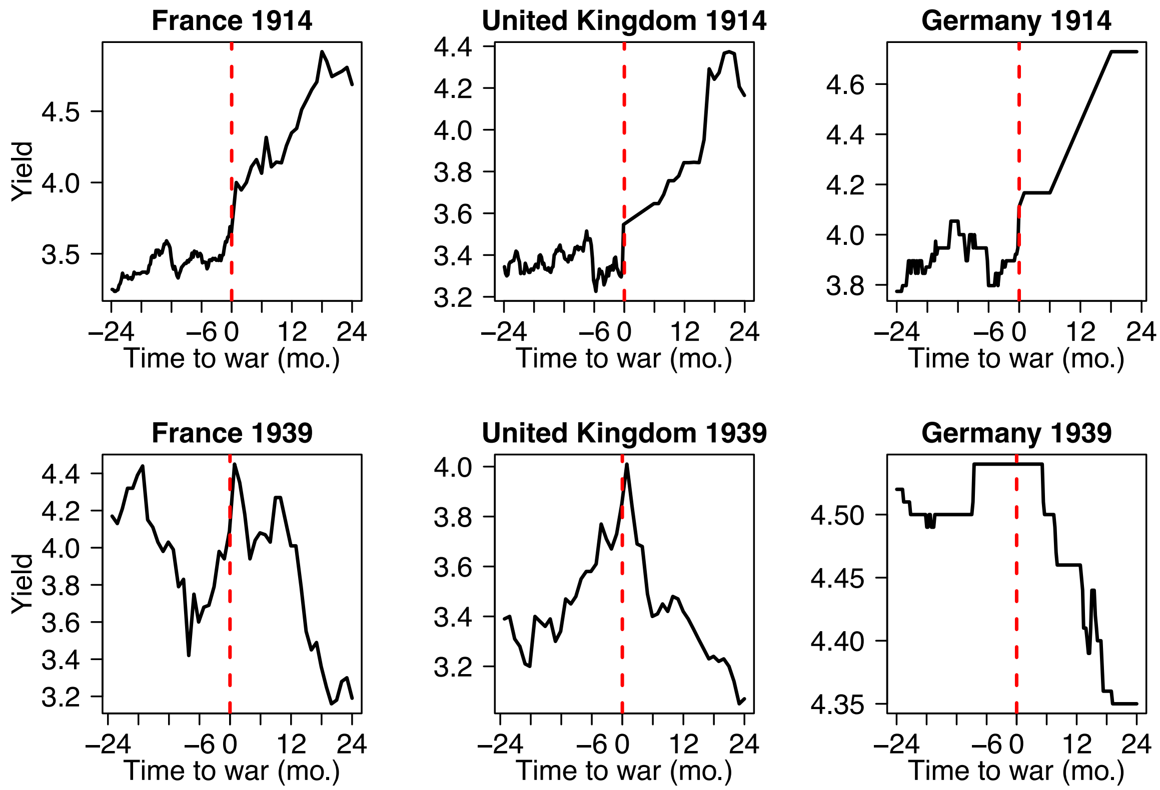


Figure 1. Evolution of government bond yields around the onset of WWI and WWII.

large wars in our sample (those with at least 10,000 deaths).[9] Overall, the pattern shows a clear jump immediately before war and following its onset (Fig. 2).[10]

We now confirm these results more formally by estimating a model in which time series of bond yields for all 2,516 conflicts (standardized as z-scores) were regressed on a dummy variable coded as zero before the onset, and one thereafter (AFTER). In addition, we include a variable measuring the number of days until or since war (TIME TO WAR).[11] Because markets may not worry about small skirmishes, we expect only severe incidents to lead to a jump and therefore interact the AFTER dummy with three different measures of conflict severity, all of which avoid the 'hindsight' fallacy that would use information that was not available at the time. For robustness purposes, we also controlled for the central bank's rate, inflation levels (see above), the country's Government debt levels as a percentage of GDP, its GDP per capita and Polity score (see page 8 for details), and the yields of countries in the network,

[9]While we include conflicts of all sizes later in the analysis, we do not expect a large reaction of the market to minor conflicts and hence not one that can be detected graphically.

[10]A simple way to test the difference between yields pre- and post-onset is to simply run a t-test or a Mann-Whitney test ($t = -11.7$ ($p < 0.001$); Mann-Whitney W = 150,339 ($p < 0.001$)). However, the null hypothesis may be rejected because of non-stationary data. Assuming for example an increasing but smooth—continuous—trend, all of these tests would conclude to a significant difference between bond yield before and after war, but not whether there is a jump.

[11]This variables takes negative values before the war (e.g. -26 means 26 days until the onset) and positive thereafter (e.g. 18 days since onset).

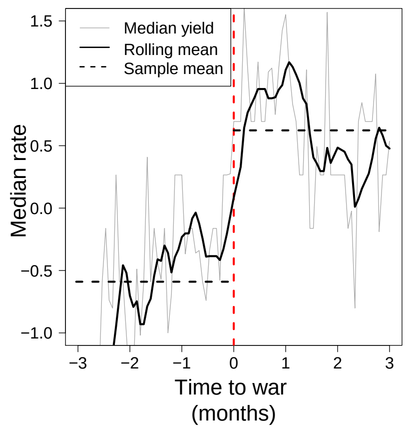


Figure 2. Median bond yields around the onset of large conflicts. Each time-series is standardized as a z-score over the $\pm$3-month span, with mean 0 and standard deviation 1.

weighted by their policy similarity ($W_i Yield_j$—see description above).

The idea behind this regression design is that if markets correctly estimate the risk of war, then we should not observe a jump in yields around the time of the onset. In other words the 'Time To War' variable would be significant—a smooth increase towards the value it takes after the war—but the 'After' war dummy would not, since there would be no jump. Yet we find the opposite: the interaction between 'after' and either of our measures of severity is strongly significant, with a substantial effect, which indicates that the onset of wars with at least some casualties does lead to a level shift in yields (Table 2).[12] This result holds for all three measures of severity, i.e., even when looking only at short wars or incidents for which the number of casualties was known from the start.

Following Central Bank Rate, 'After $\times$ Fatality dummy' is the variable with the largest effect on yields (fatality dummy here refers to whether there were any fatalities in the dispute, keeping in mind that we only look at very short—less than one week—disputes to avoid using any information that will only become available later in time).[13] Finally, we note that the effect of the onset on yields might be indirect, in the sense that the war leads the central bank to increase its rate, which in turn leads to a jump in yields. This is an issue that we address in appendix A.3. Using a mediation analysis, we show that the effect of the onset on yields is in fact even stronger than the regression results would lead to believe. Overall, then, these findings strongly support hypothesis 1 that conflicts will tend to be surprising, and hypothesis 2 that the correction is larger for more severe events. Figure 3 illustrates the predicted values of standardized bond yield values as a function of time to war, for different numbers of casualties in the first week of the conflict.

[12]Note that this method is nearly equivalent to running a Chow test comparing the fit of one single regression line against two regression lines separated by the break (Chow, 1960).

[13]For clarity, standardized ('beta') coefficients are also reported in Figure A5.

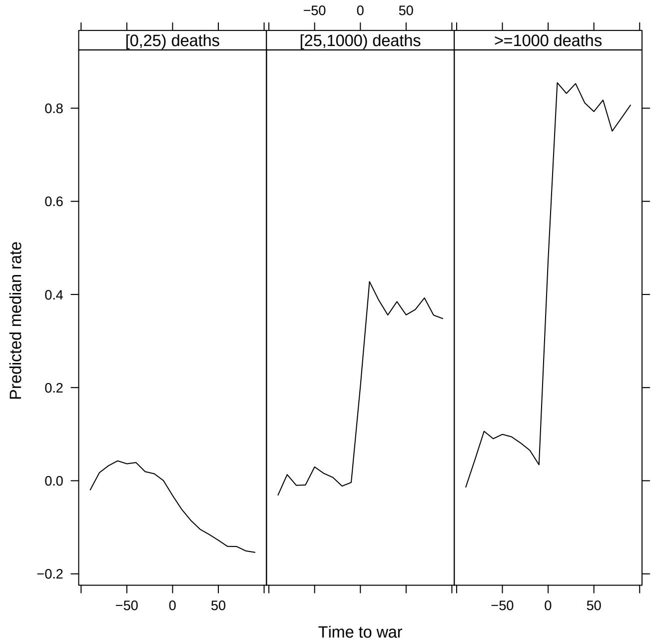


Figure 3. In-sample median of predicted values of bond yield (standardized), as a function of time to war and number of casualties in the first week of conflict. Predicted values were derived from model (1) in Table 2.

Markets underestimate the risk of war

So far we have aimed at showing the existence of a correction around the time of the onset of wars. Yet the correction that we observe may not necessarily reflect the market's underestimation of the probability of war, but rather the simple fact that the uncertainty about the onset of war was resolved. Just like a bookmaker could correctly estimate the probability of a horse winning to be 90%, even though the horse ends up losing (an outcome that we would expect to happen 10% of the time), markets may correctly anticipate the *probability* of war, and adjust the price with a correction once the war happens for certain. If the onset of war is a stochastic process, the correction may be a sign not of underestimation, but rather of a move to certainty.

To ensure that markets do in fact *underestimate* the probability of war, what we need to test is not only whether they react to the onset of war, but also whether their pre-onset estimates were actually biased. Obviously there is no way to assess whether the estimated probability of a single event was correct. If we say for example that a coin has a 20% chance of landing on Tail, and the result of a single flip is Heads, we still cannot establish that our 20% estimate was incorrect. After recording multiple flips, however, we might be able to determine whether our predicted probability was correct. What we need, in other words, is to measure the market's estimation of the risk of war in repeated samples, and to compare the *average* estimate to the actual overall probability of onset. In short, war should happen 50% of the time when the predicted probability is 50%.

We therefore estimate the predicted probability of war onset implied by the value of bonds at time t by estimating a simple model of the form:

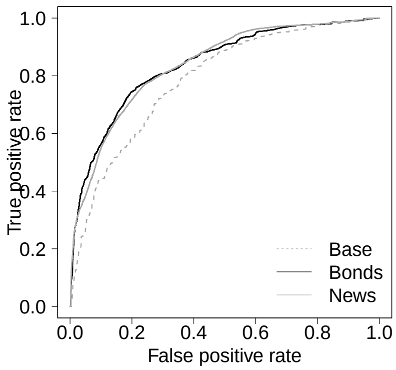
$$p(war)_{it} = \beta \text{Yield}_{it} + u_i + \varepsilon_{it},$$

where u_i is a fixed effect for state i at time t and ε_{it} is the error term. β is a parameter to be estimated. We estimated this model using a simple logistic regression and calculated out-of-sample predicted probabilities on a rolling basis (Chadefaux 2014; see also Colaresi & Mahmood 2017). For example, we use data prior to 1920 as a learning set, and calculate predicted probabilities for the following year. We therefore end up with predicted probabilities for the period 1920–2007, which we can then compare to the true probability of onset over the same period. For reference, we compare the predictions of the model based on bonds data to ones based on the number of conflict-related news (i.e., $p(war)_{it} = \gamma \text{Conflict-related news}_{it} + u_i + \varepsilon_{it}$; see Chadefaux (2014) for details on the measurement of conflict-related news) and a base model using only country-level fixed effects (i.e., $p(war)_{it} = u_i + \varepsilon_{it}$).

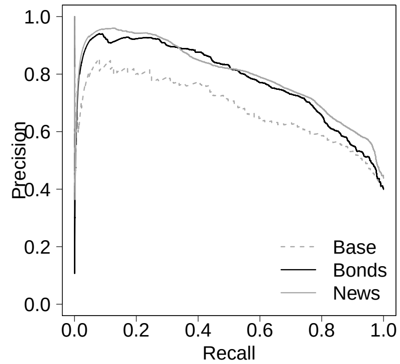
We first note that the model using bonds is good at discriminating between cases in which a war is coming and those where it is not. Indeed, its area under the Receiver-Operating Characteristic curve (ROC) is nearly indistinguishable from the one generated using data derived from news (Fig. 4a left). The area under the Precision-Recall curve is also very comparable (Fig. 4a right).[14]

However, we find that the calibration of the 'bonds' model is poor. The calibration of a test refers to its capacity to accurately predict absolute risk levels by comparing the predicted risk to the observed occurrence rate (Steyerberg et al. 2010). If markets were unbiased, we should observe that war happens 20% of the time when markets predict a 20% probability of conflict. On a calibration plot, with predicted probabilities on the x axis and actual occurrence rate on the y-axis, in other words, data points should fit neatly on the 45 degree line. Yet this is not at all what we observe. While the model based on news (a very simple model using a count of conflict-related news and fixed effects) performs well on calibration metrics, the same model based on bonds does poorly. Thus when predictions derived from the bonds model forecast a 40% probability of conflict (i.e., $\hat{Y} = .4$), the true probability is really closer to 55% (i.e, war happens in 55% of these cases—$\bar{Y} = 55\%$). This underestimation applies to nearly all levels of prediction (Fig. 4c) The average bias can be calculated simply as $\frac{1}{N}\sum_i(\hat{y}_i - y_i)$. A negative (positive) score implies underestimation (overestimation) of the probability of war onset on average. We find that both models ('news' and 'bonds') underestimate the risk of war, but that the bias is far more pronounced in the case of bonds, even though the model is the same, and this result holds for all specifications of the bonds model that we tried. On average, predicted values based on the regression using bonds are more than 7 percentage points lower than the actual risk of war. By contrast, the equivalent statistic for the model that uses news is only 3 percentage points off—less than half.

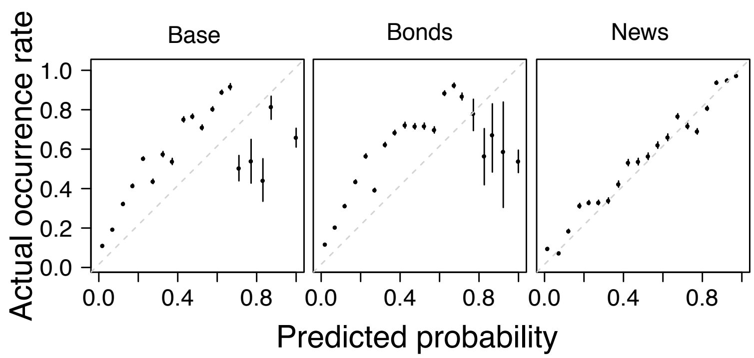
[14]The Precision-Recall curve is a better metric for the assessment of the predictive power of a model involving imbalanced data, such as conflict data.



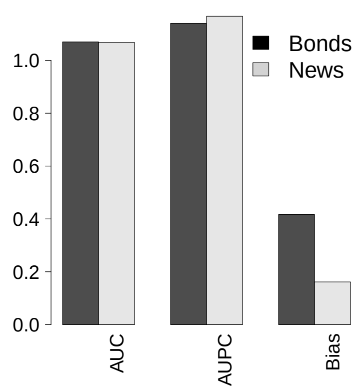
(a) Receiver operating characteristic curve



(b) Precision-recall curve



(c) Calibration plot



(d) Measures of out-of-sample performance: area under the ROC curve, area under the Precision-Recall curve; and bias ($\frac{1}{N}\sum_i(\hat{y}_i - y_i)$). All values are reported as ratios to the baseline model, which takes value 1. We note that for all metrics the model based on bonds performs approximately as well as the one based on bonds, but that the bias exhibited by the bonds model is much larger than the one for news.

Figure 4. Out-of-sample forecasting performance. For each year t, we recursively estimated logistic regressions of the onset of war in country i at time t using $[t_1$,t) as the learning set and year t as the testing set. The dependent variable is the onset of conflict in week t and country i. Three models were estimated: a baseline model including only country fixed effects; the baseline model with the addition of bond yield data; and the baseline model with the additional conflict-related news counts. We find that the bonds-based model exhibits good discrimination (a and b), but poor calibration (c and d).

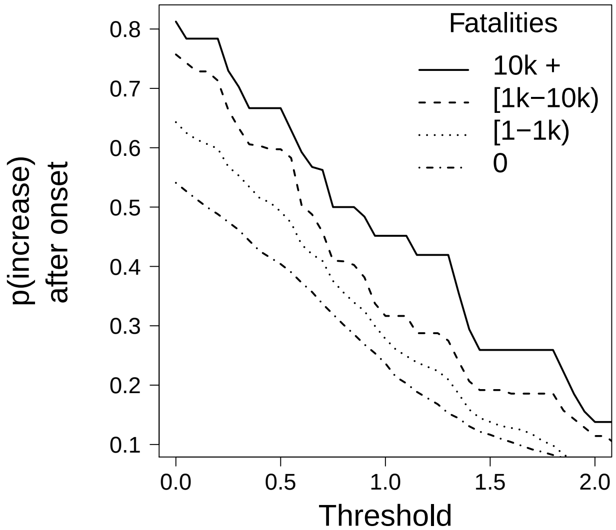


Figure 5. Probability (observed data) of a post-onset increase in bond yield as a function of the threshold selected (i.e., $p(\overline{\text{Yield}}_{\text{after onset}} - \overline{\text{Yield}}_{\text{before onset}} \geq \text{threshold})$).

This result, combined with the jump that we observe following the onset, confirms that markets tend to underestimate the risk of war. We now turn our attention to explaining the variation in that level of surprise.

Which wars are surprising?

On average, then, wars lead to a jump in bond yields. Yet many do not. How often are they surprising? The answer obviously depends on the threshold we set for a 'surprise'. Figure 5, for example, displays the proportion of cases followed by a jump in yield of at least x standard deviations. Of course the larger the threshold for a 'surprise', the fewer wars qualify as surprises. The answer also depends on the size of the war, since larger wars tend to lead to larger jumps. Wars with at least 10,000 battle deaths, for example, lead to an increase in government bond yield in more than 80% of cases (i.e., in 70 large wars out of 100, the average yield in the three months following the onset of war is larger than the yield in the preceding three months), but no conflict led to an increase of more than 1.8 standard deviations.

Regardless of the threshold we adopt, the variance in jump is puzzling. What affects whether some wars come as a surprise when others do not? We now change our dependent variable to consider the shock itself. Our dependent variable, Δ_3YIELD, is now the change in a country's government bond yields following each of the 2,516 wars for which data is available. It is obtained by subtracting the average yield in the three months that precede the onset of war from the average yield in the three months that follow it. Whereas the previous section was concerned with whether a shock occurred at all, here we are interested in the size of the shock as a function of various covariates.

Variables used to test our hypotheses include DATE, an index of time (in decades), where 1 January 1816, takes value 0 and 12 December 2007, value 19.2. We include this variable to estimate the effect of time on the severity of shocks to test hypothesis 3. Furthermore, dummy variables reference the type of war (INTER)—interstate wars are coded as one, and intrastate ones as 0. Democracies are also expected to be more reactive with a more transparent bargaining process, so that conflicts involving them should be more difficult to predict than those involving autocracies (hypothesis 4). We use the Polity IV score (Marshall, Jaggers & Gurr, 2002) (POLITY). Variable PEACE DECADES (together with its square and cube) denotes the number of decades since the onset of the previous conflict (Carter & Signorino, 2010), as we expect markets in countries with recent conflicts to be less surprised about the onset of war than countries in which conflicts are a distant memory. We also include measures of expected cost in the form of the trade ratio ($\text{Imports}_{ij}/\sum_j \text{Exports}_{ij}$) and CINC ratio ($\text{CINC}_i/\sum_j \text{CINC}_j$) and the weighted value of the change in countries that are part of i's network ($W_i\text{Change}_{j,t}$, which is defined in a similar fashion as above).

Other control variables include the size of the prior change in bond yields for that country ('Δ_3YIELD (lag)'). This variable measures the magnitude of increase or decrease in bond yields in the three months preceding the war.[15] We include this variable because we expect some time-dependence in shocks. We also control for the average change in bond yields in the world in that year, since average bond yields might decrease over time, independently of the onset of war. We also control for the country's national material capabilities (NMC) are also included using the Correlates of War's Composite Index of National Capability (Singer, 1988, v. 4.0).

Finally, we also include an index of the worldwide number of conflicts in a given year as a proxy for the level of risk associated with that period ('N conflicts this year'). A large number of conflicts may indeed indicate a dangerous system, perhaps because of multipolarity, shifts in power, or various idiosyncratic events such as the end of the Soviet Union. If conflicts are widespread, observers of international relations are less likely to be surprised at the onset of yet another one, and markets will therefore already have incorporated the risk. A larger number of conflicts in the world should therefore reduce the surprise associated with conflicts. Summary statistics are reported in Table A6.

We regressed the size of the shock following each conflict on the covariates described above. The results are reported in Table 3 (see also Fig. A6).

We find that our hypotheses are largely supported. In particular, the coefficient on date has a small and statistically insignificant effect on the size of the surprise, in support of hypothesis 3. For models in which it is significant, the coefficient is actually positive, suggesting again that our prediction record has not improved over the past two centuries.[16]

[15] I.e., $\Delta_3\text{YIELD (lag)} = \overline{\text{YIELD}}_{t\in[w-3,w)} - \overline{\text{YIELD}}_{t\in[w-6,w-3]}$, where w denotes the onset of war, and $w-3$ three months prior to that onset. $\overline{\text{YIELD}}$ is the average yield over that period.

[16] Using time dummies instead (e.g. 50-year periods) yields the same results: none of them are

Figure 6a and 6b provide visual intuition for this result by displaying the shock that followed each of the 2,516 wars in our sample, and the absence of pattern over time. This matches our conjecture that continuous learning and policy adjustments may lead to ever changing or more complex patterns prior to conflict, and hence to the fact that the wars that do occur are those that could not have been easily predicted. Just as markets are essentially random walks because participants continuously incorporate new information in such a way that no arbitrage is possible, we conjecture that leaders also adopt decisions in reaction to what they know from the past and the information available. This constant process of adjustment means that wars cannot be easily forecasted.[17]

In line with our expectations, we also find the magnitude of the shock to be larger in democratic countries than in autocratic ones, supporting hypothesis 4. This effect is strongly significant and substantial. On the other hand, while interstate wars do lead to a larger average increase in yields than intrastate wars, as expected, the effect is not significant, and hence we find little in support of hypothesis 5.

One possible concern is that the set of variables included may have been intentionally selected to support certain hypotheses, or may simply be a lucky combination. I address these concerns about model uncertainty by running a Bayesian Model Averaging (BMA), a technique used for example in Warren (2014) and Ward & Beger (2017). The outcome of the BMA are reported in appendix A.4 and strongly support our results.

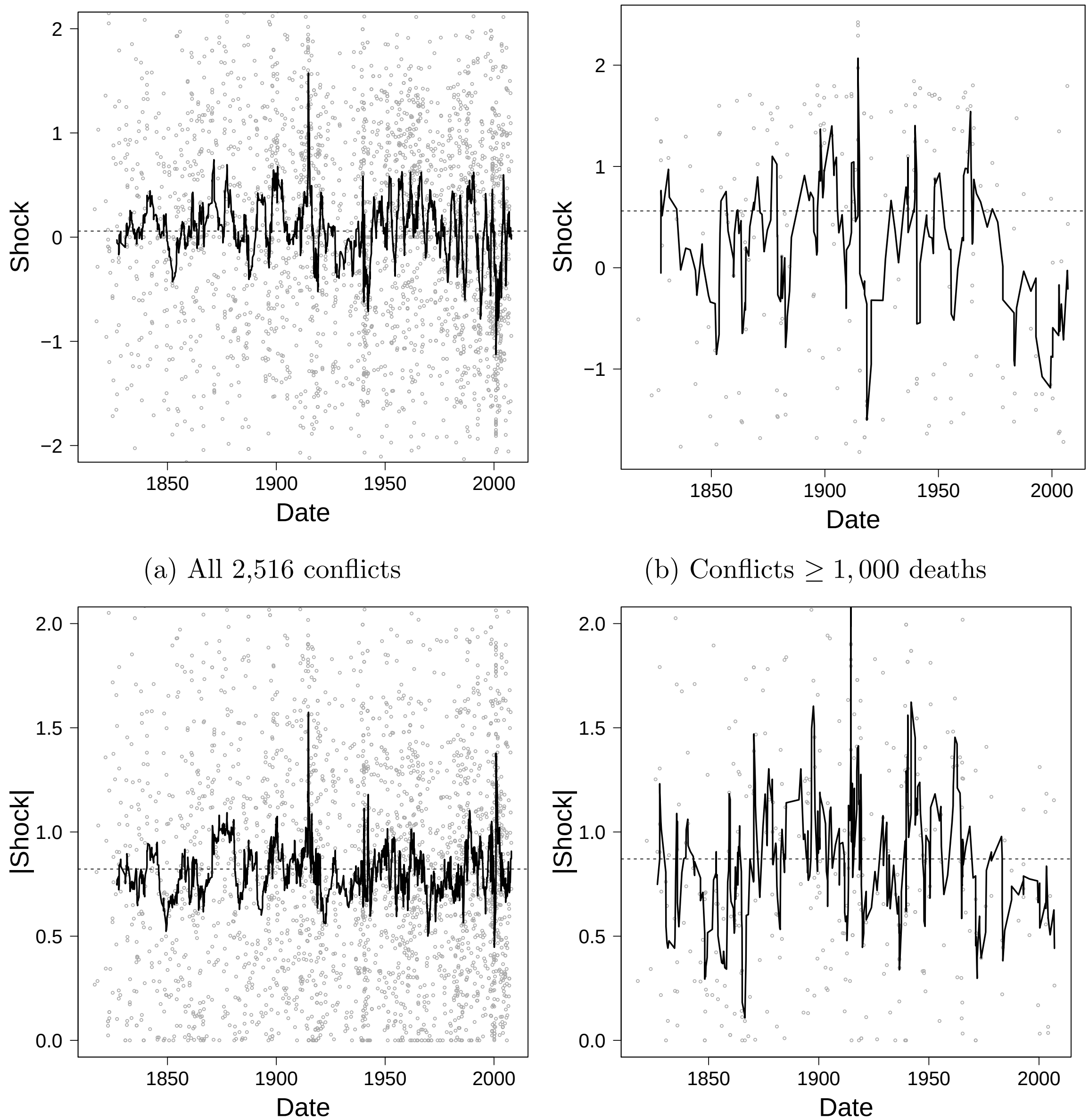
Conclusion

Policy-makers and students of international relations have long sought to anticipate and prevent the onset of conflict. Yet results presented here suggest that even those who have a financial interest in their accurate prediction have been rather unsuccessful. This does not imply that contemporaries are oblivious to the escalation of tensions (Chadefaux, 2014), but that they do tend to underestimate the risk of war.

Yet this seemingly damning result may in fact not be an indictment of markets' forecasting ability. Rather, because conflicts that are anticipated well ahead may be more likely to be avoided, only the difficult cases are left in our sample. The apparent recurrent failure to estimate the risk of war may in that sense simply be a selection effect. If policy-makers incorporate some of the available information—

significant, and in any case do not exhibit any clear pattern (not reported out of space concerns). Some of the finer-grained dummies (e.g. decade dummies) are significant, but still without any clear pattern.

[17] Negative shocks are also frequent, as is clear on Figure 6a and 6b. Negative shifts may arise for two main reasons: a) investors are 'relieved' that the large conflict they anticipated ends up just being a skirmish; b) there might be a flight to safety from investors, such that increased risk may actually push the yields down, as can be the case for safe havens such as the United States, Switzerland, or Germany. Overall, however, the aggregate pattern is one of an upward jump in yields. Moreover, even if we take all shifts, including negative ones, as evidence of a surprise, the pattern associated with the absolute values of the changes in yields shows an increasing—not a decreasing—trend over time, further supporting our hypothesis 1.



(a) All 2,516 conflicts

(b) Conflicts $\geq 1,000$ deaths

(c) All 2,516 conflicts, abs. value

(d) Conflicts $\geq 1,000$ deaths, abs. value

Figure 6. Change in yield ('shock') following the onset of conflict, and their evolution over time. Each dot plots the difference between the average yield in the three months following a given war and the three months prior to it (absolute value thereof on bottom row). The solid line is a rolling average of these shocks. The dashed horizontal line is the overall median.

including lessons from the past and forecasts given available data—then their behavior will be constantly adapting to new information, and markets may always be one step behind and will tend to be taken by surprise by policy-makers' decisions. The fact that wars are, on average, just as surprising today as they were in 1816 supports this selection process, by which only the wars that are the most difficult to predict occur.

As a corollary, countries with more transparent and possibly reactive regimes such as democracies should be better at incorporating new information, and hence predicting their behavior should actually be more difficult—a hypothesis for which we found strong support here.

Our findings may also suggest a 'policy efficiency' hypothesis. If the evidence for market efficiency is the quasi-impossibility to predict future changes in asset prices based on current patterns, then the constant inability of markets to correctly assess the risks of war may also mean that policy-makers incorporate existing information rapidly into their decisions, and hence that policy in that sense is 'efficient'. Additional work on how leaders incorporate new information and forecasts into their decisions may lead to further insights on this subject.

Wars are, at least in part, failures of predictions. They often occur when their participants fail to predict the consequences of their actions. Far from being a depressing diagnostic, then, our results show the importance of prediction as one possible instrument of conflict prevention, and the role of scholars in bridging the gap between basic and applied research.

Table 2. Evidence of a level shift after the onset of war. The dependent variable is the 10-year government bond yield for country i three months before and after the onset of war (standardized as its z-score, i.e., with mean 0 and standard deviation 1). Starred variables have also been standardized (z-score). Fixed effects for each country included.

	(1) DV: Yield_{it}	(2) DV: Yield_{it}	(3) DV: Yield_{it}
After	-0.029	-0.029	-0.072^{**}
	(0.018)	(0.019)	(0.021)
Fatalities (dispute, log)	0.035^{*}		
	(0.014)		
After × Fatalities	0.115^{**}		
	(0.019)		
Fatality dummy (dispute)		0.031	
		(0.034)	
After × Fatality dummy		0.319^{**}	
		(0.053)	
Fatalities level (incident)			0.106^{*}
			(0.049)
After × Fatalities level			0.044
			(0.024)
Gov. debt (%)	0.095	0.092	0.129
	(0.133)	(0.135)	(0.093)
GDPPC (log)	0.180^{*}	0.184^{*}	-0.103
	(0.082)	(0.085)	(0.634)
Central bank rate	0.526^{**}	0.523^{**}	0.500^{**}
	(0.074)	(0.072)	(0.050)
Polity	-0.013	-0.013	0.081^{**}
	(0.012)	(0.012)	(0.011)
Inflation*	0.013	0.013	0.030
	(0.025)	(0.025)	(0.028)
Time to war (yrs)	-0.0014^{**}	-0.0014^{**}	-0.00067^{*}
	(0.00032)	(0.00032)	(0.00030)
Time to war^2 (yrs)	-5.4×10^{-6}	-5.4×10^{-6}	-4.6×10^{-7}
	(4.1×10^{-6})	(4.1×10^{-6})	(3.4×10^{-6})
Time to war^3 (yrs)	9.7×10^{-8}	$9.8\times10^{-8*}$	$1.0\times10^{-7*}$
	(4.9×10^{-8})	(4.9×10^{-8})	(4.3×10^{-8})
$W_i\text{Yield}_{j\in J,t}$	0.00047	0.00028	0.011
	(9.4×10^{-4})	(9.8×10^{-4})	(0.011)
N	22,145	22,145	27,244
Overall R^2	0.372	0.370	0.314

$^{**}p < 0.01$, $^{*}p < 0.05$

Standard errors clustered by country in parentheses

Table 3. Factors affecting the magnitude of the post-onset correction. The dependent variable is the change in yield in country i before and after the onset (i.e., $\Delta\text{yield}_3 = \overline{Yield}_{i,t\in[-3mo,w]} - \overline{Yield}_{i,t\in[-3mo,w]}$), where $\overline{Yield}$ is the average yield over that period.

	Δyield_3	Δyield_3
Date	0.00072	0.0062^{**}
	(0.00067)	(0.0016)
Interstate conflict	−0.078	
	(0.091)	
Polity	0.024^{**}	0.030^{**}
	(0.004)	(0.007)
N conflicts this year	-0.013^{**}	0.0059
	(0.002)	(0.002)
N past wars	-0.0020^{**}	0.00027
	(0.00034)	(0.00051)
Peace decades	-0.803^{**}	−0.387
	(0.179)	(0.348)
Peace decades^2	0.342^{**}	0.418
	(0.092)	(0.281)
Peace decades^3	-0.028^{**}	−0.085
	(0.010)	(0.050)
Change_{t-1}		1.289^{**}
		(0.182)
$W_i\text{Change}_t$		0.109^{**}
		(0.018)
$W_i\text{Change}_{t-1}$		0.304^{**}
		(0.109)
ΔMILEX_i		5.4×10^{-9}
		(5.9×10^{-9})
$\sum_j \text{MILEX}_j$		7.3×10^{-9}
		(6.1×10^{-9})
GDPPC		-0.065^{**}
		(0.007)
Import ratio		0.151
		(0.210)
CINC ratio		−0.095
		(0.114)
ΔCBRATE_3		0.151^{**}
		(0.011)
Inflation		0.00010
		(0.00032)
(Intercept)	0.524^{**}	-2.038^{**}
	(0.101)	(0.347)
N	2,541	1,022
Overall R^2	0.054	0.374

$^{**}p < 0.01$, $^{*}p < 0.05$

Standard errors clustered by country in parentheses

References

Amihud, Yakov & Avi Wohl (2004) Political news and stock prices: The case of Saddam Hussein contracts. *Journal of Banking and Finance* 28(5): 1185–1200.

Beck, Nathaniel; Gary King & Langche Zeng (2000) Improving quantitative studies of international conflict: A conjecture. *American Political Science Review* 94(1): 21–36.

Beck, Nathaniel; Gary King & Langche Zeng (2004) Theory and evidence in international conflict: A response to de Marchi, Gelpi, and Grynaviski. *American Political Science Review* 98(2): 379–389.

Bhavnani, Ravi; Karsten Donnay, Dan Miodownik, Maayan Mor & Dirk Helbing (2014) Group segregation and urban violence. *American Journal of Political Science* 58(1): 226–245.

Brandt, Patrick T; John R Freeman & Philip A Schrodt (2011) Real time, time series forecasting of inter- and intra-state political conflict. *Conflict Management and Peace Science* 28(1): 41–64.

Brune, Amelie; Thorsten Hens, Marc Oliver Rieger & Mei Wang (2015) The war puzzle: Contradictory effects of international conflicts on stock markets. *International Review of Economics* 62(1): 1–21.

Campbell, John Y & Samuel B Thompson (2008) Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21(4): 1509–1531.

Carter, David B & Curtis S Signorino (2010) Back to the future: Modeling time dependence in binary data. *Political Analysis* 18(3): 271–292.

Chadefaux, Thomas (2011) Bargaining over power: When do shifts in power lead to war? *International Theory* 3(02): 228–253.

Chadefaux, Thomas (2014) Early warning signals for war in the news. *Journal of Peace Research* 51(1): 5–18.

Chadefaux, Thomas (2015a). The triggers of war: Disentangling the spark from the powder keg. Working paper (`https://ssrn.com/abstract=2409005`).

Chadefaux, Thomas (2015b). What the enemy knows: Common knowledge and the rationality of war. Working paper (`https://ssrn.com/abstract=2669555`).

Chen, Andrew H & Thomas F Siems (2004) The effects of terrorism on global capital markets. *European Journal of Political Economy* 20(2): 349–366.

Chow, Gregory (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28(3): 591–605.

Colaresi, Michael & Zuhaib Mahmood (2017) Do the robot: Lessons from machine learning to improve conflict forecasting. *Journal of Peace Research* 54(2): XXX–XXX.

De Marchi, Scott; Christopher Gelpi & Jeffrey D Grynaviski (2004) Untangling neural nets. *American Political Science Review* 98(2): 371–378.

Fama, Eugene F (1991) Efficient capital markets: II. *The Journal of Finance* 46(5): 1575–1617.

Fearon, James D (1995) Rationalist explanations for war. *International Organization* 49(03): 379–414.

Gartzke, Erik (1999) War is in the error term. *International Organization* 53(3): 567–587.

Gatignon, Hubert (2003) *Statistical analysis of management data.* New York: Springer.

Gleditsch, Kristian Skrede & Michael D Ward (2013) Forecasting is difficult, especially about the future: Using contentious issues to forecast interstate disputes. *Journal of Peace Research* 50(1): 17–31.

Guidolin, Massimo & Eliana La Ferrara (2010) The economic effects of violent conflict: Evidence from asset market reactions. *Journal of Peace Research* 47(6): 671–684.

Häge, Frank M (2011) Choice or circumstance? Adjusting measures of foreign policy similarity for chance agreement. *Political Analysis* 19(3): 287–305.

Hall, George J (2004) Exchange rates and casualties during the First World War. *Journal of Monetary Economics* 51(8): 1711–1742.

Hegre, Håvard; Joakim Karlsen, Håvard Mokleiv Nygård, Håvard Strand & Henrik Urdal (2013) Predicting armed conflict, 2010–2050. *International Studies Quarterly* 57(2): 250–270.

Leigh, Andrew; Justin Wolfers & Eric Zitzewitz (2003) What do financial markets think of war in Iraq? NBER Working Paper No. 9587.

Lucas, Robert E (1976) Econometric policy evaluation: A critique. In: K Brunner & A.H. Seltzer (eds) *The Philips Curve and Labor Markets, Carnegie-Rochester conferences on public policy* volume 1. Amsterdam: North Holland, 19–46.

Marshall, Monty G; Keith Jaggers & Ted R Gurr (2002) Polity IV project: Political regime characteristics and transitions, 1800-2010 Version p4v2010, accessed 1 April 2015.

Poole, William; Robert H Rasche & Daniel L Thornton (2002) Market anticipations of monetary policy actions. *Federal Reserve Bank of St. Louis Review 84* (`https://research.stlouisfed.org/publications/review/02/07/65-94PooleRasche.pdf`).

Preacher, Kristopher J & Andrew F Hayes (2008) Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods* 40(3): 879–891.

Ramey, Valerie A (2011) Identifying government spending shocks: It's all in the timing. *The Quarterly Journal of Economics* 126(1): 1–50.

Reinhart, Carmen M & Kenneth Rogoff (2009) *This time is different: Eight centuries of financial folly.* Princeton, NJ: Princeton University Press.

Rigobon, Roberto & Brian Sack (2005) The effects of war risk on us financial markets. *Journal of Banking and Finance* 29(7): 1769–1789.

Sarkees, Meredith R & Frank W Wayman (2010) *Resort to war: A data guide to inter-state, extra-state, intra-state, and non-state wars, 1816-2007.* London: CQ.

Schneider, Gerald; Maya Hadar & Naomi Bosler (2017) The oracle or the crowd? Experts versus the stock market in forecasting ceasefire success in the levant. *Journal of Peace Research* 54(2): XXX–XXX.

Schneider, Gerald & Vera E Troeger (2006) War and the world economy: Stock market reactions to international conflicts. *Journal of Conflict Resolution* 50(5): 623–645.

Schultz, Kenneth A (2001) *Democracy and coercive diplomacy.* Cambridge University Press.

Singer, David J. (1988) Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985. *International Interactions* 14(2): 115–132.

Steyerberg, Ewout W; Andrew J Vickers, Nancy R Cook, Thomas Gerds, Mithat Gonen, Nancy Obuchowski, Michael J Pencina & Michael W Kattan (2010) Assessing the performance of prediction models: A framework for some traditional and novel measures. *Epidemiology* 21(1): 128138.

Tetlock, Philip (2005) *Expert political judgment.* Princeton, NJ: Princeton University Press.

Ward, Michael D & Andreas Beger (2017) Lessons from near real-time forecasting of irregular leadership changes. *Journal of Peace Research* 54(2): XXX–XXX.

Warren, T Camber (2014) Not by the sword alone: Soft power, mass media, and the production of state sovereignty. *International Organization* 68(01): 111–141.

Witmer, Frank D.W.; Andrew M. Linke, John OLoughlin, Andrew Gettelman & Arlene Laing (2017) Sub-national violent conflict forecasts for sub-saharan africa, 2015-2065, using climate-sensitive models. *Journal of Peace Research* 54(2): XXX–XXX.

THOMAS CHADEFAUX, b. 1980, PhD in Political Science (University of Michigan, 2009); Assistant Professor, Trinity College Dublin (2014–). Research interests: conflict forecasting, and bargaining models of war onset.

A Online Supplementary Material

A.1 Simulated response of markets to a war surprise

A.1.1 The war correction

To illustrate how investors would respond to an anticipated war (see main text, p. 4), we simulated the yield of a typical 10-year bond under different scenarios. Suppose that a bond pays quarterly payments of $C = 1.5\%$ (i.e., 6% annual coupons) with an $r = 5\%$ market rate (we assume r is constant for simplicity, but the logic carries to a changing rate) and $M = 100$ face value (these are typical values in the data). We assume for simplicity that investors expect the cessation of all future coupons after the onset of war (i.e., $\sum_{t=war}^{N} C/(1+i)^t = 0$), though similar results apply if we instead assume a default (i.e., $M = 0$) or a less drastic reduction in payments.

We model the yield of this bond under three simple scenarios: in scenario 1, war never happens. In scenario 2, markets suddenly anticipate the occurrence of a conflict four periods later. In scenario 3, war is not anticipated by markets and happens now.

1. Scenario 1: no war. In this case, the bond pays out its coupons in every quarter (i.e., 40 times) and the principal is paid at the end of the $N = 40$ periods.

$$
\begin{aligned}
P_1 &= \frac{C}{(1+r)} + \frac{C}{(1+r)^2} + \cdots + \frac{C}{(1+r)^N} + \frac{M}{(1+r)^N} \\
&= \frac{1.5}{1+\frac{0.05}{4}} + \frac{1.5}{(1+\frac{0.05}{4})^2} + \cdots + \frac{1.5}{(1+\frac{0.05}{4})^{40}} + \frac{100}{(1+\frac{0.05}{4})^{40}} = 107.83
\end{aligned}
$$

   and hence $\text{Yield}_1 = 6/107.83 = 0.056$

2. Scenario 2: war is expected in four quarters (i.e., at $t = 4$). In this case, markets expect only the next four coupons to be paid, followed by the repayment of the principal at the end of the 40$^{\text{th}}$ quarter. When war happens at time 4, it does not come as a surprise, since it was already anticipated at time 1.

$$
P_1 = \frac{1.5}{1+\frac{0.05}{4}} + \frac{1.5}{(1+\frac{0.05}{4})^2} + \cdots + \frac{1.5}{(1+\frac{0.05}{4})^{4}} + \frac{100}{(1+\frac{0.05}{4})^{40}} = 66.66
$$

   and hence $\text{Yield}_1 = 6/66.66 = 0.090$

3. Scenario 3: war happens now. In this case no more coupons are expected, and markets therefore price the bond at the discounted value of the principal, repaid in 40 periods.

$$
P_1 = \frac{100}{(1+\frac{0.05}{4})^{40}} = 66.66
$$

   and hence $\text{Yield}_1 = 6/66.66 = 0.098$

Figure A1 illustrates this point by representing the evolution over time of the yield of a bond when investors suddenly foresee the onset of a war.

A possible concern is that the jump in yields we observe is not evidence of a surprise at the onset of war, but would rather be due to uncertainty about the precise timing of the onset. We address this concern below.

A.1.2 Timing Surprises

Here we address the concern that investors may be surprised not so much by the onset of conflict itself, but rather by its timing. An onset at time w rather than at $w + 1$ as they expected, for example, would lead to a correction in prices. We do not think that this is a major concern, however. Indeed, the fact that we use long-dated bonds alleviates this concern. Bonds with 10-year maturity should be affected by distant events, as it would be irrational for investors to wait until the very onset of war to demand a higher yield. As a result, a war that is anticipated but that occurs earlier than expected would lead to a minor jump, as compared to the shift in yield associated with the perception of a looming war. Figure A1 illustrates this point by modeling the evolution of the yield of a sample 10-year bond with quarterly payment affected first by the expectation of a coming war, and then by the earlier than expected onset. The earlier-than-expected onset is exactly equivalent to scenario 3 above, since it amounts to an unexpected war today. The expectation of war leads to a large jump, but the subsequent surprise at the earlier-than-expected onset only leads to a minor change in yield.

In fact, for surprises about timing to drive our results, two conditions would need to be met. First, only large timing errors lead to significant jumps, and hence it would need to be the case that markets systematically overestimate the time to war by a large amount such as a year or more. This seems implausible. It would also need to be that markets rarely underestimate the time to war—i.e., rarely think a war would happen soon when it really happens later. Such a systematic bias would be surprising. Overall, therefore, the long-dated bonds that form our data ensure that the jumps we observe are related to the onset of war, and not simply to their timing.

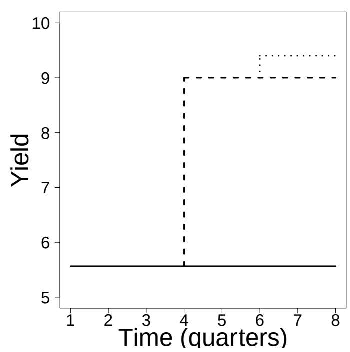


Figure A1. Simulated yield of a 10-year bond under different scenarios. Assumption: bond with face value 100 with quarterly payments of 6% coupons and 5% market rate. The yield is calculated as C/P, where C is the coupon and P is defined by Equation (1). We assume that coupons will not be paid after the onset of war, though similar results apply if we instead assume an market rate (i), or a default. Solid line: no war, and hence no change in the yield. Dashed line: at the beginning of year 1 (4^{th} quarter) investors correctly anticipate that a war will happen in one year (i.e., at $t = 8$). Dotted line: at the beginning of year 1 (4^{th} quarter) investors incorrectly anticipate that a war will happen in one year, when it really happens 6 months later.

A.2 Bond Yield Data

Bond yield data were obtained from Global Financial Data (`https://www.globalfinancialdata.com/`), which has data on government bonds from year 1520 onward, together with some indicators used here such as inflation rates and exchange rate regimes. The dataset is unfortunately not public but can be purchased and is often available through university libraries. The dataset was downloaded in June 2011. Table A1 below lists countries included in the data, together with the first year of entry in the dataset.

Table A1. Countries for which government bond yield data is available, and earliest year of record.

Country	First record	Country	First record
Argentina	1859	Netherlands	1816
Belgium	1834	New Zealand	1925
Brazil	1861	Norway	1963
Bulgaria	1993	Panama	1997
Canada	1920	Peru	1997
Chile	1839	Philippines	1997
Colombia	1899	Portugal	1851
Denmark	1880	Russia	1820
Egypt	1862	Singapore	1998
Finland	1987	South Africa	1920
France	1880	Spain	1850
Germany	1880	Sri Lanka	1951
Greece	1863	Switzerland	1899
Hungary	1997	Thailand	1979
Iceland	1993	Tunisia	1991
India	1947	Turkey	1997
Indonesia	1997	United Kingdom	1816
Ireland	1928	United States	1816
Italy	1862	Uruguay	1882
Kenya	1987	Venezuela	1914
Malaysia	1961	Zambia	1995
Mexico	1872	Zimbabwe	1965
Morocco	1996		

A.3 Mediation analysis

A concern in the section on 'Fluctuations around the onset' (see in particular p. 13) was that bond yields may jump even for a well-anticipated war. Indeed, yields may rise if the onset of conflict is associated with other changes such as an increase in central bank rates or news reports. The model above addresses some of this concern by including relevant control variables, but this could lead to underestimating the effect of the onset on yield, because 'controlling' is not sufficient to get to the causal effect. The idea is that some of the causal effect of onset on bonds is actually expressed through an effect on the central bank rate, which in turn is correlated with bond market rates (Fig. A2). This means that the effect we are estimating is likely to be underestimated if we include central bank rates as a simple control, as we do above.

To further isolate the causal path from war onset to the jump in bond yields, we therefore estimated a mediation model. Mediation models examine how X affects Y both directly and indirectly through one or more intervening variables ('mediators'). We follow in particular the technique put forward in the seminal article by Preacher & Hayes (2008), who recommend to estimate a Seemingly Unrelated Regression (SUR) model with separate equations for the mediating variables.[18] We therefore estimated the following model:

$$\begin{cases} \text{Yield}^* & = \mathbf{X_1}\boldsymbol{\beta_1} + \boldsymbol{\varepsilon_1} \\ \text{CBRATE}^* & = \mathbf{X_2}\boldsymbol{\beta_2} + \boldsymbol{\varepsilon_2} \\ \text{News}^* & = \mathbf{X_3}\boldsymbol{\beta_3} + \boldsymbol{\varepsilon_3} \end{cases}$$

Note that each equation has its own set of regressors (see Table A2). The SUR simply models the possible correlation in error terms across these equations. Standard errors, clustered by conflict, and confidence intervals for the direct and indirect effects, were obtained by bootstrapping (1,000 replications).

The results in Table A2 and Figure A2 show that the effect of the onset on yields is indeed mediated by the central bank's policy—though not by news reports. If we take into account this indirect effect, we find that the effect of a large war on yields is 1.2, rather than the 1.08 (1.12 − 0.04) that we obtain when *controlling* the for central bank rate. This means that the onset of war indeed affects more than the yield alone, but that in fact the total effect of the onset on yield is stronger than we estimated, since some of it goes through the change in central bank rate. This further strengthens our findings.

[18] See also Gatignon (2003, pp. 356–70)

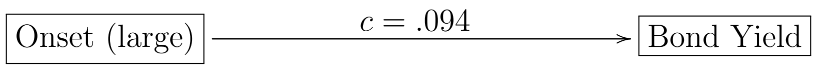
Table A2. SUR model.

	Yield	CBRATE
After	-0.028	-0.015
	(0.023)	(0.038)
After × Large	0.115^{**}	-0.007
	(0.031)	(0.014)
Fatalities (log)	0.033^{*}	-0.134^{**}
	(0.014)	(0.014)
Time to War (yrs)	-0.00014^{**}	-0.00022
	(1.4×10^{-5})	(2.2×10^{-5})
Time to War^2 (yrs)	$4.0 \times 10^{-8}{**}$	7.3×10^{-8}
	(1.5×10^{-8})	(4.5×10^{-8})
Time to War^3 (yrs)	9.3×10^{-11}	9.3×1.2^{-12}
	(1.6×10^{-11})	(2.3×10^{-12})
CBRATE*	0.522^{**}	
	(0.020)	
Inflation*	0.016	0.002
	(0.018)	(0.023)
Debt	0.058	-0.348^{**}
	(0.031)	(0.132)
GDPPC	0.016^{**}	0.025^{**}
	(0.004)	(0.008)
Polity	-0.0092	-0.014^{*}
	(0.006)	(0.007)
$W_i\text{Yield}_{j \in J,t}$	0.00062	
	(9.1×10^{-5})	
N	44,290	44,290
R^2	0.374	0.124

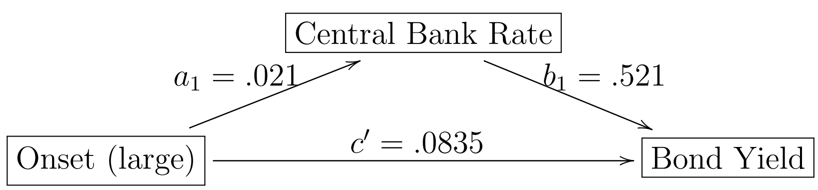
$^{**}p < 0.01$, $^{*}p < 0.05$

Standard errors clustered by country in parentheses (obtained by bootstrap)

Starred variables have been standardized over that period (i.e., $x^* = (x - \bar{x})/sd(x)$)



(a) Onset affects Bond Yield. The total effect is simply the direct effect c, estimated by OLS, including control variables (Table **??**, Model 2), but not controlling for Central Bank Rate and News in order to capture the total effect.



(b) Onset affects Bond Yield directly (c') and indirectly through the central bank rate ($a_1 b_1$). The total effect is the sum of the direct and indirect effects, i.e., $c' + a_1 b_1 \approx .094$. Estimation using Seemingly Unrelated Regression, Table A2.

Figure A2. Mediation design.

A.4 Bayesian Model Averaging

BMA models estimate models for potentially all possible combinations of the independent variables, and create a weighted average of these estimates. For K possible variables, there are 2^K possible variable combinations, and hence 2^K models to estimate. The models are then averaged using posterior model probabilities derived from Bayes' theorem, using both the researcher's prior and the model's posterior probability, which is the likelihood that the model is the right one given that we observe data D (i.e., posterior $= P(M_k|D)$).[19]

Here we address the concern that the results of section 'Which wars are surprising?' (p. 19) might be driven by our model selection. The results of the BMA are reported in Table A3 and Figures A3a and A3b. We note in particular from Table A3 that many of our variables have a very high posterior inclusion probability (PIP—right column).[20] Thus both Polity and Date have an inclusion probability at 100%. These results are not artefacts of the choice of priors. Figure A3b reports the posterior inclusion probability of each variable using binomial model priors and random (beta-binomial) priors, in addition to the common prior model probability that was assumed in Table A3 (i.e., our prior was that each model is equally likely to be the true one, with probability $p(M_k) = 2^{-K}$).

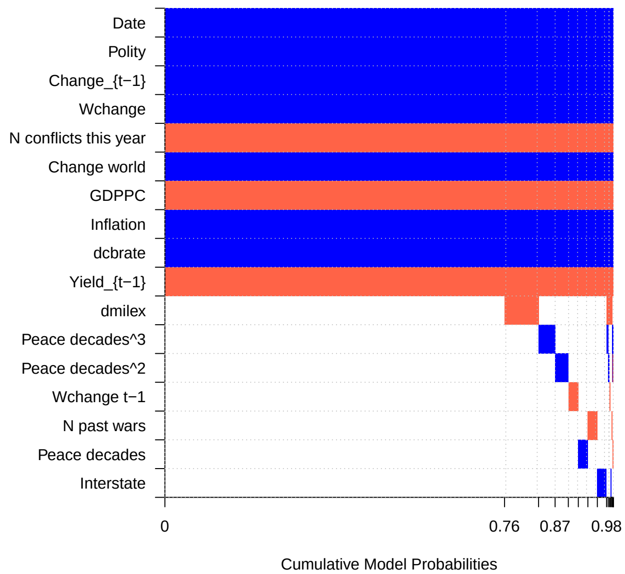
These results support our hypotheses. In particular, they show that the positive coefficient on Date is robust, which further suggests that financial markets continue to underestimate the probability of war. The coefficient on Polity is also positive, which suggests that markets in democracies are more likely to react with surprise to the onset than those in non-democracies. The variable Interstate conflict, on the other hand, is rarely included, in line with the insignificant results of Table 3. This further casts doubt on hypothesis 5.

[19]The posterior probability for model M_k is given by $P(M_k|D) = \frac{P(D|M_k)P(M_k)}{\sum_{i=1}^{K} P(D|M_i)P(M_i)}$.

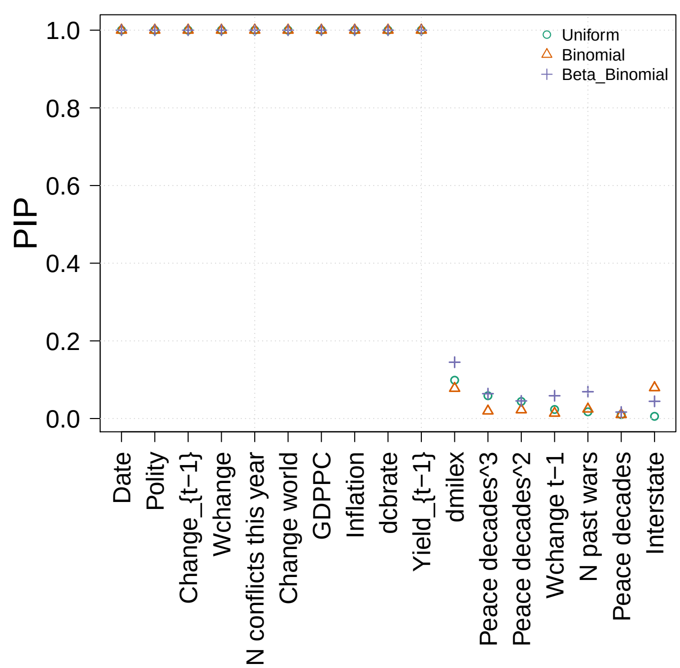
[20]The posterior inclusion probability (PIP) is the sum of posterior model probabilities for models in which a covariate was included.

Table A3. Bayesian Model Averaging. The first column, PIP (Posterior Inclusion Probabilities), reports the sum of PMPs for all models in which that variable was added. For example, all of posterior model mass relies on models that include *Date* as a covariate. The second column, 'post mean', lists the coefficients associated with a particular variable, averaged over all models.

	PIP	Post Mean	Post SD	Cond.Pos.Sign	Idx
Date	1	0.178	0.032	1	1
Polity	1	0.162	0.026	1	3
Change_{t-1}	1	0.200	0.024	1	4
$W_i$change	1	0.124	0.024	1	5
N conflicts this year	1	-0.118	0.024	0	7
Change world	1	0.257	0.024	1	9
GDPPC	1	-0.321	0.029	0	10
Inflation	1	0.197	0.029	1	11
ΔCBRATE	1	0.187	0.027	1	12
Yield_{t-1}	1	-0.273	0.037	0	13
Δmilex	0.099	-0.004	0.013	0	14
W_i Change_{t-1}	0.023	-0.0002	0.004	0	6
N past wars	0.017	-0.0001	0.004	0	8
Peace decades	0.010	0.00002	0.003	0.900	15
Interstate	0.006	0.00001	0.002	1	2



(a) Bayesian Model Averaging: Model inclusion based on best 500 models. The plot displays the 500 models with the highest posterior model probability (PMP). The plot gives a visual intuition for the variables that consistently appear in the best models.



(b) Posterior Inclusion Probability of each variable for different priors.

Figure A3. Bayesian Model Averaging

A.5 Additional Figures

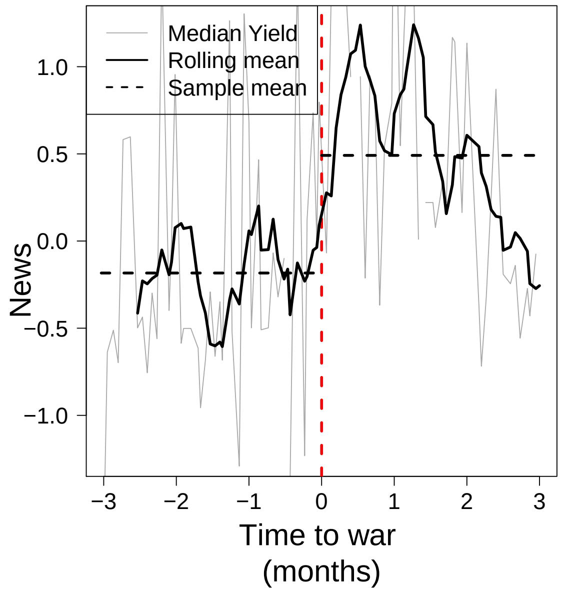


Figure A4. Median number of conflict-related news items prior to all large wars (time series are standardized as z-score with mean 0 and standard deviation 1 over the period). See Chadefaux (2014*b*) for details on the data.

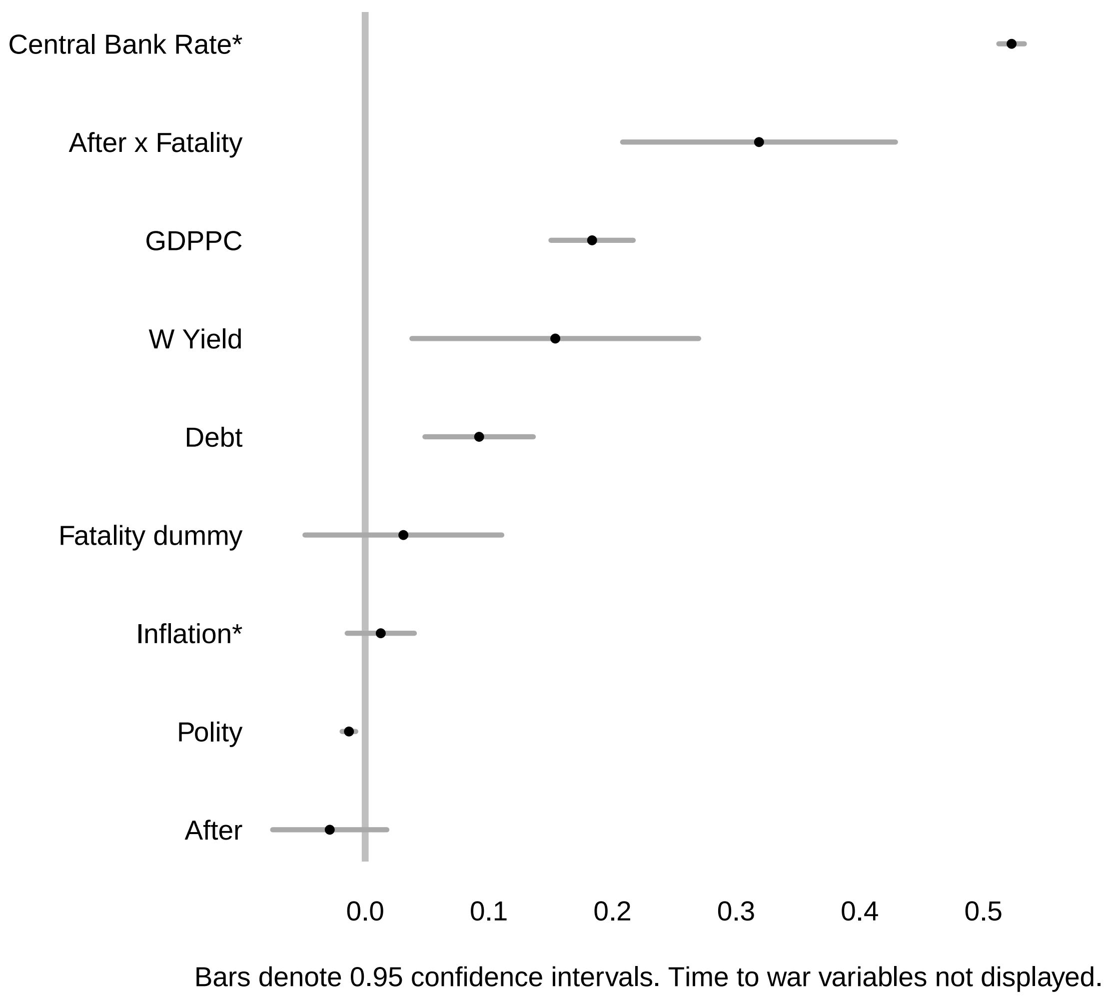


Figure A5. Standardised Coefficients. Plot corresponds to Model 2 of Table 2.

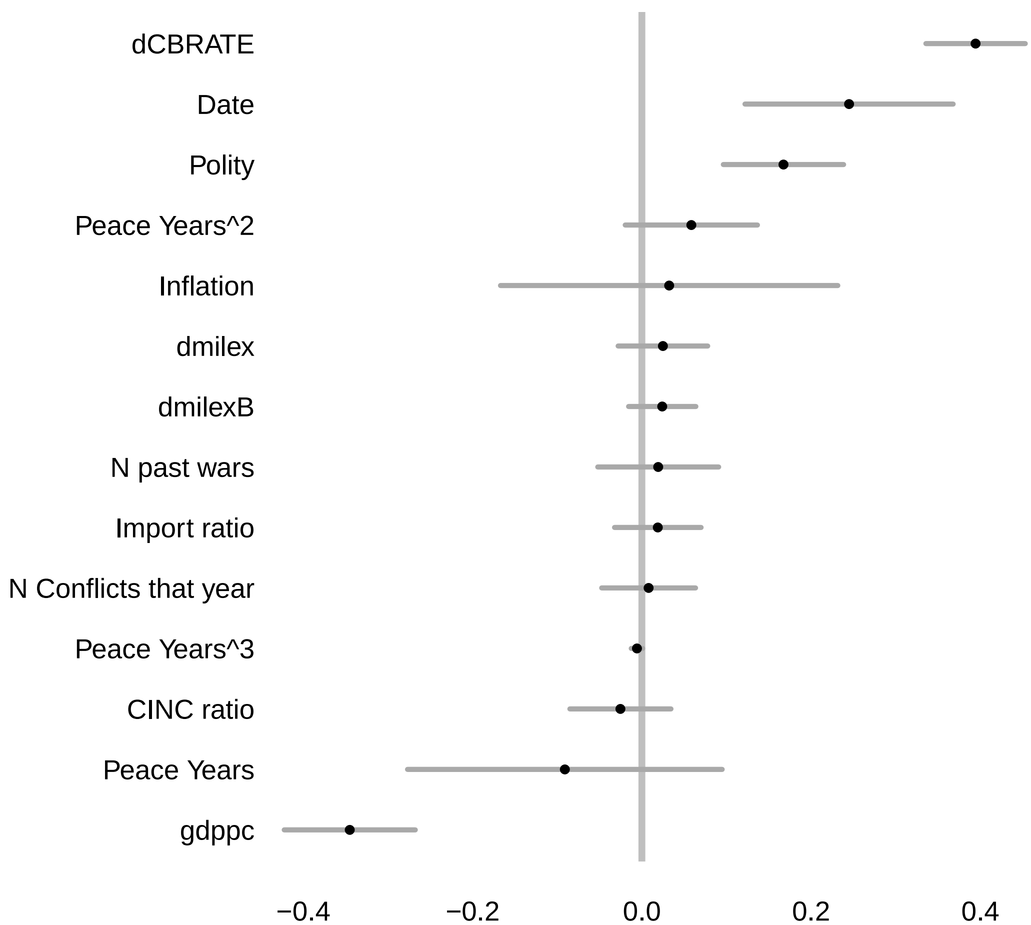


Figure A6. Standardized coefficients for Table 3. Bars denote 95% confidence intervals.

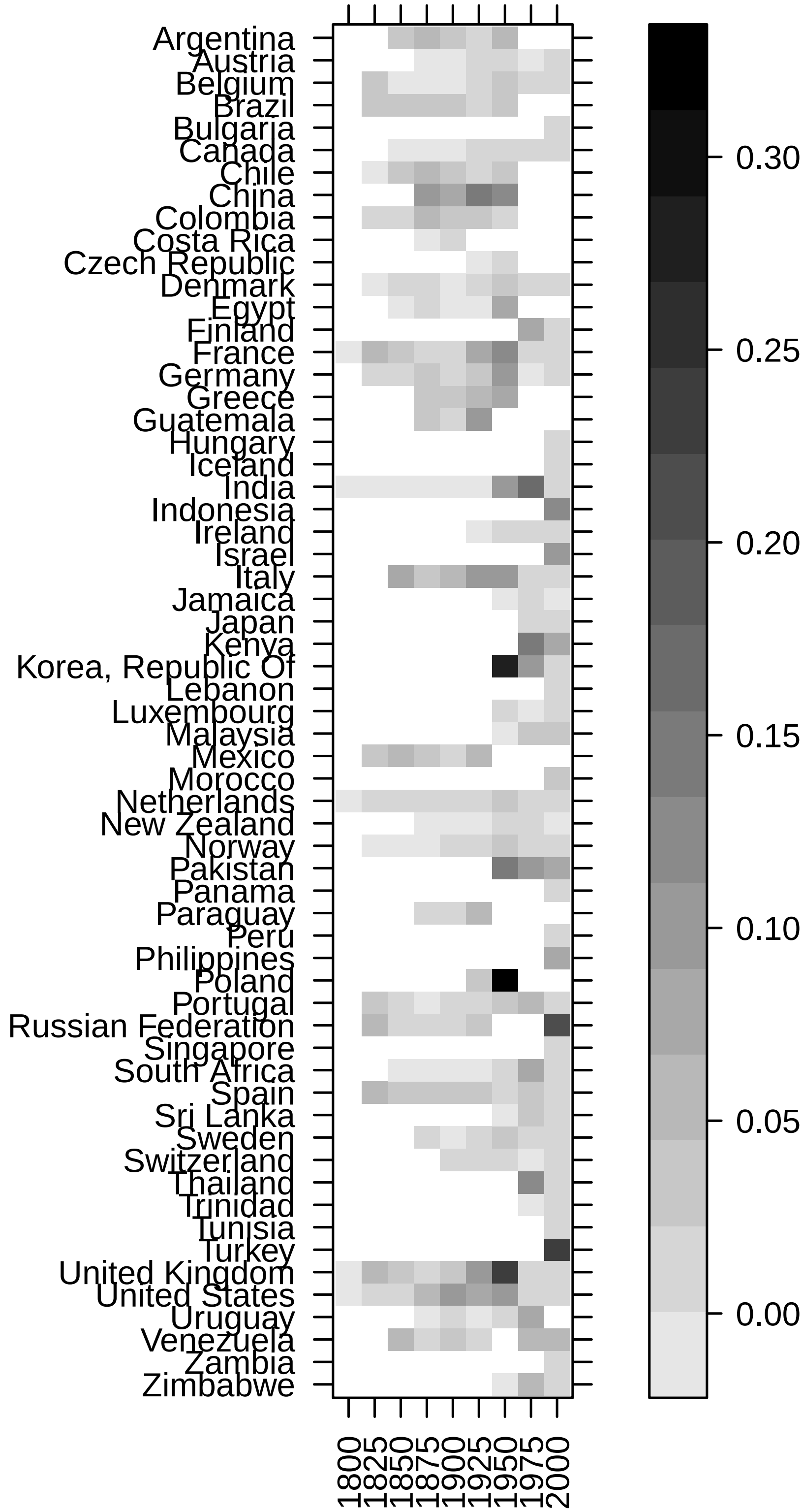


Figure A7. Incidence of conflict by country and time period. White indicates non-existing bond data.

A.6 Additional Tables

Table A4. Summary statistics for table 1.

Statistic	N	Mean	St. Dev.	Min	Max
War Onset$_{it}$	240,626	0.010	0.101	0	1
ΔYield$_{it}$ (log)	240,564	−0.0001	0.029	−3.730	2.842
ΔYield$_{world,t}$ (log)	240,565	−0.0001	0.133	−2.995	2.995
$W_i Yield_j$	240,626	0.0001	0.026	−4.094	8.494
ΔCBRATE$_{it}$ (log)	183,353	−0.0001	0.043	−2.510	2.262
ΔCPI$_{it}$ (log)	170,939	0.001	0.034	−0.232	6.949
GDPPC (log)	195,530	2.292	0.945	−0.803	3.445
Govt Debt	223,030	3.826	0.774	0.000	5.819
Polity	209,918	7.057	5.469	−10	10

Table A5. Summary statistics for table 2.

Statistic	N	Mean	St. Dev.	Min	Max
Yield$^{*}_{it}$	27,244	0.008	0.905	−3.437	3.560
After	27,244	0.499	0.500	0	1
Central Bank Rate*	27,244	−0.014	0.821	−1.769	5.722
Inflation*	27,244	−0.005	0.289	−4.941	4.469
Govt Debt	27,244	0.479	0.176	0.172	0.971
Polity	27,244	7.707	0.876	−6	10
GDPPC (log)	27,244	2.829	0.671	−0.335	3.370
Fatalities (Dispute, log)	22,145	0.068	0.492	0	7.197
Fatality dummy (dispute)	22,145	0.025	0.157	0	1
Fatalities Level (Incident)	27,244	0.355	0.717	0	2

Table A6. Summary statistics for table 3.

Variable	N	Mean	St. Dev.	Min	Max
Date	2,099	123.997	47.152	0.000	190.699
Polity	2,073	4.432	6.740	−10.000	10.000
CINC_i (log)	2,099	−3.432	1.657	−10.068	−0.964
CINC_j (log)	1,850	−4.540	2.188	−14.252	−0.349
N conflicts that year	2,099	20.253	15.397	1	65
$\Delta\overline{\text{Yield}}_{\text{World}}$	2,062	0.007	0.657	−12.731	4.835
GDPPC (log)	1,900	7.507	7.013	0.617	31.357
Inflation	1,897	1.368	14.977	−1.752	463.924
ΔCBRATE	1,686	−0.194	2.242	−40.000	13.480
$\overline{\text{Yield}}$ pre-onset	2,087	6.452	4.762	1.590	52.157
Δ Mil. expenditures	2,009	1,211,441.000	5,241,045.000	−44,866,016	56,365,000
Trade Ratio	1,343	0.477	0.161	0.000	1.000
CINC ratio	1,850	0.356	0.331	0.0001	0.998
Peace Decades	2,099	0.154	0.431	0.0003	8.193