

# **Narratives as Separating Equilibria**

## **On the Origins of the Ukraine War**

Giacomo Corneo

School of Business & Economics

Discussion Paper

Economics

2026/4

Narratives as Separating Equilibria:  
On the Origins of the Ukraine War

Giacomo Corneo \*

January 2026

---

\*I am grateful to Luca Beltrametti, Helmut Bester and Giuseppe Russo for helpful comments. Affiliations of author: Department of Economics, Freie Universität Berlin, CEPR, London, CESifo, Munich, IZA, Bonn.

Address: Department of Economics, FU Berlin, Boltzmannstr. 20, 14195 Berlin, Germany. E-mail: [giacomo.corneo@fu-berlin.de](mailto:giacomo.corneo@fu-berlin.de).

## **Abstract**

NATO enlargement and Russian annexation of Crimea marked crucial turning points. According to one narrative, the Russian occupation was part of a plan to re-establish dominion over Eastern Europe. According to a rival view, it was an attempt to counter a U.S. plan to subjugate Russia. I scrutinize the logical requirements of those narratives in a multi-stage game of incomplete information that produces equilibrium play such that first NATO is enlarged and then Russia attacks Ukraine. The two competing narratives correspond to two different separating equilibria. Conditions for their existence inform about the consistency and plausibility of the associated narratives.

*Keywords:* Ukraine war, NATO enlargement, incomplete information.

*JEL-Classification:* H8, N4, Z1.

# 1 Introduction

NATO enlargement and Russian annexation of Crimea were milestones in the intricate process that eventually led to the current war in Ukraine. Two narratives stand out. According to the first one, the Russian occupation of Ukrainian territory was part of a drive to re-establish dominion over some part of Eastern Europe. Obviously, Russia is the villain (ORIV narrative). According to the second and opposite narrative, that flagrant violation of international law was a legitimate attempt to counter an American plan to subjugate Russia. Obviously, the U.S. is the villain (OUIV narrative). While the ORIV narrative is the prevailing wisdom in Western political discourse and media, the OUIV narrative is propagated by the Kremlin. Scholarly endorsement has been provided to both views. Recent examples are, respectively, Smith (2024) and Palley (2024). According to Smith (2024, p.507), “Putin cares about NATO expansion not because NATO is a threat to Russia, but because NATO will prevent him from acting as the aggressor in the future, effectively closing the door on his imperialist ambitions.” By contrast, according to Palley (2024, p.21-22), “[t]he Neocon objective is US global hegemony, and that objective has driven both eastward expansion of NATO and interference in former Soviet Republics ... it became a bi-partisan US consensus. ... The short-term plan was NATO expansion; the medium-term plan was turning Ukraine against Russia and detaching it from Russia; the long-term plan was dismembering Russia.”

These two narratives are not merely of scholarly interest; they entail major, and opposite, policy implications.<sup>1</sup> I scrutinize the logical requirements of these narratives in a multi-stage game of incomplete information that produces equilibrium play such that first NATO is enlarged and then Russia attacks Ukraine. Several scholars have pointed out that incomplete information about Russian motivations was a main feature of the policy debate since the early 1990s, at the time the Clinton administration launched NATO expansion. By way of an example, in a September 1994 memorandum for the U.S. Defense Secretary, his adviser and deputy secretary remarked “that NATO expansion will, when it occurs, by definition be punishment, or “neo-containment” of the bad Bear.” (Sarotte, 2019, p.29). Qualifying the Bear as bad reveals the possibility of a good Bear.

---

<sup>1</sup>The spectrum of scholar interpretations of the origins of the Russian attack includes approaches that are disregarded in this paper. Among them, the constructivist approach claims that Russia’s self-understanding as a great power, recreated through habits and discursive practices in society, explains its expansionism. Other approaches focus on Putin’s personality and his need to draw voters’ attention away from political failures at home. See Götz (2017) for a critical survey and Rosa and Benati (2025) for an update and extension.

In the model of this paper, the “bad Bear” is one type of Russia and the “good Bear” is the other type. The U.S. can similarly be an imperialist type or a status quo type. Types are private information. In the model, the two competing narratives correspond to two different separating equilibria. The ORIV narrative obtains as an equilibrium that reveals that Russia is the bad Bear. The OUIV narrative corresponds to an equilibrium that reveals that the U.S. is greedy. Conditions for existence of these equilibria inform about the consistency and plausibility of the associated narrative.

Examining the consistency of narratives about the origin of the Ukraine war has potentially far-reaching implications for Europe. If ORIV is correct and alternative explanations of the Ukraine war can be discarded, Russia is greedy for sure. Then, the only effective way to stop it from satisfying its imperialist appetite is to change the balance of military power in favor of Europe. Diplomacy only makes sense after that Europe has attained a position of force. This conclusion would confirm that current EU foreign policy is on the right track. Otherwise, current EU foreign policy toward Russia may have to be thoroughly rethought.

The main insight delivered by my analysis is that both the ORIV narrative and the OUIV narrative are consistent under quite stringent conditions. Consistency of the ORIV narrative does not require that the prior belief of the bad Bear be large. It requires that the prior of an American status quo type be large and that this type be very dovish. Consistency of the OUIV narrative does not require that the prior belief in American imperialism be large. It requires that the prior of the good Bear be relatively large and that such a prior could already be formed in the early 1990s, when NATO enlargement was announced. As argued below, it is far from evident that those logical requirements are actually met. In the light of the analysis in this paper, neither the ORIV nor the OUIV narrative looks particularly plausible. By implication, NATO expansion and Russian attack on Ukraine are more plausibly explained in terms of a pooling equilibrium.

The next Section illustrates how this paper contributes to the existing literature. Section 3 succinctly chronicles the relationship between the U.S. and Russia in the wake of the end of the Cold War. Section 4 develops and analyzes the baseline model. It shows that separating equilibria that correspond to the two competing narratives may exist. The conditions for existence depend on the prior beliefs about types in a somewhat surprising way. It also shows that a pooling equilibrium may exist that contradicts both narratives. In Sections 5, 6, and 7, extensions of the model are considered that qualify the insights from the baseline model. As prior beliefs play a crucial role in determining the existence of different equilibria, Section 8 offers some remarks about how they might be gauged. Section 9 concludes. The proofs of all Propositions are relegated in the Appendix.

## 2 Contribution to the literature

This paper follows the game-theoretic literature on international relations that has produced various analyses of war outbreak under conditions of two-sided asymmetric information. This literature has recently been surveyed and extended by Acemoglu and Wolitzky (2024). As in that literature, I adopt the unitary-actor approach, in which the players are states with well-defined preferences that may be their private information. Instead of using a canonical model, I propose a simple extensive-form game that is tailored to the particular circumstances that have embedded the relationship between the U.S. and Russia within the European strategic context as it arose after the end of the Cold War. The classical themes of the Security Dilemma and the Spiral Model - trust and mistrust, reassuring and provocative actions - drive the results obtained in the current paper. Its contribution is to apply those notions to the evaluation of the two most prominent narratives about the origins of the Ukraine war.

Closely related to the current paper is Kydd (2001) who investigates a game of NATO enlargement in which this policy option entails a dilemma between building trust and breeding mistrust. In Kydd's model, the goal of NATO enlargement for a benevolent West is to foster democratization and trust among the new Eastern European members. However, the West could also be mean, which explains why NATO enlargement can be provocative for Russia. As a possible way out of the dilemma, Kydd introduces an intermediate policy option: making NATO membership conditional on a potential new member having cooperated with the other Eastern European countries ("conditional guarantees"). Hence, in addition to Russia and the West, his model depicts potential Eastern European new members that play a multilateral trust game among them. The main conclusion is that conditional guarantees might, under some rather stringent conditions, produce a reassuring effect on Russia and avoid conflict escalation.

While Kydd (2001) was written between the first and the second wave of NATO enlargement and looked forward to future developments, the current paper is backward-looking, trying to explain observed facts as equilibrium play, so as to evaluate the consistency of prevailing interpretations. My model differs from Kydd's one in two main respects. First, I posit that Russia may be expansionist and NATO enlargement may work as a device to prevent Russian expansionism in Eastern Europe if Russia is the bad Bear. While this possibility is absent from Kydd's model, it is the main tenet of the ORIV narrative. Second, I neglect the issue of conditional guarantees to new members. Reiter (2001) argues that the whole idea that membership into NATO promotes democracy is a rather fanciful one. Furthermore, as remarked by Sarotte (2019), NATO extension basically was unconditional.

My model complements recent work that has used game theory in order to illuminate key aspects of the full-scale Russian invasion of Ukraine in 2022. Smith (2024) applies a model with lack of commitment but full information in order to explain Ukrainian and Russian moves in the lead-up to the 2022 invasion. Idrisoglu and Spaniel (2024) discuss the role of asymmetric information in causing a bargaining breakdown between Russia and Ukraine just before the 2022 invasion. Rosa and Benati (2025) compare the 2014 and 2022 invasions, viewed as distinct static three-player games with different payoff functions.

### **3 Fatal years 1994 and 2014**

In order to motivate the multi-stage game developed in the next section, this one succinctly chronicles the relationship between the U.S. and Russia in the wake of the end of the Cold War. It concentrates on those aspects that the proposed game aims to capture. Clearly, it is impossible to do justice to the complexity of the process that led to the Ukraine war within the limits of this paper.<sup>2</sup>

#### **3.1 From the Fall of the Berlin Wall to the decision to expand NATO**

The dissolution of the Warsaw Pact and the USSR in 1991 meant that the U.S. was the only global superpower and the aim for which NATO had been created (defense against a possible Soviet military attack) was achieved. At that juncture, a military conflict between the U.S. and Russia was unthinkable. The Russian political leaders Gorbachev and Yeltsin had friendly relationships with the U.S. government; the Russian economic and political crises were so serious that no capability to conduct a war against the West existed. For its part, the U.S. enjoyed an extraordinary prestige as the incontestable great power that won the Cold War; thanks to its economic and ideological supremacy the U.S. was expected to succeed in steering the countries from the former communist bloc into the civilized world of liberal democracy. Under such circumstances, the verbal promises made to Gorbachev by various Western leaders in 1989 and 1990 that NATO would not move eastward could be believed as genuinely intended.

Given that the structures of NATO inherited from the Cold War had lost their *raison d'être*, a lively debate began about how to reform the European security architecture. Among various proposals, the following five were prominent (Sarotte, 2019, 2021). First, some peace activists argued that Central and Eastern Europe (CEE) should be transformed into a demilitarized and neutral zone. A second proposal, notably advertised by

---

<sup>2</sup>See e.g. Palley (2024) and Smith (2024) for detailed accounts.

Gorbachev and Mitterrand, was to build a Pan-European security organization. A third one, especially favored by the Pentagon, was called Partnership for Peace (PfP). It would have created a comprehensive military cooperation between NATO and several states that did not belong to the Alliance, including Russia. Related to it was a fourth idea: an enlargement of NATO which should begin with the inclusion of Russia. This was proposed by Yeltsin in January 1994. Finally, the idea circulated to expand NATO, but without Russia. Central European political leaders Walesa and Havel propagated this approach.

The year 1994 turned out to be a fatal one. In January, NATO announced at its summit that the PfP would be launched. But, as made clear by Clinton, PfP would prepare some states for future NATO membership. It is in this context that Yeltsin claimed that Russia had to be the first new country to join NATO. Things turned out very differently. Shortly after the final withdrawal of all Russian troops from Germany, the German Minister of Defense R uhe declared in September 1994 that Russia could not be integrated in NATO.<sup>3</sup> This position of the West became crystal clear by December, finally leading to diplomatic clashes at the Budapest summit of the CSCE/OSCE, where Yeltsin vehemently accused the U.S. to greatly harm its relations with Russia by enlarging NATO only to CEE.<sup>4</sup> In February of the following year, the U.S. Congress passed legislation to assist CEE-countries in their transition to NATO membership.

Summing up, by the end of 1994 a major turning point had occurred in the relationship between the U.S. and Russia. Among all discussed options concerning the security architecture of Europe, the U.S. government had chosen the one least preferred by Moscow, entailing an indefinite expansion of NATO (“open-door policy”) with the exclusion of Russia - at least for the foreseeable future. This course was to be maintained over the subsequent two decades and beyond.

### **3.2 From NATO expansion to the 2014 Russian invasion**

In the successive years, but before the Russian annexation of Crimea, the following twelve states became full members of NATO: Hungary, Poland, Czech Republic (members since 1999), Bulgaria, Romania, Slovakia, Slovenia, the Baltic States (members since 2004), Albania, and Croatia (members since 2009). In 2005, NATO launched its Intensified Dialogue program with Ukraine on the latter’s aspirations to membership. In 2008, at its Bucharest summit, NATO promised Ukraine (and Georgia) membership at some undefined point in the future. Ukraine shared a border with Russia of about 2,000 km, with a minimal distance to Moscow of about 430 km, and was home to an outstanding

---

<sup>3</sup>Russia could not be integrated in the EU either, see Sarotte (2019, p.29).

<sup>4</sup>One week later, Russia launched a war against secessionist Chechnya.

component of the Russian navy: its Sevastopol base.<sup>5</sup>

NATO expansion heavily affected the relative military power of Russia by substantially increasing the disposable resources of the U.S. in terms of territory, population, and armed forces. At the same time, the Russian strategic position had been further compromised by technological developments, especially those pertaining to nuclear warfare. As recognized by nuclear planners, by the early 2000s the U.S. had virtually acquired the ability to destroy the entire Russian nuclear arsenal by a nuclear first strike (Lieber and Press, 2006). U.S. withdrawal from the ABM treaty with Russia in 2002 opened the way for further deepening of nuclear unbalance by reducing the effectiveness of a potential Russian retaliatory strike.<sup>6</sup>

It is in this evolving strategic context that the relationship between the U.S. and Russia experienced its second major turnaround. The immediate events leading to the Russian occupation of Crimea in 2014 are relatively well-known. In November 2013, the Ukrainian president Yanukovich rejected the terms of the EU Association Agreement, which included significant barriers to trade with Russia and a military cooperation between Ukraine and the EU. The subsequent Euromaidan mass protests in Kyiv precipitated the country in a deep political crisis. On February 18-20, 2014, a massacre of protesters and police forces occurred at the Maidan. Shortly after midnight of February 20, Yanukovich and the parliamentary opposition agreed on a deal brokered by Russia and the EU. It included a constitutional reform, the creation of a national-unity government, and the agreement to hold early presidential elections in December of the same year. On the evening of February 21, that deal was rejected by pivotal protesters of the Maidan, who rather doubled down by issuing a public ultimatum for Yanukovich to resign by 10 a.m. the next day. In the night, Yanukovich fled, leading to a constitutional crisis. On February 27, a new Ukrainian government was inaugurated that included the far-right Svoboda Party. On the same day, Russia invaded Ukraine in Crimea, proclaiming its annexation in March. That invasion and the insurgency in the Donbass started a military conflict that dramatically escalated with the full-scale Russian invasion of February 2022.<sup>7</sup> In that conflict, the military and financial support from the U.S. and its allies has been quintessential to enable Ukraine to resist and fight back the Russian attack.

---

<sup>5</sup>As compared to the Baltic states, the length of the common border is significantly larger and the distance to Moscow significantly shorter.

<sup>6</sup>Ellsberg (2017, ch. 20) recounts several episodes during the Cold War in which the U.S. successfully employed the threat of a nuclear first strike to obtain substantial concessions in various disputes with, respectively, the Soviet Union, China, and Vietnam. Debs (2025) offers a game-theoretic model showing that a more favorable nuclear balance can improve the terms of peace, but also increases nuclear risks.

<sup>7</sup>For an analysis of the events ranging from the occupation of Crimea in 2014 to the full-scale invasion in 2022, see e.g. Katchanovski (2025) and Rosa and Benati (2025).

### 3.3 Counterfactuals and their narratives

How would have Russia behaved, had the U.S. refrained from expanding NATO and opted for a different reform scenario, e.g. one of the four other options mentioned above? How would have the U.S. behaved, had Russia refrained from attacking Ukraine and accepted its integration into NATO? Two competing views on these counterfactuals inform the prevailing narratives about the Ukraine war.

According to the ORIV narrative, the Russian invasion of Ukraine shows that, had the U.S. refrained to expand NATO eastward and instead granted to Russia a similar level of cooperation as the one existing toward its European allies, Russia would have sooner or later revived its imperial ambitions. Proponents of this view stress that craving for an empire is part of the Russian DNA because: “In contrast to Western states, ... in Russia nation-building and empire-building proceeded concurrently” (Pipes, 1997, p.68). Most likely, Russia would have exploited Western cooperation in economic, technological and military domains to upgrade its ability to threaten CEE so as to transform those countries into vassal states. At a minimum, without being held in check by a closely positioned NATO, Moscow would have reduced the states that formerly belonged to the Soviet Union to the status of clients. At a maximum, Russia would have gained the status of a world superpower and peer competitor of the United States.

This first narrative typically includes a complementary view about the second counterfactual. Had Russia refrained from attacking Ukraine and accepted its integration into NATO, peace and prosperity would have been the most likely result. Moscow’s acquiescence would have inspired trust and prompted a benevolent attitude from the U.S. and its allies. Being obvious that Russia posed no threat to the rest of Europe, the West would have significantly intensified cooperation with Moscow, deepening economic, political and cultural ties to their mutual advantage.

The OUIV narrative interprets the main facts in a diametrically opposed way. Had the U.S. refrained to expand NATO eastward and instead granted to Russia a similar level of cooperation as the one existing toward the European allies of the U.S., Russia would have likely become a pretty normal Western country: a market economy with its own welfare state, embedded in a liberal-democratic political system, rejecting war as a means to resolve international disputes. The immense death toll suffered by Russia in two world wars had taught its citizens that such a war should never be risked again; the Russian people craved for peace and stability.

This second narrative addresses the second counterfactual by observing that NATO expansion to several CEE-countries, in combination with the promise of future Ukraine membership, definitely convinced Moscow that the U.S. did not aim at cooperation with

Russia, but at its encirclement and eventual subjugation. Had Russia refrained from attacking Ukraine and accepted its integration into NATO, Georgia and Belarus would have likely followed. In any case, having the possibility to station troops and install strategic weapons at the Russian border, the U.S. would have used its military supremacy in order to make a client state out of Russia. The CIA would have launched covert operations backed by nearby established military bases to encourage a secession of selected Russian Republics; the U.S. government would have exploited its nuclear primacy to coerce Moscow into accepting the transition toward a “loosely confederated Russia – composed of a European Russia, a Siberian Republic, and a Far Eastern Republic” (Brzezinski, 1997, p.56).

## 4 A multi-stage model

This section offers a model that helps to evaluate those two competing narratives about the origins of the Ukraine war.

### 4.1 Assumptions

Two states,  $U(SA)$  and  $R(ussia)$ , play a multi-stage game with the sequence of events depicted in Figure 1. As the game is one with incomplete information, that figure does not portray the game tree but serves for illustration. First,  $U$  decides whether to cooperate with  $R$  or expand NATO eastward. The decision to cooperate may be interpreted as an American invitation to the Russians to jointly design the novel security architecture of Europe. The decision to expand NATO stands for its gradual but unlimited enlargement: the American will to progressively integrate all CEE countries (with the exception of Russia) into NATO. In the second stage,  $R$  decides whether to cooperate with  $U$  or wage war. If  $R$  cooperates after that  $U$  cooperated, the game ends with mutual cooperation. If  $R$  cooperates after that  $U$  began to expand NATO,  $U$  decides whether to cooperate or to wage war. If  $U$  cooperates, the game ends with mutual cooperation. If  $U$  wages war, it does it after the completion of NATO expansion, which has Ukraine in NATO. Wars always end the game.

FIGURE 1 ABOUT HERE

Both states' payoffs in case of mutual cooperation are normalized to zero. In case of mutual cooperation between  $U$  and  $R$ , NATO has a negligible impact on their payoffs. NATO has a nonnegligible impact in the case of war. There are three possible wars in the sequence of events depicted in Fig. 1. War 1 is waged by  $R$  after that it has

recovered thanks to Western cooperation. War 2 is waged by  $R$  after the beginning of NATO expansion, but before its completion. War 3 is waged by  $U$  after that NATO expansion eastward has been completed. In War 1, NATO only includes the states that were members of the Alliance when the Cold War ended. In War 2, NATO is as it was in 2014, including various CEE-countries but not Ukraine. In War 3, NATO includes Ukraine. Wars are decisive and the probability of victory is monotone in the index of the war. Let us denote by  $q_j$  the probability for  $U$  to win war  $j$ ,  $j \in \{1, 2, 3\}$ . Probability  $q_j$  is strictly increasing in  $j$ .

War outcomes may be interpreted as follows. If  $U$  wins,  $R$  falls apart and  $U$  exerts hegemony over the resulting states. If  $R$  wins, it exerts hegemony over some countries of CEE, while  $U$  loses its pre-existing hegemony there. As a by-product, defeat by any side also entails an international loss of reputation for its military and intelligence, which erodes the power of the defeated state over regions of the world where it exerts some hegemony. In addition, each war entails some costs in terms of casualties, destruction, and unproductive use of scarce resources. The expected payoffs from war are type-dependent. Each player has two possible types  $t$ : greedy (the player is a revisionist state) and security-seeker (the player is a status quo state), denoted by  $t \in \{g, s\}$ . Each player has private information about own type. The prior probability for  $U$  ( $R$ ) to be greedy is denoted by  $p$  ( $\pi$ ), with  $0 < p < 1$  ( $0 < \pi < 1$ ). The payoff of  $U$  ( $R$ ) in case of victory is denoted by  $m_t$  ( $\mu_t$ ). The payoff of  $U$  ( $R$ ) in case of defeat is denoted by  $-c_t$  ( $-\gamma_t$ ). Thus, the expected payoff for  $U$  in case of war  $j$  is:  $q_j m_t - (1 - q_j) c_t$ . The analog of this expression gives the expected payoff for  $R$  in case of war.

Types differ in their valuation of the benefits from hegemony relative to the costs of war. Their preferences are defined so that in a scenario without NATO expansion, the greedy types have a preference for war, whereas the security types prefer cooperation. Since the payoff of cooperation was normalized to zero, this implies, for the greedy types of the two states:

$$q_1 m_g > (1 - q_1) c_g, \tag{1}$$

for  $U$  and

$$(1 - q_1) \mu_g > q_1 \gamma_g, \tag{2}$$

for  $R$ , with reversed inequality sign for their respective security type,  $t = s$ .

It may be noted that each player's payoff and the probability of winning do not directly depend on the opponent's type - but on the opponent's strategy, which can be type-dependent. This greatly simplifies the analysis at little loss of realism because it

avoids having to introduce conditions on off-equilibrium beliefs. For instance, condition (2) and its twin for the security type immediately determine equilibrium behavior for Russia in the case of no NATO expansion, independently of Russian beliefs about the American type.

Players' behavior is required to build a Perfect Bayesian Equilibrium (PBE). This means that strategies satisfy sequential rationality and posterior beliefs satisfy Bayes' Rule whenever possible. The aim of the analysis is to reveal under which conditions the observed actions (NATO expansion followed by Russian attack, i.e. War 2 in the game) can be rationalized as equilibrium play in a PBE.

## 4.2 Narratives and equilibria

If in equilibrium a type separates itself at a given information set, it perfectly reveals the type of the player. If types pool, no new information is conveyed and the posterior belief equals the prior. We are interested in equilibrium play that has at least one type of  $U$  choosing NATO expansion in the first stage and at least one type of  $R$  choosing war in the second stage. Refer to Figure 1; there are four candidate equilibrium configurations that generate the U1-R2 path that leads to War 2:

P-S: types *pool* at the information set that corresponds to U1 (by expanding NATO) and *separate* at the information set that corresponds to R2, with only the greedy type waging war (which implies that  $R$  is greedy if War 2 is observed);

S-P: types *separate* at the information set U1, with only the greedy type expanding NATO (which implies that  $U$  is greedy if that outcome is observed) and *pool* at R2;

P-P: types always pool;

S-S: types always separate (revealing that both are greedy).

It can easily be shown that a S-S equilibrium cannot exist (see the Appendix). The P-P equilibrium can exist, but it reveals nothing about the players' true types and is therefore uninformative about the narratives we want to assess. We mainly focus on the two remaining candidate equilibria. Equilibrium P-S corresponds to the first narrative, ORIV, according to which Russia is the villain, as it desires to use its military might to subjugate CEE. Equilibrium S-P corresponds to the second narrative, OUIV, according to which the U.S. is the villain because it is eager to use its military might to subjugate Russia.

### 4.3 P-S equilibrium

In this equilibrium, both types of  $U$  choose to expand NATO in the first stage; in the third stage, the greedy type chooses to attack, whereas the security type chooses to cooperate.<sup>8</sup> By the definition of types, at the information set R1 of the game, the greedy type of  $R$  would attack, whereas its security type would cooperate. Given that NATO has been expanded, in the second stage the greedy type of  $R$  attacks, whereas its security type chooses to cooperate.

The evolution of beliefs in this equilibrium is straightforward. Since types pool in the first stage, the posterior beliefs of  $R$  at the information set R2 are equal to its priors:  $Prob\{U : t = g|R2\} = p$ . Since the types of  $R$  separate in the second stage, their type is revealed, which implies:  $Prob\{R : t = g|U2\} = 0$ . Bayes' Rule has no bite at the information set R1 because it is off the equilibrium path.

Before stating the conditions for existence of a P-S equilibrium, it is useful to introduce a variable that represents the relative valuation of hegemony (or degree of revisionism) for the two types of the two states:  $r_t \equiv m_t/c_t$  and  $\rho_t \equiv \mu_t/\gamma_t$ , for  $t \in \{g, s\}$ . The following result can be established.

**Proposition 1.** *The P-S equilibrium exists if and only if the following parameter restrictions are jointly satisfied:*

$$r_s \leq \frac{1 - q_3}{q_3}, \quad (3)$$

$$p \leq \frac{q_2 - (1 - q_2)\rho_s}{q_3 - (1 - q_3)\rho_s}, \quad (4)$$

$$\text{if } \rho_g < \frac{q_2}{1 - q_2}, \text{ then } p \geq \frac{q_2 - (1 - q_2)\rho_g}{q_3 - (1 - q_3)\rho_g}. \quad (5)$$

This result yields various insights into what is and what is not logically required to maintain the ORIV narrative - according to which Russia is the villain. To begin with, let us mention the dog that did not bark: the prior belief  $\pi$  that Russia is greedy. It might seem counterintuitive that it does not matter for the existence of the P-S equilibrium. Why doesn't  $\pi$  affect  $U$ 's behavior? If  $U$  is the greedy type, it wants war with the highest chance to prevail, so it expands NATO irrespective of  $\pi$ . The reason why the security type of  $U$  chooses to expand NATO irrespective of  $\pi$  is that in a P-S equilibrium the

---

<sup>8</sup>Would the security type of  $U$  attack in the third stage, both types of  $R$  would attack in the second stage, and this would not be a P-S but a P-P equilibrium.

probability to end up cooperating rather than fighting is the same if NATO is expanded or not. But, the probability of prevailing in case of a war is higher for  $U$  if NATO was expanded beforehand. Hence, also the security type chooses to expand NATO irrespective of  $\pi$ .

**Corollary 1.** *Let the prior probability that  $R$  is greedy,  $\pi$ , gets arbitrarily close to 1. As soon as one of the conditions listed in Proposition 1 is violated, the P-S equilibrium does not exist.*

What is required for the ORIV narrative to be consistent, is the prior belief that the U.S. is likely to be a status quo state - see condition (4) - that attaches a low value to own hegemony - see condition (3). To grasp the intuition for the latter condition, recall that  $q_3 > q_2$ , which implies that the Russian security type is willing to cooperate at R2 only if the American security type is going to do the same at U2. Given the rise of American military power in U2, even if the U.S. is a status quo state, it may be tempted to attack Russia. This temptation is ineffective if and only if the American security type puts little value on hegemony. This explains why we need (3).

The rationale behind condition (4) comes from the incentive of the Russian security seeker to cooperate at R2. Given that the American types pooled at U1, the Russian security type prefers to cooperate if and only if,

$$p[(1 - q_3)\mu_s - q_3\gamma_s] + (1 - p)0 > (1 - q_2)\mu_s - q_2\gamma_s.$$

In words, the probability to be attacked by the U.S. after including Ukraine in NATO must be small - and this probability equals, given (3), the prior probability that the U.S. is greedy.

The last condition (5) stated in Prop. 1 is only required if the Russian greedy type has a small degree of revisionism, so that it actually dislikes War 2. It chooses to wage it nonetheless, because the risk of being attacked after acquiescing to full NATO expansion is not worthwhile being run.

#### 4.4 S-P equilibrium

Let us now turn to the S-P equilibrium, the one associated with the OUIV narrative according to which the U.S. is the villain. In this equilibrium, only the greedy type of  $U$  chooses to expand NATO in the first stage; in the third stage, the American greedy type chooses to attack, while the security type may either attack or cooperate.<sup>9</sup> By the

---

<sup>9</sup>The reason why both are admissible is that the security type has posterior probability of zero at U2 in an S-P equilibrium.

definition of types, at the information set R1 of the game, the greedy type of  $R$  attacks, whereas its security type cooperates. Given that NATO has been expanded, in the second stage  $R$  attacks, whichever its type.

The evolution of beliefs in this equilibrium is straightforward. Since the types of  $U$  separate in the first stage, the posterior beliefs of  $R$  at the information set R1 are such that  $Prob\{U : t = g|R1\} = 0$ , while at R2 we have  $Prob\{U : t = g|R2\} = 1$ . Bayes' Rule has no bite at U2 because it is off the equilibrium path.

**Proposition 2.** *The S-P equilibrium exists if and only if the following condition is satisfied:*

$$\pi \leq \frac{1 - q_2(1 + r_s)}{1 - q_1(1 + r_s)}. \quad (6)$$

The intuition is as follows. Given that types separate at U1, NATO expansion implies that  $R$  expects to be attacked at U2. Since  $q_3 > q_2$ , each type of  $R$  prefers attacking at R2. Condition (6) is required to have separation at U1. It is easy to see that the greedy type of  $U$  always prefers to expand NATO. The security type prefers not to expand NATO if and only if,

$$\pi[q_1 m_s - (1 - q_1)c_s] + (1 - \pi)0 \geq q_2 m_s - (1 - q_2)c_s.$$

In words, the probability to be attacked by  $R$  after having cooperated with it must be small - and this equals the prior probability that  $R$  is greedy.

Similarly to the competing narrative, the prior belief  $p$  that the U.S. is greedy does not matter for the existence of the S-P equilibrium. The reason is simply that NATO expansion reveals greed, independently of what the prior belief about U.S. preferences was.

**Corollary 2.** *Let the prior probability that  $U$  is greedy,  $p$ , gets arbitrarily close to 1. If condition (6) is violated, the S-P equilibrium does not exist.*

In sum, the conditions for existence of the P-S and S-P equilibrium suggest that the corresponding narratives depend on prior beliefs about types in a somewhat surprising way. The ORIV narrative can satisfy the logical requirements of a PBE only if the U.S. is a status quo state with a large probability. The prior probability of Russia being an imperialist state - e.g. because of its historical origins - is logically irrelevant. Symmetrically, The OUIV narrative can satisfy the logical requirements of a PBE only if Russia is a status quo state with a large probability. The prior probability of the U.S. being imperialist is irrelevant.

## 4.5 Multiplicity, uniqueness and nonexistence

There are parameter constellations such that the conditions (3), (4), (5), and (6) are simultaneously met. When this is the case, both the P-S and the S-P equilibrium exist. Is one more likely than the other to be selected? The standard refinement for games of this type is the concept of Sequential Equilibrium. Applying a theorem from Fudenberg and Tirole (1991), it follows that both the P-S and the S-P equilibrium survive this refinement, i.e. they are Sequential Equilibria.<sup>10</sup>

The conditions stated in Propositions 1 and 2 reveal that there also exist parameter constellations such that only one of the two equilibria exists. In order for this to be the case, priors  $p$  and  $\pi$  must be sufficiently far apart. If the prior probability that  $U$  is greedy is very low (and conditions (3) and (5) hold), while the prior probability that  $R$  is greedy is very high, then only the P-S equilibrium exists. These priors would validate the ORIV narrative and contradict the OUIV narrative. The reverse applies if the prior probability that  $R$  is greedy is very low, while the prior probability that  $U$  is greedy is very high. Then, only the S-P equilibrium exists, i.e. the priors validate the second narrative and contradict the first one.

Interestingly, it is not true that if the priors are the same (i.e.  $p = \pi$ ), either both equilibria exist or both do not exist. A counterexample is presented in the next Proposition, which deals with the plausible scenario where  $q_3 - q_2 > q_2 - q_1$ . In words, this is the scenario in which the military implications from inclusion of Ukraine into NATO are more far-reaching than those from mere enlargement to CEE.

**Proposition 3.** *Suppose  $p = \pi$ ,  $r_s = \rho_s$ ,  $q_2 = \frac{1}{2}$ , and  $q_3 - q_2 > q_2 - q_1$ . Existence of the P-S equilibrium implies existence of the S-P equilibrium, but existence of the latter does not imply existence of the former.*

It may be noted that the symmetric result does not obtain if one supposes  $q_3 - q_2 < q_2 - q_1$ , see the Appendix.

If the greedy types of both U.S. and Russia have a large prior probability, neither the P-S nor the S-P equilibrium exists. Equilibrium play can then be rationalized as supported by the P-P equilibrium. In this case, NATO expansion and the Russian attack do not reveal the types of the players and both narratives, ORIV and OUIV, are refuted. Like in a Security Dilemma, War 2 may then result from a situation in which both states

---

<sup>10</sup>The same conclusion applies to other refinements because they hinge on requirements about off-equilibrium beliefs. In this game, those beliefs do not pin down the PBE.

are security-seekers, but mistrust each other. Likelier, war is a manifestation of greed on both sides (the Orwellian view).

**Proposition 4.** *Assume that conditions (4) and (6) for existence of the P-S and S-P equilibrium are violated. Then, there exists a P-P equilibrium.*

The next three sections deal with some instructive generalizations of the setup analyzed in the current section.

## 5 More general war payoffs

In the above baseline model, wars 1, 2, and 3 stand for military conflicts under a progressively larger NATO. In War 1, NATO only includes the states that were members when the Cold War ended. In War 2, NATO is as it was in 2014, including various CEE-countries but not Ukraine. In War 3, NATO includes Ukraine. This progressive enlargement of NATO is to increase the relative military power of the U.S., which was captured by the assumption  $q_3 > q_2 > q_1$ . However, NATO expansion may additionally have an impact on the benefits of victory and costs of defeat for the U.S. and Russia. This may be so because the set of involved CEE-countries is different in the three war scenarios. This aspect was neglected in the baseline model.

Consider first the U.S. in the three possible military conflicts. The main benefit of victory, hegemony over Russia, is not likely to vary to a substantial degree across the three scenarios. By contrast, the costs of defeat, which include the loss of hegemony in some CEE countries, may vary. In the scenario of War 1, a Russian victory may shift even countries like Poland and Czech Republic from the American into the Russian sphere of influence - fear of this happening was a crucial motive behind Walesa's and Havel's push for NATO membership. In the scenario of War 2, this is no more the case because NATO article 5 protects those countries from a Russian attack. A Russian victory would however imply that Ukraine moves away from the Western orbit to be fully absorbed in the Russian sphere of influence. The costs of defeat would thus be lower for the U.S. in scenario 2 than in scenario 1. In the scenario of War 3, with Ukraine as a full NATO member, the costs for the U.S. in case of defeat are still lower, because Ukraine would not be attacked by Russia for the same reason as Poland and Czech Republic in scenario 2. The costs for the U.S. would mainly include the loss of American prestige in the international arena because a defeat would undermine the reputation of the U.S. military and intelligence.

Consider now Russia. The main cost of defeat, loss of sovereignty and subsequent dismemberment, is not likely to substantially vary across the three scenarios. In a specular

fashion to the costs of defeat for the U.S., the benefits from victory for Russia - which include the spread of its sphere of influence - can be thought to vary across the three scenarios. Russian gains would be larger in scenario 1 than in scenario 2, and larger in scenario 2 than in 3, where they would be minimal.

## 5.1 Analysis

Let us examine a generalized model that allows for changes in the appetite for revisionism across the scenarios of wars 1, 2, and 3. Following the above remarks, I posit that the costs of defeat for  $U$  and the benefit from victory for  $R$  are decreasing:  $c_t \geq c_{t2} \geq c_{t3}$  and  $\mu_t \geq \mu_{t2} \geq \mu_{t3}$ , with obvious notation. This implies:  $r_{t3} \geq r_{t2} \geq r_t$  and  $\rho_{t3} \leq \rho_{t2} \leq \rho_t$ .

I focus on the P-S and the S-P equilibrium in this more general setup. The following results correspond to Propositions 1 and 2 in the previous section.

**Proposition 5.** *The P-S equilibrium exists if and only if the following parameter restrictions are jointly satisfied:*

$$r_{s3} \leq \frac{1 - q_3}{q_3}, \quad (7)$$

$$p \leq \frac{q_2 - (1 - q_2)\rho_{s2}}{q_3 - (1 - q_3)\rho_{s3}}, \quad (8)$$

$$\text{if } \rho_{g2} < \frac{q_2}{1 - q_2}, \text{ then } p \geq \frac{q_2 - (1 - q_2)\rho_{g2}}{q_3 - (1 - q_3)\rho_{g3}}. \quad (9)$$

**Proposition 6.** *The S-P equilibrium exists if and only if the following condition is satisfied:*

$$\pi \leq \frac{1 - q_2(1 + r_{s2})}{1 - q_1(1 + r_s)} \cdot \frac{c_{s2}}{c_s}. \quad (10)$$

As compared to the model of the previous section, this generalization makes both the P-S and the S-P equilibrium less likely by imposing stricter conditions for their existence. Notably, the required priors for the greedy types must be smaller than lower thresholds. Consider the P-S equilibrium. Intuitively, if the benefits from victory for Russia are small in scenario 3 of complete NATO expansion as compared to one without Ukraine in NATO, then it is less likely that the Russian security type prefers accepting Ukraine in NATO,

i.e. it separates itself from the greedy type. Consider the S-P equilibrium. Intuitively, if the costs of a defeat for the U.S. are small in scenario 2 as compared to scenario 1, then it is less likely that its security type would refrain from expanding NATO at the outset, i.e. that it separates from the greedy type.

For the sake of simplicity, the generalizations investigated in the next two sections will neglect the differential benefits and costs of wars considered in this section.

## 6 American multiple priors

The Russian Federation was created in 1991, and the three years that separate its creation from the U.S. decision to expand NATO were a relatively short period of time. How confident could the U.S. be in 1994 that those three years enabled it to make an informed appraisal of the nature of the Russian state? Of course, by the early 1990s the U.S. had accumulated a vast amount of knowledge about the USSR. But the Russian Federation is not the USSR: their geographical, ethnic, and ideological backgrounds substantially differ. Given the dearth of relevant evidence available in 1994, it might have been impossible at that time to assign definite probabilities to the type of the Russian state. Then, the decision made by the U.S. at that juncture should rather be seen as one made by a decision maker that has too little information to form a definite prior about the nature of its opponent.

In order to examine this hypothesis, the prior  $\pi$  about Russia can be replaced with a range of probabilities that the U.S. assigned to Russia being the greedy type. This section investigates such a modified model using the multiple prior approach developed by Gilboa and Schmeidler (1989). In that article, they provide an axiomatic foundation of the maxmin expected utility decision rule.<sup>11</sup> For the sake of clarity, I am going to focus on the most extreme scenario in which the range of probabilities assigned to Russia being the greedy type is the unit interval. I refer to this assumption as the case of  $U$  having multiple priors.

**Proposition 7.** *Assume that  $U$  has multiple priors.*

- (i) *The conditions for existence of the P-S equilibrium are those stated in Prop. 1.*
- (ii) *The S-P equilibrium cannot exist.*
- (iii) *If condition (4) is violated, so that no P-S equilibrium exists, there is a P-P equilibrium.*

---

<sup>11</sup>Gilboa et al. (2008) offer a broad discussion of the multiple prior model and related models of belief formation.

The main insight delivered by this Proposition is that multiple priors kill the S-P equilibrium. As shown by Gilboa and Schmeidler (1989), in such a situation the decision maker chooses the alternative that maximizes the minimum expected utility across all possible priors. In the model of this paper, the U.S. with multiple priors always chooses to expand NATO, irrespective of its type. If the U.S. is a status quo state, NATO expansion is always chosen because it works as an insurance device against the worst case, namely War 1. Hence, consistency necessarily fails for the OUIV narrative if the U.S. acted upon multiple priors when it opted for NATO expansion.

Notice that the P-P equilibrium is the unique possible equilibrium if not only the U.S. but also Russia has multiple priors.

## 7 Limited NATO enlargement

A questionable feature of the model with American multiple priors is that U.S. behavior before the Russian attack of 2014 is mute in face of arriving information about the Russian type. An alternative view is that sufficient new information arrived in the meantime that allowed to form a unique prior  $\pi$ . In particular, this might have been possible by the time Hungary, Poland, Czech Republic, Bulgaria, Romania, Slovakia, Slovenia, and the Baltic States had become members of NATO. At that time, it would not have been possible to return to the NATO borders of the end of the Cold War. Yet, it would have been possible to proclaim that NATO's door is definitely closed; or, that further countries that formerly belonged to the Soviet Union (Ukraine, Georgia, Belarus, Moldova) cannot be considered for future NATO membership. I summarize these policy options unto the label "limited enlargement". How does this option of limited NATO enlargement affect the conclusions that were drawn from the baseline model?

Let's add to the sequence of events depicted by Figure 1 a third branch starting from U1, representing limited enlargement. If  $U$  chooses limited enlargement, a novel information set denoted by R $l$  is reached, with  $l \in (1, 2)$ . Once this information set is reached,  $R$  can choose either to cooperate or wage war, which is called War  $l$  and is won by  $U$  with probability  $q_l \in (q_1, q_2)$ . As before, the payoff from cooperation is normalized to zero.<sup>12</sup> The following result can be established.

**Proposition 8.** *Assume that limited enlargement is possible.*

---

<sup>12</sup>The determination of the PBE is the same if we posit that  $U$  first decides whether to expand NATO or not, and later, having selected the first option, whether NATO expansion is limited or unlimited (as events actually unfolded), or if we posit that  $U$  simultaneously decides among those three options (as done here for economy of space).

(i) The P-S equilibrium exists if and only if the conditions stated in Prop. 1 hold together with  $\rho_g > \frac{q_l}{1-q_l}$ .

(ii.a) If  $\rho_g > \frac{q_l}{1-q_l}$ , the S-P equilibrium exists if and only if,

$$\pi \leq \frac{1 - q_2(1 + r_s)}{1 - q_l(1 + r_s)}. \quad (11)$$

(ii.b) If  $\rho_g < \frac{q_l}{1-q_l}$ , the S-P equilibrium exists if and only if  $r_s \leq \frac{1-q_2}{q_2}$ .

(iii) Assume  $\rho_g > \frac{q_l}{1-q_l}$ . If conditions (4) and (6) for existence of the P-S and S-P equilibrium are violated, there exists a P-P equilibrium.

The option of limited NATO enlargement decreases the plausibility of the ORIV narrative and increases the plausibility of the OUIV narrative. As compared to the baseline model of Sect. 4, existence of the P-S equilibrium additionally requires that the greedy type of Russia would have waged war also after that the U.S. had committed not to let Ukraine (along with Georgia, Belarus and Moldova) into NATO. However, in that case, the S-P equilibrium is very likely to exist. War under limited NATO enlargement is likely to be conducted under military conditions that are very close to those of War 2.<sup>13</sup> This means that  $q_l$  is very close to  $q_2$ . Then, condition (11) is very likely to be met because the term on its right-hand side is close to one. If instead the greedy type of Russia would have remained peaceful after that the U.S. had committed not to let Ukraine (along with Georgia, Belarus and Moldova) into NATO, the requirement that the security type of the U.S. prefers peace over the current war is sufficient to ensure existence of the S-P equilibrium.<sup>14</sup>

A questionable feature of the model considered in this section is that it neglects possible American identity costs from abandoning NATO's open door policy. According to this criticism, such an American turnaround in the 2000s would have contradicted the tenet that the U.S. is exceptional, roughly meaning that its distinctive mission is to bring Western civilization to the rest of the world by means of global leadership - rather than compromise.<sup>15</sup> Jettisoning the open door policy for the sake of Russia's interests may

---

<sup>13</sup>NATO composition is the same in those two scenarios. The gap between between  $q_l$  and  $q_2$  may stem from diminished cooperation between NATO military and Ukrainian military before War 1 if Ukraine is known to be barred from entering NATO.

<sup>14</sup>Notice that condition (11) implies  $r_s \leq (1 - q_2)/q_2$ .

<sup>15</sup>As an illustration, Sarotte (2019, p. 36-37) reports the following extracts from a March 1995 memo to the U.S. defense secretary by his deputy: "it's Russia that must move toward us, toward our way of doing things." He admitted that, for some ..." this may be an obnoxious confirmation of our doctrine of "exceptionalism". ... Well, tough. That's us; that's the U.S. We are exceptional. ... Russia is either coming our way, or it's not, in which case it's going to founder, as the USSR did." Corneo (2024) offers a dynamic model of endogenous imperial identity.

have been perceived as tantamount betrayal of that American mission, and this betrayal would have demolished the self-image of most members of the U.S. government. In terms of modeling, if these identity costs are sufficiently large, the option of limited NATO enlargement becomes irrelevant for the determination of the PBE. Given such costs, the U.S. effectively committed to indefinite NATO expansion when it proclaimed that decision in 1994.<sup>16</sup>

## 8 Gauging prior beliefs

The plausibility of the P-S equilibrium and the S-P equilibrium crucially depends on the priors  $p$  and  $\pi$  about the type of the U.S. and Russia when the game began in the wake of the end of the Cold War. What are the main factors that might have determined their actual numerical values?

A first potential source of mutual knowledge was U.S. and Soviet behavior during the Cold War. However, as pointed out above, it may be unwarranted to assume that the identity of Russia and the one of the USSR are sufficiently similar that one can infer much about Russia from past Soviet behavior. Furthermore, scholars substantially disagree in their interpretations of the Cold War. Simplifying a lot, there are those who claim that only one side was greedy (either the U.S. or the USSR), that both of them were (the Orwellian view), and those who claim that both were security-seekers stuck in a Security Dilemma.

A second source of inference is from relevant events between the end of the Cold War and the 1994 decision to expand NATO. Arguably, the only one that may have left substantial impression is Yeltsin's coup d'état in 1993. It showed the fragility of the new Russian democracy. The Russian army was decisive to solve that constitutional crisis. The power of Russian military leaders and their inclination to revanchism likely enhanced the assessment that Russia is greedy.

After the decision to enlarge NATO was made in 1994, almost two decades elapsed before Russia crossed the Rubicon and occupied Crimea. Information about the U.S. posture between 1994 and 2014 may therefore be considered as relevant for assessing the Russian prior about the type of the U.S.. From the viewpoint of theory, this implies that some behavior of the U.S. during that period had a signaling effect and the U.S. rationally expected it; therefore, the U.S. took it into account when choosing to expand NATO in

---

<sup>16</sup>Trump's recent National Security Strategy might manifest a modification of U.S. identity, entailing disavowal of world hegemony for the sake of Western values and recognition of multipolarity, with its corollary about the necessity to compromise with other great powers if the dispute occurs outside the Western Hemisphere.

1994. Discussing possible reasons why the U.S. did not avoid that signaling is beyond the scope of this paper; arguably, the U.S. expected to benefit from those actions more than they would cost to the U.S. in terms of the signals they generated.

In the main, the U.S. posture in 1994-2014 likely increased the prior that the U.S. is greedy. In 1999, NATO bombed Serbia (a traditional Russian ally) and the U.S. established in Kosovo (a province of Serbia when NATO started its attack) a large military base. In 2001, after abandoning negotiations over the extradition of O. bin Laden, the U.S. military attacked Afghanistan in a war that lasted two decades. In 2002, the U.S. refused to ratify the Adaptation of the Treaty on Conventional Armed Forces in Europe, which would have restricted troops stationing. In the same year, the U.S. unilaterally withdrew from the ABM treaty (in 2007, the U.S. announced that it would install ABM-systems in Poland and Czech Republic). In 2003, after having put forward false claims about WMD, the U.S. attacked Iraq. In 2011, in the wake of a civil war, the U.S. and its NATO-allies militarily intervened in Libya and had its leader removed. In the same year, a civil war erupted in Syria, a country that hosted Russian military. At least since 2012, the CIA had supported Syrian rebels to topple the Assad regime (operation Timber Sycamore).

The main events between 1994 and 2014 seem to suggest that  $p$  was not very low. It is far from evident that its level was compatible with the equilibrium condition (4). This tends to speak against the P-S equilibrium and the ORIV narrative. *A fortiori*, skepticism about this narrative is warranted if limited NATO enlargement was an option. For its rival, the OUIV narrative, the crucial parameter is  $\pi$ . The relatively short experience with the Russian Federation at the time NATO expansion was announced suggests that the U.S. faced considerable ambiguity concerning Russia's type. This tends to speak against the S-P equilibrium. Furthermore, the role of the military in Yeltsin's coup d'état in 1993 likely corroborated the view that the eventuality of a bad Bear should be taken seriously. So, if a definite prior was formed, it is unclear whether it satisfied the equilibrium condition (6). It likely satisfied condition (11), but it is unclear whether limited NATO enlargement was an option.

## 9 Conclusion

This paper has developed a simple multi-stage model of incomplete information with equilibrium play that reproduces in a highly stylized fashion the U.S. decision to expand NATO in 1994 and the Russian attack against Ukraine in 2014. The narrative "obviously, Russia is the villain" requires an implausibly low prior probability that the U.S. is greedy.

It also requires that the security type of the U.S. has a negligible taste for supremacy. The narrative “obviously, the U.S. is the villain” requires that already in 1994 a definite probability could be assigned to the Russian types and that such a probability was relatively low for the greedy type. Both conditions are far from being evidently met.

Compared to these two narratives, one that grounds on the P-P equilibrium may look more plausible. Pooling by the U.S. seems plausible because in 1994 a unique prior about the Russian type might not have existed. NATO expansion could have been selected by a U.S. security type because of its insurance value. In turn, pooling by Russia seems plausible because the U.S. posture before 2014, both on European and extra-European disputes, was not reassuring in terms of beliefs. A remarkable hint of Russian pooling is William Burns’ written communication to U.S. Secretary of State Condoleezza Rice in 2008:<sup>17</sup> “Ukrainian entry into NATO is the brightest of all redlines for the Russian elite (not just Putin). In more than two and a half years of conversations with key Russian players, from knuckle-draggers in the dark recesses of the Kremlin to Putin’s sharpest liberal critics, I have yet to find anyone who views Ukraine in NATO as anything other than a direct challenge to Russian interests. ... I can conceive of no grand package that would allow the Russians to swallow this pill quietly” (Woodward, 2025, p.63).

In sum, the results obtained in this paper generate little confidence in each of the two main competing narratives about the origins of the Ukraine war. This finding is especially relevant for the EU’s current stance on Russia. If ORIV is implausible, there would clearly be some merit in trying to engage in a diplomatic dialogue with Moscow aimed at jointly redefining the security architecture of the Old Continent.

The baseline model and its three extensions examined in this paper may be complemented by further analyses of relevant aspects that could not be addressed here. Importantly, great powers with a global reach take into account the significance of their actions in a given theater for their broader, global, goals. Therefore, it would be recommendable to enrich the theoretical setup of this paper by investigating the effects of NATO expansion and Russian attack against Ukraine on the reputation for resolve of those states *vis-à-vis* third countries. In particular, with regard to China for what concerns the U.S., and with respect to other former Soviet Republics for what concerns Russia. Such a challenging task is left for future research.

---

<sup>17</sup>Burns was U.S. ambassador to Moscow and later CIA director.

## APPENDIX

### Proof of Proposition 1

Posterior beliefs stated in the main text satisfy Bayes' Rule; we now prove that strategies are sequentially rational.

Since  $q_2 < q_3$ , the security type of R prefers cooperating at R2 only if the security type of U prefers cooperating at U2, which is the case if and only if condition (3) is satisfied.

Given that each type of U expands NATO at U1, the security type of R prefers cooperating at R2 if and only if,

$$p[(1 - q_3)\mu_s - q_3\gamma_s] + (1 - p)0 \geq (1 - q_2)\mu_s - q_2\gamma_s.$$

Using  $\rho_s \equiv \mu_s/\gamma_s$ , this condition reads,

$$p[(1 - q_3)\rho_s - q_3] \geq (1 - q_2)\rho_s - q_2.$$

The term in square brackets is negative because  $\rho_s < \frac{q_1}{1 - q_1} < \frac{q_3}{1 - q_3}$ . Dividing each side by minus that term and rearranging terms yields condition (4).

Given that each type of U expands NATO at U1, the greedy type of R prefers attacking at R2 if and only if,

$$p[(1 - q_3)\mu_g - q_3\gamma_g] \leq (1 - q_2)\mu_g - q_2\gamma_g.$$

There are two cases. If  $\rho_g \geq \frac{q_2}{1 - q_2}$ , the RHS is positive and, because  $q_2 < q_3$ , the above condition is always satisfied.

If  $\rho_g < \frac{q_2}{1 - q_2}$ , both sides are negative. Rewrite that condition so as to have positive term on each side, as,

$$p[q_3 - (1 - q_3)\rho_g] \geq q_2 - (1 - q_2)\rho_g.$$

Rearranging terms, this yields condition (5).

It remains to be checked that both types of U pool at U1. The greedy type clearly prefers to expand NATO because it leads to war for sure, while cooperating leads to war with probability  $\pi < 1$  and such a war is won with a smaller probability. The security type prefers to expand NATO if and only if,

$$(1 - \pi)0 + \pi[q_2m_s - (1 - q_2)c_s] \geq (1 - \pi)0 + \pi[q_1m_s - (1 - q_1)c_s]$$

which is always satisfied because  $q_2 > q_1$ . **QED**

### Proof of Proposition 2

Posterior beliefs stated in the main text satisfy Bayes' Rule; we now prove that strategies are sequentially rational.

Given that types separate at U1, R expects to be attacked at U2. Since  $q_2 < q_3$ , each type of R prefers attacking at R2.

It remains to be checked that types separate at U1.

Given that each type of R attacks at R2, but only the greedy type attacks at R1, the greedy type of U prefers to expand NATO if and only if,

$$q_2 m_g - (1 - q_2) c_g > \pi [q_1 m_g - (1 - q_1) c_g],$$

which is always satisfied.

The security type prefers cooperating at U1 if and only if,

$$q_2 m_s - (1 - q_2) c_s < \pi [q_1 m_s - (1 - q_1) c_s],$$

which can be written as

$$q_2 r_s - (1 - q_2) < \pi [q_1 r_s - (1 - q_1)].$$

Notice that the term in square brackets is negative because  $r_s < (1 - q_1)/q_1$ . Dividing each side and rearranging terms yields,

$$\pi < \frac{(1 - q_2) - q_2 r_s}{(1 - q_1) - q_1 r_s},$$

which can be written as condition (6) stated in the Proposition. **QED**

### Proof of Proposition 3

Define  $p = \pi \equiv y$  and  $r_s = \rho_s \equiv x$ , and notice that  $x \in (0, 1)$ . Suppose that the P-S equilibrium exists. Then, by Prop. 1, condition (4) is satisfied. Using  $q_2 = 1/2$ , it can be written as,

$$y < \frac{q_2 - (1 - q_2)x}{q_3 - (1 - q_3)x} = \frac{(1 - x)/2}{q_3 - x + q_3 x} \equiv \hat{y}.$$

By condition (6) in Prop. 2, the S-P equilibrium exists if and only if,

$$y < \frac{(1 - x)/2}{1 - q_1 - q_1 x} \equiv \bar{y} \in (0, 1).$$

We are now going to show that  $\bar{y} > \hat{y}$ . Using their definition, this inequality holds if and only if

$$q_3 - x + q_3 x > 1 - q_1 - q_1 x,$$

or

$$(q_3 + q_1)(1 + x) > 1 + x,$$

or

$$q_3 - q_2 > 1 - q_1 - q_2 = q_2 - q_1,$$

which holds true by assumption.

If a P-S equilibrium exists, then  $y < \hat{y}$  and, since  $\bar{y} > \hat{y}$ , an S-P equilibrium necessarily exists because  $y < \bar{y}$  and (6) is satisfied.

If an S-P equilibrium exists and  $y \in (\hat{y}, \bar{y})$ , no P-S equilibrium exists.

Consider now the symmetric case in which  $q_3 - q_2 < q_2 - q_1$ , while the other assumptions stated in Prop. 3 are maintained. We claim that existence of the S-P equilibrium does not imply existence of a P-S equilibrium. In order to see it, suppose that condition (6) for existence of the S-P equilibrium is satisfied and  $r_s \in ((1 - q_3)/q_3, (1 - q_2)/q_2)$ . Since condition (3) is violated, the P-S equilibrium cannot exist. **QED**

#### Proof of Proposition 4

We prove the claim by constructing the conditions for existence of the P-P equilibrium. In this equilibrium, both types of U choose to expand NATO and both types of R choose to respond by launching war. It is easy to see that the equilibrium conditions for the greedy types are mechanically satisfied if the corresponding conditions for the security types are satisfied. We can therefore restrict the analysis to the security types.

Given that types pool at R2, the security type of U prefer to expand NATO if and only if,

$$q_2 m_s - (1 - q_2) c_s \geq \pi [q_1 m_s - (1 - q_1) c_s],$$

or,

$$q_2 r_s - (1 - q_2) \geq \pi [q_1 r_s - (1 - q_1)].$$

Notice that the term in square brackets is negative because  $r_s < (1 - q_1)/q_1$ . Dividing each side and rearranging terms yields,

$$\pi \geq \frac{1 - q_2(1 + r_s)}{1 - q_1(1 + r_s)}, \quad (12)$$

which holds if and only if (6) is violated.

Given that types pool at U1, the security type of R prefers to go to war for sure if  $r_s \geq (1 - q_3)/q_3$  because it would be attacked for sure at U2.

Consider now the case  $r_s < (1 - q_3)/q_3$ . The security type of R prefers to go to war if,

$$(1 - q_2)\mu_s - q_2\gamma_s \geq p[(1 - q_3)\mu_s - q_3\gamma_s],$$

or,

$$(1 - q_2)\rho_s - q_2 \geq p[(1 - q_3)\rho_s - q_3].$$

The term in square brackets is negative because  $\rho_s < \frac{q_1}{1-q_1} < \frac{q_3}{1-q_3}$ . Dividing each side by minus that term and rearranging terms yields,

$$p \geq \frac{q_2 - (1 - q_2)\rho_s}{q_3 - (1 - q_3)\rho_s}. \quad (13)$$

This is met if (4) is violated.

Summing up, the P-P equilibrium exists if and only if the following conditions are jointly satisfied: (12) and, if  $r_s < (1 - q_3)/q_3$ , (13). **QED**

### Nonexistence of S-S equilibrium

We prove it by contradiction.

Suppose there is a S-S equilibrium. At R2 the security type of R must prefer to cooperate. Given that types separate at U1, this preference implies

$$(1 - q_3)\mu_s - q_3\gamma_s \geq (1 - q_2)\mu_s - q_2\gamma_s,$$

which contradicts the assumption  $q_2 < q_3$ . **QED**

### Proof of Proposition 5

Posterior beliefs are as in Prop. 1. We now prove that strategies are sequentially rational.

Since  $q_2 < q_3$ , the security type of R prefers cooperating at R2 only if the security type of U prefers cooperating at U2, which is the case if and only if condition (7) is satisfied.

Given that each type of U expands NATO at U1, the security type of R prefers cooperating at R2 if and only if,

$$p[(1 - q_3)\mu_{s3} - q_3\gamma_{s3}] \geq (1 - q_2)\mu_{s2} - q_2\gamma_{s2},$$

which can be written as,

$$p[(1 - q_3)\rho_{s3} - q_3] \geq \frac{\gamma_{s2}}{\gamma_{s3}}[(1 - q_2)\rho_{s2} - q_2].$$

The term on the LHS is negative because  $\rho_{s3} \leq \rho_s < \frac{q_1}{1-q_1} < \frac{q_3}{1-q_3}$ . Dividing each side by minus that term, using  $\gamma_{s2} = \gamma_{s3}$ , and rearranging terms yields condition (8).

Given that each type of U expands NATO at U1, the greedy type of R prefers attacking at R2 if and only if,

$$p[(1 - q_3)\mu_{g3} - q_3\gamma_{g3}] \leq (1 - q_2)\mu_{g2} - q_2\gamma_{g2}.$$

There are two cases to consider. If  $\rho_{g2} \geq \frac{q_2}{1-q_2}$ , the RHS is positive and, since the LHS is negative, the above condition is always satisfied.

If  $\rho_{g2} < \frac{q_2}{1-q_2}$ , both sides are negative. Rewrite that condition so as to have a positive term on each side:

$$p[q_3 - (1 - q_3)\rho_{g3}] \geq \frac{\gamma_{s2}}{\gamma_{s3}}[q_2 - (1 - q_2)\rho_{g2}].$$

Rearranging terms and using  $\gamma_{s2} = \gamma_{s3}$ , this yields condition (9).

It remains to be checked that both types of U pool at U1. The greedy type clearly prefers to expand NATO because it leads to war for sure, while cooperating leads to war with probability  $\pi < 1$  and such a war is won with a smaller probability. The security type prefers to expand NATO if and only if,

$$(1 - \pi)0 + \pi[q_2m_{s2} - (1 - q_2)c_{s2}] \geq (1 - \pi)0 + \pi[q_1m_s - (1 - q_1)c_s],$$

which, using  $m_{s2} = m_s$ , can be written as,

$$(q_2 - q_1)m_s \geq (1 - q_2)c_{s2} - (1 - q_1)c_s,$$

which is always satisfied because the LHS is positive while the RHS is negative. **QED**

### **Proof of Proposition 6**

Posterior beliefs stated in the main text satisfy Bayes' Rule; we now prove that strategies are sequentially rational.

Given that types separate at U1, R expects to be attacked at U2. Then, each type of R strictly prefers attacking at R2.

It remains to be checked that types separate at U1.

Given that each type of R attacks at R2, but only the greedy type attacks at R1, the greedy type of U clearly prefers to expand NATO. The security type prefers cooperating at U1 if and only if,

$$q_2m_{s2} - (1 - q_2)c_{s2} < \pi[q_1m_s - (1 - q_1)c_s],$$

which can be written as

$$q_2r_{s2} - (1 - q_2) < \pi[q_1r_s - (1 - q_1)]\frac{c_s}{c_{s2}}.$$

Notice that the term in square brackets is negative because  $r_s < (1 - q_1)/q_1$ . Dividing each side and rearranging terms yields condition (10) stated in the Proposition. **QED**

### Proof of Proposition 7

We begin by proving claim (ii). With multiple priors, at U1 player U chooses the alternative that maximizes the minimum of its expected utility across all possible priors. First, consider the greedy type of U. If it cooperates, its minimum expected utility is 0, which obtains if the prior that Russia is a security seeker goes to one. If it expands NATO, its minimum expected utility is  $\min\{q_2m_g + (1 - q_2)c_g, q_3m_g + (1 - q_3)c_g\} = q_2m_g + (1 - q_2)c_g > 0$ , because the greedy type optimally wages war if U2 is reached.

Consider now the security type. If it cooperates at U1, its minimum expected utility is  $q_1m_s + (1 - q_1)c_s < 0$ . If it expands NATO, its minimum expected utility is  $q_2m_g + (1 - q_2)c_g$ . In order to see it, suppose first  $r_s > (1 - q_3)/q_3$ . Then, at U2 the security type would choose War 3 and its minimum expected utility if it expands NATO is the payoff of War 2. Suppose now  $r_s < (1 - q_3)/q_3$ . Then, at U2 the security type would choose to cooperate. Since War 2 is worse than War 3 and War 3 is worse than cooperation, its minimum expected utility if it expands NATO is again the payoff of War 2,  $q_2m_g + (1 - q_2)c_g > q_1m_s + (1 - q_1)c_s$ .

Therefore, both types of U choose NATO expansion at U1, which implies that there is no S-P equilibrium.

We now prove claim (i). Because of claim (ii), we only need to show that at R2 the greedy type prefers to wage war, while the security type prefers to cooperate. Since the Russian posterior belief at R2 equals the prior, the situation is equivalent to the one of Proposition 1. The conditions for the separation of types are therefore the same as those derived in its proof.

We now prove claim (iii). Suppose that (4) is violated. Then, the security type of R prefers to wage war at R2. The greedy type prefers to wage war if and only if,

$$p[(1 - q_3)\mu_g - q_3\gamma_g] \leq (1 - q_2)\mu_g - q_2\gamma_g.$$

As in the proof of Prop. 1, there are two cases. If  $\rho_g \geq \frac{q_2}{1 - q_2}$ , the RHS is positive and, because  $q_2 < q_3$ , the above condition is always satisfied.

If  $\rho_g < \frac{q_2}{1 - q_2}$ , both sides are negative. Rewrite that condition as,

$$p \geq \frac{q_2 - (1 - q_2)\rho_g}{q_3 - (1 - q_3)\rho_g}.$$

Since  $\rho_g > \rho_s$  and the RHS is strictly decreasing in  $\rho_g$ , this condition is mechanically met if (4) is violated. Hence, the greedy type prefers to wage war at R2, and there is a P-P equilibrium. **QED**

### Proof of Proposition 8

Notice as a preliminary step that the branch leading to R1 dominates the one leading to R2 for each type of U. Thus, we can eliminate the latter branch.

Claim (i). Suppose first that  $\rho_g > \frac{q_1}{1-q_1}$ . Then, the greedy type of R prefers to wage war at R1, and the game is formally equivalent to the game studied in Section 4. The conditions for existence of the P-S equilibrium are those stated in Prop. 1.

Suppose now that  $\rho_g < \frac{q_1}{1-q_1}$ . Then, both types of R prefer to cooperate at R1. As shown in the proof of Prop. 1, in order to have separation at R2, it is necessary that condition (3) be satisfied. Since  $q_3 > q_2$ , then, the security type of U prefers cooperation over War 2. This makes a P-S equilibrium impossible because the security type of U would profit from deviating to limited enlargement, which gives the payoff of cooperation for sure, which is better than War 2.

Claim (ii.a). Assume  $\rho_g > \frac{q_1}{1-q_1}$ . Then, the greedy type of R prefers to wage war at R1, and the game is formally equivalent to the game studied in Section 4 where subscript “1” has to be replaced with “l”. This yields condition (11) for existence of the S-P equilibrium.

Claim (ii.b). Assume  $\rho_g < \frac{q_1}{1-q_1}$ , which implies that the greedy type of R prefers to cooperate at R1. Given that types separate at U1, it is optimal for each type of R to wage war at R2. Consider now how types decide at U1, given pooling at R2. Since limited NATO expansion leads to cooperation, it is optimal for the greedy type of U to choose unlimited NATO expansion - because it prefers War 2 over cooperation. The optimal decision of the security type depends on its degree of revisionism. If  $r_s < (1 - q_2)/q_2$ , it prefers cooperation over War 2; hence, it chooses limited expansion and the S-P equilibrium exists. If  $r_s > (1 - q_2)/q_2$ , it prefers War 2 over cooperation; hence, it chooses unlimited expansion and the S-P equilibrium does not exist.

Claim (iii). Since  $\rho_g > \frac{q_1}{1-q_1}$ , the game is formally equivalent to the game studied in Section 4 where subscript “1” has to be replaced with “l”. We can then apply Prop. 4 and its proof to prove claim (iii) of this Proposition. **QED**

## References

- Acemoglu**, D. and A. Wolitzky, 2024, Mistrust, Misperception, and Misunderstanding: Imperfect Information and Conflict Dynamics, in Dube, O., Morelli, M. and D. Ray (eds.), *Handbook of the Economics of Conflict*, Vol. 1, Amsterdam: North-Holland.
- Brzezinski**, Z., 1997, A Geostrategy for Eurasia, *Foreign Affairs* 76, 50-64.
- Corneo**, G., 2024, Economic Growth and Imperialism, CEPR DP 19680.
- Debs**, A., 2025, On Nuclear Superiority and National Security, *Journal of Conflict Resolution*, forthcoming.
- Ellsberg**, D., 2017, *The Doomsday Machine*, New York: Bloomsbury Publishing.
- Fudenberg**, D. and J. Tirole, 1991, Perfect Bayesian Equilibrium and Sequential Equilibrium, *Journal of Economic Theory* 53, 236-260.
- Gilboa**, I., A. Postlewaite and D. Schmeidler, 2008, Probability and Uncertainty in Economic Modeling, *Journal of Economic Perspectives* 22, 173-188.
- Gilboa**, I. and D. Schmeidler, 1989, Maxmin Expected Utility with Non-Unique Prior, *Journal of Mathematical Economics* 18, 141-153.
- Götz**, E., 2017, Putin, the State, and War: The Causes of Russia's Near Abroad Assertion Revisited, *International Studies Review* 19, 228-253.
- Idrisoglu**, I. and W. Spaniel, 2024, Information Problems and Russia's Invasion of Ukraine, *Conflict Management and Peace Science* 41, 514-533.
- Katchanovski**, I., 2025, *The Russia-Ukraine War and Its Origins*, Cham: Palgrave Macmillan.
- Kydd**, A., 2001, Trust Building, Trust Breaking: The Dilemma of NATO Enlargement, *International Organization* 55, 801-28.
- Lieber**, K. and D. Press, 2006, The End of MAD?, *International Security* 30, 7-44.
- Palley**, T., 2024, *The Ukraine-Russia War Explained: How the US Exploited Internal Fractures in the Post-Soviet Order (Plus Lessons for Georgia)*, Economics for Democratic and Open Societies.
- Pipes**, R., 1997, Is Russia Still an Enemy?, *Foreign Affairs* 76, 65-78.

**Reiter, D.**, 2001, Why NATO Enlargement Does Not Spread Democracy, *International Security* 25, 41-67.

**Rosa, P.** and S. Benati, 2025, The Alliance Dilemma and the Onset of War. A Game-Theoretic Analysis of the Russia-Ukraine Conflict, *Peace Economics, Peace Science and Public Policy* 31, 409-437.

**Sarotte, M.E.**, 2019, How to Enlarge NATO, *International Security* 44, 7-41.

**Sarotte, M.E.**, 2021, Not One Inch, New Haven: Yale University Press.

**Smith, B. C.**, 2024, Commitment Problems and Russia's Invasion of Ukraine, *Conflict Management and Peace Science* 41, 494-513.

**Woodward, B.**, 2025, War, London: Simon & Schuster.

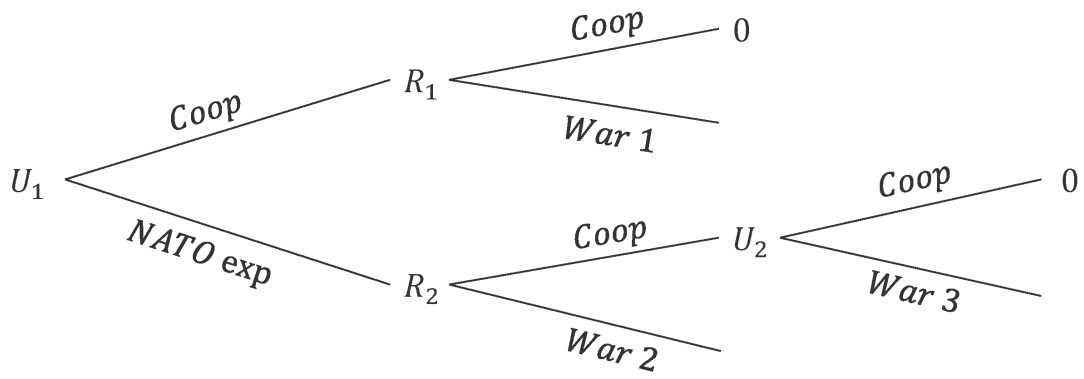


Figure 1: Sequence of events.

**Diskussionsbeiträge - Fachbereich Wirtschaftswissenschaft - Freie Universität Berlin**  
**Discussion Paper - School of Business & Economics - Freie Universität Berlin**

2026 erschienen:

- 2026/1      Hundsdoerfer, Jochen; Löwe, Maren: How Do Value Added Taxes Affect  
Wages and Labor?  
*FACTS*
- 2026/2      Gril, Lorena; Rendtel, Ulrich: Mapping High-Income Taxpayers in Berlin Using  
Kernel-Smoothed Proportions from Aggregated Georeferenced Data  
*Economics*
- 2026/3      Gril, Lorena; Hossain, Md Jamal; Tzavidis, Nikos; Rendtel, Ulrich: Kernel  
density estimation under masking of geolocations with applications to DHS  
data  
*Economics*