Expert Survey of Conflict Risk between the Russian Federation and any EU state

Nils-Christian Bormann

Executive Summary

I present the result of an expert survey of political scientists, economists, historians, and area study experts at German Universities (n=210). The goal of the survey was to elicit a probability estimates of armed conflict risk between the Russian Federation and any member state of the European Union within the next five years. Data collection took place between 7 July and 10 September 2025 (i.e., before the Russian drone incursion into Polish airspace). The mean probability estimate of a fatal militarized interstate dispute is 36.37 per cent (median: 30 per cent). The mean probability of war is 26.38 per cent (median: 15 per cent). Political scientists report lower risks than members of other disciplines, as do professors relative to junior researchers. There are no meaningful differences between self-reported gender or preferred methodological approaches.

Conflict risk assessments

The survey asked two substantive questions:

1. How likely do you think it is (on a scale of 0 to 100) that Russia will launch a military attack resulting in fatalities on an EU member state in the next five years?

2. How likely do you think it is (on a scale of 0 to 100) that there will be a war with at least 1,000 fatalities between the Russian Federation and member states of the European Union in the next five years?

These questions capture standard political science conflict categories, a fatal militarized interstate dispute (MID) as measured by the Correlates of War (COW) project, and a high-intensity interstate war as defined by most large scale conflict data collection efforts (COW, Uppsala Conflict Data Program).

Figure 1 plots the mean probability estimates across all respondents (n=210) including 95 per cent confidence intervals. The overall expert assessment reveals mean estimates of 36.4 per

cent (CI: 32.9, 39.9) for the probability of a fatal MID and 26.4 per cent (CI: 23.1, 29.6) for the risk of a high-intensity war.
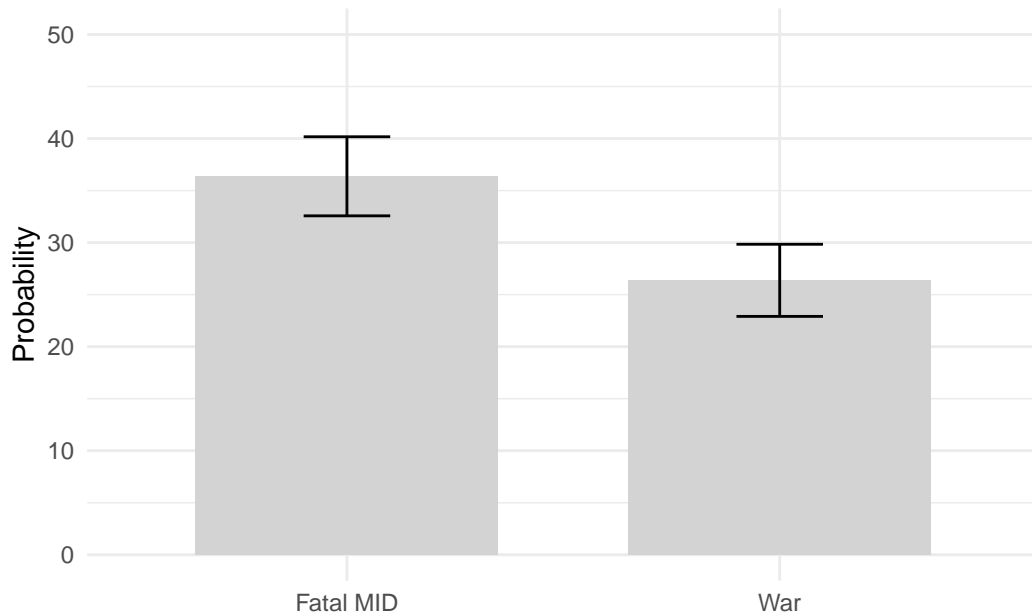

Figure 1: Mean probability assessments of conflict risk between Russia and any EU state by 210 social scientists.

The distributions are clearly tail-heavy and median estimates decrease to 30 per cent for a fatal MID and 15 per cent for interstate war respectively as Figure 2 indicates.

Yet even these median estimates surpass historical frequencies of fatal MIDs and high intensity war over the last two hundred years by several orders of magnitude.[1]

_____

[1]To simplify, I assumed that the annual probability estimate from the survey is simply the median estimate divided by the number of years. For fatal MIDs, this would be 6 per cent. I further assumed that the estimated probability includes all Russia-EU member state dyads, since Russia can be considered a great power. Simply dividing the overall probability by the number of EU member states is unlikely to result in an accurate estimate as individual dyads are not independent. For example, it is likely that a Russian attack on Lithuania will bring other EU member states into the conflict. Yet as a lower baseline estimate, this simple division serves well: the annual probability of conflict for any of the 27 dyads is then 0.22 per cent (6/27). The historical frequency of fatal MIDs is about 1 per year. Focusing on so-called relevant dyads, i.e. neighboring states and any dyad involving a great power, the baseline for potential fatal MIDs is about 1,200 today. The historical frequency of fatal MIDs is thus roughly .0008 per cent (1/1200). The calculation for the probability of wars is analogous, although the difference between expert estimates and the historical incidence is even greater.
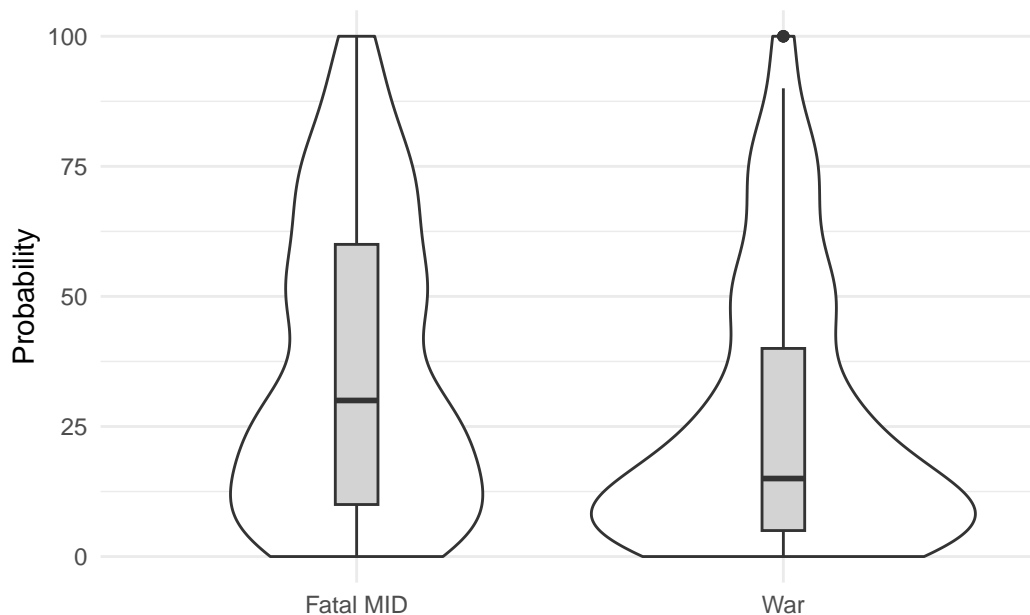
Figure 2: Median and distribution of conflict risk assessments between Russia and any EU state by 210 social scientists.

Heterogeneity of risk assessments

How credible are these assessments? Past research characterizes expert judgments as no more predictive of future events than random guesses (Tetlock, 2005). Yet studies critical of expert knowledge evaluate the accuracy of individual forecasts. This survey brings together a large number of experts, and thus draws on the wisdom of the crowd.

One way to assess the credibility of the assessments is to evaluate the spread of the answers, which is substantial (see Figure 2). The large variance of estimates reflects broad differences in individual and disciplinary risk assessments (Figure 3). Within disciplines the individual probability estimates of a fatal MID vary between 0 and 100 per cent for historians, political scientists, and economists. This large spread suggests different mental/theoretical models between individuals.

Across disciplines political scientists estimate the lowest risk of a fatal MID along with the lowest standard deviation, followed by economists, area specialists (though the number of responses is very small), those who did not disclose a disciplinary background, and historians. Notably, historians seem to be split into two camps, one suggesting a lower probability and another a higher probability of a fatal MID.

What could explain these differences? One plausible answer could be that political scientists, who most frequently study political violence, are familiar with the low historical incidence of armed conflicts and low-probability quantitative model predictions (Beck et al., 2000). They
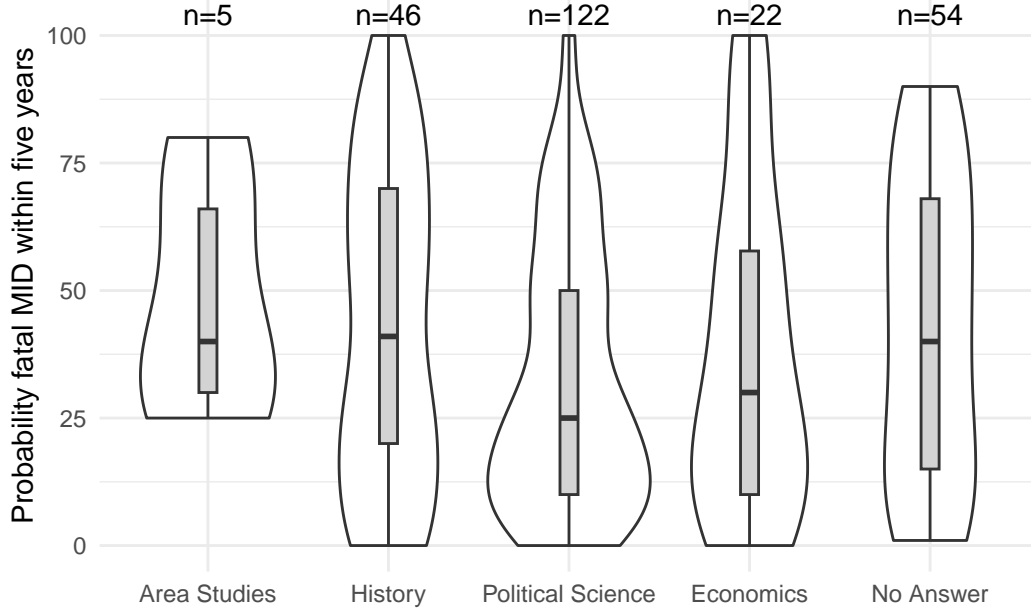
Figure 3: Median and distribution of conflict risk assessments between Russia and any EU state by disciplinary background.

thus appropriately include the low baseline risk in their forecasts. However, the absence of major observable differences between different methodological backgrounds sheds doubt on this interpretation (Figure 4). Arguably, political scientists using quantitative methods should be more familiar with the statistical predictions of large-N research, and therefore adjust their baseline accordingly. Yet the data do not reveal any difference between the broad methodological approaches used by survey participants.[2]

Finally, the survey reveals meaningful differences in risk assessments across career stages. Generally, researchers who were either Junior Professors or Professors estimated lower probabilities than early-career researchers (Figure 5). Whether this is a result of different age cohorts (socialization effects), or different theoretical or methodological priors is hard to say, not least because a plurality of respondents chose not to answer the question. The survey revealed no differences across reported gender in risk estimates.

Although the survey reveals systematic differences in conflict risk estimates across disciplines and career progression, the larger differences within disciplines and career groups suggests that much of the variation stems from unobserved characteristics such as different priors and different information/information sources. Diverse information backgrounds and sources across respondents would increase the credibility of the mean estimates. Different interpretations of

---

[2]Further distinguishing preferred methodological approaches into experimental, qualitative interpretative, qualitative observational, and quantitative observational approaches does not change this interpretation. Neither do the estimates of full scale war.
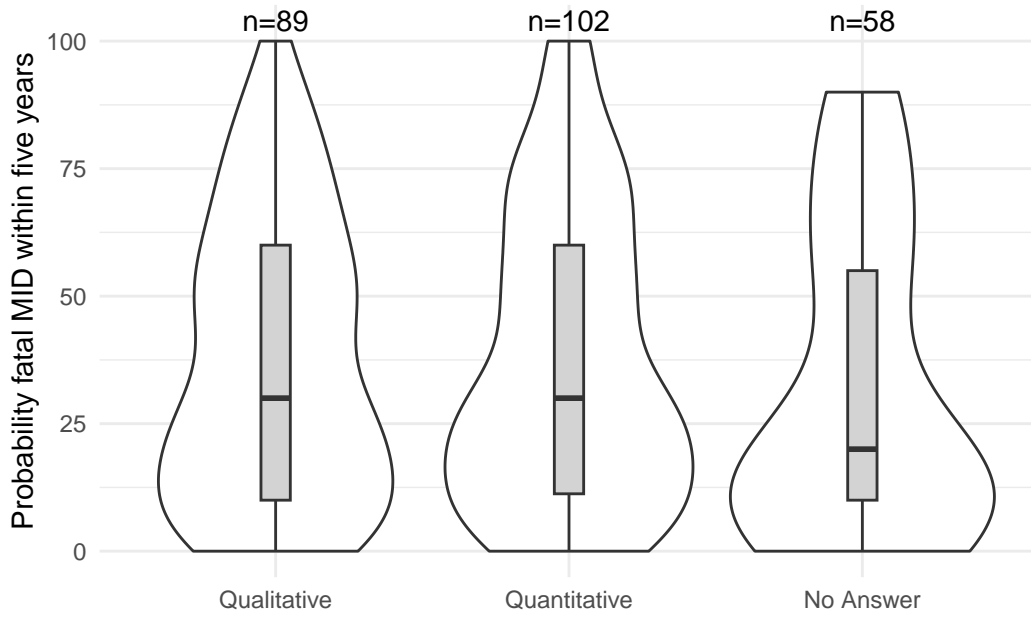
Figure 4: Median and distribution of conflict risk assessments between Russia and any EU state by preferred method.
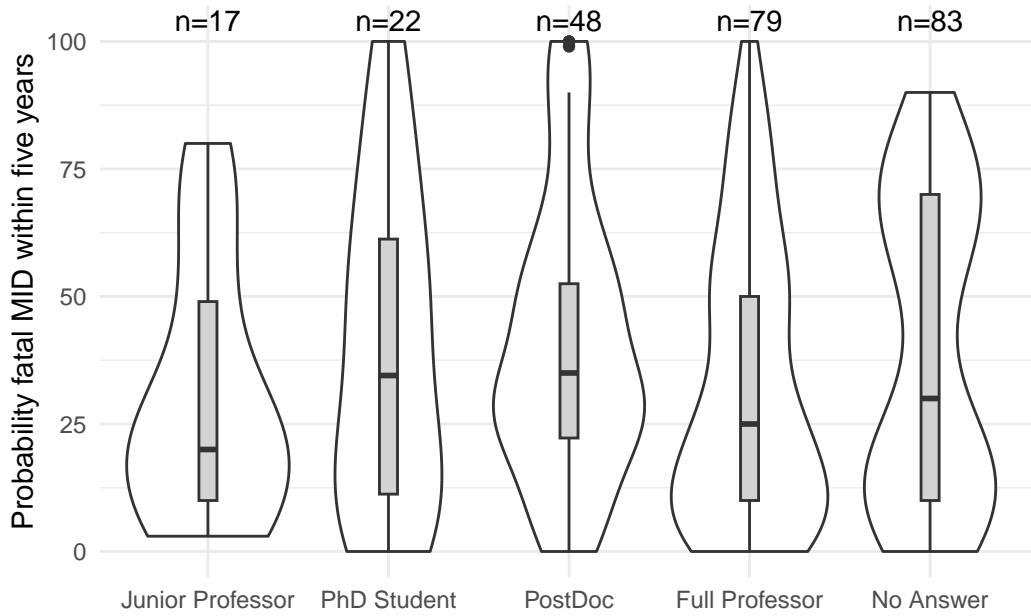

Figure 5: Median and distribution of conflict risk assessments between Russia and any EU state by career progression.

the same information across diverse theoretical backgrounds might also increase the credibility of the risk estimates, depending on the appropriateness of the theoretical models employed.

Statistical literacy

Another way to probe the credibility of the overall assessment is to evaluate whether individual predictions show characteristics that are known to result in good forecasts. If the average probability estimates marked by high-quality individual characteristics differs significantly from the average estimate that do not meet such quality criteria, the credibility of the forecasts would be lower.

One quality characteristic is a proper application of the laws of probability. The survey tested the understanding of one of these laws by adding a question about the conditional probability of a fatal MID/war in case the war between Russia and Ukraine would end at the front lines that were in place during the time the survey was conducted or more favorably for Russia. The laws of probability dictate that this conditional probability is at most as large as the overall probability of fatal MID/war risk.[3]
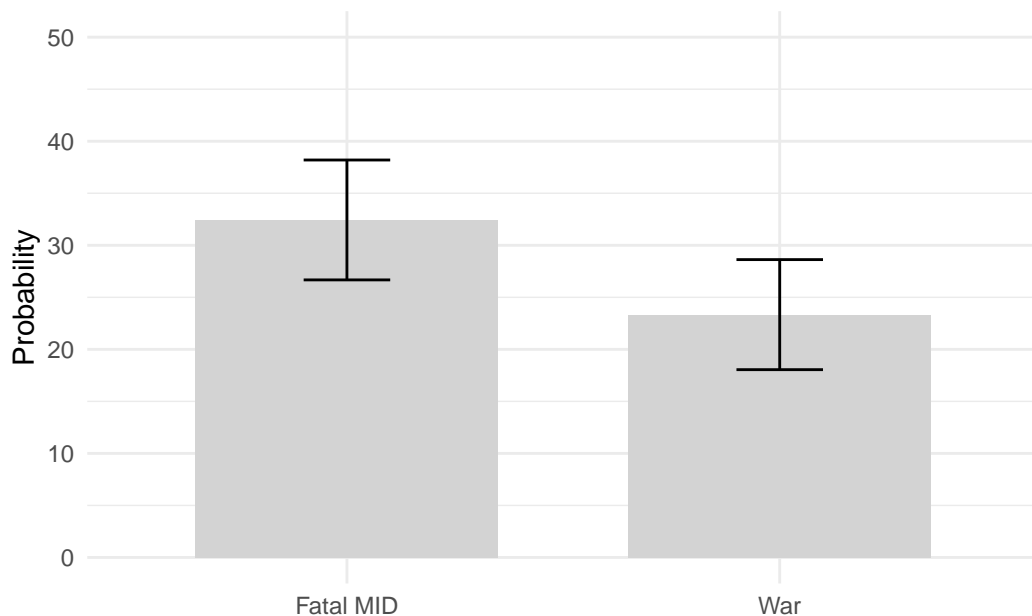
Figure 6: Mean conflict risk assessments between Russia and any EU state. Only statistically literate answers.

Slightly over half of all respondents (110/210~52.3 per cent) gave answers that respected this law. Figure 6 only plots these "statistically literate'' responses. Reassuringly, the average

---

[3]The latter also includes the risk of Russia-EU conflict in case the Russian/Ukraine war ends unfavorably for Russia.

conflict estimates do not differ in statistically significant ways from the averages of all respondents.[4]

Conclusion

The survey of social scientists at German universities revealed substantively higher average risks of a fatal militarized dispute/war between any EU state and the Russian Federation than predicted by the historical frequency of such events or predictions by quantitative models. The mean estimates of the likelihood of a fatal militarized interstate dispute over the next five years is slightly larger than one in three. The mean probability estimate of full-scale war is slightly larger than one in four.

The large number of responses and the diversity of backgrounds of respondents suggests that the estimates are informed by a variety of information sources and/or differences in theoretical/empirical priors. Yet systematic differences across disciplines and career progression do not form a consistent overall picture of what these differences in priors might be. In case future surveys are fielded, these should include questions about information sources and more specific information on theoretical backgrounds.

The utility of surveys like this will have to be measured against rival approaches, such as quantitative prediction models (Hegre et al., 2025; Schincariol et al., 2025). While these models show considerable promise in forecasting active conflict trajectories they struggle in predicting the onset of new violence. Prediction markets constitute a second alternative. However,I could not find any open questions on the likelihood of a military conflict between the Russian Federation and an EU state on common prediction markets at the time of writing. Expert surveys might be a useful, low-cost alternative. To better gauge their accuracy, the number of potential conflicts should be expanded to different geographic contexts (India-Pakistan, China-Taiwan, etc.) and the temporal period to be forecast should be reduced (months rather than years) to enable greater specificity in forecasts.

Finally, the goal of (public) forecasts of human behavior is not always accuracy but sometimes to change the trajectory of human interactions before major crises occur. Economists' forecasts of key economic statistics such as GDP growth, inflation, and unemployment show a decidedly mixed record in terms of accuracy. Nevertheless these surveys might provide important signals of the necessity of intervention to policy-makers. Political scientists might provide a similar public service on their topics of expertise.

---

[4]One sample t-tests that respect the conditional nature of the two distributions yield p-values of 0.1784 and 0.2568 for a fatal MID and interstate war respectively. Similarly, permutation tests yield non-statistical differences. Comparing the statistically literate and non-literate estimates yields a statistically significant difference for the fatal MID risk (p=0.0314) but not for war risk (p = 0.0852).

References

Beck, N., King, G., and Zeng, L. (2000). Improving quantitative studies of international conflict: A conjecture. *American Political science review*, 94(1):21–35.

Hegre, H., Vesco, P., Colaresi, M., Vestby, J., Timlick, A., Kazmi, N. S., Lindqvist-McGowan, A., Becker, F., Binetti, M., Bodentien, T., et al. (2025). The 2023/24 views prediction challenge: Predicting the number of fatalities in armed conflict, with uncertainty. *Journal of Peace Research.*

Schincariol, T., Frank, H., and Chadefaux, T. (2025). Accounting for variability in conflict dynamics: A pattern-based predictive model. *Journal of Peace Research*, page 00223433251330790.

Tetlock, P. (2005). *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press, Princeton, NJ.