

Stopping in Time: Prospect Theory and The Failure of Deterrence Policy

Shiyu Chen *

Department of Political Economy, King's College London, London, The United Kingdom

*Corresponding Author: Shiyu Chen

ABSTRACT

The failure of deterrence policies in minimizing ongoing threats often occurs when traditional frameworks, which predominantly rely on cost-benefit calculations, prove inadequate. Most existing studies are anchored in these static models, focusing primarily on the immediate calculations of cost and benefit. However, such an approach overlooks the importance of a more dynamic understanding of decision-making processes, which are inherently fluid and responsive to shifts in environmental perceptions. This research seeks to address this gap by providing a dynamic analysis of Japan's decision to attack Pearl Harbor. By utilizing secondary sources from archival materials, this study aims to offer a more comprehensive explanation of how changing perceptions influenced Japan's strategic choices, contributing to the broader discourse on the limitations of deterrence policies.

KEYWORDS

Prospect Theory; Deterrence Policy; Japan; WWII

1. INTRODUCTION

When does deterrence policy fail—specifically, when the targeted party defies the policy sender's intent despite facing significant risks of failure? Existing research largely interprets such risky decision-making as a singular event, driven by a straightforward cost-benefit analysis. However, this view is overly simplistic, as it neglects the ongoing, complex interactions and the evolving environmental context that shape decision-making over time. Attributing deterrence failure solely to short-term and long-term cost-benefit calculations risks overlooking the unique considerations and motivations of the threatened party. This research endeavors to address this oversight by incorporating Prospect Theory to introduce a more dynamic and nuanced framework. By doing so, it seeks to provide a systemic analysis of risk-taking behavior, offering deeper insights into the complexities of deterrence policy failure.

The research unfolds in a structured sequence, beginning with a critical review of existing scholarship on the failure of deterrence policies. This section identifies and critiques the limitations of these studies, particularly their reliance on static, cost-benefit frameworks, and proposes an alternative approach to address these gaps. Following this, the study introduces and elaborates on the core concepts of Prospect Theory, explaining its application within the context of deterrence policy. This theoretical foundation is crucial for the subsequent analysis, as it provides the tools to understand the dynamic and psychological underpinnings of risk-laden decision-making. In the third section, the research applies this theoretical framework to a historical case study: Japan's decision to attack Pearl Harbor. Drawing on archival evidence, the study meticulously traces how the interactions between the United States and Japan created a psychological framing of loss for Japan. Within this frame, Japan perceived that adhering to the status quo would lead to intolerable costs, making the attack on

Pearl Harbor appear as an inevitable, rational choice under the circumstances. Consequently, the U.S. deterrence policy was structurally destined to fail. The final section synthesizes the findings, highlighting the implications of this dynamic approach for understanding deterrence policy failures. It also identifies potential areas for future research, suggesting that further exploration of the interplay between psychological framing and international relations could yield valuable insights into the complexities of deterrence and conflict escalation.

2. EXISTING STUDIES AND THEIR LIMITATIONS

Deterrence, a cornerstone of international security strategy, has been extensively studied and critiqued, particularly concerning its potential to fail under various conditions. This literature review examines key scholarly works that explore the limitations and potential failures of deterrence, providing a comprehensive understanding of the complexities involved in implementing and sustaining deterrence policies.

2.1. Literature on Failure of Deterrence Policy

Wieseltier critically examines the precarious nature of nuclear deterrence in the post-Cold War era, emphasizing its reliance on human rationality and the assumption that mutual destruction will always prevent conflict [1]. He argues that deterrence is not infallible and that the very foundations of nuclear deterrence—complacency, technological reliability, and rational decision-making—are vulnerable to failure. Wieseltier's work underscores the moral and strategic ambiguities of deterrence, suggesting that its failure could have catastrophic consequences, particularly as modern warfare evolves and expectations of "antiseptic" conflict management rise.

Weinstein extends this critique by identifying ten specific reasons why nuclear deterrence might fail. He argues that a false sense of security, the emergence of new dimensions of deterrence involving non-state actors, and the compartmentalization of national security all contribute to the fragility of deterrence [2]. Furthermore, the operational decline and strategic mismatch between the U.S. nuclear arsenal and contemporary threats highlight the potential for deterrence to fail when faced with asymmetric challenges. Weinstein's analysis calls for a reassessment of U.S. nuclear policies to address these vulnerabilities and ensure the credibility of deterrence in an increasingly complex global security environment.

Lebow revisits historical instances of deterrence failure, arguing that miscalculations by challengers, rather than weaknesses in deterrence strategy, often lead to conflict. He critiques assertions that deterrence fails due to poor execution by defenders, emphasizing instead the role of misperceptions and domestic pressures in driving aggressors to challenge deterrence. Lebow's analysis of crises such as the Korean War and the Cuban Missile Crisis underscores the complexity of deterrence and the difficulty of maintaining its credibility in the face of diverse and unpredictable challenges [3]. In a different vein, Huth and Russett focus on the dynamics that lead to the failure of deterrence and the subsequent escalation of conflict. Their analysis of 58 historical cases of deterrence reveals that deterrence success is closely tied to the short-term balance of military forces and the credibility of the defender's commitments [4]. When deterrence fails, the likelihood of escalation to full-scale conflict is influenced by factors such as alliance commitments, geographic proximity, and the defender's bargaining strategy.

Arquilla shifts the focus to regional deterrence, highlighting its limitations in the post-Cold War era. He argues that political, economic, and military constraints, combined with the U.S.'s gradualist approach to crisis management, undermine the effectiveness of regional deterrence. Arquilla critiques the reliance on multilateralism and the delays it introduces, which can compromise the credibility of deterrence. Moreover, he identifies the rise of asymmetric threats, such as mass movements driven by nationalism or religious fervor, as significant challenges that traditional deterrence strategies are

ill-equipped to address. The article concludes that regional deterrence is likely to fail unless the U.S. adopts a more flexible and proactive approach, including the possibility of unilateral action [5].

Quackenbush provides a comprehensive evaluation of deterrence theory, distinguishing between classical and perfect deterrence theory. He argues that while classical deterrence theory was foundational during the Cold War, it is increasingly challenged by the more nuanced perfect deterrence theory, which accounts for credibility issues and the limitations of rationality. Quackenbush emphasizes the need for a more sophisticated understanding of deterrence, particularly in the context of asymmetric conflicts and the evolving nature of warfare. He also highlights the difficulties in empirically testing deterrence theory, which complicates efforts to develop a universally applicable deterrence strategy [6].

Di Lonardo and Tyson introduce the concept of political instability within aggressor states as a critical factor that can undermine deterrence. They develop a formal model showing how internal political dynamics, such as leadership vulnerability and ideological conflicts, can lead to deterrence failure. The authors identify two key scenarios: the “instability curse,” where a leader’s fear of being replaced by a more dovish successor leads to reckless provocations, and the “deterrence curse,” where a dovish leader provokes a crisis to prevent hawkish elites from gaining power. Their analysis highlights the fragility of deterrence in politically unstable states, suggesting that traditional deterrence strategies may not be effective in such contexts [7].

2.2. Limitations and Reflections

The reviewed literature on the failure of deterrence provides valuable insights into the factors that contribute to the breakdown of deterrence strategies. However, a critical reflection on these works reveals a common limitation: each study tends to focus on independent and single-level factors that trigger war and cause deterrence to fail, either from the defender’s or the attacker’s perspective. While these factors—ranging from the credibility of threats, internal political instability, and the balance of military power—are undeniably important, they often fail to capture the full complexity of deterrence as a dynamic process. Deterrence is not a static condition but a form of balancing that evolves over time, influenced by the interaction of multiple factors across different levels and timeframes.

The traditional approach in these studies treats the causes of deterrence failure as discrete, one-time events rather than elements of an ongoing, dynamic process. This perspective overlooks how the interplay between defender and attacker evolves, as well as the changing strategic environment, can gradually erode the effectiveness of deterrence. For instance, factors such as the perceived threat from the defender to the attacker or the evolving domestic pressures on a state’s leadership are not isolated incidents; they are part of a continuous feedback loop that shapes decision-making and can ultimately lead to the breakdown of deterrence. Therefore, to fully understand why deterrence fails, it is essential to adopt a more holistic and dynamic approach that considers how these factors interact over time and influence each other, rather than viewing them as static and independent causes.

Despite the aforementioned shortcomings, the existing body of research significantly contributes to the foundation of my study, particularly in identifying critical variables that influence the success or failure of deterrence. These studies have meticulously detailed how specific factors, such as the credibility of the defender’s threats, the internal stability of the attacker, and the strategic calculations based on perceived threats, play pivotal roles in deterrence dynamics. For example, Weinstein (2000) and Quackenbush (2011) highlight the importance of perceived credibility and rationality in deterrence strategies, which are crucial to understanding the broader mechanisms at play. Similarly, Di Lonardo and Tyson (2022) bring attention to the internal political instability within aggressor states as a significant determinant of deterrence outcomes, offering valuable insights into how domestic factors intersect with international security concerns.

These contributions are conducive to my research as they provide a solid framework for examining how perceived threats from the defender influence the attacker’s decision-making process,

particularly in a dynamic context. By integrating the identified factors from these studies with Prospect Theory, this research aims to go beyond static analyses and explore how these variables evolve over time, interact with each other, and collectively shape the deterrence landscape.

3. PROSPECT THEORY AND DETERRENCE POLICY

Prospect Theory, formulated by Daniel Kahneman and Amos Tversky, fundamentally challenges the traditional assumptions of rational decision-making that underpin much of classical deterrence theory. While conventional models of deterrence assume that actors are fully rational and consistently seek to maximize their expected utility, Prospect Theory introduces a more nuanced understanding of decision-making under risk. This theory highlights that individuals and, by extension, state actors are often influenced by cognitive biases and psychological factors, particularly when they face uncertainty and potential losses [8]. Thus, one can propose that the risky actions taken by states facing instability could be elucidated by Prospect Theory.

The first step in applying Prospect Theory to deterrence is to identify the reference points that both the defender and the attacker use to evaluate their positions. In Prospect Theory, decision-makers assess outcomes relative to a specific reference point, typically the status quo or another baseline of perceived normalcy or entitlement. For states involved in deterrence, these reference points might include maintaining territorial integrity, preserving strategic influence, or avoiding regime change. By identifying these reference points, we can better understand how states perceive their positions and how deviations from these reference points (perceived as losses or gains) influence their behavior. For instance, if a state perceives a potential territorial loss as a deviation from its reference point, it may react more aggressively than if it perceived the situation as maintaining the status quo [9]. Once reference points are established, the next step is to assess how loss aversion impacts the decisions of states. Prospect Theory posits that decision-makers are typically more sensitive to losses than to equivalent gains, a phenomenon known as loss aversion. This aspect of the theory is particularly relevant to deterrence because states often frame challenges to their security or strategic position as potential losses relative to their reference points. As a result, they may adopt more aggressive or risk-seeking behaviors to avoid these losses, even when such actions increase the overall risk of conflict. For example, a state facing the prospect of losing strategic territory may be willing to escalate a crisis, taking on significant risks to prevent what it perceives as an unacceptable loss.

Framing effects are another critical component of this framework. According to Prospect Theory, the way in which choices are presented—or framed—significantly influences decision-making. In the context of deterrence, how a state frames its security situation can determine whether it perceives its options as involving potential gains or losses. For instance, if a deterrence strategy is framed internally as essential to preventing a catastrophic loss, state leaders may view any retreat or compromise as a failure, leading to a rigid and escalatory posture. Conversely, framing the same situation as an opportunity to maintain stability might lead to more diplomatic or flexible approaches [10]. Analyzing the framing of security policies and strategic decisions provides insights into why states might choose more aggressive or more restrained responses in deterrence situations. The framework also includes an analysis of how risk preferences shift when states are dealing with potential gains versus losses. Prospect Theory suggests that decision-makers are risk-averse when it comes to securing gains but become risk-seeking when trying to avoid losses. This shift in risk preference is particularly relevant in deterrence scenarios, where the stakes are high and the perceived risks of inaction (e.g., strategic losses) can drive states to adopt riskier strategies. For example, a state that perceives itself to be losing strategic ground might engage in risky escalation or brinkmanship, believing that the potential losses of inaction outweigh the risks of confrontation.

By incorporating insights from Prospect Theory above, one can expect that the failure of deterrence policies can be addressed dynamically. The classical model assumes that states act rationally, calculating the costs and benefits of their actions to avoid conflict. However, Prospect Theory reveals

that decision-making under risk is influenced by biases that can lead to irrational or suboptimal choices. By accounting for these biases, Prospect Theory helps explain why deterrence might fail even when the strategic balance appears to favor the defender. One of the most significant contributions of Prospect Theory is its explanation of miscalculations in deterrence. States may escalate conflicts or engage in risky behavior not because of a lack of capability or rational strategic calculation, but because of psychological factors such as loss aversion, framing effects, and skewed risk preferences.

4. CASE STUDY: JAPAN'S ATTACK ON PEARL HARBOR

The attack by Japan was meticulously planned and executed with the aim of delivering a crippling blow to the US Navy. On the morning of December 7, 1941, Japanese aircraft launched from six aircraft carriers and descended upon Pearl Harbor in a series of waves. The assault resulted in significant damage to the Pacific Fleet: eight battleships were damaged, with four being sunk, along with the destruction of over 300 aircraft. However, crucially, the US aircraft carriers were not in port at the time and escaped destruction.

Despite the tactical success of the attack, the strategic outcomes were far less favorable for Japan. The assault galvanized American public opinion, which had been divided over involvement in the war, and unified the US in its resolve to enter the conflict. The following day, the US declared war on Japan, leading to a protracted and ultimately devastating conflict for Japan [11]. The attack on Pearl Harbor, while initially a military success, failed to achieve its strategic objective of neutralizing the United States as a Pacific power and instead set the stage for Japan's eventual defeat in World War II. Given the significant risk of provoking the US into full-scale war, one might question why Japan pursued such a hazardous course of action. The following analysis grounded in Prospect Theory provides a compelling explanation: Japan's decision was heavily influenced by its perception of loss, which made any further erosion of Japan's strategic position intolerable.

First, the given fact about Japan, no matter it had triggered the Pacific war or not, is it always has a sense of insecurity rooted deeply in its national identity. As an island nation with limited natural resources, Japan has long viewed its dependency on foreign imports as a strategic vulnerability, especially in the face of potential conflicts that could disrupt supply lines. It was further exacerbated by the forced opening to Western trade in the mid-19th century, which exposed Japan to the threat of colonization and led to a national resolve to avoid the fate of neighboring China. Consequently, Japan rapidly modernized its military and adopted a preemptive defense strategy that prioritized securing territory beyond its borders as essential to its survival. The rise of militarism and the perceived encirclement by powerful neighbors and Western powers reinforced this expansionist agenda, ultimately leading Japan to pursue aggressive territorial acquisitions to safeguard its sovereignty and strategic interests [12].

Japan's trajectory leading up to the Pacific War was marked by its increasing role as a risk-taking actor in the Asia-Pacific region. This period was characterized by heightened tensions with the US, largely stemming from Japan's aggressive expansion in Southeast Asia, which directly threatened American interests in the region [13]. The US viewed Japan's territorial ambitions as a destabilizing force that could undermine its strategic and economic interests in territories such as the Philippines and its broader influence in the Pacific. In an effort to curb Japan's expansionist policies, the US initiated a series of economic and diplomatic actions aimed at constraining Japan's military capabilities. These measures included cutting off vital supplies of oil and steel, freezing Japanese assets within the US, and imposing comprehensive economic sanctions. The oil embargo, in particular, was perceived by American policymakers as the most potent non-military instrument to pressure Japan, as the nation relied on the US for approximately 60% of its crude oil imports. Given that Japan's existing oil reserves were only sufficient for 12 to 18 months of wartime operations, the

embargo placed Japan in a precarious position, severely limiting its capacity to sustain prolonged military engagements [14].

From the Japanese perspective, these actions were not just economic sanctions; they were existential threats that compounded the nation's strategic dilemma. For a nation that had long prioritized national dignity and military strength, retreating in the face of American pressure was not a viable option. Such a concession would have been tantamount to a repudiation of decades of expansionist policy, undermining the very foundation of Japan's imperial ambitions. Moreover, the situation was exacerbated by cognitive biases that influenced Japan's strategic decision-making. The sunk cost effect played a significant role, as Japan had already invested heavily in its expansionist endeavors. Abandoning these efforts in the face of US sanctions would have meant acknowledging the futility of these investments, a psychologically untenable position for Japanese leadership. Concurrently, the frame of loss further distorted Japan's perception of the situation. The narrative of American suppression had long been ingrained in the Japanese psyche, reinforcing a sense of victimization and inevitability, which led to an escalation in Japan's strategic calculations, as the leadership increasingly viewed war as the only means to preserve national honor and secure access to the resources necessary for survival.

These cognitive biases, coupled with the realpolitik of the time, pushed Japan towards a confrontational path. The combination of economic desperation and the psychological weight of perceived American hostility left Japan with few strategic alternatives. Ultimately, these factors contributed to Japan's decision to launch a preemptive strike against the US at Pearl Harbor, a move that marked the beginning of the Pacific War.

5. CONCLUSION

The case study of Japan's decision to attack Pearl Harbor provides a compelling illustration of the limitations inherent in traditional deterrence policies that rely predominantly on cost-benefit analysis. Japan's aggressive response, despite facing overwhelming odds, underscores the need to reconsider how we conceptualize state decision-making in the context of international conflict. The traditional framework of deterrence, which assumes rational actors making calculated decisions to avoid high costs, fails to account for the complex psychological factors that often drive state behavior. As demonstrated in this study, Japan's strategic choices were not merely the result of a rational assessment of risks and benefits but were deeply influenced by cognitive biases such as loss aversion, the sunk cost fallacy, and the framing effect. These biases led Japan to perceive its situation as one of inevitable decline unless drastic action was taken, rendering the US deterrence policy structurally ineffective. This research calls for the integration of Prospect Theory into the analysis of deterrence policy, particularly in understanding why and how such policies may fail. By incorporating insights from Prospect Theory, scholars and policymakers can gain a more nuanced understanding of the decision-making processes that underpin state behavior. The case of Pearl Harbor thus serves as a vital lesson in the broader discourse on international relations and security studies. It highlights the dangers of underestimating the role of psychological factors in strategic decision-making and the potential for these factors to lead to the failure of deterrence policies. Future research and policy formulation must therefore consider the psychological dimensions of deterrence, ensuring that strategies are not solely based on assumptions of rationality but also on an understanding of the complex human factors that influence state behavior in high-stakes situations. By doing so, we can better anticipate and mitigate the risks of deterrence failure, ultimately contributing to more effective and stable international security policies.

REFERENCES

- [1] Wieseltier, L. (1985). When deterrence fails. *Foreign Affairs*, 64(3), 525-546. <https://doi.org/10.2307/20042292>

- [2] Weinstein, J. M. (2000). Ten reasons why nuclear deterrence could fail: The case for reassessing US nuclear policies and plans. *Small Wars & Insurgencies*, 11(3), 21-34. <http://dx.doi.org/10.1080/09592310008423275>
- [3] Lebow, R. N. (1987). Deterrence failure revisited. *International Security*, 12(1), 197-213. <https://doi.org/10.2307/2538925>
- [4] Huth, P., & Russett, B. (1988). Deterrence failure and crisis escalation. *International Studies Quarterly*, 32(1), 29-45. <https://doi.org/10.2307/2600411>
- [5] Arquilla, J. (1995). Bound to fail: Regional deterrence after the Cold War. *Comparative Strategy*, 14(3), 215-228. <http://dx.doi.org/10.1080/01495939508403019>
- [6] Quackenbush, S. L. (2011). Deterrence theory: Where do we stand? *Review of International Studies*, 37(2), 741-762. <https://doi.org/10.1017/S0260210510000896>
- [7] Di Lonardo, L., & Tyson, S. A. (2022). Political instability and the failure of deterrence. *The Journal of Politics*, 84(1), 220-233. <https://doi.org/10.1086/715258>
- [8] Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I* (pp. 99-127). https://doi.org/10.1142/9789814417358_0006
- [9] Werner, K. M., & Zank, H. (2019). A revealed reference point for prospect theory. *Economic Theory*, 67(4), 731-773. <https://doi.org/10.1007/s00199-017-1096-2>
- [10] Boettcher III, W. A. (2004). The prospects for prospect theory: An empirical evaluation of international relations applications of framing and loss aversion. *Political Psychology*, 25(3), 331-362. <https://doi.org/10.1111/j.1467-9221.2004.00375.x>
- [11] The National WWII Museum. (n.d.). The path to Pearl Harbor. <https://www.nationalww2museum.org/war/articles/path-pearl-harbor/>
- [12] Barnhart, M. A. (1987). *Japan prepares for total war: The search for economic security, 1919–1941*. Cornell University Press.
- [13] Department of State. (1943). *Peace and war: United States foreign policy 1931–1941*. United States Government Printing Office. (Original work archived May 12, 2008). Retrieved August 12, 2024, from <http://www.ibiblio.org/hyperwar/Dip/PaW/>
- [14] Cox, S. (n.d.). *The battle of the Atlantic, September 1939 - May 1943: Battle of the North Atlantic*. Retrieved August 18, 2024, from <http://www.cv6.org/1941/btlord1/btlord1.htm>