Do Leaders Matter? International Conflict, Leader Traits, and Leader Selection

Robert Bognar
Lincoln College
University of Oxford

Thesis submitted in partial fulfilment of the requirements for the degree of DPhil in International Relations in the Department of Politics and International Relations at the University of Oxford

Hilary 2023

Word Count: 67,957

Abstract

Why do leaders start conflicts? The bulk of IR tends to look outside of leaders, focusing on institutional and structural forces. Where there is more uncertainty is whether the factors within leaders might also influence their decision to start conflicts. Thus, this thesis asks the question: To what extent do the personal traits of leaders influence their decision to start international conflicts? Recent literature has found that several traits of leaders strongly predict conflict initiation, such as military experience and age. However, these findings may be biased if leaders with these traits are more likely to be selected into environments where the probability of conflict is greater. Yet, this problem of selection has not been given enough attention in previous research. To deal with the problem of selection, I develop a unique strategy, which I call the "counterfactual leader strategy." The core idea of this strategy is that each time a leader is selected in reality, there is a counterfactual world in which someone else was selected instead. I call these individuals "counterfactual leaders" and the traits of these individual can be used as a "placebo" to indirectly test for bias. Since the counterfactual leaders' traits should not be associated with conflict initiation, any association reveals bias in the traits of leaders. Using an original dataset on leaders and counterfactual leaders between 1945 and 2004, I find that several traits of counterfactual leaders predict conflict in the same way as those traits among leaders. I argue this is because states that frequently engage in conflict are more likely to have candidates with these traits due to demographics and thus have a higher probability of selecting a leader with one, even if by chance. Therefore, I argue traits do not influence leaders' decision to start conflicts, and what matters instead is the environment they are selected into.

For my children, Cameron and Lily, and my loving wife Lucy.

Table of Contents

1 INTRODUCTION 12

1.1 GAPS 14

1.2 CORE ARGUMENT 16

1.3 CONTRIBUTION 16

1.4 CONCEPTS 19

1.4.1 INTERNATIONAL CONFLICT 19

1.4.2 LEADERS 19

1.4.3 HIGH-RISK TRAITS/LEADERS 20

1.4.4 THE SELECTORATE 20

1.4.5 SELECTION EVENTS 21

1.4.6 COUNTERFACTUAL LEADERS 21

1.5 OVERVIEW 22

1.6 RESEARCH DESIGN 26

1.6.1 CASE SELECTION STRATEGY 26

1.6.2 EMPIRICAL ANALYSIS 28

1.6.3 WHY NO CASE STUDY? 28

1.7 MAIN FINDINGS 29

2 DO LEADERS MATTER? 31

2.1 LEADERS AS CAUSAL VARIABLES: TRAIT EXPLANATIONS OF INTERNATIONAL CONFLICT 41

2.2 WHICH TRAITS EXPLAIN WHY LEADERS INITIATE INTERNATIONAL CONFLICTS? ... 42

2.2.1 MILITARY EXPERIENCE 45

2.2.2 WAR EXPERIENCE 48

2.2.3 AGE 49
2.2.4 BIOLOGICAL SEX 51
2.2.5 POLITICAL IDEOLOGY 54
2.3 HOLISTIC MEASURES OF A LEADER'S WILLINGNESS TO USE FORCE 56
2.3.1 LEADER RISK 57
2.3.2 HAWKISHNESS 59
2.4 OTHER TRAITS 60
2.4.1 EDUCATION 60
2.4.2 OCCUPATION AND UPBRINGING 62
2.5 CONCLUSION 63

3 THE PROBLEM OF SELECTION 64

3.1 LEADER SELECTION – A THEORETICAL FRAMEWORK 69
3.1.1 LEADER SELECTION EFFECTS 72
3.1.2 CANDIDATE SELECTION EFFECTS 80
3.2 SELF-SELECTION EFFECTS 87
3.3 CONCLUSION 89

4 A WAY FORWARD 90

4.1 THE FUNDAMENTAL PROBLEM 90
4.1.1 MEASUREMENT ERROR 92
4.1.2 BALANCE OF FORCE 96
4.1.3 PAST CONFLICT 99
4.1.4 SELF-SELECTION 100
4.1.5 OMITTED VARIABLES 101

4.2 BRINGING IT ALL TOGETHER 101

4.3 HOW TO GET AROUND THE FUNDAMENTAL PROBLEM 103

4.4 THE COUNTERFACTUAL LEADER 107

4.5 THE STRATEGY 109

4.6 EVIDENCE FOR THE LEADER SELECTION EFFECT EXPLANATION 117

4.7 CANDIDATE SELECTION EFFECT 118

4.8 EVIDENCE SUPPORTING THE TRAIT PERSPECTIVE 120

4.9 CANDIDATE AND LEADER SELF-SELECTION EFFECTS 122

4.10 KEY ASSUMPTION - COUNTERFACTUAL LEADER TRAITS AS PLACEBO 123

4.11 CONCLUSION 125

5 HYPOTHESES 126

5.1 MILITARY EXPERIENCE 129

5.1.1 MILITARY EXPERIENCE – LEADER SELECTION EFFECT EXPLANATIONS 130

5.1.2 MILITARY EXPERIENCE – CANDIDATE SELECTION EFFECT EXPLANATIONS 135

5.1.3 MILITARY EXPERIENCE – SELF-SELECTION EXPLANATION 138

5.2 NON-COMBAT MILITARY EXPERIENCE 140

5.2.1 NON-COMBAT MILITARY EXPERIENCE – LEADER SELECTION EFFECT EXPLANATIONS 141

5.2.2 NON-COMBAT MILITARY EXPERIENCE – CANDIDATE SELECTION EFFECT EXPLANATIONS 143

5.2.3 NON-COMBAT MILITARY EXPERIENCE – SELF-SELECTION EXPLANATION 144

5.3 WAR-WIN EXPERIENCE 145

5.3.1 WAR-WIN EXPERIENCE – CANDIDATE SELECTION EXPLANATIONS 147

5.4 AGE 149

5.4.1 AGE – LEADER SELECTION EFFECT EXPLANATIONS 149

5.4.2 AGE – CANDIDATE SELECTION EXPLANATIONS 154

5.5 FEMALE LEADERS 154

5.5.1 FEMALE LEADERS – CANDIDATE SELECTION EFFECT EXPLANATION 155

5.6 RIGHT-WING IDEOLOGY 156

5.6.1 RIGHT-WING IDEOLOGY – LEADER SELECTION EFFECT EXPLANATIONS 157

5.6.2 RIGHT-WING IDEOLOGY: CANDIDATE SELECTION EFFECT EXPLANATION 159

5.6.3 RIGHT-WING IDEOLOGY – LEADER SELF-SELECTION EXPLANATIONS 160

5.7 HOLISTIC MEASURES 161

6 METHODOLOGY 162

6.1 KEY TERMS OF REFERENCE 162

6.2 SCOPE 163

6.3 UNITS OF ANALYSIS - LEADER SELECTION EVENTS 163

6.4 THE COUNTERFACTUAL LEADER 164

6.5 OPERATIONALISING LEADER SELECTION EVENTS AND COUNTERFACTUAL LEADERS 164

6.5.1 ELECTIONS 165

6.5.2 APPOINTMENTS 170

6.5.3 SUCCESSIONS 172

6.5.4 SELF-SELECTIONS 173

6.6 SOURCES 175

6.7 LEADER AND COUNTERFACTUAL LEADER TRAITS 176

6.7.1 AGE 177

6.7.2 MILITARY EXPERIENCE 177

6.7.3 MILITARY EXPERIENCE TYPE 178

6.7.4 WAR WIN/LOSS EXPERIENCE 178

6.7.5 Right-Wing Ideology 179
6.7.6 Hawkishness 179
6.7.7 Conflict Initiation 180
6.8 Supporting Hypotheses Explanatory Variables 182
6.8.1 Past International Conflict Involvement 182
6.8.2 Veteran Population Proportion 183
6.8.3 Gender Equality 185
6.8.4 Life Expectancy 186
6.8.5 Incumbency 186
6.8.6 Past War Win / War Loss 186
6.9 Data Analysis 187
6.10 Testing the Selection Effect Hypotheses 187
6.10.1 Testing the Placebo Assumption 188
6.10.2 Testing the Supporting Hypothesis 189

7 FINDINGS 192

7.1 A Note on Datasets Used 193
7.2 LSD Summary Statistics 194
7.3 Leaders and Counterfactual Leaders: Summary Statistics 196
7.3.1 Geographic Statistics 196
7.3.2 Selection Mechanism Statistics 198
7.3.3 Conflict Initiation Statistics 199
7.4 A Note on Information Bias 202
7.5 Prior Military Experience 203
7.5.1 Military Experience: Summary Statistics 203
7.5.2 Military Experience: Selection Effect Hypotheses Tests 205

7.5.3 Military Experience: Main Findings 206
7.5.4 Military Experience: Placebo Assumption Test 208
7.5.5 Military Experience: Supporting Hypotheses Tests 209
7.5.6 Military Experience: Conclusion 213
7.6 Military Experience Type 213
7.6.1 Military Experience Type: Summary Statistics 213
7.6.2 Military Experience Type: Selection Effect Hypotheses Tests 214
7.6.3 Military Experience Type: Main Findings 215
7.6.4 Military Experience Type: Placebo Assumption Test 218
7.6.5 Military Experience Type: Supporting Hypotheses Tests 218
7.6.6 Military Experience Type: Conclusion 221
7.7 War Win Experience 222
7.7.1 War Experience Type: Summary Statistics 222
7.7.2 War Experience Type: Selection Effect Hypotheses Tests 222
7.7.3 Leader War Experience Type: Main Findings 223
7.7.4 War Experience Type: Placebo Assumption Test 225
7.7.5 War Experience Type: Supporting Hypotheses Tests 226
7.7.6 War Experience Type: Conclusion 229
7.8 Age 229
7.8.1 Age: Summary Statistics 229
7.8.2 Age: Selection Effect Hypotheses Tests 230
7.8.3 Leader Age: Main Findings 231
7.8.4 Age: Placebo Assumption Tests 233
7.8.5 Age: Supporting Hypotheses Tests 233
7.8.6 Age: Conclusion 236
7.9 Biological Sex 236

7.9.1 Sex: Summary Statistics 236
7.9.2 Sex: Selection Effect Hypotheses Tests 236
7.9.3 Biological Sex: Main Findings 237
7.9.4 Sex: Supporting Hypotheses Tests 239
7.10 Ideology 241
7.10.1 Ideology: Summary Statistics 241
7.10.2 Ideology: Selection Effect Hypotheses Tests 242
7.10.3 Leader Ideology: Main Findings 243
7.10.4 Ideology: Supporting Hypotheses Tests 245
7.11 Hawkishness 246
7.11.1 Hawkishness: Summary Statistics 246
7.11.2 Hawkishness: Selection Effect Hypotheses Tests 248
7.11.3 Leader Hawkishness: Main Findings 249
7.12 Conclusion 251

8 DISCUSSION AND CONCLUSION 252

8.1 Anticipating Objections: Counterfactual Leaders 255
8.2 The Problem with Leader Risk 258
8.3 Issues Regarding International Conflicts as the Outcome 260
8.3.1 Measurement Bias: Militarized Interstate Disputes as Outcome Variable .. 261
8.3.2 The Problem of Predecessors 263
8.4 Are International Conflicts Really Conflicts? 264
8.5 Endogeneity in Conflicts 265
8.6 Future Research 266
8.7 Conclusion 268

1 Introduction

Leaders make decisions, not states. Yet, throughout International Relations (IR) scholarship, major theories and prominent researchers tend to ask why *states* behave the way they do. Why do states go to war? Why do states cooperate, fear each other, arm, escalate, or start international conflicts short of war? Since we are consistently asking questions about states, does this mean we are also asking questions about leaders—the individuals who ultimately decide on and enact those behaviours, or are they considered less consequential? Of course, no one argues that leaders do not matter at all when asking these important questions. But exactly where leaders fit in these questions is less clear. The bulk of IR scholarship tends to privilege the institutional and structural forces pushing and pulling states towards war and conflict, building on the traditions of Waltzian "structural" theories of international relations, so we might think that leaders are merely causal mechanisms that only serve to translate these powerful forces into action. In other words, leaders matter in IR the same way cogs matter in a machine. On the other hand, leaders appear to fundamentally change the world – for better or worse – so it is hard to believe their importance is limited to being necessary but irrelevant causal mechanisms. Indeed, it is difficult to imagine that Germany would have embarked on a similar road of conquest and destruction had Hitler died in 1933. Thus, the puzzle is not why states start conflicts, but rather why leaders start conflicts.

This general puzzle can be approached in many ways but ultimately forces us first to consider two things. First, we would look *outside* leaders to the environments they inhabit and consider the extent to which their decisions to start conflicts simply mirrors the institutional and structural realities they face, or the broader groups of individuals to which they belong or answer to. Second, we would then look *within* leaders and consider the extent to which their individual traits, such as their backgrounds, life experiences, personality, and biology, influence those same decisions. In other words, we would juggle between two competing explanations – the environment and their traits – to consider which one better explains their decisions, especially

important high-stakes decisions like those to start conflicts. Since the bulk of IR already focuses on factors associated with the leaders' environment, such as relative power, interstate rivalries, and regime types, environmental explanations are already well understood. Where there is considerably more uncertainty, however, is the extent to which individual traits explain when and why leaders start conflicts. Thus, the central research question of this thesis is: **To what extent do the personal traits of leaders influence their decision to start international conflicts**?

While leaders and their traits have been traditionally overlooked in the literature, there has been a growing interest in exploring how and why the backgrounds and attributes of leaders might matter in IR, especially when trying to understand their link to conflict and war.[1] The limited but growing research so far has found that several leader traits, such as prior military experience and age, are strongly associated with conflict initiation, and more so than some of the more usual suspects in IR, such as regime type and the concentration of power (i.e., how much relative power a state has among others).[2] Some scholars even call leaders with a certain combination of these traits "high-risk" leaders.[3] It appears, then, that the growing view in IR seems to be that leader traits have an underappreciated and very important role in explaining why leaders start conflicts. This has important implications for IR theory and the causal forces in world politics, as well as implications for the conduct of international affairs and how we anticipate and deal with other states. There are only a handful of these types of studies, however, so our understanding of the

[1] Michael C. Horowitz and Matthew Fuhrmann, 'Studying Leaders and Military Conflict: Conceptual Framework and Research Agenda', *Journal of Conflict Resolution* 62, no. 10 (2018): 2072–86; Daniel Krcmaric, Stephen C. Nelson, and Andrew Roberts, 'Studying Leaders and Elites: The Personal Biography Approach', *Annual Review of Political Science* 23 (2020): 133–51; Stephen Benedict Dyson and Thomas Briggs, 'Leaders and Foreign Policy: Surveying the Evidence', in *Oxford Research Encyclopedia of Politics*, 2017.

[2] See Chapter Two for a full discussion of the research on which traits are associated with conflict. I focus primarily on military experience, military experience type, war experience, age, sex, ideology and two holistic measures based on a variety of traits.

[3] Michael C. Horowitz, Allan C. Stam, and Cali M. Ellis, *Why Leaders Fight* (Cambridge University Press, 2015); Jeff Carter and Charles E. Smith, 'A Framework for Measuring Leaders' Willingness to Use Force', *American Political Science Review* 114, no. 4 (2020): 1352–58.

research question is in its infancy and more research is needed to uncover the puzzle of why leaders start conflicts.

1.1 Gaps

Despite the growing number of studies in IR demonstrating leader traits help to explain why leaders start conflicts, one particular problem inherent in these studies has not received enough attention – selection. How are leaders selected into those positions in the first place, and from what pool of alternatives? As we will see, this can change everything. While the existing findings on leader traits and international conflict are robust to various statistical and methodological concerns, they rely entirely on observational data, which are notoriously prone to bias. Leaders are not randomly selected, so these findings could be spurious if leaders with certain traits, such as military experience, are disproportionately selected into leadership positions, especially if that happens in environments where the probability of conflict is already higher. While nearly all leader trait studies assume that leader selection effects exist, they try to control them away instead of investigating them with any meaningful rigour. Rather than expanding our knowledge about leader selection, these avenues of inquiry so far do the opposite; black boxing the problem and leaving our knowledge of selection effects sparse. Thus, if leaders' traits are important causal variables as research suggests, it is imperative we fully investigate whether selection explains these findings. To date, however, there has been no systematic investigation as to when, where, how, and why leaders with these traits are selected for office in the first place.

By considering the role of leader selection—how a leader comes to find themselves in charge – there are several implications to understanding not only the role of leaders in IR but for those who select them as well. Suppose people, such as voters, are choosing certain types of leaders in more dangerous places and times. If this were the case, there is a potentially alarming feedback loop compounding the risk of conflict – when conflict is more likely, people might be, perhaps unintentionally, choosing individuals who are likely to bring about more conflict rather than peace.

Indeed, an Israeli father who lost his son in war once reflected: "maybe by always choosing warriors, we doom ourselves always to be in wars."[4] However, the corollary to this feedback loop is that maybe by always being in wars, we also doom ourselves to always have warrior leaders.

Alternatively, it could mean that high-risk leaders might not really be "risky" at all, but rather just selected into more high-risk environments. If this were the case, our current understanding of the causal effect of leader traits on international conflict could be entirely backward if people choose leaders with certain traits over others in response to increased international conflict.

This study asks several other questions about leaders, conflict, and selection. Do people choose for certain leaders in response to their security environment? If so, what does this tell us about the people who select leaders? Do they yearn to assert power and choose leaders they believe are better suited toward this aim? While there is abundant research on how ordinary citizens can influence international politics by supporting certain foreign policy aims, little is known about how these same citizens can influence foreign policy by selecting certain leaders. Equally, it could be that leaders with certain traits are selected into high-risk environments because demographic factors correlate with future conflict. For example, if ex-military leaders are more likely to initiate conflict than non-ex-military leaders, then this might only be because conflict-prone states tend to have more veterans in the population, and hence a greater likelihood of selecting one as leader, even if by chance. Yet, there is also a shortage of research in understanding how demographics might shape leader selection within and across states and how either could be an effect of the international system. In essence, we do not know why or how high-risk leaders are selected and what consequences this might have on international politics. This study seeks to remedy that problem.

[4] Rachel Cooke, 'David Grossman: "I Cannot Afford the Luxury of Despair"', The Observer, 28 August 2010, https://www.theguardian.com/books/2010/aug/29/david-grossman-israel-hezbollah-interview.

1.2 Core Argument

Based on an empirical analysis of international conflict data, I argue that leaders with "high-risk" traits are not really risky at all. Rather, leaders with these traits only appear conflict-prone because they are just more likely to be selected into high-risk environments. In other words, we could switch around leaders at random and international politics would look very much the same. This is not to say that leader traits do not matter at all when trying to understand international political outcomes. Extreme leader personalities, such as Hitler and Hussein, probably explain their states conflict behaviour better than any other institutional or structural factor. Indeed, if we could swap Hitler and Hussein's personality with any other leader at random, both countries would probably have been very less conflict-prone than in reality. Rather, I argue more specifically that age and prior military experiences have no causal effect on leaders' decision to start conflicts. Ronald Reagan, for example, likely would have had an aggressive foreign policy even if he was younger and never served in the military, and might never have been elected in 1980 had he been left-wing. Reagan's hawkishness is likely better explained by his personality rather than his age and military experience, but older and ex-military leaders like Reagan just tend to be selected for during dangerous times and in powerful conflict-prone states, like the United States, Russia, China, and Israel, rather than weaker conflict-averse states like Costa Rica, Barbados, Singapore, and Cape Verde. This selection effect for ex-military and older leaders in powerful conflict-prone states is largely because of demographics – they simply have more veterans in the population and higher life expectancies than weaker conflict-averse states.

1.3 Contribution

My dissertation makes several contributions to IR. First and more importantly, my research adds to the ongoing "Do leaders matter?" debate in IR. I help clarify the uncertainty about whether leaders are important causal variables in explaining international political outcomes. My findings strongly suggest that selection explains nearly all previous findings on why leaders with "high-risk"

traits appear more likely to initiate conflict. These findings, therefore, cast doubt on the idea that leaders are important causal variables, at least in terms of the variation in military experience and age, and to a lesser extent, sex and ideology. Understanding how and why leaders are selected is necessary to improve future quantitative research on leaders in IR to address the critical problem of endogeneity. Without doing so, leadership scholars might be misled by their results and make inappropriate causal inferences about the relationship between leaders and international conflict. We might even get international politics and diplomacy wrong if we focus too much on leaders and not enough on the environments and constraints in which counties find themselves—a direct problem already in our perceptions because of the fundamental attribution error.

Second, my research also adds to a growing "first image reversed" literature in IR – how systemic-level factors can directly or indirectly cause variation in the types of leaders that enter office. For instance, I find that candidates are more likely to be ex-military as the veteran population share increases, which is an effect of the international system due to wars and conscription. To my knowledge, no other studies have demonstrated a clear, although indirect, relationship between leader selection and international conflict. Furthermore, I find no evidence that past international conflicts cause people to vote for candidates with high-risk over low-risk traits, suggesting that a first image reversed theory (i.e., the international system causes voters' preferences for leader to change) does not appear to explain why certain types of leaders are selected over others, at least in response to international conflict.

I also demonstrate the problems with using leaders' natural deaths as a strategy to remove endogeneity from selection. Using natural deaths in this way is growing in IR as many commend this as a "godsend" to first-image scholars since leader transitions in these cases are thought to be as-if random.[5] However, while the timing of these events might be random, who takes over as the

[5] Robert Jervis, 'Do Leaders Matter and How Would We Know?', *Security Studies* 22, no. 2 (2013): 158; Benjamin F. Jones and Benjamin A. Olken, 'Do Leaders Matter? National Leadership and Growth since World War II', *The Quarterly Journal of Economics* 120, no. 3 (2005): 835–64.

subsequent leader is still non-randomly determined. Many are still previously elected in Vice Presidential elections, immediately elected in party leadership elections, or even directly chosen by Heads of State.

While my findings primarily advance the theoretical understanding of how leader traits might matter toward explaining international conflict, my findings also have real-world implications as well. IR research is often used to inform foreign policy, so previous findings on the importance of leader traits could be used by decision-makers as a quick heuristic to predict or explain other leaders' foreign-policy behaviour. Considering that many studies find leader traits have effect sizes comparable to more notable institutional and structural factors, decision-makers might be misled into thinking that older and ex-military leaders, for instance, are more "aggressive" and take unnecessary security precautions as a result, which may in themselves elevate tensions or be seen as provocative. In essence, a leader-focused view of the world may worsen the security dilemma if our findings our fundamentally flawed. Likewise, since IR research also informs the general public – although to a lesser degree – if people are led to believe that a given electoral candidate might be more conflict-prone because they have certain traits, then there is a chance they could vote against that candidate (or fail to put them forward among political parties) when in fact it may be in their interest to do so. My findings can be used to directly challenge any of these misleading preconceptions about leaders that political elites or voters might have due to the previous findings.

Lastly, I create an original dataset of 2,680 events in which leaders are selected (i.e., via elections, appointments, successions, and self-selections such as coups and failed coups) between 1945 and 2004. These data include the traits of 2,273 individuals within these events, which are not only leaders, but individuals not previously included in leader-level datasets, such as the runners-up in elections and leaders of coup attempts. By providing a database of these various individuals and their traits, leader-level IR scholars can ensure their findings are not susceptible to bias from selection. These data are especially beneficial for those utilising regression discontinuity designs based on close elections where this bias can be removed entirely, at least in theory.

Relatedly, these data also allow scholars to identify which individuals could be considered viable "counterfactual" leaders, which could be beneficial to indirectly test for bias (as I do here), or for a case selection strategy in counterfactual analyses. Another fruitful avenue of research is for scholars to build on this dataset and assess candidates' electability for other relevant attributes not studied in this dissertation, such as personality.

1.4 Concepts

Before describing how this project will unfold, several key concepts I use throughout this dissertation must first be clarified for context.

1.4.1 International Conflict

I define international conflict synonymously with militarised interstate disputes as "an interaction between two states where one or both sides use military force."[6] This definition encompasses many incidents that are often classified as "short of war", such as the Cuban Missile Crisis, aerial dogfights, border skirmishes and naval engagements. Since every conflict has the potential to escalate toward war, we might expect people to be concerned with these events when choosing their leader, regardless of their severity.

1.4.2 Leaders

I define the leader, following existing research, as the individual who is ultimately responsible for the foreign policy of their state.[7] In constitutional monarchies and semi-presidential systems with a Head of State and Head of Government, only the latter (Prime Minister

[6] Douglas M. Gibler, *International Conflicts, 1816-2010: Militarized Interstate Dispute Narratives*, vol. 2 (Rowman & Littlefield, 2018).
[7] Henk E. Goemans, Kristian Skrede Gleditsch, and Giacomo Chiozza, 'Introducing Archigos: A Dataset of Political Leaders', *Journal of Peace Research* 46, no. 2 (2009): 269–83; Cali Mortenson Ellis, Michael C. Horowitz, and Allan C. Stam, 'Introducing the LEAD Data Set', *International Interactions* 41, no. 4 (2015): 718–41.

or equivalent) is considered to be the leader of the state. In many of these types of governments, such as in Austria, the role of the President is largely ceremonial, and so has little to no influence on foreign policy.[8]

1.4.3 High-Risk Traits/Leaders

I refer to high-risk traits as those that have been found to be associated with conflict in previous research. Since I limit my discussion to only five of these traits, a "high risk" trait can mean either military experience, non-combat, experience war-win experience, old age, or being female. This is not to say that these traits do indeed cause leaders to be risky; rather, I only use it as shorthand when referring to them together. Accordingly, when I refer to "high-risk leaders," I similarly use this as shorthand for the leaders that are ex-military, non-combat veterans, war-winners, female, and relatively older. Again, I do not imply that these leaders are "risky" or that people perceive them to be, but rather they are a specific leader prototype that previous research has found to be among the most likely to initiate conflict.

1.4.4 The Selectorate

Another concept I wish to clarify upfront is that of the selectorate. This term is largely confined to IR and was only recently introduced by Bruce Bueno de Mesquita in his 2003 book, *The Logic of Political Survival*.[9] The selectorate are those individuals who have a government-granted say in selecting leaders. These individuals can number anywhere from one or more political elites,

[8] While there are certainly debateable cases as to which individual can be considered "leader" in a given state and time, for the most part these are exceptions that prove the rule and do not readily affect my results. Most of these debateable cases occur during coups or revolutions and the debate is more or less in regard to the date at which there is a transition from one individual to another.

[9] Bruce Bueno de Mesquita, The Logic of Political Survival (Cambridge, Massachusetts: MIT Press, 2003).

such as a Head of State when appointing a Prime Minister, or millions of voters in elections. While I borrow a key term from selectorate theory, it should be noted my project's aims are entirely unrelated. Selectorate theory is primarily concerned with how leaders stay in power, whereas I am more concerned with why and how selectorates might choose some leaders over others.

1.4.5 Selection Events

Selection events are a new concept that I create and are broadly defined as an actual or potential leader transition. In other words, any time a leader comes into power through one of four selection mechanisms: elections, appointments, successions, or self-selections. The fundamental purpose of selection events is that to understand why leaders are selected, we need to first identify when they were selected, how they were selected, and who was *not* selected. I explain in more detail how I conceptualise selection events in Chapter Three and operationalise the concept in Chapter Six.

1.4.6 Counterfactual Leaders

Relatedly, another concept I introduce is that of the counterfactual leader, which is defined as the individual who likely would have been leader had the eventual leader not been selected. Put differently, they are the individuals who were *almost* selected, such as runners-up in elections, deposed incumbents in coups, or leaders of failed coups. A short summary of this complex concept likely raises more questions than answers. However, the fundamental importance of the counterfactual leader is that these individuals are the closest we can get to observing whether a state could have selected a leader with the opposite trait, which is essential to ensure that selection effects are not confounding the relationship between leader traits and conflict initiation. I more thoroughly explain the concept of counterfactual leaders, the underlying assumptions, and why we

can infer selection effects using their traits at length in Chapter Four and operationalise the concept in Chapter Six.

1.5 Overview

If we could switch the military power of Mexico and the United States, we would probably have very little doubt that this would result in drastic changes to world politics. Conversely, if we could switch the Presidents between the two countries, we might have considerably more doubt as to whether this would cause any meaningful impact on world politics. Again, when considering whether leaders matter, we can hold one of two competing perspectives: On the one hand, we might believe that leader traits are largely unimportant in IR and that the variation in their environment can entirely explain any variation in leaders' foreign policy behaviour. This view is the "environmental perspective" and argues that leaders are just causal mechanisms and not important causal variables (a cog in the larger machine). Conversely, we might believe that the traits of leaders are much more important and likely explain a not insignificant part of international political outcomes. This view is the "trait perspective" and argues that leaders are not just causal *mechanisms* but important causal *variables* in themselves.

Existing research on whether traits explain why leaders initiate conflict lends supports to the trait perspective. While the research is in its infancy, several studies have found that traits such as military experience, age, sex, and ideology are strongly associated with conflict initiation.[10] Furthermore, a select number of studies have grouped all these traits into a single index and similarly find that leaders with a combination of these traits are indeed more likely to initiate conflict. Thus, the literature suggests that leaders' particular backgrounds, experiences and attributes are important to understanding why leaders start conflicts.

[10] See Chapter Two for a full review of the literature on these traits.

While these findings are important and help us better understand why conflicts occur, the fundamental problem is that leaders are not randomly selected. In essence, these findings may be biased if leaders with high-risk traits were more likely to be selected into environments where the probability of conflict is greater. Such a confound would undermine our confidence in those previous results. As I discuss in more detail in Chapter Three, there are generally three ways selection can more specifically confound the relationship between leader traits and conflict initiation. The first is a "leader selection effect" – any reason candidates with high-risk traits could be disproportionally selected as leaders in environments with a greater probability of conflict. For example, in a Presidential election between a military veteran and a non-veteran, voters in more conflict-prone states might be more likely to choose the veteran in the belief they are better equipped to deal with interstate threats. The second way the trait-conflict association can be confounded is through a "candidate selection effect" – any reason individuals with high-risk traits could be disproportionately selected as *candidates* in environments with a greater probability of conflict (from which the leader will inevitably be drawn). For example, if voters preferred ex-military leaders in more dangerous times, then political parties might also be more likely to nominate them as candidates in elections in those same times. Alternatively, if high-risk leaders are generally older and military veterans, with the increase in life expectancy in these countries and prior wars, there could merely be a surplus of such individuals within the population to run for office. As such, conflict-prone states could be more likely to have military veterans than conflict-averse states, even if candidates and leaders were selected by chance. The last way selection can confound these findings is through a *self*-selection effect – any reason individuals with high-risk traits could be more likely to attempt to self-select (e.g., via a coup to take power) or be more successful at doing so. If leaders with high-risk traits are only more likely to initiate conflicts after self-selecting, then we know this is just the causal effect of the coup rather than their trait.

The fundamental problem with these potential selection effects, though is that they are extremely difficult to control for since the causal variables required to test them are unobservable

or measured with considerable error. For example, we cannot accurately measure the probability of conflict when leaders are selected since the two more important facets of the environment, relative power and past conflict, suffer from measurement error, and so are not sufficient to control for or test selection effects. This problem is discussed at length in Chapter Four to identify a novel methodological solution. Furthermore, the leader preferences of the selectorate are unobservable as well since there exists no data on whether people tend to vote for candidates with high-risk traits in response to their international security concerns. Lastly, many demographic variables needed to test for candidate selection effects do not exist, such as the share of veterans in the population across states. Compounding this fundamental problem of unobservable confounders, many of the conventional methods in IR we could consider as a solution, such as differences-in-differences, regression discontinuity using close elections, or a natural deaths approach, are ultimately insufficient to control for selection effects since they will likely suffer from low statistical power due to the nature of the data or are based on questionable theoretical assumptions. One of the challenges, and contributions, of this thesis is therefore to devise an alternative way to test for the presence of selection effects.

Accordingly, I develop a unique strategy to get around the fundamental problem and test whether selection potentially confounds the association between traits and conflict initiation, which I call the "counterfactual leader strategy." The core idea of this strategy, as discussed in Chapter Five, is that each time a leader is selected in reality, there is a counterfactual world in which someone else was selected instead. In most cases, these individuals are easy to observe; they are the runners-up in elections (e.g., Donald Trump in the 2020 U.S. election), the person who the Head of State almost appointed (e.g., Rab Butler following the 1957 resignation of Anthony Eden), the third-in-line in successions (e.g., John McCormack following John F. Kennedy's 1963 assassination), the deposed leaders in coups and revolutions (e.g., Mohamed Reza in the 1979 Iranian Revolution), or the leader of failed coup attempts (e.g., Hugo Chavez in the 1992 Venezuelan coup attempt). I call these individuals "counterfactual leaders" because they are the

closest we can get to observing a counterfactual world where a state could have selected someone else as leader. Now, the critical innovation here is that the *traits of counterfactual* leaders can be used as a "placebo" to indirectly test for selection. That is, the counterfactual leaders' traits should *not* be associated with conflict initiation because their bearers are not the leader, and thus do not have a meaningful causal effect on conflict initiation.[11] By deduction, this means that if the counterfactual leaders' trait *is* associated with conflict initiation, then this can only occur through a non-causal (incidental) path. Given that the leaders' and counterfactual leaders' traits both result from the same selection process, the most likely non-causal path responsible is one of the three selection effects. Importantly, sometimes their traits will be the same (or similar), and sometimes they will be different, which offers a way for us to differentiate the effects of selection on *leaders* (alone), from the selection on *candidates*.

Furthermore, counterfactual leaders reveal whether a state "almost" received the counterfactual treatment – a leader with the opposite trait. So, if the effect of a leader's trait on conflict depends on the counterfactual treatment (i.e., trait of the counterfactual leader, whether the same or different to that trait in the leader), then we know this can only be explained by one of the selection effects. Whether the effect is positive or negative, after conditioning on whether the counterfactual treatment would have been the same or different, will reveal one or other of the selection effects. For instance, if leaders with military experience are more likely to initiate conflict than leaders without military experience only when the counterfactual leader has the same military experience, then we know this can only be explained by candidate selection. In other words, the effect of military experience is *only* an effect of selecting ex-military candidates in the first place (from which only an ex-military leader was, and could be, selected). Conversely, if

[11] The assumption that counterfactual leader traits are a valid placebo depends on several conditions that I more fully discuss in Chapter Five. Namely, this assumption could be violated if counterfactual leaders were previous leaders (or close advisors), since their trait could still have causal effect on past conflicts and hence made conflict more likely in the future. In any case, I test this possibility in my findings chapter, and confirm that the placebo assumption is satisfied.

leaders with military experience are more likely to initiate conflict than leaders without military experience only when the counterfactual leader's military experience is different, we know, for the opposite reasons, that only leader selection effects can explain this. In this case, the effect of military experience is *only* an effect of selecting "against" non-military candidates for leader. Lastly, if ex-military leaders are more likely to initiate conflict than non-ex-military leaders regardless of the counterfactual treatment, then we know that the trait (military experience) likely has a genuine causal effect on conflict since it does not depend on the counterfactual treatment.

1.6 Research Design

Since existing leader datasets (LEAD and Archigos) only include actual leader transitions – and not potential leader transitions – I create an original dataset, which I call the Leader Selection Dataset (LSD) that more specifically identifies the who, what, when, where, and how leaders are selected. The fundamental units of observation in the LSD dataset, or cases, are leader selection events, defined above. For each leader selection event, I first identify the selection mechanism (elections, appointments, successions, and self-selections), and then I identify the counterfactual leader, which is usually the runner-up in elections, the individual the Head of State almost chose, the next-in-line in appointments, the incumbent leader in successful self-selections, or the coup leader in failed self-selection attempts.

1.6.1 Case Selection Strategy

When choosing which selection events to include for analysis, I rely on two initial criteria – that the selection event determines the state's leader, who is, as defined previously, the individual who is ultimately responsible for foreign policy. In this way, elections for ceremonial presidents that have no impact on foreign policy are excluded. Second, the selection event must be an

election, appointment, succession, or self-selection (whether a success or failure). I exclude auto-coups from my cases since it is not clear who the counterfactual leader is in these cases.

For practical and methodological reasons, I limit my analyses to the leader and counterfactual leader, as opposed to third-place candidates in elections. While selection events such as elections are not strictly contests between two candidates, in practice however, voters are often heavily incentivised to vote for one of the two most competitive options.[12] Furthermore, since runners-up are the closest individual to being the counterfactual leader, there is little research benefit to include third-place candidates.

I also limit the scope of my analysis to the years between 1945 and 2004, primarily due to data availability. Many of the candidates before this period lack sufficient biographical information, especially those in smaller states, and including them would require significant additional research effort for little theoretical and statistical benefit. Additionally, many important demographic variables, such as gender equality for sex and life expectancy for age, are missing across many countries before 1945. The end date of 2004 is because leaders of states are taken from the LEAD dataset which ends in 2004. Including more individuals after this date would be ideal, but the latest is 2015 since that is the last observation year in the Militarized Interstate Disputes dataset, which is where the outcome – conflict initiation – is drawn. While expanding the scope to years before 1945 and after 2004 would increase the external validity of my findings, limiting my scope to between these years was simply more research efficient and made the project feasible within the DPhil timeframe and resources. There would likely be a significant amount of missing information on candidates before 1945, and so would ultimately be excluded in the main statistical analysis anyway. Including years after 2004 would require me to identify each leader in every state, which could result in a difference between LEAD and Archigos in how leaders are identified and coded. Furthermore, this would only add ten years to the data, and there is little

[12] Gary W. Cox, *Making Votes Count: Strategic Coordination in the World's Electoral Systems* (Cambridge University Press, 1997).

reason to suspect substantial changes in leader selection during this ten-year period that might significantly alter the results.

Ultimately, the LSD dataset includes 2,680 leader selection events across 1,738 individual leaders. I determined the selection event in 98% of cases and identified the counterfactual leader in 63% (n=1,709) of cases. Most missing counterfactual leaders were in the appointment mechanism since Heads of State rarely acknowledge who they almost chose.

1.6.2 Empirical Analysis

I utilise quantitative methods to test whether traits or selection better explain why leaders start conflicts. More specifically, I use statistical regression techniques using the trait of the counterfactual leader as a "negative control," or placebo, to infer whether there is evidence for any of the selection effects. If the association between leader traits and conflict initiation depends on the trait of the counterfactual leader, then this is considered sufficient evidence that traits on their own does not explain why leaders initiate conflict. To further strengthen the evidence for any selection effects, I also use statistical methods to investigate whether candidates are more likely to have high-risk than low-risk traits conditional on several explanatory variables, such as past international conflict and demographics. Furthermore, I also use statistical techniques to investigate whether candidates with high-risk traits are more likely to be chosen over those with low-risk traits.

1.6.3 Why No Case Study?

In mixed method approaches, such as in a nested analysis, quantitative results are often complemented with one or more case studies to more thoroughly investigate the causal mechanism between the explanatory variables and outcome variables. Standard practice is to conduct "on-the-line" cases that support key hypotheses in the regression results, and sometimes also extreme

or deviant cases that do not follow the broader pattern.[13] However, the strongest finding in my analysis is that demographics often predict the selection of candidates with high-risk traits. For example, candidates are more likely to be ex-military in conflict-prone states, because these states tend to have larger share of veterans in the population (meaning leaders turn out to be more likely to be veterans by chance). Thus, the key explanatory variable of interest here, the percentage of veterans in the population, is not an appropriate causal mechanism to investigate using traditional case study methods. That is, it does not make sense to conduct interviews or process-trace a population percentage, especially when many of these cases can be easily explained by the data on conscription policies throughout this period, such as in Israel and France. Even in states with limited conscription years, such as the United States, there is unlikely to be sufficient information to process-trace whether party elites or voters wanted ex-military candidates during this period, among all the other factors feeding into such decisions. Thus, since the causal mechanism is a probability (i.e., the share of the population with a high-risk trait increases the probability of candidates from which they are drawn also having a high-risk trait, which then increases the probability of selecting a leader with a high-risk trait), there is no reason to conduct a case study on such an obvious causal mechanism. In short, the research questions can be sufficiently answered, and the causal mechanism accounted for, using quantitative methods alone.

1.7 Main Findings

I find leaders with high-risk traits likely have no effect on conflict initiation, in contrast to previous research. Rather, it appears that states that frequently engage in conflict are more likely to have candidates with high-risk traits due to demographics and thus have a higher probability of selecting a high-risk leader, even if by chance. More specifically, leaders with high-risk traits are more likely to initiate conflicts than low-risk leaders *only* when the counterfactual leader has the

[13] Evan S. Lieberman, 'Nested Analysis as a Mixed-Method Strategy for Comparative Research', *American Political Science Review* 99, no. 3 (2005): 435–52.

same trait (suggesting candidate selection), but not when the counterfactual leader has the opposite trait (suggesting no trait effect). This finding means that when states could have had a candidate with either a high-risk or a low-risk trait, the individual that ultimately became leader did not cause any significant difference in conflicts. For example, in elections where the top two candidates were a veteran and non-veteran, electing the former does not result in more conflicts than the latter, since the environment generates the level of conflict, not the leader. These results are strongest for several military experience variables and age, but weaker for ideology due to too few observations. Furthermore, I find no selection effects for female leaders, although again, there are too few observations to be confident in these findings. In any case, considering the expectation that leader traits should be associated with conflict regardless of the counterfactual leader's trait, I argue this is sufficient evidence that nearly all of the high-risk traits likely have no effect on conflict initiation and are instead confounded by selection. Furthermore, I also find that selectorates are not more likely to choose high-risk over low-risk candidates in more conflict-prone states, but this lack of an effect is mostly because there are few opportunities for them to do otherwise. Again, candidates in the same selection events tend to have the same traits due to demographics, so voters rarely have a choice in the matter. As a result, even in a counterfactual world where all the runners-up were selected instead, we would likely still see a disproportionate number of leaders with "high-risk" traits in states where the baseline probability of conflict is greater. Essentially, conflict in the international system brings warriors into office, not the other way around.

2 Do Leaders Matter?

When understanding how leader traits might influence leaders to start conflicts, we must first understand why and how leaders and their traits matter across international politics more generally. In other words, the question of whether leader traits matter for international conflict fits within the larger debate in IR summed up as "Do leaders matter?". Given that this is the overarching debate this dissertation seeks to contribute, in this chapter I help to clarify how leaders may or may not matter in IR.

No IR scholar argues that leaders have no role to play when trying to understand some international political outcome. After all, leaders are usually – but not always – the ultimate decision-making authority for a state's foreign policy,[14] and take on what Renshon and Renshon term the "magisterial role" – "presiding at the apex of the decision structure, allocating resources and ultimately, making the final decision."[15] Beyond directly making foreign-policy decisions, leaders can also set the strategic vision of their foreign policy that subordinates then carry out or simply carry on their predecessors' policies.[16] In other words, nearly every critical foreign policy outcome is traceable to a leader's decision-making.

Nowhere, though is leaders' decision-making more influential than in war and international conflict. Every international conflict begins with a single use of force by military troops, and this is nearly always because of a direct order by their superiors and so on up the chain of command until it eventually stops at the leader. Outside of directly ordering conflict, as typically seen in war declarations, leaders can start conflicts through more indirect orders, such as setting an

---

[14] Margaret G. Hermann and Charles F. Hermann, 'Who Makes Foreign Policy Decisions and How: An Empirical Inquiry', *International Studies Quarterly* 33, no. 4 (1989): 361–87.
[15] Jonathan Renshon and Stanley A. Renshon, 'The Theory and Practice of Foreign Policy Decision Making', *Political Psychology* 29, no. 4 (2008): 526.
[16] Many authors make this point that leaders should be considered the fundamental unit of analysis. Scott Wolford, 'The Turnover Trap: New Leaders, Reputation, and International Conflict', *American Journal of Political Science* 51, no. 4 (2007): 772–88; Horowitz and Fuhrmann, 'Studying Leaders and Military Conflict'.

aggressive force posture that inevitably results in a border skirmish. Indeed, if the leader did not explicitly or implicitly approve (i.e., not condemning or apologising for) such incidents, they would not be considered an international conflict.[17] All this means is that, at minimum, leaders are a necessary condition in the complex causal chain leading to conflict initiation. Or, in short, leaders are causal mechanisms.

While no IR scholars would dispute that leaders are causal mechanisms, where there is considerable uncertainty is the extent to which leaders are causal variables. If we could switch around leaders randomly, would we expect to see the same political outcomes? Of course, we can never know this. But, even if we could, it still would not tell us why Adolf Hitler or Saddam Hussain, for instance, appeared so much more willing to use force than other leaders. So, we are ultimately interested in the variety of factors that influence leaders' decision-making to start international conflicts, either through direct orders or indirectly through changing toward a more aggressive force posture. Figure 2-1 below visualises this framework within the context of the central puzzle of this dissertation.

[17] Daniel M. Jones, Stuart A. Bremer, and J. David Singer, 'Militarized Interstate Disputes, 1816–1992: Rationale, Coding Rules, and Empirical Patterns', *Conflict Management and Peace Science* 15, no. 2 (1996): 168.

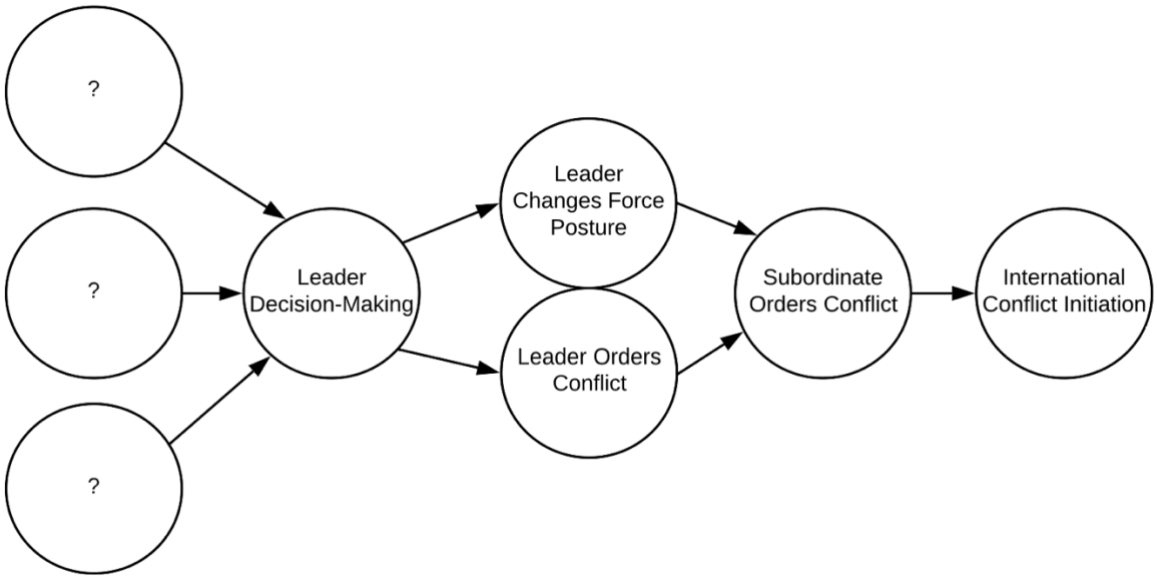


Figure 2-1: Directed Acyclic Graph (DAG) of the research question.

While numerous possible sources influence a leader's decision-making, generally they can fit within one of two distinct and independent groups. The first group of sources springs from the leader's environment – the broad range of factors that are external, or exogenous of the leader and are essentially the more important institutional and structural factors underlying many IR theories, such as relative power, alliances, regime type, norms, and rivalries. More specifically, these various environmental factors influence a leader's decision-making through the incentives and constraints they create. Military capabilities, for instance, are inert. Still, their quantity and quality can incentivise leaders to use force (e.g., military-industrial complex) or constrain them (e.g., less likely to start conflicts if the military is weak or abolished). Similarly, incumbent leaders might be motivated to start conflicts if it helps their re-election chances due to a rally around the flag effect. In essence, it is difficult to imagine how the various factors of the environment cannot influence a leader's decision to start conflicts in some form or another.

The second source of a leader's decision-making is their traits – the broad range of factors that are internal, or endogenous, to the leader, such as their biological attributes, life experiences, and personality. Concerning starting international conflicts, traits primarily influence leaders'

decision-making by shaping their risk tolerance. Since every international conflict has the potential to escalate to war, starting them is obviously quite risky for leaders, given that failure can ultimately result in their removal from office or even eventual death. Accordingly, we might expect leaders with traits associated with risk-taking more generally, such as young age, military experience, or psychological pathologies, to adopt more risky foreign policy decisions, like starting conflicts and wars.

Of course, how each source influences a leader's decision-making is complex, and the various traits and environmental factors are not always completely independent. I will get to how both can influence each other through selection in the next chapter. For now, it is only necessary to highlight that these two sources have an independent effect on a leader's behaviour, in theory. That is, if we could randomly assign individuals from across the world to be leaders of any state, their age and their state's military power, for instance, would produce two independent causal effects on their decision-making. One does not directly or indirectly cause the other. All this means is that leader traits and environmental factors are, in causal inference terms, two groups of competing exposures. They both exert an independent causal effect on a leader's decision to start conflicts, but to different degrees. However, the fundamental issue in this debate is not whether the environment influences a leader's decision-making more than their traits – that much seems to be agreed upon by all IR scholars – but rather the extent to which leader traits even matter. More specifically, in our case, we are interested in the effect of leader traits on conflict initiation regardless of environmental factors. Thus, the "do leaders matter" debate in IR is more appropriately a debate as to whether *leader traits* matter.

Where people may side in this debate is best exemplified by how they might answer a hypothetical question for any given conflict. If leader *A* had been replaced with individual *B*, would conflict *Y* still occur? This hypothetical question lays at the heart of the debate as it forces us to choose between a dichotomy of yes/no based on what we believe are the most influential factors that ultimately shape a leader's behaviour. To illustrate with a more famous

example of this question: had Al Gore won the 2000 U.S. Presidential Election instead of George Bush, would the US still have invaded Iraq? Given that all relevant environmental factors are essentially the same in this counterfactual (ignoring potential party effects for the moment), we would only need to consider two things: whether there was an important difference in the traits between Gore and Bush and whether any difference in their traits would be powerful enough to overcome the institutional and structural forces pushing the U.S. towards war. Considering that most people tend to believe the Gore-peace counterfactual in the belief that Gore had a more dovish personality than Bush,[18] the trait perspective might appear to be widespread in IR, at least in this specific case.

Yet, we might hold a different perspective when contemplating the explanatory power of leader traits across states. Consider another question: During the Cold War, why did U.S. Presidents initiate fifty-nine conflicts and Costa Rican Presidents only initiate one?[19] Most IR scholars would probably believe their different environments can explain this disparity. The US was a superpower, vied with the U.S.S.R. for economic and political supremacy, and hence had various interstate rivals, which likely motivated American leaders to adopt a more interventionist and aggressive foreign policy. Conversely, Costa Rica was a small state that had abolished its military at the outset of the Cold War and had little reason to be overly aggressive toward its neighbours or any major powers. Even if we could re-run history and switch the Presidents between the two countries, it is difficult to imagine that Costa Rican Presidents would have initiated fifty-nine conflicts and U.S. Presidents only one. The leaders in the two countries came to power in two entirely different environments, and this is probably the simplest explanation for the disparity in conflict initiations. Thus, we might believe that any variation in the traits of the Presidents between both countries likely only explains very little, if any.

---

[18] Frank P. Harvey, *Explaining the Iraq War: Counterfactual Theory, Logic and Evidence* (Cambridge University Press, 2011).
[19] According to MIDB v5.0. See: Jones, Bremer, and Singer, 'Militarized Interstate Disputes, 1816–1992'.

The above examples help to illustrate two ontological positions in the debate as to whether leader traits matter. First, we might believe that leader traits are primarily irrelevant and have little to no explanatory power. This view is what I refer to as the "environmental perspective." To explain differences in leaders' behaviour, institutional and structural factors are necessary and sufficient. Gore would have probably still invaded Iraq, and the U.S. would likely have initiated around fifty-nine times more conflicts than Costa Rica even if both countries' Presidents were randomly selected. Of course, the environmental perspective still views leaders as necessary, but their explanatory power is limited mainly to being a mechanism that transmits the causal energy from various external sources into action. In other words, the environmental perspective views leaders as important causal mechanisms rather than causal variables, although perhaps outside of extreme personalities such as Hitler.

Conversely, we might believe traits have a much more significant role in explaining why leaders start conflicts. This view is what I call the "trait perspective." We may not necessarily believe that traits have more explanatory power than environmental factors, but rather only believe that leader traits can help to "fill in the gaps" where an environmental perspective would struggle. Gore might not necessarily have invaded Iraq, and there could have been a more peaceful U.S. or warlike Costa Rica if the presidents between the two were switched. Perhaps the strongest examples of the trait perspective are many of the more aggressive leaders in history, like Napoleon and Hitler, as it is hard to imagine that France or Germany would have embarked on a similar road of conquest and destruction had someone else been leader during the time. Thus, the trait perspective views leaders not only as causal mechanisms but essential causal variables.

Framing this debate into two competing perspectives follows what others have observed and similarly framed as "mirroring/steering" and "situational/dispositional."[20] Regardless of

[20] Dyson and Briggs, 'Leaders and Foreign Policy'. Joshua D. Kertzer, *Resolve in International Politics*, vol. 2 (Princeton University Press, 2016).

specific terminology, all these distinctions (and mine) try to illustrate that both perspectives exist at polar ends of a spectrum and compete to explain better why leaders start conflicts. However, environmental and trait explanations are not necessarily mutually exclusive since both can interact with each other in complex ways. Indeed, other leadership disciplines outside of IR use similar trait-environment frameworks, yet the common perspective is the interactionist view that any explanation of human behaviour can only be understood as the interaction of factors between these two groups.[21] In leadership theories of management, for example, contingency theory (also called "it all depends" theory of leadership) is the dominant perspective, which naturally assumes then that leaders' behaviour is the product of complex interactions between their traits and their situation.[22] However, interactionism is more an ontological position as it does not necessarily mean that environmental factors will always have the same causal effect as trait factors and vice versa. Rather, the interactionism framework is important in the sense that any study of human behaviour must first consider both trait and environmental factors together before we can parse out their respective causal effect.

Of course, most IR scholars would probably not place themselves entirely at the environmental or trait end but rather somewhere in between, although I suspect most would probably identify more towards the environmental end of the spectrum.[23] In any case, some authors frame the debate over whether leader (traits) matter as one where the two sides strongly disagree, although I do not find this to be the case.[24] Instead, those that might otherwise agree

---

[21] Timothy A. Judge and Cindy P. Zapata, 'The Person–Situation Debate Revisited: Effect of Situation Strength and Trait Activation on the Validity of the Big Five Personality Traits in Predicting Job Performance', *Academy of Management Journal* 58, no. 4 (2015): 1149–79; John F. Kihlstrom, 'The Person–Situation Interaction', 2013.

[22] Roya Ayman, 'Situational and Contingency Approaches to Leadership.', 2004; Francis J. Yammarino and Shelley D. Dionne, 'Leadership and Levels of Analysis', *What's Wrong With Leadership*, 2018.

[23] Dominic Johnson makes a similar point more generally that social scientists appear more willing to accept that human behaviour is entirely the product of socialisation rather than biology. Dominic DP Johnson, 'Survival of the Disciplines: Is International Relations Fit for the New Millennium?', *Millennium* 43, no. 2 (2015): 756.

[24] For example, see: Horowitz and Fuhrmann, 'Studying Leaders and Military Conflict'.

with the environmental perspective are largely ambivalent about the explanatory power of leaders.[25] In most quantitative IR analyses, leaders and their traits are not the main explanatory variables, let alone included as controls. Although, to be fair, large-*n* data on leaders was unavailable until the mid-2000s, so their absence may be due to this limitation rather than an ontological belief of their unimportance. Nevertheless, we can visualise our central research question in the context of the trait vs environment debate below in Figure 2-2 (simplified version of Figure 2-1 above). International conflict is an outcome of a leader's decision decision-making, jointly determined by both the environment and their traits.

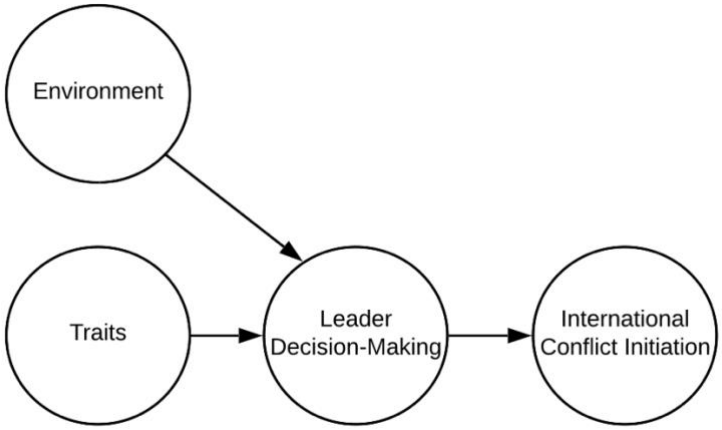


Figure 2-2: Causal Diagram of the trait vs environment framework.

Within this trait vs environment debate, we can distinguish between two types of studies that add evidence to one or the other. Environmental explanations find strong relationships between institutional and structural variables and conflict initiation, whereas trait explanations find convincing and important relationships between leader traits and conflict initiation.

Most research on international conflict in IR appears to support environmental explanations. Although, the environmental explanations are better understood as "null" explanations – as opposed to being in direct opposition to trait explanations – because they are

[25] Jervis, 'Do Leaders Matter and How Would We Know?'

the dominant focus of most IR research, as leader-level variation is rarely included or compared against. Indeed, between 1980 and 2017, only 13% of the 9,116 articles in the top IR journals even focused on the individual level, and only 5% of these had the word "leader" somewhere in the title.[26] In comparison, the systemic level was the focus of 58% of articles. Even though these studies often frame their theories in terms of "state behaviour", it is otherwise just shorthand for either a leader's decision-making, or the aggregate decision making of political elites.[27] In other words, they argue that states start conflicts because of some institutional and structural constraint/incentive, but this is essentially synonymous with leaders since the only agency states might have been through the actions of people.

Even studies where leaders are the central unit of analysis are still better understood as an environmental explanation. The key causal variables in these studies are not leaders but the domestic or structural factors that ultimately trigger leaders to act in predictable ways. Most of the studies in this group revolve around leader survivability, which generally investigates whether incumbents start conflicts to remain in office. Selectorate theory, for instance, argues that leaders in states with large winning coalitions (i.e., leaders in pluralistic democracies) face more accountability from those who keep them in power than leaders with small winning coalitions (i.e., dictators) and so should be more selective, and hence less likely, to start conflicts and wars.[28] Likewise, audience costs argue that leaders with larger audiences (i.e., millions of voters compared to a cadre of elites) should be less likely to initiate conflicts because backing down comes with greater political costs to the incumbent.[29] Lastly, diversionary theory and its sister

---

[26] 'Teaching, Research, and International Policy Project', TRIP Journal Article Database Release (Version 3.3), accessed 22 March 2023, https://trip.wm.edu/data/dashboard/journal-article-database.

[27] Ross James Gildea, 'Psychology and Aggregation in International Relations', *European Journal of International Relations* 25th Anniversary Special Issue (2020): 1–18.

[28] Bruce Bueno De Mesquita et al., 'Testing Novel Implications from the Selectorate Theory of War', *World Politics* 56, no. 3 (2004): 363–88.

[29] For the general theory of audience costs, see: James D. Fearon, 'Domestic Political Audiences and the Escalation of International Disputes', *American Political Science Review*, 1994, 577–92. For empirical support of audience costs, see: Shuhei Kurizaki and Taehee Whang, 'Detecting

theories, rally around the flag and gambling for resurrection, essentially argue that incumbent leaders should be more or less likely to initiate conflict if they believe it increases their re-election chances.[30] While these types of theories and studies explicitly model leaders into their analyses, they assume that all leaders have roughly the same motivation to stay in power, and so are not causal variables. Indeed, even if leaders were selected at random, these theories would expect to achieve the same findings. Thus, their key causal variables are not leader traits but environmental factors that constrain or incentive leaders to start conflicts.

Another vein of leader-centric research in IR that is better understood as an environmental explanation are cognitive approaches to leader conflict behaviour. Neoclassical realists, for example, often argue that various cognitive biases, such as misperceptions about interstate threats, can have large impacts on leaders' behaviour and hence foreign policy decisions about starting conflicts.[31] While it might appear strange to consider these types of studies as an environmental rather than trait explanation, the cognitive mechanisms are often assumed to be constants or universals across all leaders and only triggered into action by their situations. While the strength of the cognitive biases does, of course, vary among leaders, that variation is not explicitly operationalised or measured in their analyses. Rather, the authors tend to focus on how these cognitive biases manifest in the same way across a variety of leaders. However, since all leaders are assumed to be susceptible to these errors or biases, the causal variables here are the environments in which certain biases or errors manifest, not the biases *per se*.

Classical realist arguments about human nature are also counterintuitively best understood as an environmental explanation. The more refined versions of the human nature argument are based on evolutionary psychology, arguing that humans, and hence leaders, have evolved

---

Audience Costs in International Disputes', *International Organization* 69, no. 4 (2015): 949–80. For opposing findings, see: Casey Crisman-Cox and Michael Gibilisco, 'Audience Costs and the Dynamics of War and Peace', *American Journal of Political Science* 62, no. 3 (2018): 566–80.

[30] Giacomo Chiozza and Hein Erich Goemans, *Leaders and International Conflict* (Cambridge University Press, 2011).

[31] Robert Jervis, *How Statesmen Think* (Princeton University Press, 2017).

behaviours that can predispose them to be more willing to use force, but the central assumption is that these evolved behaviours are only triggered in certain environments.[32] Since both human nature and an evolved psychology are by definition human universals, they then cannot be causal variables. Again, while the strength of the evolved predispositions undoubtedly varies between leaders, this variation is not explicitly modelled and is instead assumed to affect all leaders in roughly the same way, much like the cognitive bias approaches. Thus, even in cognitive and evolutionary approaches to IR, where a leader's psychology is placed front and centre, leaders are not analysed as causal variables but rather only causal mechanisms that convey universal decision-making errors or consistency from cognitive biases or evolved predispositions.

2.1 Leaders as Causal Variables: Trait Explanations of International Conflict

Those who believe leaders are powerful causal variables do not reject the greater importance of the environment in explaining leaders' decision-making but posit that different leaders placed into the same environments will behave quite differently and have a much larger influence than what the majority of IR otherwise assumes. Perhaps the key difference between environmental and trait perspective, then, is that the former tends to assume all leaders have roughly similar rationality, whereas the latter tends to assume leaders' rationality is significantly bounded by their traits. This is not to say those who might favour the trait perspective believe that leaders are not rational. Rather, leaders will obviously weigh the costs and benefits of starting conflicts given their environment, but these calculations will also be quite distorted by any biological or socialised predispositions toward risk-taking, aggression, or any other proclivities that might make them more willing to use force. Likewise, those who favour

[32] Dominic DP Johnson and Monica Duffy Toft, 'Grounds for War: The Evolution of Territorial Conflict', *International Security* 38, no. 3 (2013): 7–38; Anthony C. Lopez, 'The Evolutionary Psychology of War: Offense and Defense in the Adapted Mind', *Evolutionary Psychology* 15, no. 4 (2017): 1474704917742720; Bradley A. Thayer, *Darwin and International Relations: On the Evolutionary Origins of War and Ethnic Conflict* (University Press of Kentucky, 2009).

environmental explanations do not reject that leaders' decision-making is bounded by their various traits in some way. Only that any potential causal effect traits have on a leader's decision-making will be largely drowned out by the effects of the various structural and domestic variables stemming from their environment.

2.2 Which Traits Explain Why Leaders Initiate International Conflicts?

In a perfect world where we could observe and measure every facet of a leader's mind, we would ideally only be interested in one trait, which we might otherwise call their "willingness to use force." This would be best understood as a latent variable that is itself a complex product of many aspects of a leader's personality that cause leaders to be more likely to start conflicts, such as aggressiveness, risk-proneness, or resolve, to name a few. Now, if we could measure this trait without error, then we could perfectly test the extent to which leader traits matter by simply regressing it against conflict initiation (assuming there are no selection effects – more on that in the next chapter). If we could run this hypothetical design with perfectly measured variables, the environmental perspective would predict that leaders with a low willingness to use force in powerful states should be more likely to initiate conflict than leaders with a high willingness to use force in weak states because the constraint of low relative power is too great for aspiring aggressive leaders to overcome. Put differently, Hitler probably would have been an irrelevant footnote of history had he been the leader of Luxembourg, and Gandhi probably would not have been remembered as being a paragon of peace had he been the leader of the Soviet Union during the Cold War. Conversely, the trait perspective might predict that Luxembourg Hitler's ambitions would have been more constrained than German Hitler's ambitions, but probably would still be remembered as an aggressive leader, and Soviet Union Gandhi would have been remembered as being one of the more peaceful Soviet leaders.

While this hypothetical experiment of a leader's "willingness to use force" would tell us whether leader traits matter, we are still kicking the can down the road as to why some leaders

would even have a higher willingness to use force than others to begin with. That is, a leader's willingness to use force has to be caused by *something*, so we need to find out what those factors are. The fundamental problem, however, is that a leader's willingness to use force and many underlying traits are unobservable or extremely difficult to measure at best, so leader trait studies tend to use instead a variety of observable and easier-to-measure traits that are often empirically correlated with aggression or hawkish beliefs among people more generally. The logic here is obvious; if these traits cause ordinary people to be more aggressive – broadly defined – then we should also expect to see a correlation between these traits among state leaders and "aggressive" actions in the realm of IR, such as starting international conflicts and wars.

The key aim for trait explanations, then, is to identify which backgrounds and attributes of leaders might be associated with starting conflicts depending on what psychological research suggests. That is, given that there must be a theoretical justification to hypothesise why a certain trait would cause leaders to be more conflict-prone, trait studies often rely heavily on the voluminous literature across psychology to justify their research aim and hypotheses. Accordingly, one important assumption for the trait perspective is that the underlying psychological mechanisms for a given trait act the same way among both ordinary people and state leaders. Scholars in other disciplines can run experiments to get a better sense of the underlying causal mechanism for many traits and can often directly observe the neurological and neuroendocrinological effects. However, even with the strength of these experimental designs, science is still hotly debated on exactly how many of the more common traits affect human behaviour. The simple association between higher levels of testosterone and aggressiveness, for instance, is well documented across countless experimental studies in animals and humans, yet there is still debate as to how exactly testosterone "causes" aggression or whether it even does at all.[33] In IR, we are even more limited in understanding the mechanistic link from trait to action

[33] Angela S. Book, Katherine B. Starzyk, and Vernon L. Quinsey, 'The Relationship between Testosterone and Aggression: A Meta-Analysis', *Aggression and Violent Behavior* 6, no. 6 (2001):

since we cannot run experiments or use electroencephalography devices (measure brain activity) when leaders make foreign policy decisions. Furthermore, most of the non-IR literature on the trait-to-action causal path is based on everyday social interactions where the participants act within eyesight of each other, which is obviously an entirely different environment than when leaders decide whether to start a conflict or war against a foreign state. So, to even theorise why certain traits should be correlated with international conflict or any foreign policy outcome, we can only draw from psychological research and must assume that the underlying psychological mechanism affects leaders in roughly the same way as ordinary people. For now, I make no judgements on whether this extrapolation of the psychological mechanisms from ordinary people to leaders is appropriate. Rather, I only wish to highlight the fundamental assumptions the trait perspective relies upon before we begin to interrogate exactly what previous research says on the matter.

The last issue with trait explanations I wish to bring up is the methodological split in studying leader traits due to measurement limitations. Those interested in the perhaps more theoretically relevant measures of leader traits, such as those often found in Leadership Trait Analysis (LTA) – personality, trustworthiness, leadership style, and "need for power" – rely on small-*n* qualitative methods due to the difficulty in measuring such abstract concepts.[34] Conversely, nearly all of the existing quantitative research on leader traits and international conflict is limited to the perhaps less relevant but easier-to-measure demographic variables such as age, gender and military experience. The reason for this split is that LTA requires the use of

579–99; John Archer, 'Testosterone and Human Aggression: An Evaluation of the Challenge Hypothesis', *Neuroscience & Biobehavioral Reviews* 30, no. 3 (2006): 319–45; Robert M. Sapolsky, *Behave: The Biology of Humans at Our Best and Worst* (Penguin, 2017); Mike Martin, *Why We Fight* (Oxford University Press, 2018).

[34] Two such examples of the use of LTA in IR include: Samuel R. Rohrer, 'What Makes a Prime Minister Great?: A Leadership Trait Analysis of the Effectiveness of British Prime Ministers from 1902 to 2004', *Research & Politics* 1, no. 3 (2014): 2053168014558970; Juliet Kaarbo, 'Prime Minister Leadership Style and the Role of Parliament in Security Policy', *The British Journal of Politics and International Relations* 20, no. 1 (2018): 35–51.

content analysis from written sources and speeches, which are simply not available to most leaders. As a result, LTA is rare in quantitative studies of international conflict since it is far more practical for IR scholars to examine the easier-to-find demographic variables that also might have large effects on leader behaviour, in theory at least. Accordingly, since the question of whether leader traits matter is essentially a question of generalisability, I limit my focus for the rest of this review to only those traits that are supported by large-*n* quantitative analyses. Furthermore, I limit the scope of my review here to five of the more relevant individual traits across the literature: military experience, war experience, ideology, age, and sex, and two holistic traits that attempt to measure the combined effect of several dozen traits, the leader risk score and hawkishness index. While it would be ideal for investigating more traits, there are several reasons for limiting my analysis to these seven traits. First, these seven traits have the most well-documented theoretical justification in psychology for why we might expect variation in these traits among leaders to be correlated with international conflict initiation. Second, each of these traits has a variety of supporting evidence across a range of studies and authors. Lastly, there are other traits that scholars have investigated besides these seven, but they are generally not significantly related to conflict initiation or are not coded with explicitly stated criteria. In the following sections, I provide a more thorough review of the findings for each of the seven traits in turn.

2.2.1 Military Experience

Military experience has been perhaps the most researched leader trait in IR for obvious reasons; the military is one of the only socially acceptable professions where individuals specialise in the art of killing other humans. We might suspect, then, that being trained in the military likely has profound impacts on a leader's decision to start conflicts. Exactly how this experience can cause leaders to be more conflict-prone revolves around two competing hypotheses. The

first is the militarism hypothesis; given that the military socialises individuals to be more familiar and comfortable with military force, state leaders that have served in the military might view the military as an effective foreign policy tool and therefore be more likely to use military force abroad. This argument is not entirely surprising considering that not only is it intuitive, but there is sufficient empirical evidence from studies conducted among the general population.[35] Insofar as there is a hawkish bias among veterans, however, another key question is whether this bias is due to self-selection or socialisation. That is, whether more hawkish individuals tend to join the military or whether individuals tend to adopt more hawkish views because of being in a pro-military social environment. While the evidence is mixed, it does appear as though both sources are working in tandem; while more hawkish people do tend to self-select into the military, veterans also become more hawkish after their service than before.[36] In short, there is plenty of theoretical evidence to back up the militarism hypothesis that ex-military leaders should be more likely to start conflicts than leaders without previous military experience.

The counterargument, known more generally as "military conservatism theory", instead hypothesises that ex-military leaders should be less likely to start conflicts than leaders without prior military experience. Although this hypothesis should really be considered a subset of the militarism argument since it argues that only combat veteran leaders will be more conflict-averse than non-veteran leaders, but non-combat ex-military leaders will still be more conflict-prone than either.[37] The reason for this is intuitive; since combat veterans have seen the destructive

[35] For the pro-hawkish bias among veterans, see: Jonathan D. Klingler and J. Tyson Chatagnier, 'Are You Doing Your Part? Veterans' Political Attitudes and Heinlein's Conception of Citizenship', *Armed Forces & Society* 40, no. 4 (2014): 673–95. For studies that do not find a hawkish bias among veterans, see: Ole R. Holsti and James N. Rosenau, 'The Domestic and Foreign Policy Beliefs of American Leaders', *Journal of Conflict Resolution* 32, no. 2 (1988): 286.
[36] Jerald G. Bachman et al., 'Distinctive Military Attitudes among US Enlistees, 1976-1997: Self-Selection versus Socialization', *Armed Forces & Society* 26, no. 4 (2000): 561–85; David Stadelmann, Marco Portmann, and Reiner Eichenberger, 'Military Careers of Politicians Matter for National Security Policy', *Journal of Economic Behavior & Organization* 116 (2015): 142–56.
[37] Travis W. Endicott, 'Combat Experience and the Foreign Policy Positions of Veterans', *Social Science Quarterly* 101, no. 4 (2020): 1413–29.

horrors of war first-hand, they more acutely understand the great potential costs of using military force and so are probably more hesitant to start armed conflicts that could escalate.[38] This hypothesis has some empirical support among studies of the general population, albeit with slightly mixed findings and limited country-level variation. For instance, some studies have found support for military conservatism theory in the U.S.,[39] whereas others have found the opposite in Israel.[40] Additionally, in the United States, a higher percentage of veterans in the executive and legislative branches has led to a decrease in conflict initiation but also an increase in conflict escalation.[41]

Returning to how military experience might affect state leaders, the existing evidence largely supports the militarism argument, albeit with some nuance. The early work on the relationship between military experience and conflict initiation focused on regimes and not necessarily individual leaders, finding that states led by military regimes (and by default military leaders) are more likely to initiate conflict than those led by civilian regimes,[42] although one study has found that military regimes are less conflictual than states with weak civilian control of the military.[43] While it is difficult to assess in these studies whether having a military leader or being a military regime that is causally responsible for the association, more recent studies separating leaders from their regimes have found similar results. Across all leaders between 1945 and 2004,

---

[38] Samuel P. Huntington, *The Soldier and the State: The Theory and Politics of Civil–Military Relations* (Cambridge, Massachusetts: Harvard University Press, 1957).
[39] Richard K. Betts, *Soldiers, Statesmen, and Cold War Crises* (Cambridge, Massachusetts: Harvard University Press, 1979).
[40] Guy Grossman, Devorah Manekin, and Dan Miodownik, 'The Political Legacies of Combat: Attitudes toward War and Peace among Israeli Ex-Combatants', *International Organization* 69, no. 4 (2015): 981–1009.
[41] Christopher Gelpi and Peter D. Feaver, 'Speak Softly and Carry a Big Stick? Veterans in the Political Elite and the American Use of Force', *American Political Science Review* 96, no. 4 (2002): 779–93.
[42] Brian Lai and Dan Slater, 'Institutions of the Offensive: Domestic Sources of Dispute Initiation in Authoritarian Regimes, 1950–1992', *American Journal of Political Science* 50, no. 1 (2006): 113–26; Jessica L. Weeks, 'Strongmen and Straw Men: Authoritarian Regimes and the Initiation of International Conflict', *American Political Science Review*, 2012, 326–47.
[43] Todd S. Sechser, 'Are Soldiers Less War-Prone than Statesmen?', *Journal of Conflict Resolution* 48, no. 5 (2004): 746–74.

ex-military leaders who have never been in combat tend to be more conflict-prone than combat veteran and non-veteran leaders, although combat veterans are more likely to initiate conflict than non-combat and non-veteran leaders in autocracies and military regimes.[44] While some evidence suggests that non-combat veterans are indeed more likely than combat veterans to initiate conflict, both are still more likely than non-veterans, which casts doubt on the military conservatism argument.[45] In other words, the evidence largely supports the militarism argument since leaders with prior military experience, regardless of combat experience, tend to be more conflict-prone than leaders that have never served in the military. Although, there is some limited support for the military conservatism argument since some studies find that while combat veterans are more likely to initiate conflict than leaders with no military experience, the relationship is not significant when controlling for various environmental factors.

2.2.2 War Experience

Related to combat experience, whether a leader has ever fought in a successful or failed war appears to influence their decision to start conflicts as well. While this particular trait is less investigated than military experience (only two studies), leaders that have fought in a successful war are significantly more likely to initiate a conflict than leaders who fought in a lost war, as well as leaders who never fought in a war.[46] The causal logic here is essentially an extension of the military conservatism argument – leaders who have previously fought in a successful war do not perceive the costs and dangers of war to be as high when compared to those who have lost, so they are less risk-averse with using military force in the same way as non-combat veterans.

---

[44] Michael C. Horowitz and Allan C. Stam, 'How Prior Military Experience Influences the Future Militarized Behavior of Leaders', *International Organization* 68, no. 3 (2014): 527–59.
[45] Jaron Wharton, 'Three Essays on Decisions to Use Military Force' (PhD Dissertation, Duke University, 2018), 85; Horowitz, Stam, and Ellis, *Why Leaders Fight*.
[46] Horowitz, Stam, and Ellis, *Why Leaders Fight*, 67; Horowitz and Stam, 'How Prior Military Experience Influences the Future Militarized Behavior of Leaders'.

2.2.3 Age

Whether a leader's age influences their decision to start conflicts has also been of interest to IR scholars, given the important effect of age on human behaviour more generally. There are two competing hypotheses here. The first is the "young and reckless" hypothesis: younger leaders should be more likely to initiate conflict, given the well-documented cross-cultural evidence that younger individuals tend to commit the lion's share of violence and aggression more generally.[47] For example, the 20-29 age group has consistently been the majority of violent crime offenders in Europe for the last few centuries;[48] the majority of insurgents and terrorists around the world tend to be under the age of 40,[49] and there is a significant and positive relationship globally between the onset of civil war and the proportion of males between the ages of 15 and 29.[50] Thus, regardless of whether the underlying causal mechanisms are biological, experiential, or social, there is a significant volume of evidence to initially suspect the young and reckless hypothesis should also apply to state leaders as well.

Despite the large amount of evidence outside of IR for the young-and-reckless hypothesis, however, there are several reasons to suspect that older leaders might be more likely to start conflicts. The first is the "legacy hypothesis" – since older leaders are closer to death, they feel a more urgent need to fulfil a lasting legacy, which they might believe can be accomplished

---

[47] Margo Wilson and Martin Daly, 'Competitiveness, Risk Taking, and Violence: The Young Male Syndrome', *Ethology and Sociobiology* 6, no. 1 (1985): 59–73.
[48] Manuel Eisner, 'Long-Term Historical Trends in Violent Crime', *Crime and Justice* 30 (2003): 83–142.
[49] Edwin Bakker, 'Characteristics of Jihadi Terrorists in Europe (2001-2009)', *Jihadi Terrorism and the Radicalisation Challenge. European and American Experiences. Furnham, Ashgate*, 2011, 141; Sebastian Chávez Wurm, 'Arming the Revolution. The Shining Path and Its Struggle for Resources (1980–1993)', *Jahrbuch Für Geschichte Lateinamerikas* 49, no. 1 (2012): 261–84; Brent L. Smith and Kathryn D. Morgan, 'Terrorists Right and Left: Empirical Issues in Profiling American Terrorists', *Studies in Conflict & Terrorism* 17, no. 1 (1994): 39–57.
[50] Paul Collier, Anke Hoeffler, and Dominic Rohner, 'Beyond Greed and Grievance: Feasibility and Civil War', *Oxford Economic Papers* 61, no. 1 (2009): 1–27.

through engaging and winning international conflicts.[51] The second is the "political experience" hypothesis – since older leaders generally have more political experience than younger leaders, they will tend to face fewer constraints from political elites when crafting foreign policy generally and starting conflicts in particular.[52]

While there are only a few studies that directly investigate the link between age and conflict, the evidence appears to support the legacy hypothesis. Several studies find that older leaders are more likely to initiate conflict than younger leaders, especially in democracies.[53] The effect of a leader's age on conflict initiation appears to be quite large as well; an 80-year leader is around 200% more likely to initiate conflict than a 30-year-old leader.[54] Furthermore, leaders with less political experience are more likely to initiate conflict, at least in the case of U.S. and U.K. leaders, casting doubt on the political experience hypothesis.[55]

Beyond the association between advanced age and conflict initiation, however, there is little supporting evidence for the legacy hypothesis. Horowitz and Stam, for instance, rely solely on anecdotes about how Ronald Reagan and Menachim Begin's decreasing health contributed to their foreign-policy aggression, yet according to their own data, both leaders initiated fewer conflicts as they got older, not more.[56] Additionally, if leaders do indeed have the desire to leave

---

[51] Michael Horowitz, Rose McDermott, and Allan C. Stam, 'Leader Age, Regime Type, and Violent International Relations', *Journal of Conflict Resolution* 49, no. 5 (2005): 661–85; Horowitz, Stam, and Ellis, *Why Leaders Fight*.

[52] Horowitz, McDermott, and Stam, 'Leader Age, Regime Type, and Violent International Relations'; Horowitz, Stam, and Ellis, *Why Leaders Fight*.

[53] Horowitz, McDermott, and Stam, 'Leader Age, Regime Type, and Violent International Relations'; Daehee Bak and Glenn Palmer, 'Testing the Biden Hypotheses: Leader Tenure, Age, and International Conflict', *Foreign Policy Analysis* 6, no. 3 (2010): 257–73; Philip BK Potter, 'Does Experience Matter? American Presidential Experience, Age, and International Conflict', *Journal of Conflict Resolution* 51, no. 3 (2007): 351–78; Costel Calin and Brandon Prins, 'The Sources of Presidential Foreign Policy Decision Making: Executive Experience and Militarized Interstate Conflicts', *International Journal of Peace Studies* 20, no. 1 (2015): 17–34.

[54] Horowitz, McDermott, and Stam, 'Leader Age, Regime Type, and Violent International Relations'.

[55] Potter, 'Does Experience Matter?'; Calin and Prins, 'The Sources of Presidential Foreign Policy Decision Making'.

[56] According to the WLF data, Reagan and Begin did not initiate any conflicts in their last year in office.

a legacy before they exit office, then this should affect democratic leaders of all ages due to short-term limits. French Prime Ministers during the Third Republic, for instance, left office after 288 days, on average, so there should be increased pressure for French leaders of any age to leave a legacy, given that they knew they probably had little time to do so.[57] Also, it is not entirely clear why leaders should think pursuing conflict rather than peace would lead to a more lasting personal legacy. Some of the most aggressive leaders in recent history, such as Hitler and Saddam Hussein, are not favourably remembered, to put it lightly. Thus, there is sufficient reason to question the association between advanced age and conflict. Of the few studies that have found a link between advanced age and conflict initiation, the proposed causal mechanisms for this relationship are either contradicted by evidence or simply illogical. Given the apparent large effect of advanced age on conflict initiation, it is therefore important to investigate other potential explanations for this counterintuitive association.

2.2.4 Biological Sex

Another major trait that IR scholars have investigated is biological sex. Given the numerous sex differences in human behaviour across nearly every human activity, it is reasonable to theorise we might expect similar sex differences among leaders and starting conflicts. Like the other traits, there are again two competing hypotheses here. The first is that male leaders should be more likely to initiate conflict than female leaders – the "aggressive male" hypothesis. This hypothesis is perhaps more intuitive considering that not only do males tend to engage in more violence and aggressive behaviour than females across nearly every socio-cultural context, but there is cross-national evidence that males are more likely than females to support militaristic policies as well.[58]

[57] Data on French Prime Ministers taken from the LEAD dataset and is the mean tenure of all leaders who came into office between 1870 and 1940.
[58] Richard C. Eichenberg, *Gender, War, and World Order: A Study of Public Opinion* (Cornell University Press, 2019).

Some IR scholars even argue that democratic peace theory is at least partly explained by the increase in women's suffrage and increased political representation.[59] The causal mechanisms underlying the aggressive male hypothesis are often based on a combination of biological mechanisms, such as evolved neuroendocrinological differences, and/or sociological mechanisms, such as socialisation into gender roles.[60] Regardless of the underlying causal mechanism, though, the existing evidence seems to suggest we should expect to see male leaders being more likely to start conflicts than female leaders. Indeed, as Boris Johnson recently said, "If Putin was a woman…I really don't think he would have embarked on a crazy macho war of invasion and violence the way that he had."[61]

A minority view, however, argues that female leaders could be more likely to start conflicts. This is the "female posturing" hypothesis, which more specifically argues that since international politics is male-dominated, female leaders feel extra pressure to signal their more masculine traits, such as assertiveness and toughness, and are thus more likely to start conflicts than male leaders.[62]

In general, the evidence largely supports the female posturing hypothesis, albeit with slightly mixed results. Several separate studies find that female leaders are more likely to initiate conflict than male leaders, although this relationship is often not statistically significant due to the low sample size of female leaders.[63] Additionally, among historical European Monarchs

---

[59] Francis Fukuyama, 'Women and the Evolution of World Politics', *Foreign Affairs*, 1998, 24–40; Joslyn N. Barnhart et al., 'The Suffragist Peace', *International Organization* 74, no. 4 (2020): 633–70.
[60] For a more sociological explanation, see: Joshua S. Goldstein, *War and Gender: How Gender Shapes the War System and Vice Versa* (Cambridge University Press, 2003). For a more biological explanation, see: Stephen Peter Rosen, *War and Human Nature* (Princeton, NJ: Princeton University Press, 2005).
[61] Christian Sievers, 'Johnson: Putins Krieg Hat Europa Gestärkt', ZDF, 28 June 2022, https://www.zdf.de/nachrichten/politik/boris-johnson-g7-gipfel-gemeinschaft-100.html.
[62] Jane S. Jaquette, 'States Make War', *Foreign Affairs* 78, no. 1 (1999): 129.
[63] Jeff Carter and Timothy Nordstrom, 'Term Limits, Leader Preferences, and Interstate Conflict', *International Studies Quarterly* 61, no. 3 (2017): 721–35; Horowitz, Stam, and Ellis, *Why Leaders Fight*; Abigail S. Post and Paromita Sen, 'Why Can'ta Woman Be More like a Man? Female Leaders in Crisis Bargaining', *International Interactions* 46, no. 1 (2020): 1–27.

between 1480 and 1913, Queens were more likely to start wars than Kings.[64] However, there is some limited evidence for the aggressive male hypothesis. Female leaders are more conflict-prone only in states with low levels of women's political representation and empowerment.[65] Additionally, Caprioli finds that female leaders are less likely to initiate force in conflicts than their male counterparts.[66] Although the reason for this lone contradictory finding appears to be due to a difference in samples, whereas other studies compare all leaders, Caprioli only compares dyads in which military force was used and thus excludes the more numerous cases where male leaders did not use force. This is not necessarily a problem as it can be considered a form of matching, but it highlights there is at least some uncertainty in trying to understand the effect of biological sex on a leader's decision to use military force abroad.

If the female posturing hypothesis is right, then this points to an important and puzzling relationship between female leaders and conflict. On the one hand, the effect size of female leaders on conflict initiation is relatively large, given stereotypes about male aggression and violence more generally. For instance, female leaders are more likely than military leaders to initiate conflict,[67] and also increase a state's likelihood of initiating conflict to a greater extent than being a major power.[68] On the other hand, there is an abundance of evidence showing that more gender-equal states are less likely to initiate conflict,[69] yet we would only really expect

---

[64] Oeindrila Dube and S. P. Harish, 'Queens', *Journal of Political Economy* 128, no. 7 (2020): 2579–2652.
[65] Jonathan Powell and Karina Mukazhanova-Powell, 'Demonstrating Credentials? Female Executives, Women's Status, and the Use of Force', *Journal of Women, Politics & Policy* 40, no. 2 (2019): 241–62; Madison Schramm and Alexandra Stark, 'Peacemakers or Iron Ladies? A Cross-National Study of Gender and International Conflict', *Security Studies* 29, no. 3 (2020): 515–48; Horowitz, Stam, and Ellis, *Why Leaders Fight*.
[66] Mary Caprioli, 'Gender Equality and State Aggression: The Impact of Domestic Gender Equality on State First Use of Force', *International Interactions* 29, no. 3 (2003): 195–214.
[67] Carter and Nordstrom, 'Term Limits, Leader Preferences, and Interstate Conflict'.
[68] Michael T. Koch and Sarah A. Fulton, 'In the Defense of Women: Gender, Office Holding, and National Security Policy in Established Democracies', *The Journal of Politics* 73, no. 1 (2011): 1–16.
[69] Caprioli, 'Gender Equality and State Aggression'; Katie Heaney, 'Gender Inequalities and International Military Aggression: The Role of Feminism in Achieving Peace', *Res Publica-Journal*

female leaders to come into power in states with higher levels of gender equality. Why, then, would female leaders be more likely to initiate conflict than male leaders, but gender-equality be negatively associated with conflict initiation? It is therefore important to investigate other possible explanations for this paradoxical finding.

2.2.5 Political Ideology

Another leader trait that scholars have investigated is political ideology. Obviously, a leader's political ideology influences their beliefs and foreign policy decision-making in many ways. Though, when it comes to using military force abroad, conventional wisdom and numerous studies suggest that right-wing leaders should be more likely to start conflicts. This is the "right-wing hawkishness" hypothesis and is perhaps the more dominant view given that right-wing beliefs tend to be associated with hawkish policy preferences more so than centre or left-wing beliefs.[70] So it is only logical that leaders with a political ideology that is generally associated with hawkishness will then make more hawkish foreign policy decisions, such as starting international conflicts. Relatedly, audience costs might also influence why right-wing leaders should be more conflict-prone than left-wing and centre leaders. That is, right-wing leaders understand that they could be punished by their right-wing supporters for what could be perceived as more dovish actions.[71]

While the alternative argument is that left-wing leaders should be more likely to initiate conflict, there are no scholars who put forth a "left-wing hawkishness" hypothesis. Rather,

---

of Undergraduate Research 14, no. 1 (2009): 9; Mary Caprioli and Peter F. Trumbore, 'Human Rights Rogues in Interstate Disputes, 1980–2001', *Journal of Peace Research* 43, no. 2 (2006): 131–48.

[70] Calin Costel, 'Hawks versus Doves: The Influence of Political Ideology on the Foreign Policy Behavior of Democratic States' (University of Tennessee, Knoxville, 2010).

[71] To my knowledge there are no sources which directly test this, although the audience costs logic seems to suggest this is a possibility. Joshua D. Kertzer and Ryan Brutger, 'Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory', *American Journal of Political Science* 60, no. 1 (2016): 234–49.

some have suggested that incentives to moderate might push liberal leaders to be more hawkish and conservative leaders to be more dovish.[72] If this were true, then we could see left-wing leaders being more conflict-prone if the moderation incentives were strong enough, although this has not been directly tested. In any case, outside of this sole argument, there do not appear to be any scholars or studies arguing that we should see left-wing leaders being more likely to initiate conflict than right-wing leaders.

The majority of research supports the hawkish right-wing leader hypothesis, unsurprisingly. The clearest evidence of a link between right-wing leaders and conflict initiation is Bertoli *et al.'s* regression discontinuity design based on close elections, finding that "barely" electing right-wing leaders (i.e., by winning an election with less than 2% of the total vote share over a left-wing/centre candidate) lead to a 25% increase in a state initiating a high-level conflict than barely electing left-wing or centre leaders.[73] Beyond leaders, several studies find that right-wing governments are more likely to start conflicts than left/centre governments.[74] So, considering that right-wing governments lead to increased conflict, we can think this relationship will probably manifest among state leaders as well, assuming that there is little meaningful difference between leaders' and parties' ideology. However, there are only a few studies that investigate the link between leaders' ideology and conflict, which is presumably due to the difficulty in measuring such an abstract concept, especially across hundreds and thousands of leaders. Existing attempts use indicators based on a variety of speeches, membership in political

[72] Kenneth A. Schultz, 'The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?', *International Organization* 59, no. 1 (2005): 1–38.
[73] Andrew Bertoli, Allan Dafoe, and Robert F. Trager, 'Is There a War Party? Party Change, the Left–Right Divide, and International Conflict', *Journal of Conflict Resolution* 63, no. 4 (2019): 950–75.
[74] Glenn Palmer, Tamar London, and Patrick Regan, 'What's Stopping You?: The Sources of Political Constraints on International Conflict Behavior in Parliamentary Democracies', *International Interactions* 30, no. 1 (2004): 1–24; Philip Arena and Glenn Palmer, 'Politics or the Economy? Domestic Correlates of Dispute Involvement in Developed Democracies', *International Studies Quarterly* 53, no. 4 (2009): 955–75.

organisations, political platforms, and policy decisions.[75] Accordingly, there is some subjective element to coding for a leader's ideology that leads to cross-validation issues between datasets, although this only seems to be mostly a problem for leaders who can be considered centre.[76] Otherwise, there is consistent agreement as 91% of leaders are coded the same in the two most all-encompassing datasets on leaders' ideology, Herre's Identifying Ideologues and Brambor and Lindval's Heads of Government datasets.[77] Thus, the limited but consistent existing research above suggests that right-wing leaders are more likely to start wars and conflicts, although more research is needed to understand the extent to which ideology helps to explain why leaders start conflicts.

2.3 Holistic Measures of a Leader's Willingness to Use Force

While most leader trait studies investigate one trait, two studies have attempted to capture leaders' willingness to use force in a single index based on a wide variety of traits: Horowitz and Stam's "leader risk score" and Carter and Smith's "hawkishness index."[78] In essence, both of these measures try to capture the combined effect of various leaders' measurable traits in relation to either holding a hawkish foreign policy in general (hawkishness index) or starting conflicts in particular (more for Leader Risk Score). Given that both are unobservable latent variables, it is first necessary to explain how each is measured more thoroughly before we can understand how they are associated with conflict initiation.

---

[75] Thomas Brambor and Johannes Lindvall, 'The Ideology of Heads of Government, 1870–2012', *European Political Science* 17, no. 2 (2018): 211–22; Bastian Herre, 'Identifying Ideologues: A Global Dataset on Political Leaders, 1945–2020', *British Journal of Political Science*, 2022, 1–9.
[76] The Heads of Government dataset is above 90% in agreement with other datasets for left and right-wing leaders, but only 47% for centre. See: Brambor and Lindvall, 'The Ideology of Heads of Government, 1870–2012'.
[77] Ibid.; Herre, 'Identifying Ideologues'.
[78] Horowitz, Stam, and Ellis, *Why Leaders Fight*; Carter and Smith, 'A Framework for Measuring Leaders' Willingness to Use Force'.

2.3.1 Leader Risk

Leader risk is discussed at length in Horowitz and Stam's book *Why We Fight* (WLF) and is an index of a leader's latent conflict-propensity based on the additive effect of over three dozen of their traits, as illustrated in the columns of Table 2.2 below. To construct each leader's risk score, the authors use the predicted probability that a state initiates a conflict based on every trait in a leader-year monadic model, take the mean predicted probability across every year the leader was in power, and then turn that mean value into a single "leader risk score." At face value, the leader risk score seems to reasonably reflect who we tend to think of as extremely aggressive and conflict-prone leaders; Ayatollah Khomeini, Ronald Reagan and Mao Zedong all have risk scores within the top-ten highest since 1875. Leaders with higher scores are what Horowitz and Stam refer to as "high-risk leaders" and I use the term similarly as noted in the introduction. More specifically, a high-risk leader is a leader who have a combination of traits that result in a higher average predicted probability of conflict using Horowitz and Stam's model.

To demonstrate the explanatory power of their leader risk score, the authors compare it to a similar "systems risk score" – the same model and method above but with domestic and international variables included instead of leader traits. The goal here is simple: to determine whether a single index of multiple leader traits is better at predicting international conflict than a single index of multiple systemic factors. Since Horowitz and Stam cannot test the predictive power of either score using the same methods to construct them, they convert both into percentiles and then compare their accuracy in predicting the top 2% of leaders with the most conflict initiations. Table 2.1 below illustrates the logic among some of the more notable leaders in history: whichever risk score is closer to 99% is a better predictor of reality.

Table 2.1: Example of leader risk vs systems risk comparison in WLF.

Country	Leader Name	Leader Risk Prediction (%)	System Risk Prediction (%)	Conflict Initiations
Iran	Ayatollah Khomeini	99	94	64
Germany	Adolf Hitler	98	99	52
China	Mao Zedong	99	98	45
Iraq	Saddam Hussein	98	94	44
Soviet Union	Josef Stalin	89	91	38

While Table 2.1 above is just a sample of a few leaders, Horowitz and Stam find that their leader-risk score outperforms the systems-risk score in about half of the top 40 most conflict-prone leaders, suggesting that leader's traits are certainly an important piece to the puzzle of why leaders start conflicts. Although it should be stressed that they do not claim the leader risk score is meant to have more explanatory power than systemic factors, only that it "shows the potential explanatory power of our approach."[79] Indeed, a combined model with both systems and leader-level variables outperforms either model independently, which again seems to be fairly strong evidence that leader traits do matter; otherwise, we would expect to see no difference. However, the leader risk index is relatively new (published in 2015), so there is little literature to judge its explanatory power beyond WLF. Only one such study outside of WLF using the risk score exists, finding that higher-risk leaders are also more likely to pursue nuclear weapons,[80] which is additional evidence supporting the central finding in WLF that the leader risk score does indeed reflect some underlying propensity for leaders to pursue "riskier" foreign policies, like starting conflicts.

[79] Horowitz, Stam, and Ellis, *Why Leaders Fight*, 79.
[80] So Yeon (Ellen) Park, 'Why Do Leaders Matter?: Exploring Leaders' Risk Propensities and Nuclear Proliferation' (PhD Dissertation, Santa Barbara, University of California, 2021).

2.3.2 Hawkishness

Like WLF's leader risk, Carter and Smith have also attempted to measure a leader's willingness to use force through Bayesian methods, termed the "Hawkishness index."[81] As Table 2.2 below shows, the hawkishness index uses a narrower set of leader traits than the risk score – mostly those related to a leader's past military experience and more subjective personality trait assessments based on other research.[82] To construct each leader's hawkishness value, they use a standard Rasch model, which relies on two priors, Adolf Hitler and the Dalai Lama, reasonably reflecting who we tend to think of as being the most and least aggressive leaders in history, respectively. Carter and Smith's hawkishness index passes a battery of statistical tests and, like the leader risk score, appears quite reasonable at face value. For example, Winston Churchill and George W. Bush are considered more hawkish than Neville Chamberlain and Bill Clinton, which is what people might otherwise agree upon. Like the risk score, the hawkishness index is also relatively new (published in 2020), so, unfortunately, there is very little literature to judge its explanatory power further. Nevertheless, a leader's hawkishness score is positively and significantly associated with conflict initiation, especially when term-limited.[83] Thus, we again see here that a holistic index of leader traits appears to robustly reflect some underlying risk propensity to start conflicts.

Table 2.2: The Relevant Traits Included for Leader Risk and Hawkishness

Traits used to measure….		
Leader Risk Score	Both	Hawkishness Index

[81] Carter and Smith, 'A Framework for Measuring Leaders' Willingness to Use Force'.
[82] Carter and Smith use four different "hawkishness" models by varying the amount and type of traits included, but their second model is the best performing measure and so is used in this dissertation when referring to leader hawkishness.
[83] Carter and Nordstrom, 'Term Limits, Leader Preferences, and Interstate Conflict'.

age, education, number of spouses, married, married in power, divorced, number of children, parental status, legitimate child, royalty, teacher, journalist, lawyer, medical, religious, activist, politician, creative, business, aristocrat, police, science/engineer, blue collar, years since last conflict initiation	military service, military combat experience, war veteran - won, war veteran – lost, military career, military education, rebel, rebel experience, rebel – won, rebel – lost, irregular entry, gender	right-left wing, international peace, hawk-dove, ideology, constraint challenger

2.4 Other Traits

While the above traits appear to be the most important and are the focus of the rest of this dissertation, several other traits that I do not focus on are worth briefly mentioning. Again, I exclude these traits primarily because they either suffer from measurement issues related to construct validity and cross-coder reliability or have already been found not to be significantly related to conflict initiation.

2.4.1 Education

The role of a leader's educational background has also been of interest to IR scholars due to the perceived socialising effects of higher education. That is, since universities tend to lean liberal and promote norms of conflict aversion and resolution, we might therefore expect higher-

educated leaders to be less conflict-prone than lower-educated leaders.[84] Only a few studies have investigated the link between a leader's education and conflict and generally find that higher-educated leaders are slightly more likely to initiate conflict, although the association is not statistically significant.[85] However, one study by Joan Barceló finds that a leader's past education depends on where they live and go to university; non-Western leaders who have been previously educated in a Western university are much less likely to initiate conflict than those who have no education and those who have graduated from a non-Western university.[86]

Originally, I had planned to investigate whether selection also explains why Western-educated leaders appear less likely to initiate conflict. However, during data collection, I discovered that Barceló's study is fundamentally flawed due to miscoding. Roughly 62 leaders are coded as having prior Western education differently than the original Western education dataset on which Barceló's study relies.[87] Compared to the original study, Barceló miscodes 39 leaders as not having Western education when they should and miscodes 23 leaders as having Western education when they should not. This miscoding issue is particularly problematic because the former 39 miscoded leaders initiated 75 conflicts, whereas the latter 23 miscoded leaders only initiated 10. For instance, Jewerhal Nehru initiated 17 conflicts during his time as Indian Prime Minister – the 9th highest of the 901 leaders in the dataset – and he is coded in Barceló as not having prior Western education despite getting a bachelor's degree from Cambridge. As a result, the effect of not having Western education on conflict initiation is artificially inflated because most conflict-prone leaders are miscoded as not being Western educated when they should be. There is no relationship with

[84] Joan Barceló, 'Are Western-Educated Leaders Less Prone to Initiate Militarized Disputes?', *British Journal of Political Science*, 2018, 1–32.
[85] Horowitz, Stam, and Ellis, *Why Leaders Fight*. Nicholas Carnes and Noam Lupu, 'What Good Is a College Degree? Education and Leader Quality Reconsidered', *The Journal of Politics* 78, no. 1 (2016): 35–49.
[86] Barceló, 'Are Western-Educated Leaders Less Prone to Initiate Militarized Disputes?'
[87] Original coding is in: Thomas Gift and Daniel Krcmaric, 'Who Democratizes? Western-Educated Leaders and Regime Transitions', *Journal of Conflict Resolution* 61, no. 3 (2017): 671–701. It is also very difficult to accurately compare the miscoding between the two studies as Barceló does not include the leader's name, only the country code and year.

conflict initiation when replicating Barceló's original model with the correctly coded Western education variable. Thus, I will not further investigate education or Western education in the remainder of this dissertation since there is no relationship for selection to potentially explain.

2.4.2 Occupation and Upbringing

In WLF, Horowitz and Stam also investigate numerous occupational characteristics and their childhood upbringing. However, most are not significant and have small effects besides a few.[88] For instance, the trait "creative" is significant and has a reasonably large effect size compared to other traits, but this appears to be driven primarily by Adolf Hitler and Ronald Reagan, who are presumably coded as creative due to the former's painting career and the latter's acting career. Additionally, leaders with medical careers appear less likely to initiate conflict, although only 69 leaders are considered to have medical backgrounds. For childhood upbringing, whether a leader's parents were divorced appears to lead to more conflict, although the relationship is weak, and leaders born "illegitimately" are significantly less conflict-prone. Lastly, other studies find that more politically experienced leaders are less likely to start conflicts, as mentioned above in the section on age.[89] While the relationships between these traits and conflict initiation are interesting, I do not further investigate them for two reasons. First, these are all derived from the Leader Experiences and Attributes Dataset (LEAD),[90] yet it does not provide clear coding criteria, and since they are quite subjective, it is unlikely I would be able to code similarly, resulting in construct validity issues from intercoder unreliability. Second, it is also unlikely I will be able to find most of the information for these traits on other candidates since they have significantly less biographical information available than leaders, especially

[88] Horowitz, Stam, and Ellis, *Why Leaders Fight*, technical appendix.
[89] Calin and Prins, 'The Sources of Presidential Foreign Policy Decision Making'; Potter, 'Does Experience Matter?'
[90] Ellis, Horowitz, and Stam, 'Introducing the LEAD Data Set'.

farther back in time and weaker and smaller states. Thus, the interrelated problems of subjectivity, unclear coding criteria, and lack of information will almost certainly result in heavily biased measures with little theoretical importance to gain at the expense of significant research effort.

2.5 Conclusion

From the review of the literature above, we have stepped closer to understanding the extent to which traits influence leaders to start conflicts. This research question fits within a larger debate in IR as to whether leaders matter more generally, but more specifically as to whether leader *traits* matter. Some might say no, believing the environment explains most of a leader's foreign-policy decision-making. Others might say yes, believing that traits are an influential and important piece to the puzzle of why leaders start conflicts. The existing evidence suggests that we have reason to be more confident in the latter perspective; studies increasingly show that several traits are associated with the initiation of conflict: military experiences, age, biological sex, and ideology. However, one problem remains – selection. In the next chapter, I more specifically discuss why we might be sceptical that these traits are as important as existing research implies.

3 The Problem of Selection

When we set out to answer whether traits influence a leader's decision to start conflicts, we need to be sure that our answers to this question are not biased. While previous research suggests that several traits do indeed cause leaders to be more conflict-prone, these findings could be spurious if leaders with these traits tend to be selected into conflict-prone environments more than others. In other words, we need to first consider how and why leaders are selected to get closer to understanding the extent to which traits influence leaders to start conflicts. In this chapter, I more specifically discuss the primary way the relationship between leader traits and conflict initiation could be spurious – leader selection. Ultimately, we are interested in the causal effect of leader traits on conflict. While the authors of the various leader trait studies to date are cautious in their language and are hesitant to make definitive causal conclusions about the effect of traits on conflict, they do implicitly make them else they would not first hypothesise why we might expect to see these associations in the first place. So, the big question is whether their findings are indeed supportive of a causal effect.

One way we can begin to investigate the causal logic of the trait perspective is through counterfactual reasoning. Indeed, every causal claim is fundamentally based on counterfactual thinking.[91] To help illustrate, consider the effect of prior military experience since this trait appears to have one of the more robust relationships with conflict initiation. Also, consider one of history's more conflict-prone ex-military leaders, Ronald Reagan. The central causal logic of the trait perspective ultimately rests on a counterfactual claim: "Had Ronald Reagan not served in the military, he would have initiated fewer conflicts." That is, when some individual is given the "treatment" of serving in the military, we would expect them to be more likely to initiate conflict as leaders than had they never served in the military due to the theorised socialisation

[91] Guido W. Imbens and Donald B. Rubin, *Causal Inference in Statistics, Social, and Biomedical Sciences* (New York: Cambridge University Press, 2015).

mechanisms mentioned in the last chapter. Accordingly, there are two potential outcomes here; the total number of conflicts Reagan initiated in the real world where he had prior military experience $Y_{Reagan}(T_1)$ and the total number in the counterfactual world where he did not $Y_{Reagan}(T_0)$. If we had God-like powers and could observe both worlds, we could perfectly estimate the causal effect of military experience by subtracting the number of conflicts Reagan initiated in the world where he had military experience from the number of conflicts he initiated in the counterfactual world where he did not. Unfortunately, we cannot go into the past, undo Reagan's military service and re-run history to see if he would have been less aggressive. This problem reflects the more commonly known fundamental problem of causal inference: the counterfactual world we need to compare is obviously unobservable, so we can never know the true causal effect of military experience on conflict initiation.

Even though overcoming the fundamental problem of causal inference is impossible, we can at least use various methods to get close as possible to approximating the key counterfactual logic. Instead of trying to estimate the unobservable individual treatment effect of military experience within the same leader, we can instead estimate the observable Average Treatment Effect (ATE) of ex-military leaders across states. The ideal way to do this would be to take a random set of individuals, randomly conscript them, randomly assign them as state leaders, and then compare whether those who are military veterans initiated more conflicts than non-veterans. In this perfect hypothetical experiment, the ATE would simply be the difference in the number of conflict initiations between states led by veterans (treatment group) and non-veterans (control group). Since we randomised treatment assignment (selecting an ex-military leader), we would have much more confidence military experience is the culprit for any large difference in conflicts between the two groups. Thus, even though we cannot estimate the individual causal effect of leader traits, we can estimate their ATE, at least in theory.

As much as IR scholars would love to do these kinds of experiments, they are impossible. We can only use historical observational data and compare the frequency of conflict initiation

between previously selected leaders. This is what any leader trait study is forced to do. As a result, the unit of analysis is not high-risk traits but rather leaders with high-risk traits. So, the counterfactual logic then shifts from "had Reagan not served in the military, he would have initiated fewer conflicts" to "had the U.S. selected *someone else* without military experience, they would have initiated fewer conflicts." This means that leader trait studies are not directly estimating the causal effect of military experience on conflict initiation but rather indirectly as an effect of selecting ex-military leaders on conflict. This distinction is key because leaders are not randomly selected, and so the fundamental challenge for the trait perspective is to ensure that ex-military leaders like Reagan are not systemically selected into particular domestic and international environments where the probability of future conflict is higher. In causal inference terms, we can think of leader selection as analogous to an assignment mechanism – "the process that determines which units receive which treatments, hence which potential outcomes are realised and thus can be observed, and, conversely, which potential outcomes are missing."[92] In our case, the units are states, and the treatment is a leader with a high-risk or low-risk trait.

To help illustrate why treatment assignment is important for our purposes, consider a hypothetical world where only two states exist, *A* and *B*, and leaders are chosen in Americanesque Presidential elections where only two candidates can reasonably win. Our main explanatory variable here is prior military experience, *T*, which is binary with two values: T_0 and T_1. Our measured outcome, *Y*, is the number of conflict initiations during the tenure of each leader following the Presidential elections. Assume that everything about each state and leader is the same except for *T* and *Y*. Now, when voters cast their ballots, there are two potential outcomes for each state: the number of conflict initiations after electing either a veteran, $Y_i(T_1)$, or a non-veteran $Y_i(T_0)$ as President. If a veteran were elected in State A and a non-Veteran in State B, we observe and have data for the potential outcomes $Y_a(T_1)$ and $Y_b(T_0)$, respectively.

[92] Ibid., 31.

Conversely, there are two other potential outcomes where the runner-up candidates were elected instead, $Y_a(T_0)$ and $Y_b(T_1)$. Yet, these are the counterfactual worlds for which we do not have data and are essentially a missing data problem.

State	Winner Trait *T*	Runner-Up Trait *T*	$Y_i(T_1)$	$Y_i(T_0)$
A	0	1	NA	0
B	1	0	2	NA

Figure 3-1: Example of potential outcomes in a hypothetical world.

If we were to simply compare the average difference in *Y* in our hypothetical world, then it would appear that the ex-military leader is indeed more likely to initiate conflict. This finding, however, depends on the assumption that the veteran in state *B* was not elected for any reason that might also cause them to initiate more conflicts in the future. While we know this assumption holds in our hypothetical world since everything else in states *A* and *B* was assumed away, in the real world, however, leaders are selected for any number of reasons and one of those reasons could be causally related to future conflict. So, we need to be sure that all states have the same probability of selecting a leader with trait T_0 or T_1. In other words, we need to control for any institutional and or structural factors (i.e., the environment) that might cause states to both select leaders with T_1 and initiate conflict. In Figure 3-2 below, this is easy to understand. Without controlling for selection, any estimated causal effect for leader traits and conflict initiation would then be biased, and we would be misled into thinking that ex-military leaders with traits like Reagan are more aggressive when they are just more likely to end up in office where the baseline probability of conflict is higher.

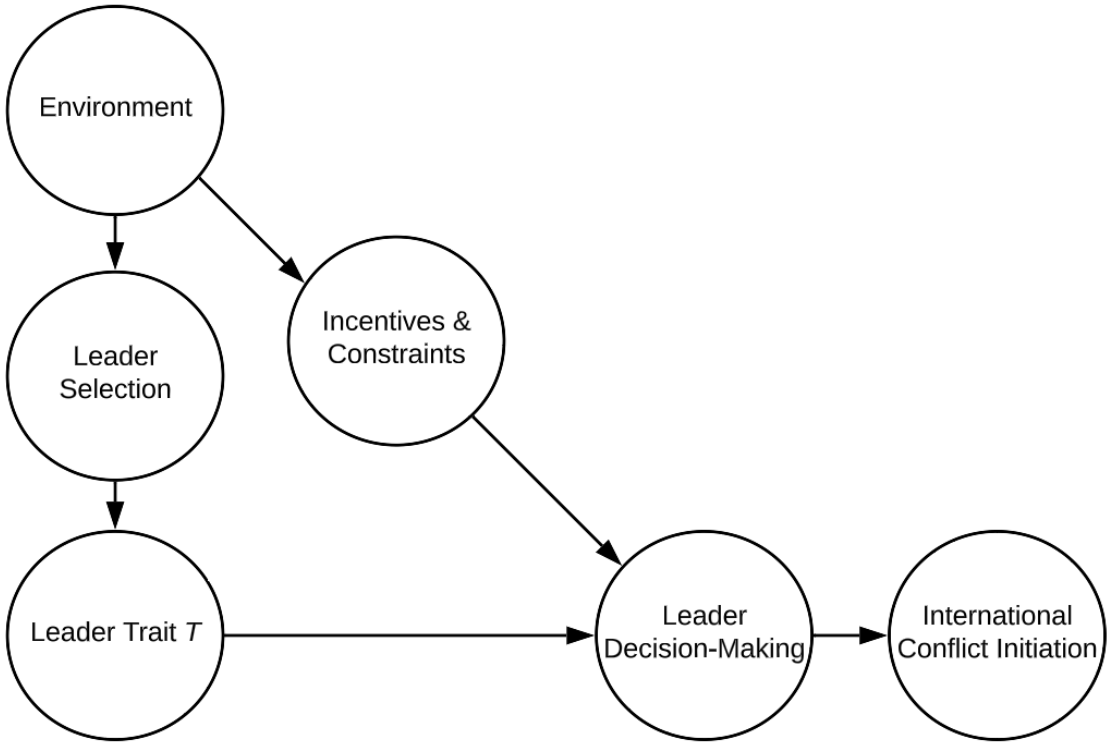


Figure 3-2: DAG illustration of how the trait-conflict relationship can be confounded by leader selection.

In essence, the central challenge to answering the research question is that leaders with certain traits could be more likely than others to be selected into conflict-prone environments.[93] However, leader selection is a complex process, and we have not yet explicitly discussed the specific causal mechanisms for why and how leader selection could more specifically confound the association between traits and conflict. Unfortunately, there is little existing research to base any theoretical framework of leader selection on, as most leader trait studies assume selection is a potential problem but just control for it using institutional or structural variables as proxies.[94] Furthermore, the two large-n leader datasets include only two variables related to leader

[93] It should be noted the problem of selection here is not the same as selection bias, which is when certain observations are included, and others excluded from the analysis.
[94] One study directly investigates the selection of leaders with political experience globally, but they only regress regime type on a leaders' prior political experience, and do not provide any detailed discussion as to how and why this leader selection process occurs. See; Alexander Baturo and Johan A. Elkink, 'What Countries Select More Experienced Leaders? The PolEx Measure of Political Experience', *British Journal of Political Science*, 2021, 1–10.

selection, whether or not a leader entered office "irregularly" and whether the previous leader died of natural causes, which, while useful, do not fully capture the nuances of leader selection.[95] Therefore, to more acutely understand how selection potentially confounds the relationship between leader traits and conflict initiation, we need a new theoretical framework of exactly how and why high-risk leaders could end up in environments where conflict is more likely. However, before we consider why leaders are selected, we must first consider how they are selected.

3.1 Leader Selection – A Theoretical Framework

We know several things about how leaders are selected. First, we know that leader *X* became *de facto* leader at some point in time. This is simply the moment they enter office, otherwise known in leader-level datasets as their "start date". In short, these are actual leader transitions. We also know that leaders do not show up randomly and start barking around orders. Every time an actual leader transition occurs, it must be a result of some selection mechanism, or a process that formally or informally grants them the mandate of being the ultimate decision-making authority of their state. When considering what these selection mechanisms are, we come to the realisation that every time a leader transition occurs, it is because the new leader has either won an election, been appointed by the Head of State (when the leader is the Prime Minister), inherited the position through succession, or self-selected through coups or wars. These are the only four selection mechanisms that can cause an actual leader transition to occur. While there may be some overlap between them (e.g., parliamentary systems where the new leader previously won a general election but is technically appointed by the Head of State), these four are the only mechanisms that provide new leaders with their legitimacy to be the ultimate decision-making authority. Grouped together, these four selection

[95] Both variables are in the Archigos dataset. "Irregular" entry indicates whether a leader entered office by "subverting the established rules and procedures of leader selection in their state." See: Goemans, Gleditsch, and Chiozza, 'Introducing Archigos'.

mechanisms are what I term "leader selection events" – or more specifically defined as any time an individual is elected, appointed, inherits, or self-selects the position of state leader. Leader selection events can occur at a specific time, such as in elections, or over a lengthier time, such as in autocracies, when one individual increasingly wrestles political control away from an incumbent. Furthermore, they can occur several months before the leader formally takes power, as is often the case following elections, or at the exact moment they take power, such as when a coup leader unexpectedly arrests an incumbent and proclaims themselves "Saviour" of the nation.

Second, we know that there are always other individuals within selection events that could have been leader but ultimately were not deemed legitimate enough to be given the mandate. These individuals are either the runner-up in elections, the person who the Head of State almost appointed, third-in-line in successions, or the deposed incumbents following coups. For simplicity, I refer to both the runner-up and eventual leader in selection events as "candidates." While the term candidate tends to be more narrowly used in the sense of individuals competing for some position, I use the term in the broader sense of "one that is qualified for office" since this is more inclusive of instances where individuals are not formally competing for leader but nonetheless still possibly could be leader. In free-and-fair elections, for example, two or more candidates compete for votes (i.e., nominees or party leaders). In coups, there are at least two candidates always fight to acquire or maintain the position of leader through force (i.e., coup leader and incumbent). In succession, where it may appear as though there is only one "candidate", multiple other individuals are ranked within the order of succession but simply do not contest the established rules of succession. Even though these individuals may not contest the leadership position, they are the next in line and at least have a legitimate claim to be the leader.

We also know that which candidate becomes the leader in selection events is not randomly determined. In elections and appointments, the new leader is determined by the selectorate –

any individual with a government-granted say in the selection of leaders, whether it be millions of voters or just one individual like the Head of State.[96] Selectorates either choose new leaders directly, such as by being elected, or appointed by the Head of State, or indirectly, such as when an elected Vice-President inherits the Presidency after the latter's sudden death. In self-selections, who becomes leader is determined through force, whereas in monarchical successions, who becomes leader is determined by bloodlines. Thus, who eventually becomes leader and who does not is always a result of non-random causes, and it is these causes which could result in different types of leaders being selected into different types of environments.

However, we also know that there are *potential* leader transitions in two of these selection mechanisms, elections and self-selections, even though no actual transition has occurred. For instance, voters can re-elect presidents or prime ministers, or incumbents can thwart coup attempts. In these two cases, leaders are non-randomly selected despite no leader transition occurring. Thus, if we are to consider how and why leaders are selected, we must also consider potential leader transitions since who became leader in these instances is due to the same factors that result in actual leader transitions.

Lastly, we also know that the candidates in leader selection events are not randomly determined either. They are similarly selected in non-random ways through candidate selection – which is any reason why an individual became candidate in the selection event. In elections, candidates are often first selected in either presidential primaries or party leadership elections in parliamentary systems. In appointments, candidates are all the individuals who the Head of State considers to be a viable replacement. In successions, candidates are similarly selected into the pool of potential leaders through bloodlines. In self-selections, however, there is no candidate selection. Rather, we can imagine that those who attempt coups have reached a position where

[96] Bueno de Mesquita, *The Logic of Political Survival*.

they feel they could overthrow the incumbent, which reflects the same general idea that there are non-random reasons for why candidates are in the running to be the new leader.

The fundamental importance of understanding how leaders are selected is that it allows us to better understand why they are selected. We are not, however, interested in every single reason why, say, a veteran might be elected over a non-veteran. Rather, we are only interested in selection effects: any reason why leaders with high-risk traits could be disproportionately selected into environments with a greater baseline probability of conflict. If this were the case, then leader selection effects potentially confound the relationship between traits and conflict. There are three different forms of selection effects, with different subsets, which I will discuss in turn: leader selection, self-selection, and candidate selection effects.

3.1.1 Leader Selection Effects

The first way the trait-conflict association can be confounded is through leader selection effects. Recall that in leader selection events the outcome always results in one individual who became leader and another who became runner-up. A leader selection effect is thus any reason why candidates with high-risk traits are more likely to be leader than runners-up in environments where the probability of conflict is greater. At this point, we do not necessarily need to be concerned with exactly what those reasons could be. As long as we observe a leader selection effect, then this is the minimum we need to observe that the effect of leader traits on conflict could be confounded. Nevertheless, it is important to have some theoretical justification for why we might expect to see a leader selection effect for any of the high-risk traits. Accordingly, I believe three types of leader selection effects can confound the trait-conflict association: A first-image reversed effect, a demographic effect, and a self-selection effect.

3.1.1.1.1 The First Image Reversed Leader Selection Effect

The most intuitive leader selection effect is that people might be more likely to choose high-risk over low-risk candidates for leader precisely because of a belief (right or wrong) that those with high-risk traits are more skilled at managing international conflict. This is what I refer to as a "First-Image Reversed leader selection effect" (hereafter FIR – leader selection effect) since the environment (second and third-image variables) causes variation in people's preferences (first-image variable) for leaders with certain traits.

First image variables, such as a leader's age or prior military experience, have been traditionally conceptualised among IR scholars as causing variation at the international level from the bottom up.[97] Yet, the causal logic between the three images does not only flow in one direction. Peter Gourevitch first inverted this traditional analytical focus of IR and instead sought to explain how domestic politics is not only a cause but also an effect of international politics.[98] More recently, several scholars have extended this inversion one step further to investigate how individuals' beliefs and behaviours can also be an effect of international politics – the "first image reversed (FIR)" literature.[99] In essence, the FIR framework simply flips the directional logic of the traditional first image framework on its head; international politics causes variation among first image variables.

The FIR leader selection effect looks like it fits nicely within the more traditional FIR logic; leader traits are not only an explanatory variable of conflict but also an outcome variable. However, scholars tend to use the FIR approach quite differently. In essence, they explore how

---

[97] Kenneth N. Waltz, *Man, the State, and War: A Theoritical Anylsis* (New York: Columbia University Press, 1959).
[98] Peter Gourevitch, 'The Second Image Reversed: The International Sources of Domestic Politics', *International Organization* 32, no. 4 (1978): 881–912.
[99] Joshua D. Kertzer and Dustin Tingley, 'Political Psychology in International Relations: Beyond the Paradigms', *Annual Review of Political Science* 21 (2018): 319–39; David H. Bearce and Thomas R. Cook, 'The First Image Reversed: IGO Signals and Mass Political Attitudes', *The Review of International Organizations* 13, no. 4 (2018): 595–619; Caleb Pomeroy, 'The Psychology of Power: The First Image Reversed and International Security', 2021.

individuals' beliefs and behaviours, whether elites or the mass public, change in response to receiving information about international politics.[100] Since international conflict cannot directly cause leaders to change from being, say, male to female, it might appear as though the FIR framework does not apply in our case. However, international conflict can cause the beliefs of the selectorate to change. So, the outcome variable for the FIR leader selection effect is the preferences of the selectorate, which then causes leader trait variation via selection.

But this just begs the question of whether there is any theoretical or empirical evidence for this FIR leader selection effect. Indeed, the argument that selectorates would choose high-risk over low-risk candidates in response to a dangerous international environment depends on two fairly strong assumptions: that the selectorate not only cares about potential future international conflicts when choosing their leader but also care about the specific backgrounds and attributes of candidates. A few of the authors of the leader trait studies outright reject these two assumptions. Horowitz and Stam, for instance, claim that any proponents of this selection argument "miss[es] the fact that nearly all countries select leaders based on their domestic policy preferences, rather than beliefs about how the background experiences will shape their military behaviour in office."[101] Likewise, Barceló argues that bias from selection "is plausible but very unlikely" because leader selection is ultimately important for domestic outcomes but "less central in studies of international outcomes."[102] However, these are untested assumptions and several studies even show that people tend to prefer leaders with seemingly politically irrelevant facial cues, such as attractiveness, competence, and dominance.[103] Furthermore, more recent literature

[100] Bearce and Cook, 'The First Image Reversed'.
[101] Horowitz, Stam, and Ellis, *Why Leaders Fight*, technical appendix, 11.
[102] Barceló, 'Are Western-Educated Leaders Less Prone to Initiate Militarized Disputes?', 550–51.
[103] Mark Van Vugt and Allen E. Grabo, 'The Many Faces of Leadership: An Evolutionary-Psychology Approach', *Current Directions in Psychological Science* 24, no. 6 (2015): 484–89; Anthony C. Little et al., 'Facial Appearance Affects Voting Decisions', *Evolution and Human Behavior* 28, no. 1 (2007): 18–27.

under the FIR umbrella suggests sufficient evidence to suspect that people's preferences for leader might change in response to concerns about their security environment.

One key assumption of the FIR framework is that the selectorate receives information about international conflicts, which is more or less accurate. This information is what some FIR theorists refer to as "international-level signals" as they originate from the international level and are then transmitted to individuals through some medium, such as news or social media. While scholars have used the FIR framework to explore various research questions in IR, the most obvious question for our purposes is the extent to which foreign policy influences the ballot choices of ordinary citizens. The early work on this question was unambiguously cynical. Through what came to be known as the "Almond-Lippman consensus," many scholars converged around the belief that the mass public is too uninterested in international politics and their support or dissent for any particular foreign policy depended entirely on their emotional "mood" at the time.[104] Following the Vietnam War, however, this view shifted as surveys continually demonstrated that ordinary citizens have at least a rational and coherent understanding of international politics[105] The prevailing consensus among electoral scholars today is that foreign issues are often equally and sometimes even more important for voters than domestic issues.[106] Thus, voters notice and care about what goes on outside their state and cast ballots accordingly.

The selectorate, then, has what Rathbun considers "international paradigms in their minds."[107] Even if they are completely unaware of the specific differences between competing IR theories, they intuitively understand their more basic disagreements. Indeed, many scholars have noticed

---

[104] Gabriel Almond, *The American People and Foreign Policy*, 8th ed. (United States of America: Fredereick A. Praeger, 1960); Walter Lippmann, *Essays in the Public Philosophy* (Transaction Publishers, 1955).
[105] Ole R. Holsti, 'Public Opinion and Foreign Policy: Challenges to the Almond-Lippmann Consensus', *International Studies Quarterly* 36, no. 4 (1992): 439–66.
[106] Ibid.
[107] Brian C. Rathbun, 'It Takes All Types: Social Psychology, Trust, and the International Relations Paradigm in Our Minds', *International Theory* 1, no. 3 (2009): 345–80.

the parallels between IR realism and social conservatism or other "right-wing" belief systems, as well as IR liberalism, social progressivism, and other "left-wing" belief systems.[108] Along the same lines, we can think of the selectorate as holding "folk theories" of international politics. These are distinguished from their more scientific IR counterparts in that their proponents do not explicitly identify them and are held as a belief system rather than an analytic lens.[109] Like the contrast between realism and idealism in IR, folk theories can be juxtaposed between "folk realism" and "folk idealism." Folk realism shares the fundamental pessimism of IR realist thought and instinctively understands its central tenets but is largely unaware of the more technical points (e.g., Balance-of-Power, the security dilemma).[110] Indeed, as Drezner pointed out regarding American's tendency to hold folk realist beliefs:

> "Realists assert that uncertainty about other countries' intentions forces states to act in ways that bolster their security. Perhaps Americans, when asked about foreign affairs, react in a similar manner – focusing on prudence and self-defence as risk-averse responses to an environment they find complex and opaque."[111]

Folk realism, then, shares the central prediction of IR realism of how relations between groups are likely to trend toward conflict given no overarching authority.[112] Indeed, Johnson and Thayer argue that some form of "anarchy" plagued humans well before 1648, so the fundamental realist

[108] For realism and conservative, see: Piki Ish-Shalom, 'The Triptych of Realism, Elitism, and Conservatism', *International Studies Review* 8, no. 3 (2006): 441–68.
[109] Benjamin Martill and Adrian Rogstad, 'The End of Consensus? Folk Theory and the Politics of Foreign Policy in the Brexit Referendum', *Global Affairs* 5, no. 4–5 (2019): 347–67.
[110] Joshua D. Kertzer and Kathleen M. McGraw, 'Folk Realism: Testing the Microfoundations of Realism in Ordinary Citizens', *International Studies Quarterly* 56, no. 2 (2012): 245–58.
[111] Daniel W. Drezner, 'The Realist Tradition in American Public Opinion', *Perspectives on Politics* 6, no. 1 (2008): 62.
[112] Steven E. Lobell et al., *Neoclassical Realism, the State, and Foreign Policy* (Cambridge University Press, 2009), 14.

propositions of fear, self-interest, and concerns over relative power between groups are likely part of an evolved psychology stretching back millions of years.[113]

This folk realism is contrasted by "folk idealism" – optimism about the future of the international order instead of pessimism, favouring cooperative foreign policy rather than self-interest, and generally more trusting of other states' intentions instead of fearful. Many FIR studies also fit within this folk idealist perspective. One dominant strand, for instance, explores how international human rights organisations broadcast of states' violations of international legal norms helps to mobilise popular dissent and reduce public support for drone strikes, torture and military intervention.[114] Furthermore, the social media presence of many IGOs, such as the UN during international crises or the World Health Organization during the COVID-19 pandemic, is evidence these organisations believe that signalling international norms to the mass public can change people's minds and influence policy change. Thus, the international-level signals here are international norms, and the evidence is clear that selectorates do notice and care whether their state is adhering to them on the world stage.

Insofar that the mass public holds folk theories of international relations, however, research suggests that they tend to be more aligned with folk realism than idealism. Martill and Ragstad, for instance, frame Brexit within the folk realist and idealist divide, of which the former won out.[115] Furthermore, Brewer *et al.* find that many more Americans distrust other nations than trust them, leading them to conclude that "most Americans see the realm of international relations as resembling the 'state of nature' described by Hobbes. Put more simply, they see it as a 'dog-eat-

[113] Dominic DP Johnson and Bradley A. Thayer, 'The Evolution of Offensive Realism: Survival under Anarchy from the Pleistocene to the Present', *Politics and the Life Sciences* 35, no. 1 (2016): 1–26.
[114] Sarah E. Kreps and Geoffrey PR Wallace, 'International Law, Military Effectiveness, and Public Support for Drone Strikes', *Journal of Peace Research* 53, no. 6 (2016): 830–44; Geoffrey PR Wallace, 'International Law and Public Attitudes toward Torture: An Experimental Study', *International Organization*, 2013, 105–40; Christoph Mikulaschek, 'Cueing Foreign Elite Consensus or Divisions: The Effect of Unanimity in International Organizations on Public Opinion', in *Political Economy of International Organization Conference*, 2018.
[115] Martill and Rogstad, 'The End of Consensus?'

dog' world."[116] Exactly why the mass public appears to be more folk realist than idealist might be due to negativity bias. That is, people get information about international affairs through the mass media, which tends to choose and frame stories more negatively.[117] This negativity bias is best illustrated in the real world under the media's axiom, "If it bleeds, it leads." For example, the Israeli media's negative focus on violence and casualties in the Israeli-Palestine conflict has acted as a "peace spoiler," yet this was not a concerted effort by the press so much as an attempt simply to tell stories that captured their readers' interest.[118] As a result, fears about the intentions of powerful adversaries are probably amplified among ordinary citizens. Stephen Van Evera describes such effects as threat inflation and "a pervasive feature of international politics."[119] Threat inflation due to negativity bias is largely a bottom-up phenomenon in that people tend to be more attentive to negative information than positive, which is largely an unconscious response out of our control.[120] Threat inflation can also be orchestrated from the top down. Perhaps the most famous example of this is how ordinary Americans bought into the Bush administration's misrepresentation of Iraq's capabilities and intentions to target the U.S. with nuclear weapons.[121] Although, if voters do indeed adopt folk realist beliefs due to an inflated threat perception from receiving negative news about world affairs, it is probably entirely conditional on the state's security environment. In the same way realists often only limit their theories of state behaviour to great powers, so too might a folk-realist perspective.[122] We would not expect the selectorate in, say, Cape Verde to adopt folk

---

[116] Paul R. Brewer et al., 'International Trust and Public Opinion about World Affairs', *American Journal of Political Science* 48, no. 1 (2004): 105.
[117] Dominic DP Johnson and Dominic Tierney, 'Bad World: The Negativity Bias in International Politics', *International Security* 43, no. 3 (2018): 96–140.
[118] Tamir Sheafer and Shira Dvir-Gvirsman, 'The Spoiler Effect: Framing Attitudes and Expectations toward Peace', *Journal of Peace Research* 47, no. 2 (2010): 205–15.
[119] Stephen Van Evera, 'Foreward', in *American Foreign Policy and the Politics of Fear: Threat Inflation since 9/11*, ed. Trevor A. Thrall and Jane K. Cramer (New York: Routledge, 2009), xi.
[120] Timothy D. Wilson, *Strangers to Ourselves* (Harvard University Press, 2004), 63–64.
[121] Chaim Kaufmann, 'Threat Inflation and the Failure of the Marketplace of Ideas: The Selling of the Iraq War', *International Security* 29, no. 1 (2004): 5–48.
[122] Mearsheimer, for example, only refers to "great powers" when outlining the assumptions of his offensive realist theory: John J. Mearsheimer, *The Tragedy of Great Power Politics* (WW Norton & Company, 2001), 100.

realist beliefs, whether it was 1939 or 2020, as there were no immediate foreign threats to Cape Verdeans. But voters in France in 1939 had far more reason to adopt folk realist beliefs about the international order than they do today. Thus, in an extraordinarily fearful time, such as the late 1930s or the Cold War, we would expect the realist portion of the population to become more entrenched in their folk realist beliefs, especially so in more conflict-prone states.

3.1.1.2 International Conflict and the Selectorate

The next big question for the FIR leader selection mechanism is the extent to which selectorates care about international conflict. Just as states see international conflict as an existential threat, so too do their citizens. International conflicts are likely to make headlines, so the selectorate will often receive signals about recent military conflicts their state has engaged in. As noted in the introduction, many IR theories rely heavily on this assumption and find strong evidence this is the case. Audience costs, rally around the flag effect, and diversionary theory all demonstrate that signals about international conflict appear to influence voters' choice to re-elect incumbents, depending on contextual factors.[123] Furthermore, most of these studies use militarised interstate disputes as their outcome, which include many "short of war" conflicts that voters might not care about, such as shows of force and fishing disputes.[124] If these events do indeed shift voters' preferences for the incumbent because of some underlying perception of their ability to engage in conflict, then this might apply to other leader traits, such as military experience and age. Consequently, the FIR leader selection effect could confound the trait-conflict association through the following back-door non-causal path: Leader Trait $T \leftarrow$ Selectorate

[123] Michael Tomz, 'Domestic Audience Costs in International Relations: An Experimental Approach', *International Organization*, 2007, 821–40; Robert F. Trager and Lynn Vavreck, 'The Political Costs of Crisis Bargaining: Presidential Rhetoric and the Role of Party', *American Journal of Political Science* 55, no. 3 (2011): 526–45; Stephen Chaudoin, 'Promises or Policies? An Experimental Analysis of International Agreements and Audience Reactions', *International Organization*, 2014, 235–56.
[124] Gibler, *International Conflicts, 1816-2010*.

Preferences ← Environment → Incentives & Constraints → Leader Decision-Making → International Conflict Initiation. Figure 3-3 below visualises how the FIR leader selection effect could confound the trait-conflict association.

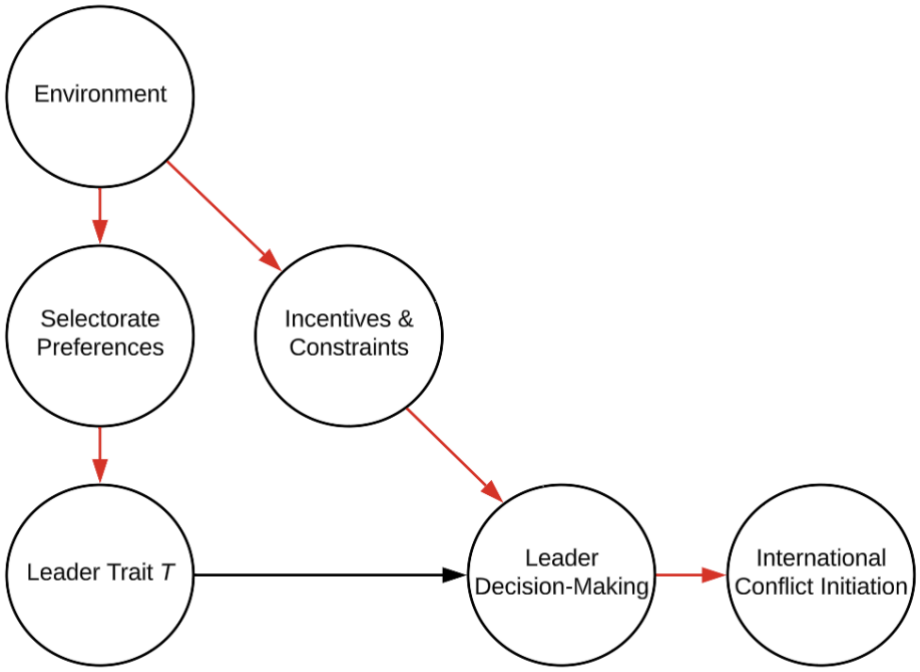


Figure 3-3: DAG illustrating confounding due to an FIR leader selection effect (highlighted in red)

3.1.2 Candidate Selection Effects

Candidate selection effects are any reason why candidates with high-risk traits are disproportionately selected into environments where the probability of conflict is greater. Candidate selection effects produce essentially the same confounding as leader selection effects, except before the leader selection event. While there are, of course, an endless number of potential reasons why we might see a candidate selection effect for a particular trait, I only review three of the most likely: a similar First-Image Reversed effect as above, a demographic effect, and candidate constraint effect.

3.1.2.1 The FIR Candidate Selection Effect

The first way that candidate selection effects can confound the trait-conflict initiation relationship is an extension of the FIR argument above but applied to choosing candidates rather than the leader. That is, if people consistently choose a high-risk over low-risk candidate for leader in response to their security concerns, then we would probably expect that same preference would manifest when first choosing candidates. Considering the logic is exactly the same as the FIR leader selection effect, it is unnecessary to provide another thorough review of the underlying evidence of why selectorates might nominate or choose high-risk over low-risk individuals for candidate. Where the FIR candidate selection effect slightly differs, however, is that even if voters care very little about leader traits in response to their security environment, political parties in conflict-prone states might still be more likely to nominate high-risk over low-risk candidates. More specifically, party elites could still believe (rightly or wrongly) that candidates with high-risk traits are more electable and thus be more likely to nominate them when international conflict is looming. In any case, the FIR candidate selection effect can confound the association between leader traits and conflict initiation through the following back-door non-causal path (Figure 3-4): Leader Trait T ← Candidate Pool of Trait T ← Selectorate Preferences$_{ts-1}$ ← Environment$_{ts-1}$ → Environment$_{ts0}$ → Environment$_{ts+1}$ → Leader Decision-Making → International Conflict Initiation). Since there are now two separate selection periods, I distinguish the environment and selectorates preferences into two periods, where *ts0* is the time of selection, *ts-1* is the period prior when selecting candidates. To account for the likelihood a leader's trait could influence the environment post-selection, I denote *ts+1* as the period during their term.

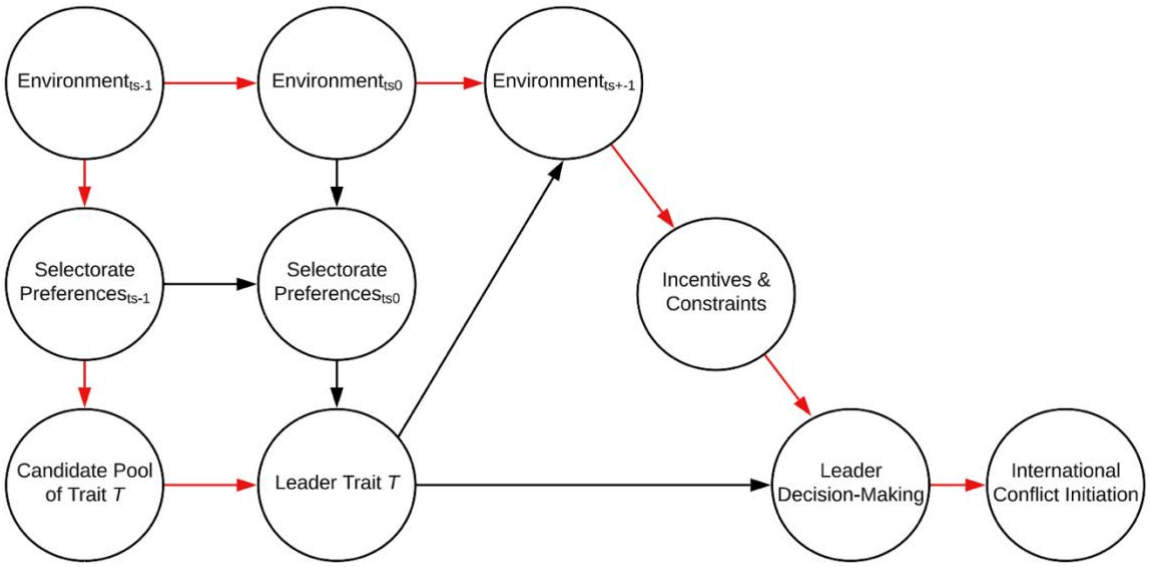


Figure 3-4: DAG of confounding due to an FIR candidate selection effect (highlighted with red arrows).

Furthermore, this potential elite-driven FIR candidate selection effect is not limited to elections and party elites. For instance, when it is time for a Heads of State to appoint a new leader, they might only consider individuals with high-risk rather than low-risk traits for similar reasons. Consider the 1940 appointment of 66-year-old combat veteran Winston Churchill U.K. Prime Minister at the outbreak of the Second World War. It is perhaps obvious that Churchill's extensive military and political experience was desirable when Parliament and King George VI considered who they would accept as Prime Minister. However, the only alternative individual that seemed to be supported by Parliament was Edward Wood, who was also relatively older (59) and a military combat veteran.[125] Thus, it is likely that King George VI would never have even considered a non-veteran and inexperienced (i.e., younger) alternative candidate to Churchill. If elites did prefer older veterans during dangerous times more generally, then they would probably be more likely to only consider older veterans in their pool of candidates when contemplating

[125] Jeff Wallenfeldt, 'Edward Frederick Lindley Wood, 1st Earl of Halifax', Brittanica, 9 March 2023, https://www.britannica.com/biography/Edward-Frederick-Lindley-Wood-1st-earl-of-Halifax.

who to choose as leader. So, if there were a preference for candidates with high-risk traits among elites in more conflict-prone states, then there would be a larger number of high-risk candidates, and hence there would be a higher probability these states would select high-risk leaders.

3.1.2.2 Demographic Candidate Selection Effect

The second way candidate selection can potentially confound the trait-conflict initiation relationship is through demographics. That is, if conflict-prone states tend to have a higher population proportion of some high-risk trait than conflict-averse states, then we would expect the candidate pool between the two to reflect that same proportion even if candidates were selected at random. Accordingly, this would result in a greater or lesser probability a state will have a leader with certain traits even if there were no FIR candidate or leader selection effect. Since the candidate selection effect is simply based on probability, it does not necessarily rely on any assumptions outside probability theory, which is much easier to accept compared to the many assumptions needed for the FIR selection effects. However, there are two separate ways in which demographics can be associated with the environment.

The first is that the environment directly causes demographic changes, which I refer to as a direct demographic candidate selection effect – where the environment has a direct causal effect on demographics, or through a mediator of some other domestic factor, which then causes differences in that leader trait across states. For example, states that frequently engage in conflict are more likely to be involved in wars and so probably have a larger surplus of veterans in the population, and hence have more ex-military candidates than conflict-averse states. Accordingly, the selectorates in more conflict-prone states would have little choice but to select a veteran as leader because there simply is a high proportion of veteran candidates. Consider Israel, one of the more conflict-prone states during the period, and Costa Rica, one of the least. Given that Israel has had universal conscription since its independence in 1949, it is no surprise

that 14 out of 17 of their leaders between 1949 and 2004 were ex-military. In contrast, Costa Rica abolished their military in 1949 and has not had an ex-military leader since.[126] As a result, Israeli selectorates had little choice but to select an ex-military leader since nearly all males (and females to a lesser extent) have served, whereas Costa Rican selectorates could never have chosen an ex-military leader even if they wanted to since none of the candidates in elections were veterans. Thus, the demographic selection effect confounds can confound the trait-conflict relationship through the following back-door path (Figure 3-5): Leader Trait *T* ← Candidate Pool of Trait *T* ← Demographics of *T* ← environment (S_{t-1})→ environment (S_{t0}) → environment (S_{t+1}) → conflict initiation.

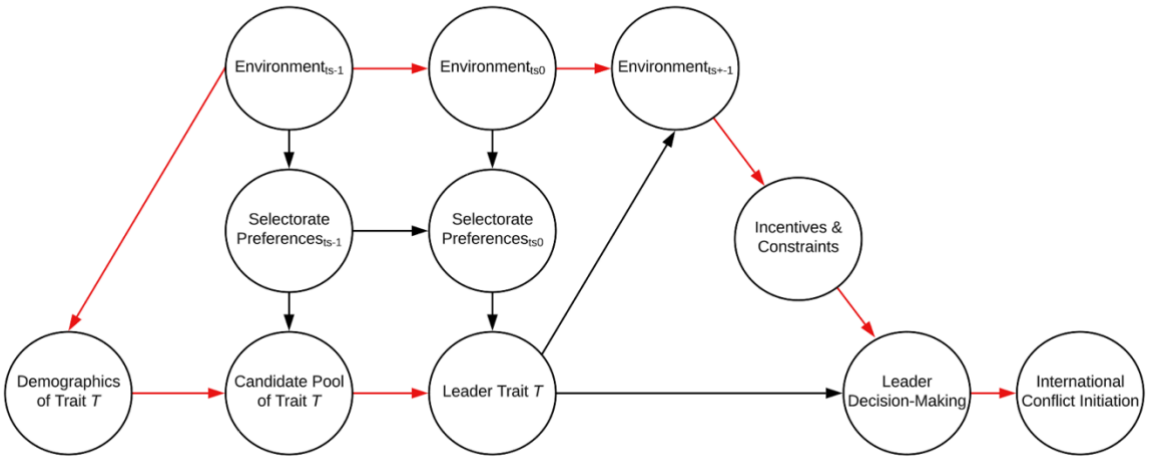


Figure 3-5: DAG of confounding due to direct demographic candidate selection effects (arrows highlighted in red).

The second is that the environment and demographics are only correlated due to being a common effect of another factor, which I refer to as an indirect candidate selection effect. Age is perhaps the best trait to illustrate how this indirect confounding could work. We know that powerful states are more likely to initiate conflict but are also probably more likely to be richer and have higher life expectancies. Consequently, older leaders might only appear more likely to

[126] According to the LEAD dataset, see: Ellis, Horowitz, and Stam, 'Introducing the LEAD Data Set'.

initiate conflict because older individuals are likelier to become candidates in powerful rather than weaker states. More generally, though, the confounding for any leader trait susceptible to this confounding effect would go through the following back-door path (Figure 3-6): Leader Trait *T* ← candidate pool of trait *T* ← population demographics ← Domestic Factor Z ← (or →) environment (S_{t-1})→ environment (S_{t0}) → environment (S_{t+1}) → conflict initiation.

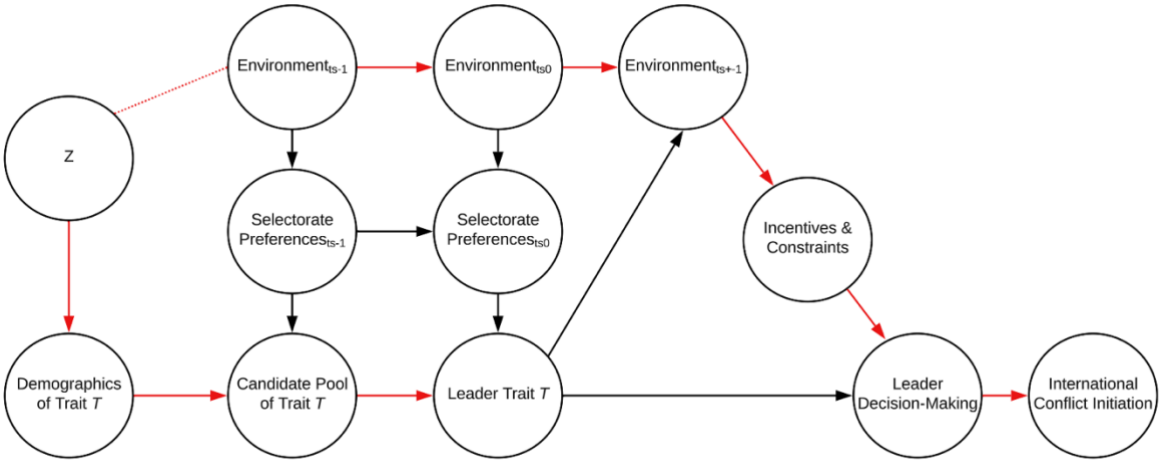


Figure 3-6: DAG of confounding due to indirect demographic candidate selection effects (arrows highlighted in red).

Another possible confounding bias due to demographics is affinity voting – the tendency for people to vote for candidates with similar demographics; regardless of whether the environment-demographic association is direct or indirect, even if their reasons for doing so have nothing to do with the environment or international politics. In this way, the environment does not cause the selectorate preferences to change, but rather demographics cause selectorates preferences to change, and demographics are just correlated with the environmental factors that directly cause future conflict (Figure 3-7). Regardless of whether the environment and demographics are directly or indirectly associated, demographics can cause selectorates to be more likely to choose candidates and then leaders with similar traits.

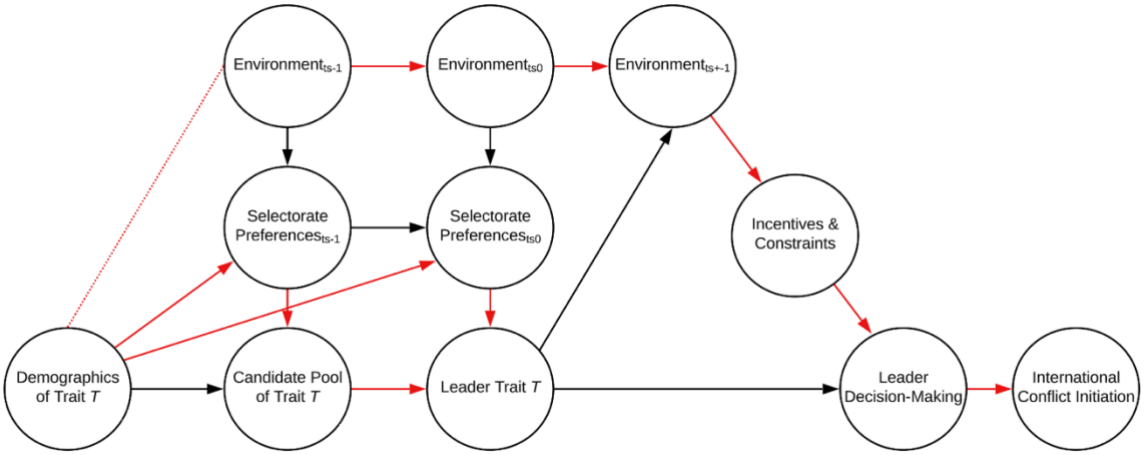


Figure 3-7: DAG of confounding due to affinity demographic candidate selection effects (arrows highlighted in red)

3.1.2.3 Candidate Constraint Selection Effect

The third way candidate selection can potentially confound the trait-conflict initiation relationship is through what I call a "candidate constraints" selection effect. The central idea here is that conflict-prone states could have a common constraint, such as legislative rules or cultural norms, which prevent individuals with certain traits from being allowed to be candidates in the first place (Figure 3-8). For example, Israel has no restrictions on females being Prime Minister, yet only males can be Head of State in Brunei. Given that Israel has historically been significantly more conflict-prone than Brunei, it is unsurprising that Golda Meir initiated more conflicts than any male Brunei leader. Now, if this were the case more generally (that conflict-prone states are less likely to restrict female candidates than conflict-averse states), then it would be self-explanatory why female leaders *appeared* more likely to initiate conflict in previous research – they are just less restricted from being candidates in states where the probability of conflict is higher.

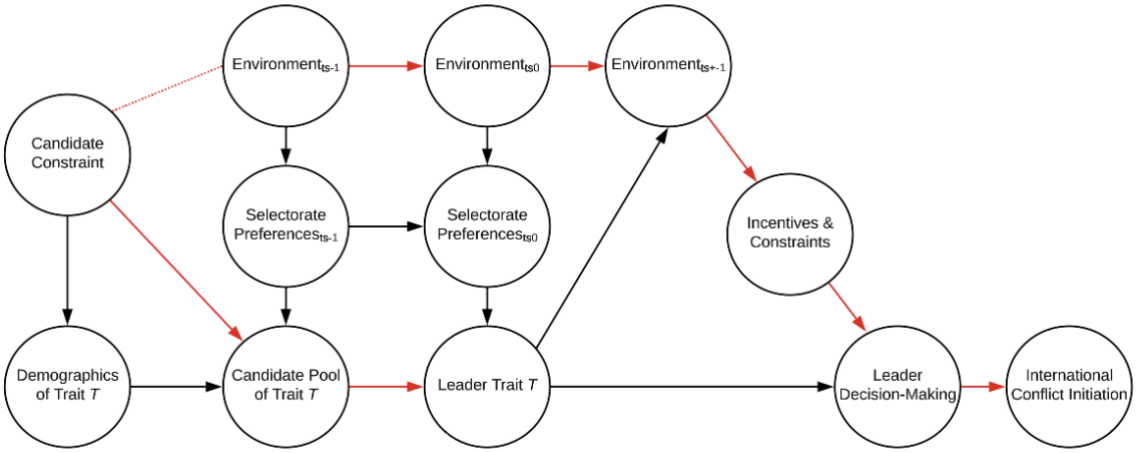


Figure 3-8: DAG of confounding due to candidate constraint selection effect (highlighted by red arrows).

3.2 Self-Selection Effects

Another type of selection effect that can confound the trait-conflict association is self-selection. Again, by self-selection I mean any time a leader installs themselves into office through force. Individuals with high-risk traits might be both more likely to attempt self-selection and be more successful at doing so than individuals with low-risk traits primarily because of their traits. For instance, coups are regularly conducted and led by members of the military, so it is likely that coup leaders with military experience, especially those high-ranking members, have more support from the military and populace and hence have a higher probability of success than leaders with low-risk traits such as no military experience or young age. In this way, self-selection confounds the trait-conflict initiation association differently than the above selection effects. Instead of a back-door path, self-selection could entirely mediate the relationship between a high-risk trait and conflict. For example, given that coups almost always involve the military, it is obvious that those who enter office through coups are more likely to be ex-military, and since coups shake up their state's domestic and international political environment, it could be that the causal effect of military experience on conflict initiation might be entirely mediated by self-selection (Figure 3-9). In other words, ex-military individuals might

be more likely to self-select into power through a coup, which then causes a shock to the status quo, damaging relations with allies as well as creating new rivals, and so the probability of future conflict likely increases. Conversely, military coup members might appoint a civilian leader to restore democracy, decreasing the probability of future conflict. As a result, two very different effects of self-selection could correlate with leader traits. Military leaders are probably more likely to self-select through coups and steer their state toward personalist regimes (i.e., Gaddafi). In contrast, civilian leaders appointed after coups are probably more likely to steer their state towards democratic ideals since the lack of being in the military is likely one of the key reasons pro-democratic coup leaders choose them (e.g., Jose Linhares).[127] Thus, it could be that the association between high-risk traits and conflict initiation is only because leaders with that trait are more likely to self-select, and the increase in conflict is only due to the coup rather than the trait (Figure 3-9).

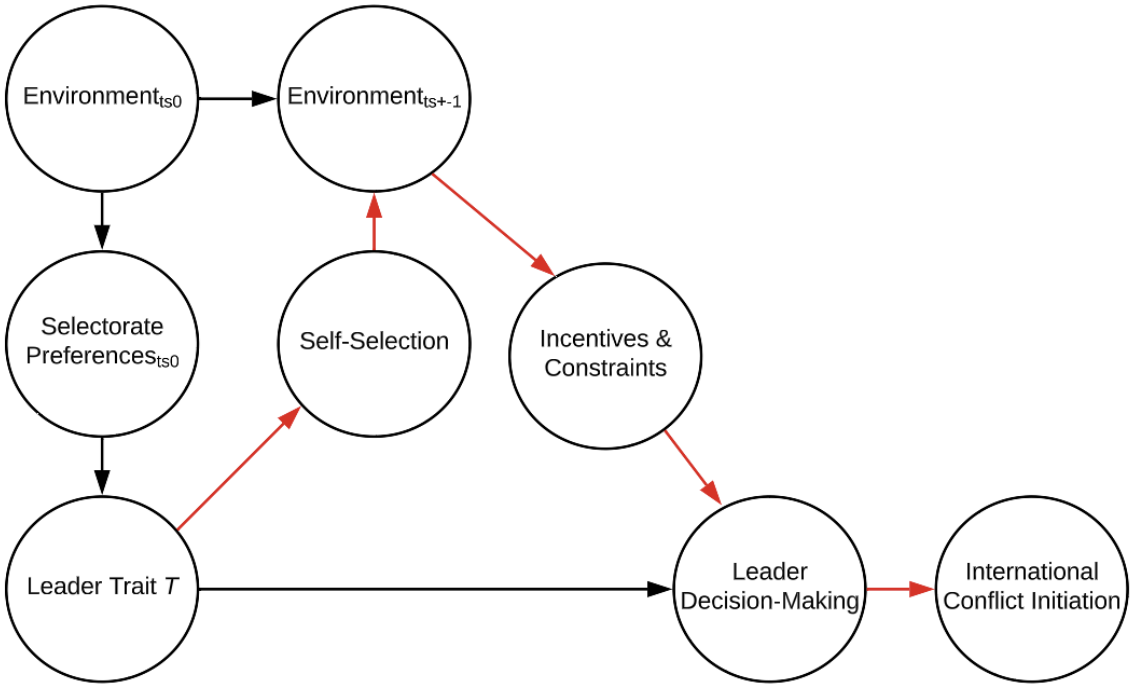


Figure 3-9: DAG illustrating confounding due to leader self-selection effects.

[127] While some coups lead to more democratic regimes, most lead to more authoritarian and violent regimes: George Derpanopoulos et al., 'Are Coups Good for Democracy?', *Research & Politics* 3, no. 1 (2016): 2053168016630837.

3.3 Conclusion

In this chapter I have outlined the various ways that selection could bias the relationship between leader traits and conflict initiation. Namely, leaders with high-risk traits might be more likely than those with low-risk traits to be selected into environments where the probability of conflict is greater. There are three ways this can happen. First is leader selection – that selectorates might be more likely to choose leaders with high-risk traits over those with low-risk traits when international conflicts are especially salient, directly putting these leaders into office in dangerous times. Second, a candidate selection effect could produce a disproportionate number of candidates with high-risk traits into environments where conflict is more likely. This candidate selection effect could be because political parties are more likely to nominate high-risk over low-risk candidates in more dangerous times, or because conflict-prone states tend to have similar demographics, and so candidates within these states are more likely to have high-risk than low-risk traits. Lastly, individuals with high-risk traits might be more likely than those with low-risk risk traits to attempt coups, and be more successful at doing so, which then causes conflicts to occur due to the effect of the coup rather than their trait. In other words, without controlling for these potential selection effects, we cannot get closer to understanding the extent to which traits influence a leader's decision to start conflicts. In the next chapter, I develop a way forward that I believe can get around the problem of selection.

4 A Way Forward

The problems with selection effects highlighted in the previous chapter reveal the central challenge toward understanding the extent to which traits influence leaders' decisions to start conflicts. There is only one obvious solution; to find a way to remove or at least mitigate most of the bias from any possible selection effect. However, in this chapter, I show the difficulties in trying to control for selection effects due to unobservable and mismeasured variables that we could otherwise control for in an ideal world. I then discuss why other common approaches used in IR are ultimately not suitable to control for selection, and why a new strategy is warranted. I then provide my own strategy, which I call the "counterfactual leader strategy" using hypothetical data to help illustrate the central assumptions and logic. The core of the strategy is that the traits of runners-up in selection events can be used as a placebo to indirectly test for bias due to the selection effects.

4.1 The Fundamental Problem

The three selection effects in the previous chapter highlight exactly how the association between any leader trait and conflict initiation can be *theoretically* confounded. However, what we can do about them in practice is extremely limited since the estimated causal effect of leader traits on conflict initiation is ultimately unidentifiable. To help illustrate how and why this is a fundamental problem, consider a hypothetical world where we set out to investigate whether ex-military leaders are more likely than non-ex-military leaders to initiate conflict (Figure 4-1). Fortunately, we can observe the important confounders of leader and candidate selection effects: the environment, a selectorate preference for veterans, and the number of veterans in the population. For now, assume that we can observe and measure each of these confounders without error – the environmental risk is the latent probability of conflict based on every single factor outside of the leader; selectorate preferences are the average probability of all voters when asked whether they would choose a veteran over non-veteran as leader; and the veteran

population percentage is the share of veterans in the population. Lastly, we also know the veteran status of candidates and the outcomes of the primary elections in each state. Assume everything else in our hypothetical world is the same.

State	Total T=1 Candidates in Winning Political Party Primary	Total T=1 Candidates in Runner-up Political Party Primary	Leader Trait T	Runner-up Trait T	Conflict Initiations	Environmental Risk	Selectorate Preferences	Veteran Population %
A	0/2	0/2	0	0	0	0%	0%	0%
B	0/2	0/2	0	0	0	0%	0%	0%
C	0/2	1/2	0	1	1	10%	10%	10%
D	0/2	1/2	0	1	1	10%	10%	10%
E	1/2	0/2	1	0	2	20%	20%	20%
F	1/2	0/2	1	0	2	20%	20%	20%
G	2/2	2/2	1	1	3	30%	30%	30%
H	2/2	2/2	1	1	3	30%	30%	30%

Figure 4-1: Hypothetical world example showing how selection effects can confound trait explanations of conflict.

Now, if we were only to compare the number of conflict initiations between leaders of each state, it would indeed appear veterans initiate more conflicts on average. Veterans initiated 2.5 conflicts on average, whereas non-veterans initiated only 0.5 on average. However, we can see the major problem when we look at the confounders: veterans are more likely to be elected over non-veterans and be elected candidates as all three increase. So, here we have evidence of both leader and candidate selection effects; as the environmental risk of conflict increases, veterans are more likely to be candidates and win general elections over non-veterans.

In our hypothetical world, we could easily control for these selection effects by simply controlling for the environmental risk, since that is minimally sufficient to remove all back-door paths from leader traits to conflict initiation. In the real world, though, controlling for the

environmental risk is not so easy. While the authors in the existing leader trait studies somewhat acknowledge, anticipate and address the problem of selection, they have largely "black-boxed" leader selection. That is, they do not control for selection directly, but rather only control for the downstream institutional and structural factors that potentially cause selection effects, such as relative power, system concentration of power, recent conflict history, and regime type, among others. Again, this makes sense in theory since controlling for the environment is minimally sufficient. However, the primary problem with this design choice is that "black boxing" selection and instead controlling for relevant confounders of the environment is ultimately insufficient to control for selection effects due to two common problems of observational studies: measurement error and omitted variables.

4.1.1 Measurement Error

While many institutional and structural factors controlled for in the existing leader trait studies are common in IR, the first problem is that they are all measured with some error. This is not to disparage the tremendous work and effort by other IR scholars but rather to highlight the limitation when including these variables in statistical models of international conflict that are quite sensitive to construct validity. Indeed, international conflicts are rare, so even slight mismeasurements of important variables can fundamentally alter the findings.[128] The problem of measurement error is clearest for relative power, which is perhaps the most relevant confounder for a fairly obvious reason: states with more military capabilities, such as China, should be much more likely to initiate conflicts than states with fewer capabilities, such as Cape Verde, regardless of who is in charge simply because they have more capabilities to do so. Therefore, leader trait

[128] Saddam Hussein, for instance, is considered a non-combat military veteran in WLF, yet it appears he just dressed the part and never actually served in the military. If Hussein is changed to having no military experience in WLF, the effect size of military non-combat experience on conflict initiation decreases by about 27% because Iraq started so many conflicts in the 1980's.

analyses must ensure that any relationship between leader traits and international conflict is not simply because more powerful states select different types of leaders than weaker states.

To deal with the potential confounder of relative power, every leader trait study includes relative state power using the Correlates of War's Composite Index of National Capabilities (CINC) score. The CINC score is the most widely used measure of states' relative power in IR and is based on the proportion of a state's national material capabilities (NMC) within the international system each year.[129] However, the problem with using CINC as a proxy for relative power is that it is measured in a way that results in a counterintuitive relationship with conflict initiation. Consider the CINC scores over time in the U.K., U.S., and China (Figure 4-2).

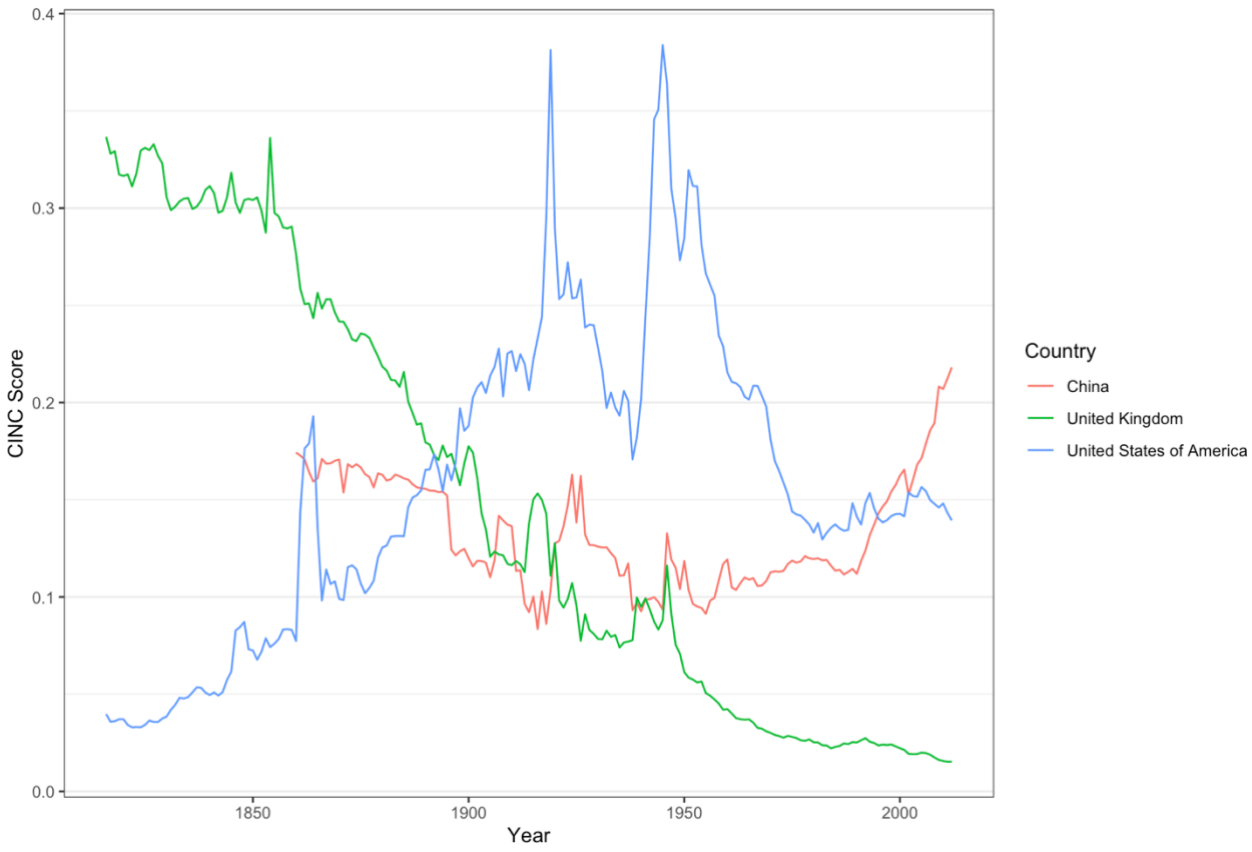


Figure 4-2 CINC score over time in China, the United Kingdom, and the United States

[129] The NMC is the total sum of six indicators: military expenditure, military personnel, energy consumption, iron and steel production, urban population, and total population. J. David Singer, 'Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816–1985', *International Interactions* 14, no. 2 (1988): 115–32.

If the CINC scores are supposed to be linearly related to the probability of conflict initiation – as leader trait studies imply – then this suggests that U.K. leaders in 1850 were around twice as likely to start a conflict than U.S. and Chinese leaders in 1980, given their differences in relative power. This is unlikely, of course, given that militaries' size, technological capabilities, and force projection have dramatically increased, resulting in a much higher probability of conflict over time, regardless of leader. Using CINC scores within states is also problematic for similar reasons. The U.S. has a CINC score nearly twice higher in 1909 than in 2000. Yet, the post-Cold War period is generally regarded as a period in which the US has been the sole hegemon in a unipolar system,[130] so it seems counterintuitive to consider William Taft (low-risk leader) as being selected in a time in which the US was twice as powerful than George W. Bush (high-risk leader). The main problem the examples above point out is that CINC scores are sensitive to changes in state membership in the international system – as new states increasingly enter the international system, every state's share of military capabilities decreases.[131] As a result, CINC scores generally tend to be higher in the past (outside of the two World Wars), even though the size and capabilities of militaries are increasing and we would expect the probability of conflict to increase, not decrease.

While it might then seem reasonable to instead use the raw total across the six indicators of military capabilities, this would ultimately have the opposite problem. Present states have significantly more capabilities than the past and would likely overestimate the probability of conflict as time increases. For instance, Panama in 2012 would be considered nearly twice as powerful as Germany in 1913 using the raw score, even though the former had no standing military at the time and the latter was a great power.[132] Naturally, we might also consider alternative relative

[130] G. John Ikenberry, Michael Mastanduno, and William C. Wohlforth, 'Unipolarity, State Behavior, and Systemic Consequences', *World Politics* 61, no. 1 (2009): 1–27.
[131] Kelly Kadera and Gerald Sorokin, 'Measuring National Power', *International Interactions* 30, no. 3 (2004): 211–30.
[132] According to the NMC dataset, Germany's raw NMC total in 1913 is 389,672, whereas Panama's raw NMC total in 2012 is 558,285.

power measures, but these are often susceptible to the same problems, if not more. GDP, for instance, is the second-most common way to measure relative power but is very similar to CINC scores, leading to the same issues.[133] Other less common but proposed alternative measures suffer from even more problems, such as theoretical inconstancies or limited temporal coverage, and are unlikely to alleviate the concerns of leader selection.[134]

The 1940s U.S. peak exemplifies another problem with using CINC scores, as seen in Figure 4-2 above. The highest CINC scores in the dataset are generally among the countries that fought during the two World Wars. While this is what we might expect, the problem is that if two states are engaged in an ongoing war, no additional conflicts can occur between the two countries. Since states increase their military capabilities during times of war for obvious reasons, leaders of the most powerful states in the system, as measured by CINC scores, are among the *least* likely to initiate a conflict simply because they are already engaged in a war with most of their adversaries in the first place. Germany in 1943, for instance, has one of the highest CINC scores in the dataset, yet since Germany was at war with nearly all its neighbours and the strongest world powers, there were essentially no more states left for Hitler to start a conflict against. Likewise, the U.S. also has some of the highest CINC scores in the dataset in 1942, yet they had no conflict initiations that year because they were either in a war with, or allied to, most other major powers. Thus, the overwhelming number of conflicts are started by weaker states, and the proportion of conflict initiations (i.e., years where leaders initiated a conflict divided by years where they did not) is relatively lower in the most powerful states (Figure 4-3).

[133] Jacek Kugler and Marina Arbetman, 'Choosing among Measures of Power: A Review of the Empirical Record', *Power in World Politics*, 1989, 49–77; Douglas Lemke and Suzanne Werner, 'Power Parity, Commitment to Change, and War', *International Studies Quarterly* 40, no. 2 (1996): 235–60.

[134] For a good discussion of the pros and cons of the different ways of measuring relative power, see: Carsten Rauch, 'Challenging the Power Consensus: GDP, CINC, and Power Transition Theory', *Security Studies* 26, no. 4 (2017): 642–64.

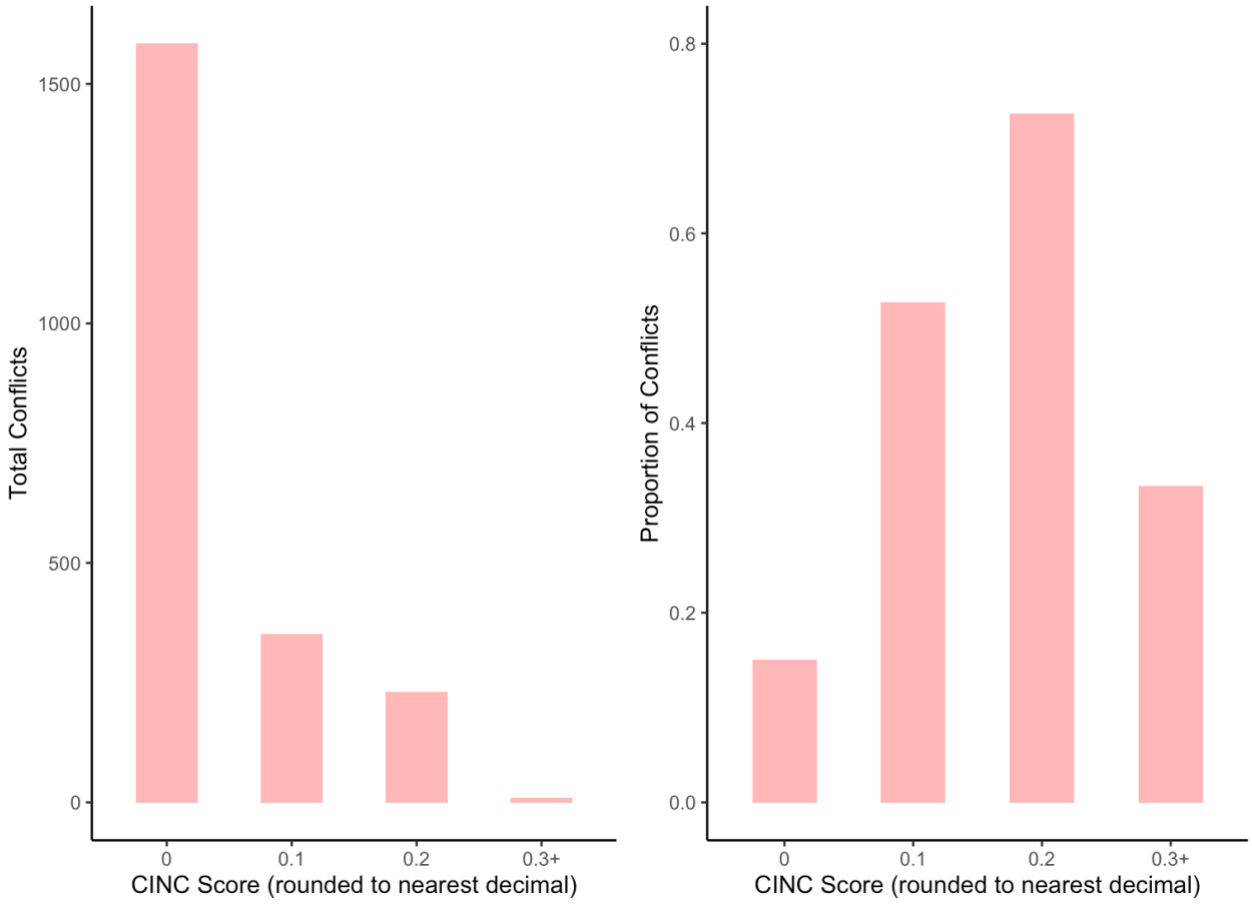


Figure 4-3: Total and proportion of conflict initiations by rounded CINC score

Therefore, not only are CINC scores sensitive to membership changes in the international system, but they are also sensitive to how conflict is measured. Countries with the highest CINC scores often have the fewest opportunities, methodologically, to initiate a conflict in the first place. Thus, given that CINC scores are inflated in the past and are highest when conflicts are unlikely to occur, including CINC scores as a control cannot fully account for the possibility that powerful states might be disproportionally selecting high-risk leaders.

4.1.2 Balance of Force

CINC scores in dyadic models have many of the same problems. In these analyses, relative power is controlled for using the "Balance of Force" between two states in a dyad. This is calculated as the proportional difference in CINC scores between side A and side B in a dyad-year

and results in a variable between 0 and 1, with .5 being equal, 1 being very favourable to Side A, and 0 to side B. While the balance of force measure better captures the fine-grained differences in relative power between two states, this is no panacea to the problem of selection. The primary problem here is that the balance of force does not consider the relative power of a given state compared to all states in the international system. For example, in 1988 the United States has an unfavourable balance of force of .45 against the Soviet Union, whereas Costa Rica has an extremely favourable balance of force of .99 against Dominica. Even though the United States had a significantly lower balance of force against the Soviet Union than Costa Rica did against Dominica, the former dyad is still very much more likely to result in a conflict than the latter given that the U.S. and USSR were the two most powerful states in the system that year and neither Costa Rica nor Dominica even had standing militaries at the time. Because powerful states tend to initiate more conflicts against each other as well as have more equal dyads, the proportion of conflicts to non-conflicts is biased toward equal dyads (.5 balance of force). Indeed, when regressing conflict initiation using a nonlinear quadratic term of CINC and balance of force (instead of a linear measure as in all the leader trait studies), both are significant and have the highest probability toward the middle-ranged values (Figure 4-4).

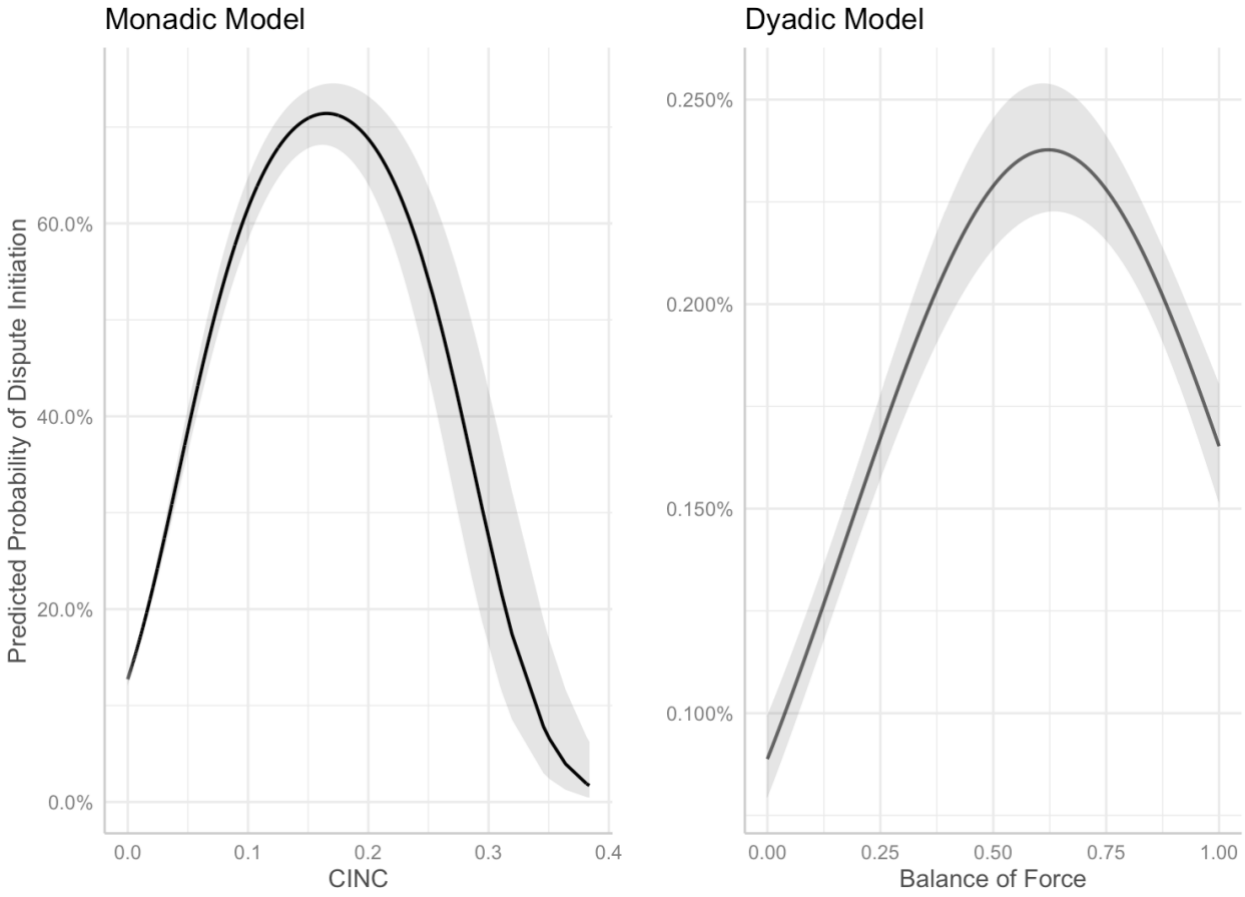


Figure 4-4: Predicted probability of conflict initiation by quadratic CINC and Balance of Force.

The point I am trying to get across is that relative power is perhaps the most relevant confounder for the trait perspective, but the use of CINC scores or any other alternative measures do not adequately address the nuances of relative power over time – a crucial element to consider its effect in any analysis of leaders and international conflict. These measures cannot fully account for the possibility that high-risk leaders might be disproportionately selected into more powerful states, and hence into environments where conflict is more likely. In the monadic models, CINC scores are inflated farther back in time and only ensure that leader selection is not occurring in more powerful states of the past. Whereas in the dyadic models, controlling for balance of force only ensures that leader selection is not occurring in equal dyads rather than powerful states. Thus, until IR develops more precise relative power measures, existing measures are unlikely to account for selection fully.

4.1.3 Past Conflict

Another important confounder that suffers from measurement error is past conflict. This control is particularly important to include because past conflicts probably lead to future conflicts, so leader trait studies need to ensure that high-risk leaders are not disproportionately selected into states where conflict is more frequent. Past conflict is measured using a peace-year spline – the total number of years elapsed since the state last initiated a conflict, cubed – and is the most influential out of all the institutional and structural variables in the WLF systems-risk model.[135] While including the peace-year spline is reasonable and has long been recommended by IR scholars studying international conflict,[136] the primary problem is that states which achieved independence must start at 0, so they are also considered to have initiated a conflict in the past year in their models. But because the peace-year spline strongly correlates with conflict initiation, many small nations that achieved independence during decolonisation after the Second World War have unrealistically higher predicted probabilities of conflict than states where we would expect conflict to be more likely. For instance, Lyndon Pindling (low-risk leader) became the Prime Minister of the Bahamas upon independence in 1973, and the WLF systems risk model predicts he entered an environment with a probability of conflict twice higher than Charles de Gaulle of France in 1945. While France certainly was at a lower risk of conflict in 1945 than say, 1939, they were still a great power with global geopolitical interests and were soon to be embroiled in decolonisation and the Cold War, and thus likely were to be involved in many future conflicts. Conversely, the Bahamas' interests were largely limited to their coast, had no militarily powerful rivals, and so they were less

[135] The peace year spline has a variable importance factor of 19. Compare this to some of the more relevant ones: CINC (14), system concentration of power (5.5), democracy (3), and autocracy (2.5).
[136] Nathaniel Beck, Jonathan N. Katz, and Richard Tucker, 'Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable', *American Journal of Political Science* 42, no. 4 (1998): 1260–88.

likely to be involved in future conflicts. Thus, controlling for the peace-year spline is not adequate to remove bias from selection because it often overestimates the likelihood of conflict in small conflict-averse states with low-risk leaders, such as Pindling in the Bahamas, and underestimates the likelihood in large conflict-prone states with high-risk leaders, such as De Gaulle in France.

4.1.4 Self-Selection

The various leader trait studies do somewhat attempt to control for self-selection. However, this is controlled for using a binary "irregular" entry variable, which is defined as subverting the established rules and procedures of leader selection in their state.[137] While this might seem to be a reasonable control for self-selection, the primary problem is that irregular entry describes a broad range of ways these leaders are irregularly selected. Indeed, 17% of all leaders in Archigos are considered to have entered office irregularly. Concerning the problem of self-selection, we only need to control for relevant ways of irregular entry that could be confounders by causing a disproportionate number of high-risk leaders to self-select compared to low-risk leaders. Namely, we only need to control for those who self-select themselves because that is the only "irregular" entry that is likely a confounder. The irregular entry variable does not pick up this potential selection effect since it also includes many of the more peaceful civilian leaders appointed after coups by others instead of self-selecting themselves. As explained previously, pro-democratic leaders who are appointed after coups are probably much less likely to initiate conflict than dictators who forced their way into office, and the latter events are also probably much more likely to cause a disruption in the environment and increase the risk of conflict compared to the former. Thus, to get closer to the true causal effect of military experience on conflict we need to more specifically control for self-selection instead of the broader irregular entry variable.

[137] Goemans, Gleditsch, and Chiozza, 'Introducing Archigos', 272.

4.1.5 Omitted Variables

While the inclusion of the various environmental factors like relative power and peace years in their models does reduce at least some bias, these controls alone cannot entirely remove bias from selection without the impossible task of accounting for every single factor that could cause states to both select leaders with high-risk traits and initiate conflict. In other words, there are likely to be other important factors that cause both the selection of certain leaders as well as future conflict that we simply cannot quantify or include in any reasonable way. Compare the selection of Adolf Hitler of Germany in 1933 (high-risk leader) and Ibrahim Nasir of the Maldives in 1965 (low-risk leader). According to the systems risk model in WLF, Nasir was selected into an environment with a 15% chance of initiating conflict, whereas Hitler was selected with only an 8% chance. However, do we really believe the Maldives in 1965 were twice as likely to initiate conflict than Germany in 1933? Unlikely. One possible explanation is that German elites and the mass public wanted a leader with traits like Hitler to return their territorial, political and economic losses after their loss in the First World War. Conversely, Maldivian elites and the mass public likely had little reason to select a high-risk aggressive leader since militarised international confrontations were much less likely. What this particular example is pointing out is that largely unobservable and difficult-to-quantify factors, such as Germany's revanchism, might also confound the relationship between leader traits and conflict, yet these kinds of factors cannot be easily included in leader trait analyses. As a result, even if previous studies controlled for every relevant observable environmental factor, they cannot account for other important but unobservable ones.

4.2 Bringing it All Together

What all this means, unfortunately, is that the estimated causal effect of any given leader trait on conflict is ultimately unidentifiable because the key confounders are all unobservable or measured with error (Figure 4-5). Again, the environment is unobservable because we cannot possibly include every relevant component and even those we can observe and include, such as

relative power and past conflict, are largely measured with error (denoted as environment* in Figure 4-5 below). Alternatively, we would not need to control for the environment if we could control for selectorate preferences, demographics, and any relevant candidate constraint *Z* since these would close both back-door paths through the environment anyway. But here, we run into similar problems since selectorate preferences are also unobservable. Ideally, we would need some kind of feeling thermometers across time and states on whether their citizens choose leaders because of their traits and in response to increased conflict. No such data exists, and even attempting to collect such data is too herculean of a task for any number of researchers, so will not be attempted here. Similarly, controlling for demographics might appear easier but ultimately is problematic for comparable reasons. For example, states that frequently engage in conflict might have more ex-military leaders because these states tend to have a more pro-military culture or "pressure to serve", which we obviously cannot measure with any accuracy. While we could use the proportion of veterans in the population as a proxy, only one state – the United States – has census records detailing the percentage of veterans in the population. So again, using demographics as a proxy to control for a domestic candidate selection effect is ultimately inadequate, and proxies do not fully alleviate bias even if they are measurable.

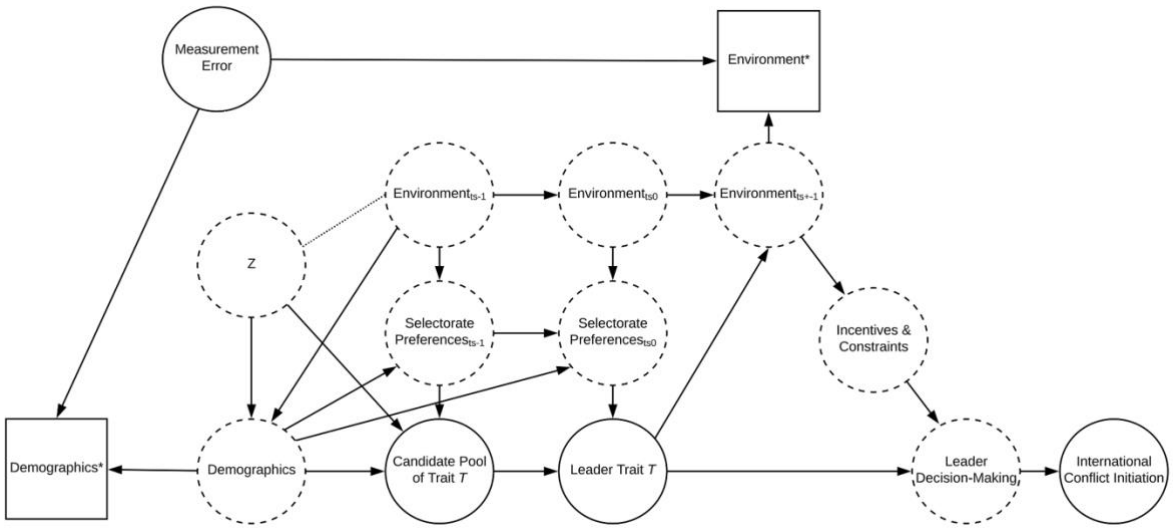


Figure 4-5: DAG illustrating the fundamental problem of the trait perspective.
Note: Dashed circles mean variable is unobservable. Squares mean variable is controlled for.

Even though we have more thoroughly identified the possible confounding due to the selection effects, we are still ultimately unable to get an unbiased estimate of the true causal effect of a given leader trait on conflict initiation. However, endogeneity is a problem to be solved, not just shrug our shoulders and give up. The aim of the rest of this chapter is to formulate a strategy that can indirectly test whether the effect of a given leader trait on conflict is likely biased due to the selection effects.

4.3 How to get around the fundamental problem

So, what do we do about the problem of selection? Our first instinct, perhaps, would be to try and use an instrumental variables (IV) approach. The fundamental idea of the IV approach is first to predict the treatment with an exogenous variable (i.e., instrument) that is only correlated with the outcome through the treatment, and then only use the part of the treatment explained by the instrument to predict the outcome.[138] For our purposes, however, the central problem with the instrumental variables approach is that it is extremely difficult to find an exogenous variable that predicts the variation of leader traits globally that is not also in some way correlated with conflict initiation through other ways. In one study on military experience, for example, the authors use conscription as an instrument, claiming that this is appropriate since conscription occurs too far in the past to influence conflict during a future leader's tenure.[139] However, according to their own data, whether a state implements conscription is still positively and significantly associated with conflict initiation even 40 years into the future.[140] This is

[138] Erik Igelström et al., 'Causal Inference and Effect Estimation Using Observational Data', *J Epidemiol Community Health* 76, no. 11 (2022): 965.
[139] Horowitz and Stam, 'How Prior Military Experience Influences the Future Militarized Behavior of Leaders'.
[140] Ibid. I simply regressed conflict initiation on a 40-year lag of conscription using their same data and model. Part of the problem is that the conscription variable is miscoded. Many U.S. Presidents, for example, served in the Second World War when the US implemented the draft, yet they are not coded as having served during conscription. Data on conscription taken instead from Nathan Toronto, 'Military Recruitment Data Set', *Ft. Leavenworth, KS: Foreign*, 2007.

probably because states that utilised conscription in the past have done so because they have been traditionally more likely to engage in wars and conflicts throughout history and so likely continue to do so even far into the future. Thus, even variables in the distant past are still correlated with future conflict initiation in some form or another, which violates the central assumption of the IV design.

Another potential strategy is to use a Differences-in-Differences (DiD) approach. The fundamental idea of the DiD approach is to measure the outcome in cases before and after a treatment is introduced and then compare it to the outcome in cases where the treatment is never introduced. In our case, we could compare the frequency of conflicts in a state before and after an opposite change in a leader trait, and the difference would ideally reveal the estimated causal effect. For example, if veterans were more likely to initiate conflict, we should expect to see more conflict initiations after a state goes from a non-veteran to a veteran leader than when they go from a veteran to a non-veteran leader. However, the central problem for the DiD strategy is the low number of leaders with different traits in the same country and the rarity of international conflicts. According to LEAD, For instance, 53% (n=340) of all veteran leaders succeeded a non-veteran, but only 35% (n=118) of these even initiated a conflict, and their first one occurred after 1139 days in office, on average.[141] Furthermore, the frequency of conflicts globally increases over time, so if ex-military leaders also increase over time globally, then we would see more conflicts after states change from ex-military to non-military leaders, even if there were no true causal relationship or selection effect. Lastly, the DiD strategy does not overcome the problem of the potential First-Image Reversed leader selection effect because it could still be that veterans are more likely to be chosen over non-veterans in dangerous times. Thus, not only is there a lack of variation in conflict initiations immediately before and after the

[141] Data on leader traits from WLF. Data on conflict initiations taken from MIDB v5.0.

change in leader traits, but there is also the potential of both temporal correlation and selection effects that ultimately undermines the utility of a DiD strategy.

Another strategy we could consider is to use natural deaths. Using this method, we would compare the difference in conflict initiations between leaders who came into office after the natural death of the previous leader, such as when U.S. Vice President Calvin Coolidge became President following Warren Harding's death from a heart attack. The key logic here is that since natural deaths are unrelated to anything that might also cause future conflicts, then there are no potential confounders to the trait-conflict relationship. This strategy is increasingly becoming more common in leadership studies as it is thought to sufficiently remove bias from selection effects since leaders enter office "as-if" randomly.[142] A few leader trait studies even use the natural deaths strategy for robustness tests, replicating their original findings and reinforcing their conclusions.[143] However, the primary problem with the natural death strategy is that those who eventually replace a naturally deceased leader are still selected in various ways, most often being elected in leadership or Vice-Presidential elections or being directly appointed by a Head of State, and so selection is still a possible confounder. Thus, bias from the potential selection effects is still not removed even when leaders enter office after the natural death of their predecessors.

The last common strategy we could consider as a solution to the selection problem is to use close elections. The logic of the close election strategy is simple: in elections where the top-two candidates are separated by less than 2% of the total vote, who ultimately wins is thought to be close to random since the vote count could shift either way by pure chance alone due to the

[142] Jost Sieweke and Simone Santoni, 'Natural Experiments in Leadership Research: An Introduction, Review, and Guidelines', *The Leadership Quarterly* 31, no. 1 (2020): 101338.
[143] Horowitz and Stam, 'How Prior Military Experience Influences the Future Militarized Behavior of Leaders'; Horowitz, McDermott, and Stam, 'Leader Age, Regime Type, and Violent International Relations'; Barceló, 'Are Western-Educated Leaders Less Prone to Initiate Militarized Disputes?'

complexities in counting millions of votes.[144] Applied to our case, we could look at close elections where the top-two candidates had a high-risk and low-risk trait and then use a regression discontinuity technique to see if there is a significant difference in the number of conflict initiations within the 2% cut-off. The primary limitation of this strategy, though, is statistical power. Close elections are already quite rare and probably even more so where the top-two candidates have opposing traits, especially if demographics cause many candidates with the same trait. Furthermore, the assumption of randomness under the 2% vote-difference threshold is not only questionable but also unsuitable for Parliamentary elections since seats rather than votes determine them, and so it is less likely that a 0-2% difference in the total number of seats could be due to random vote counting complexities. Even with these problems, though, the close election strategy is still entirely feasible for presidential elections as long as there are an appropriate number of observations necessary for sufficient statistical power. However, it would be unwise to rely entirely on the close election strategy when it depends on an ample number of observations that are unlikely to be available under the conditions outlined above.

The above review ultimately points out that we have a very difficult challenge to overcome. Most existing strategies cannot fully control for selection effects when trying to understand the extent to which traits influence leaders to initiate conflict. The close election strategy is perhaps the nearest we can get to control for selection effects, but as noted, this depends on several questionable assumptions and, even then, is likely to result in low statistical power due to a lack of observations. Nevertheless, I believe the next best strategy available is to use what I term the "counterfactual leader strategy." Because the concept of the counterfactual leader is being introduced here for the first time, it is first necessary to begin by explaining what

[144] For a good example of this method, see: Bertoli, Dafoe, and Trager, 'Is There a War Party?'

the counterfactual leader is to help clarify why and how we can use these individuals to indirectly test for selection effects.

4.4 The Counterfactual Leader

To understand whether some trait causes conflict we ideally want to know the counterfactual outcome – the number of conflict initiations a state would have initiated if they had selected a leader with the opposite trait instead. While we cannot observe the counterfactual outcome across states, we can observe whether a state came close to having the counterfactual *treatment* – a leader with the opposite trait. That is, in every leader selection event there is always at least one other individual who "almost" became leader but ultimately did not. Accordingly, we can consider these individuals as the "counterfactual leader" since they are the *closest* we can get to observing a world in which a state selected someone else. In the Methodology Chapter, I more specifically detail how to identify counterfactual leaders across the different types of selection mechanisms, but to briefly summarise here, counterfactual leaders are typically runners-up in elections and political appointments, the new heir apparent in successions, the incumbent in coups, or the coup leader in failed coups (or whom the coup leader wanted to install).

The purpose of identifying the counterfactual leader is that we need to understand why a given state selected a leader with T_1 rather than T_0, and vice versa. That is, we cannot assume states would have initiated fewer conflicts if they had a leader with trait T_0 instead of trait T_1 if we would have to completely rewrite history to assume that scenario could have even happened. Thus, the critical assumption for determining the counterfactual leader is who could have been the leader with the least rewriting of history. More specifically, we need to adhere to the "minimal-rewrite rule" – to even assume the counterfactual treatment could have occurred (i.e., a leader with the opposite trait), we cannot assume away too much of history to where it also

affects the likelihood of future conflict as well.[145] For example, if we were to question whether a war-winner would have started more conflicts than a war-loser in Germany in the last half of the 20th century, we would run into an obvious problem given the only two wars Germany fought previously to this period were both losses. Accordingly, it was extremely unlikely a German leader with war-winning experience could have been leader during this time. So, to even assume a counterfactual world in which a war-winner could have been the leader of Germany is to assume a world in which Germany won either World Wars or another war that did not happen but reasonably could have. Yet, if we were to assume an alternate history where either of these counterfactual worlds could have occurred, then this would fundamentally change the environment any counterfactual German leader would face as well. That is, had Germany won either war, it is likely that they probably would have been more conflict-prone regardless of leader war experience simply because they would not have had their militaries largely dismantled and foreign territories taken as they did in the aftermath of both wars. Accordingly, any German leaders would end up in an environment where they had more means to use military force abroad, and hence war winners would likely appear more conflict-prone than others as well. This problem is exactly what we are trying to discover – whether the environment might cause different states to have different probabilities of selecting leaders with different traits.

However, we may not necessarily be able to say that the counterfactual leader would have been the leader instead if we could rewrite history. Indeed, if we needed to rewrite history to assume a counterfactual world in which the runner-up was selected, we would need to assume away the very reasons why they were not selected in reality. For instance, Walter Mondale was the runner-up to Ronald Reagan in the 1984 U.S. election. However, given that Mondale lost the election in a landslide (525 to 13 electoral votes), to assume he could have been the leader

[145] Richard Ned Lebow, 'What's so Different about a Counterfactual?', *World Politics* 52, no. 4 (2000): 550–85; Philip E. Tetlock and Aaron Belkin, 'Counterfactual Thought Experiments in World Politics', *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*, 1996, 1.

instead we would need to assume away the very reasons why he lost by such a large margin. Accordingly, it could be argued that a minimal rewrite of history would have actually led the Democratic Party primary runner-up, Gary Hart, to become leader rather than Mondale, since that contest was relatively closer than the Presidential election (1,929 to 1,164 delegate count). However, we would then need to rewrite history at more places in times to assume Hart could have beat Mondale, and then could have beat Reagan to become U.S. President. Therefore, we can consider runners-up like Mondale as the *most likely* counterfactual leader, given a minimal rewrite of history. That is, we know Mondale was the "next best" option, so to assume anyone else could be the counterfactual leader, we would also have to rewrite away the very reasons why he was good enough to be nominated for President but not good enough to win. Accordingly, runners-up in selection events are the closest we can get to observing the potential counterfactual treatment while adhering to the minimal rewrite rule, and without needing to conduct an in-depth counterfactual analysis for each leader selection event.

4.5 The Strategy

With the counterfactual leader conceptualised, we can now return to the strategy for dealing with selection effects. The fundamental importance of the counterfactual leader is that their trait can be used as a proxy to test for any selection effects since the traits of the leader and the counterfactual leader are both causal effects of the same data-generating process. That is, without even needing to observe any other factors, we know that candidate selection effects will tend to result in both the leader and counterfactual leader having the same trait (e.g., conscription causes all candidates in an election to be veterans). In contrast, leader selection effects will tend to result in each having opposite traits (e.g., voters prefer veteran leaders, so choose a veteran candidate over a non-veteran candidate in the election). Of course, these simple correlations are insufficient to tell us whether selection effects are occurring since we would need to see a conditional association between them and some other causal variable of

interest. But again, the fundamental problem is that we cannot observe any of the causal variables potentially responsible for the selection effect. So, how can we not only test for the selection effects but infer the association is causal, or at least not spurious? Given this fundamental problem, I believe the best we can do is to infer selection effects by testing whether the effect of leaders' traits on conflict depends on whether their state could have received the counterfactual treatment.

To help illustrate why and how this strategy allows us to infer the various selection effects, let us consider a hypothetical study where our aim is to investigate whether leaders with trait T_1 are more likely than leaders with trait T_0 to initiate conflict. Now, we have data from three hypothetical worlds; one that appears to show support for the trait perspective (Trait World), one that appears to show candidate selection effects (Candidate Selection World), and another that appears to show leader selection effects (Leader Selection World). To simplify the hypothetical study, let us assume that leaders in all three worlds have been selected in 200 Americanesque presidential elections. The only observable difference between them is the leaders' and counterfactual leaders' traits and the number of conflict initiations. In other words, we have the same problem of unobservable confounders in all three worlds as in reality.

The first step is to find the treatment effect of leader trait T on conflict initiation across the entire population of leaders. We will call this the "population treatment effect," which allows us first to confirm an association between trait T and conflict initiation without any controls. Indeed, we would not need to test for selection effects and bias if there were no potential causal effect of interest. However, since we are interested in traits that are already known to have a positive association (those mentioned in Chapter Two), we would ultimately just be replicating previous research at this point, but this step is necessary to ensure that any evidence for selection we may find is not simply due to a different data generating process. So, let us assume we confirm with previous research that there is a significant and positive association between trait T and conflict initiation in our population of leaders. In all three worlds, the population treatment

effect is the same (Table 4-1); leaders with T_1 initiated four times as many conflicts as leaders with T_0.

Table 4-1: Simple comparison between leader trait T and conflict initiations for each hypothetical world.

n	Leader Trait T	Number of Conflict Initiations
100	0	10
100	1	40

While we might reasonably conclude that this disparity is evidence that T_1 likely causes leaders to be more conflict-prone in each world, we also know this could instead be due to either candidate or leader selection effects mentioned in the last chapter. However, since the variables potentially causing the selection effects are unobservable, we are in a bind as to how to tell whether selection effects are confounding the trait-conflict relationship in all three worlds.

We can get around the problem of selection effects by including the trait of the counterfactual leader when each leader was selected. Next, we find the effect of counterfactual leader trait T on conflict initiation. We will call this the "population *placebo* treatment" effect because the counterfactual leader trait should not have any meaningful causal effect on conflict initiation. Of course, there may be reasons to question this assumption if the counterfactual leader had a prominent role in the government prior to or after selection. I will discuss this assumption more later, but for now, I would like just to get the main idea across. So, assume that the counterfactual leader in each case has not had any significant role in the government before or after selection, and so cannot possibly have a causal effect on conflict initiation.

Next, we then compare the population treatment effect and the placebo treatment effect. In doing so, there are three possible outcomes; the population placebo effect can be the same, null or opposite to the population treatment effect. Since we know that leader trait *T* is positive and significantly associated with conflict initiation in all three worlds, we ultimately want to see if the counterfactual leader trait *T* is either positive (the same), zero (null), or negative (the opposite). We run this step for each of the three hypothetical worlds and find the results in Table 4-2 below.

Table 4-2: Population treatment effect and placebo treatment effect in each of the three worlds.

Trait World

n	Leader Trait T	Number of Conflict Initiations
100	0	10
100	1	40

n	Counterfactual Leader Trait T	Number of Conflict Initiations
100	0	25
100	1	25

Candidate Selection World

n	Leader Trait T	Number of Conflict Initiations
100	0	10
100	1	40

n	Counterfactual Leader Trait T	Number of Conflict Initiations
100	0	10
100	1	40

Leader Selection World

n	Leader Trait T	Number of Conflict Initiations
100	0	10
100	1	40

n	Counterfactual Leader Trait T	Number of Conflict Initiations
100	0	40
100	1	10

Since the trait perspectives assumes there are no selection effects, the counterfactual leaders' trait must be a random variable and so the placebo treatment effect should be null. If, however, the placebo treatment effect is associated with conflict initiations in the same way as the population treatment effect, then this is our first evidence that the trait-conflict initiation association is likely biased due to a candidate selection effect. We see how this would look in

Candidate Selection World since the effect of counterfactual leader trait T is the same as leader trait T. Conversely if the placebo treatment effect is negative, then this is initial evidence that the population treatment effect is biased due to a leader selection effect, which we see in Leader Selection World. However, it is necessary to move on to the next step to understand how we can infer whether the positive placebo effect results from candidate or leader selection effects in these two worlds.

To figure out which selection effect is responsible, we need to break down each leader selection event into four strata based on whether a state almost had the counterfactual treatment – a leader with the opposite trait. Since counterfactual leaders are the closest to observing the potential counterfactual treatment, there are four possibilities for each selection event (Table 4-3): a leader with T_0 and a counterfactual leader with T_1 (stratum 1), a leader with T_0 and a counterfactual leader with T_1 (stratum 2), a leader with T_1 and a counterfactual leader with T_0 (stratum 3), or a leader with T_1 and a counterfactual leader with T_0 (stratum 4). Without yet considering why one was selected over the other, we can see two baseline probabilities that a state was *almost* assigned the counterfactual treatment. If both the leader and counterfactual leader have the same trait, then we know the probability of almost being assigned the counterfactual treatment equals 0%. Conversely, if the leader and counterfactual leader have the opposite trait, then we know the probability of almost being assigned the counterfactual treatment is greater than 0% and less than 100%.

Table 4-3: Table of leader traits, counterfactual leader traits, and counterfactual treatment probability.

Strata	Leader Trait *T*	Counterfactual Leader Trait *T*	Counterfactual Treatment Probability
1	0	0	0
2	0	1	0<1
3	1	0	0<1
4	1	1	0

The importance of breaking it down into these four strata is that if selection effects did not bias the treatment effect of leader trait *T*, then the observed outcome should not differ depending on the probability of receiving the counterfactual treatment (0 or 0<1). Conversely, if the observed outcome does depend on the probability of receiving the counterfactual treatment, then the only way this disparity can make theoretical sense is through some selection effect (as long we accept the placebo assumption), either those that I have posited previously or some others that are unknown. For example, why one election has zero ex-military candidates (i.e., strata 1) but another has two ex-military candidates (i.e., strata 4) can only have one non-random cause – candidate selection. Indeed, what other reason could produce this besides randomness? In these two elections, we also know that why a non-veteran is elected in the former and a veteran is elected in the latter is not caused by any leader selection effect because both candidates had the same trait. So, if the veteran were to initiate more conflicts than it can only be a *potential* candidate selection effect since it cannot be explained by leader selection. We can therefore define a candidate selection effect as the effect of T_1 within strata 1 and 4.

Consider two other elections, one where a non-veteran won against a veteran (i.e., strata 2) and another where a veteran won against a non-veteran (i.e., strata 3). We know the difference between the two is not caused by candidate selection since there is only veteran and non-veteran

candidate in each. Furthermore, why the veteran or non-veteran veteran won can only have one non-random cause – leader selection. Accordingly, if the veteran were to initiate more conflicts than it can only be a *potential* leader selection effect since it cannot be explained by candidate selection. So, we can define a leader selection effect as the effect of T_1 within strata 2 and 3.

What we ultimately want to do with the two selection effects is think about why and how each can come about and what any alternative findings theoretically mean. But to do this, we first need to unpack the necessary assumptions about our unobservable key confounder, the environmental risk, for either the trait or environmental perspective to make sense. We know that conflict initiations will not be entirely explained by leader trait *T for all three worlds*, so the remaining variation must come from somewhere in the environment. The missing variation in conflict initiations must also not come from other leader traits either, because that would necessarily result in confounding. Accordingly, the trait and environmental perspective require that the causal effect of the environment (denoted hereafter as *CEE*) must adhere to a set of minimal conditions. For the trait perspective, the *CEE* must explain less than 100% of the variation in conflicts among leaders with T_1, and all of the variation among leaders with T_0 ($Y(T_0)$ > CEE and CEE = $Y(T_1)$). Additionally, the causal effect of trait T_1 must explain the same proportion of the variation in conflicts regardless of the *CEE*. For example, assume that trait T_1 causes an additional 50% of conflicts in any environment. If there were two states, A and B, and the *CEE* of state A was ten conflicts and the *CEE* of state B was twenty conflicts, then adding a leader with T_1 to both states must result in a total of fifteen and thirty conflicts, respectively, because the effect of T_1 is assumed to have the same causal effect on leaders all else being equal. Conversely, for the environmental perspective, the *CEE* must explain all of the variation in conflicts among T_1 and T_0 ($Y[T_{(0,1)}]$ = CEE). So, what is ultimately important to remember is that when we compare the difference in conflict initiations between strata, we will need to assume the *CEE* explains some minimum amount of variation in conflict initiations without

violating the assumptions required by the trait or environmental hypothesis to make theoretical sense.

Returning to our hypothetical study, we apply the counterfactual leader trait *T* to all three worlds and find the results below (Table 4-4).

Table 4-4: Example of ideal findings to support each explanation.

Trait World

Strata	n	Leader Trait T	Runner-up Trait T	Number of Conflict Initiations
1	50	0	0	5
2	50	0	1	5
3	50	1	0	25
4	50	1	1	25

Candidate Selection World

Strata	n	Leader Trait T	Runner-up Trait T	Number of Conflict Initiations
1	50	0	0	0
2	50	0	1	10
3	50	1	0	10
4	50	1	1	40

Leader Selection World

Strata	n	Leader Trait T	Runner-up Trait T	Average Number of Conflict Initiations
1	50	0	0	5
2	50	0	1	5
3	50	1	0	45
4	50	1	1	5

4.6 Evidence for the Leader Selection Effect Explanation

The leader selection effect explanation is when the effect of T_1 is positive and significant *only* when the counterfactual leaders' trait is not the same. We can see visualise how this would look in a regression in Figure 4-6 below; The effect of leader trait T_1 is only greater than leader trait T_0 when the counterfactual leader has the opposite value of T. We can also see there is no candidate selection effect (left facet in Figure 4-6) since the effect of two candidates with T_0 is the same as two candidates with T_1. Furthermore, this tells us the direct causal effect of selecting a candidate with T_1 over a candidate with T_0 to be leader. While we cannot know whether it is the trait or the leader selection effect that is the real causal variable here, we can reasonably conclude that it is likely due to the leader selection effect and not the trait considering that the is no significant difference between T_1 and T_0 when the counterfactual leader has the same trait.

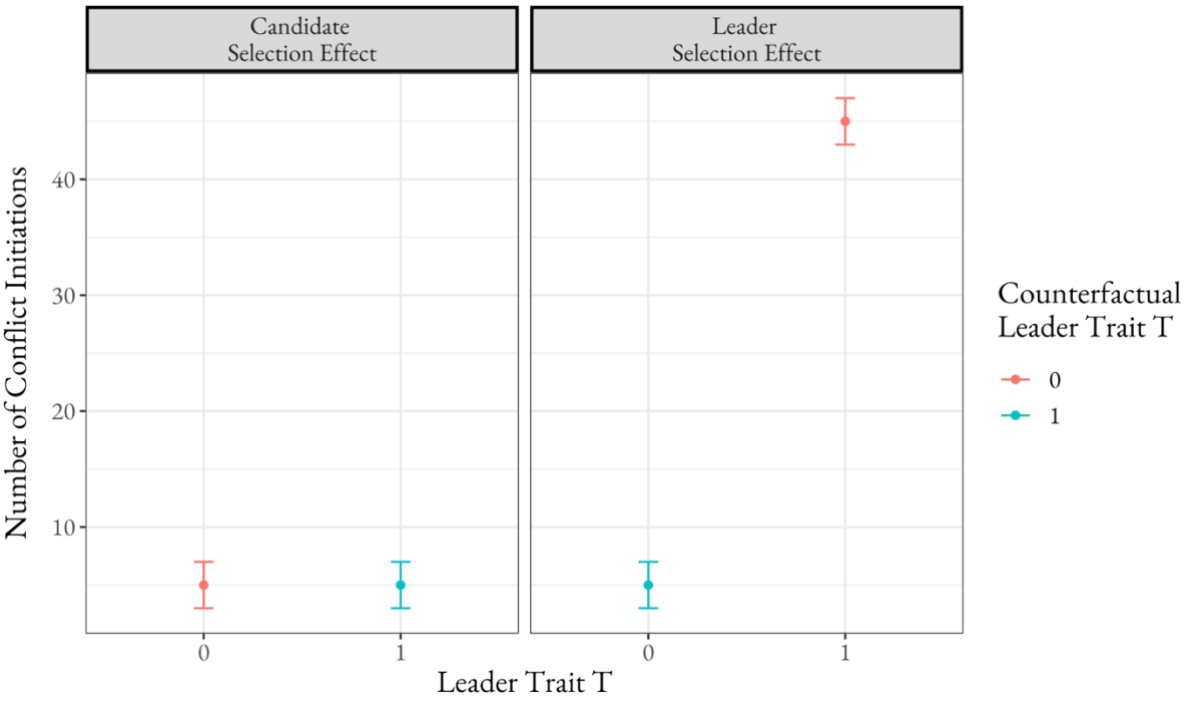


Figure 4-6: Example of ideal evidence to support leader selection effect explanation.

Of the three explanations, the leader selection effect explanation is the simplest to explain the findings in Leader Selection World because the trait and candidate selection effect

explanation two require significantly more assumptions. When we consider the assumed *CEE* for each explanation (Table 4-5), to maintain the trait explanation can still explain these findings, we would need to clarify why candidates with T_1 are more likely to be selected leaders over those with T_0 as the *CEE* increases, but also why they are less likely to be candidates than those with T_0 as the *CEE* increases, but also explain why states are less likely to have candidates with T_1 than with T_0 as the *CEE* increases. Likewise, the candidate selection explanation is harder to accept since the number of candidates with T_1 does not increase with the *CEE*.

Table 4-5: Leader Selection World example findings with CEE assumptions included.

Strata	n	Leader Trait T	Runner-up Trait T	Number of Conflict Initiations	Assumed CEE	
					Trait Perspective	Environmental Perspective
1	50	0	0	5	5	5
2	50	0	1	5	5	5
3	50	1	0	45	22.5	45
4	50	1	1	5	2.5	5

4.7 Candidate Selection Effect

The candidate selection effect is when the effect of T_1 is positive and significant *only* when the counterfactual leaders' trait is the same. This means that the treatment effect of T_1 is likely driven by the number of candidates with T_1 rather than the effect of T_1. We can visualise why this is the case in Figure 4-7 below. Leaders with T_1 only initiate more conflicts when the counterfactual leader trait is the same, and the number of conflicts increases as one additional candidate with T_1 is included. Furthermore, there is no leader selection effect since the effect size is the same and confidence intervals overlap.

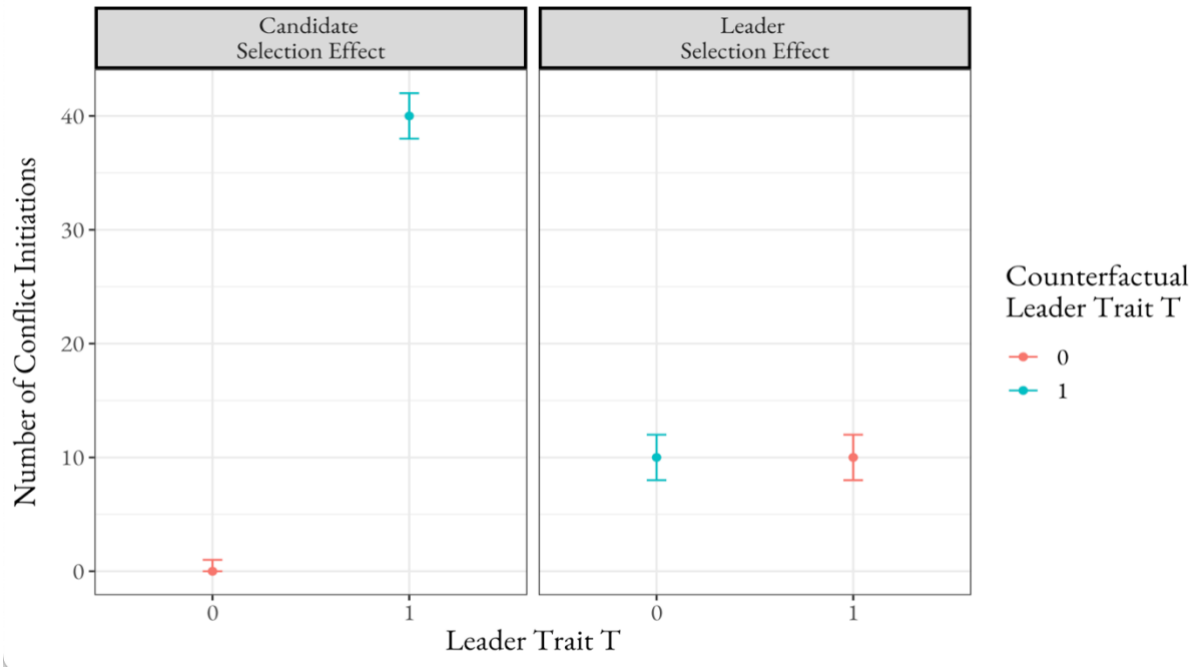


Figure 4-7: Example of ideal evidence to support candidate selection effect explanation.

Again, when we include the assumed CEE, the trait and leader selection explanations are difficult to explain the findings (Table 4-6). For the trait perspective to remain valid, we would have to explain why a leader with T_1 results in so many more conflicts when the counterfactual leader has T_1 than T_0. To use military experience, for example, why would prior military experience socialise leaders to be more conflict-prone when runners-up have military experience than when runners-up do not have military experience? There are no clear reasons for this other than the hypothesis that candidates are more likely to be ex-military in environments where conflict is more likely. We can also simply reject the leader selection explanation because candidates with T_1 are less likely to be selected leaders than candidates with T_0 as the *CEE* increases.

Table 4-6: Candidate Selection World example findings with *CEE* assumptions included.

Strata	n	Leader Trait T	Runner-up Trait T	Number of Conflict Initiations	Assumed CEE	
					Trait Perspective	Environmental Perspective
1	50	0	0	0	0	0
2	50	0	1	10	10	10
	50	1	0	10	5	10
4	50	1	1	40	20	40

4.8 Evidence Supporting the Trait Perspective

For evidence to support the trait perspective that traits cause leaders to initiate conflict, there must be no significant difference between candidate and leader selection effects. Again, this is because the trait perspective rejects selection, and so the treatment effect of T should not depend on the counterfactual leader trait T as long as it has no possible causal effect on the outcome. We can better visualise how the findings from Trait World support the trait perspective hypothesis in Figure 4-8 below. We can see that the effect of leader trait T_1 does not depend on the counterfactual leader trait T since the effect of T_1 is the same regardless of whether the counterfactual leader has the same value of T.

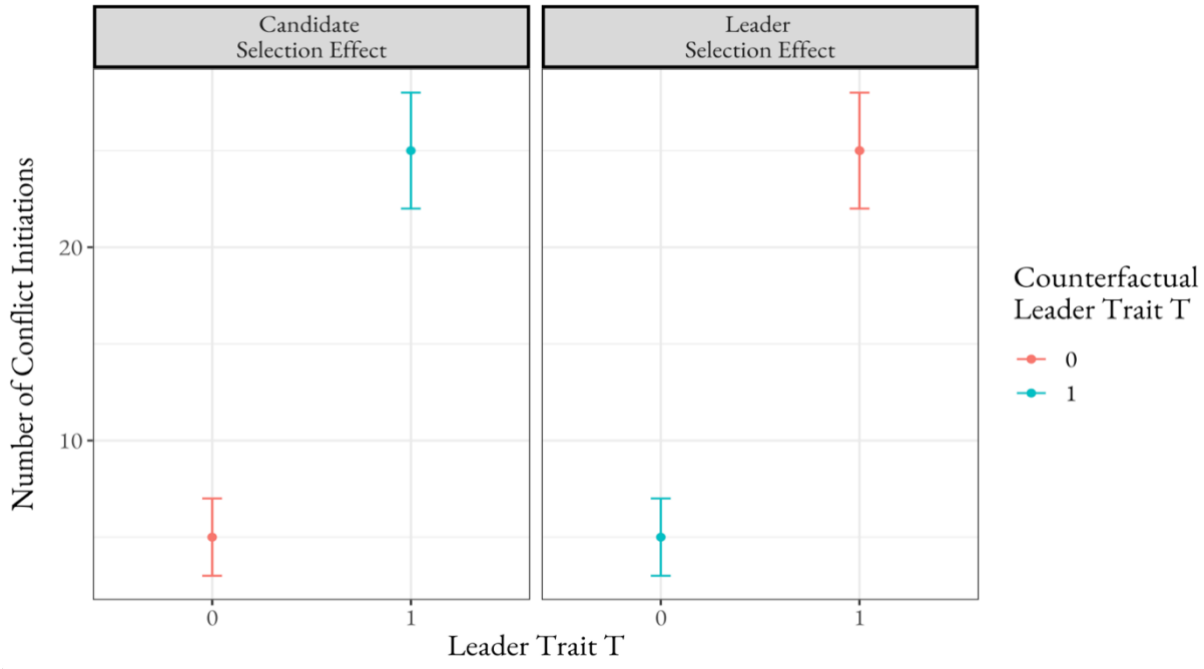


Figure 4-8: Example of ideal evidence to support trait explanation.

While the candidate and leader selection effects are both positive, the environmental perspective is harder to accept when we consider the minimum *CEE* required for it to make sense (Table 4-7). The *CEE* for T_I is the same for strata 3 and 4, which means that we cannot say the population treatment effect is entirely due to a leader selection effect since the effect size of T_I is the same when leader selection is controlled for (strata 3 and 4). Likewise, candidate selection effects cannot explain why the *CEE* does not increase when adding one additional candidate with T_I going from strata 1 and 2, and also from strata 3 and 4. That is, if traits did not matter, then the *CEE* should be the same for strata 2 and 3.

Table 4-7: Trait World example findings with CEE assumptions included.

Strata	n	Leader Trait T	Runner-up Trait T	Number of Conflict Initiations	Assumed CEE	
					Trait Perspective	Environmental Perspective
1	50	0	0	5	5	5
2	50	0	1	5	5	5
3	50	1	0	25	12.5	25
4	50	1	1	25	12.5	25

4.9 Candidate and Leader Self-Selection Effects

I used elections in the above example to simplify and explain how and why we can infer bias by using the counterfactual leader traits as a negative control, or placebo treatment. In the real world, however, leaders can self-select into power and those who do so are probably more similar to each other than those who do not. Thus, we also need to consider self-selection since it could be that any association between a leader trait and conflict is just the causal effect of self-selecting (e.g., coups, revolutions), and leaders with high-risk traits might be more likely to attempt or be successful at self-selecting than leaders with low-risk traits.

Yet, the central logic of the strategy above can still be used to tell us the same thing even when considering self-selection, just with an extra step. That is, we would split leaders into the four strata as above but then split them further into two groups – those who were the incumbent/challenger in coups and those who were not. If there is a candidate selection effect for leaders with T_1 only following self-selection events, then this is evidence that T_1 does not explain conflict initiations since leaders with that trait are more likely to attempt to self-select, and self-selecting only causes conflicts (else T_1 would be positive elsewhere). This is what we can call a "candidate self-selection effect." Likewise, suppose there is only a leader selection effect among leaders with T_1 in self-selection events. In that case, this suggests that leaders with T_1 are more likely successful at self-selecting, but it is just the causal effect of the coup, since the effect of T_1 is not positive in any other strata.

4.10 Key Assumption - Counterfactual Leader Traits as Placebo

The above strategy relies on only one assumption: any association between a counterfactual leader trait and conflict initiation occurs through a back-door non-causal path through the same selection effect that determines the leader trait. Put differently, this assumption can be violated if a counterfactual leader's traits are associated with conflict initiation through any front-door causal path. Figure 4-9 below visualises this assumption within the context of the full causal model from the previous chapter.

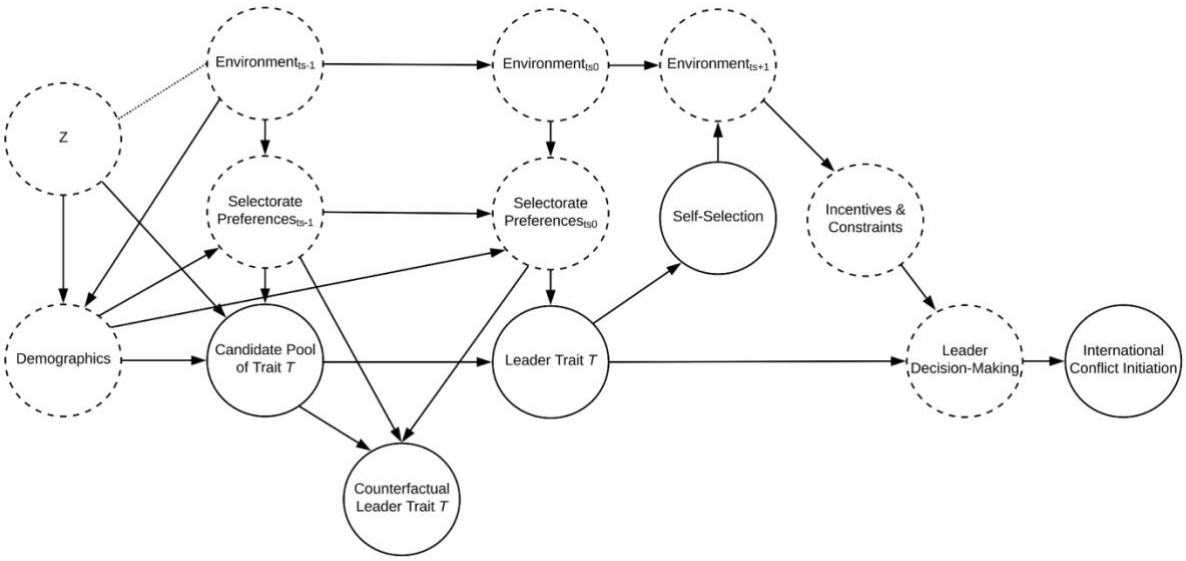


Figure 4-9: DAG illustration of the counterfactual-leader-trait-as-placebo assumption.

The only realistic front door path is if counterfactual leaders regularly have had a prominent role in pre-selection or post-selection governments. For example, if military experience causes leaders to be more conflict-prone, then it could be that an incumbent with military experience started more conflicts, which then created a more dangerous environment when the new leader took over (e.g., Counterfactual Leader Trait → Environment$_{st-1}$ → Environment$_{st0}$ → Conflict Initiation). Alternatively, even if counterfactual leaders were not

incumbent, those with military experience could still have been given a prominent ministerial position in the new leader's government, such as being Secretary of State, and so could directly increase the likelihood of conflict through their direct influence on the leader's foreign policy decision-making (e.g., Counterfactual Leader Trait → Minister Trait → Leader Decision-Making → Conflict Initiation). If either of these were the case more generally across leaders, their traits could still have a direct causal influence on conflict (Figure 4-10).

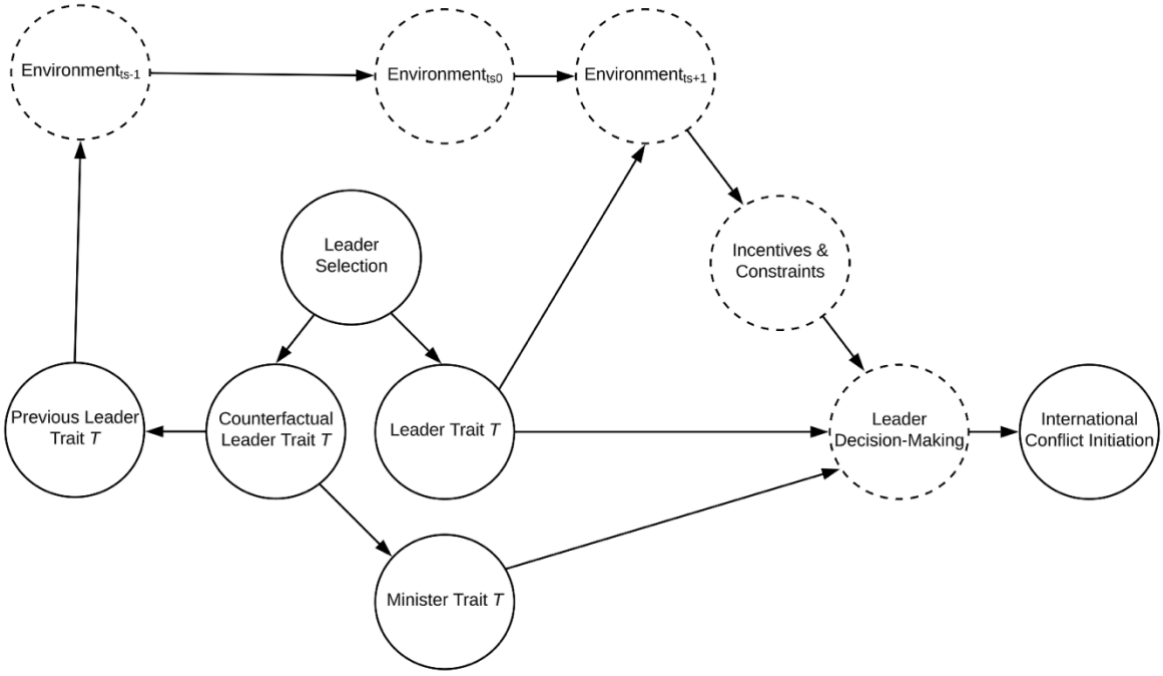


Figure 4-10: DAG illustration of how the counterfactual-leader-trait-as-placebo assumption can be violated.

There are several reasons to reject this possibility. In the counterfactual leader strategy, we only need to compare the likelihood of assumptions rather than strictly meet them. That is, if the association with conflict initiation were the same between the counterfactual leader and leader trait *T*, there are three possible explanations: 1) the association is evidence of selection for candidates with trait T, or 2) the association is evidence that counterfactual leaders with T_t become ministers and cause leaders with T_t to initiate conflict, or 3) the association is because counterfactual leaders with trait T_t were previous leaders, more conflict-prone, and so created an

environment where conflict was more likely when the new leader took over. Given a large enough sample, I argue that it is much easier to accept the first explanation than the second and third. I suspect most runners-up are from a different political party to the leader and likely have no official role in the future government. In some cases, however, the runner-up can be appointed Leader of the Opposition during the leaders' tenure, which means they could still have a causal effect on leaders' decision-making. Nevertheless, it is still easier to assume that any association between counterfactual leader traits and conflict initiation is more likely to be a placebo effect due to selection than a direct causal effect. Opposition Leaders are not always the counterfactual leader, limited to a few countries, change frequently, and do not greatly affect a leader's foreign policy. It is more reasonable to argue that the counterfactual-leader-as-placebo assumption is violated by being incumbent, but we can easily test for this, as explained in the Methodology Chapter.

4.11 Conclusion

In this chapter, I have highlighted the difficulties in disentangling the important but complicated effects of leaders and their environment on conflict initiation. Having rejected the utility of conventional methods (IV, DiD etc.) to resolve this endogeneity problem, I have identified a novel strategy that, while not able to perfectly differentiate leader trait and selection effects, offers a way forward that will allow us to adjudicate, on the basis of empirical evidence, the likelihood of leader traits versus selection effects in explaining the association and causal mechanisms between leaders and conflict. For the first time, I have revealed the thorny problems that have so far undermined this literature and designed a new research strategy to advance our knowledge on the matter. In the following chapter, I use the strategy here to develop hypotheses more thoroughly for why each of the selection effects could potentially explain why leaders with high-risk traits appear more likely to initiate conflict than those with low-risk traits.

5 Hypotheses

Recall there are two competing arguments to the research question of whether traits influence leaders to initiate conflict. First, the environmental perspective suggests that leader traits do not matter and that the positive associations between traits and conflict found in previous research are spurious due to selection effects. Second, the trait perspective suggests leader traits do indeed matter and that the positive associations between leaders with high-risk traits and conflict initiation are causal and robust to selection effects. However, we cannot test these competing explanations given the fundamental problem outlined in the previous chapter. Though, in a perfect world where we could observe and measure everything we needed, these two competing theories would be formulated into the following two testable hypotheses:

Trait Perspective Hypothesis:

> Leaders with high-risk traits are more likely to initiate conflict, controlling for selection effects.

Environmental Perspective Hypothesis:

> Leaders with high-risk traits are not more likely to initiate conflict, controlling for selection effects.

While we cannot test these hypotheses directly, we can indirectly test their central underlying argument using proxy variables that are at least observable and measurable without significant error. More specifically, we can use the counterfactual leader strategy to infer whether selection effects likely explain any positive association between a given leader trait and conflict initiation. Since the strength of this strategy is that we do not need to observe and measure anything other than the leader and the counterfactual leader's trait to test for bias, we can initially test five hypotheses for each individual trait. Doing so will clarify which effect occurs and when they do not. These five hypotheses address the following effects: candidate selection, leader

selection, candidate self-selection, leader self-selection, with the addition of the trait hypothesis. Table 5-1 below is a summary of the five hypotheses, and the treatment group and comparison group for clarity (assume that, for a given trait, *T,* the perceived high-risk version is T_1 and the low-risk version is T_0).[146]

Table 5-1: Summary table of the five initial selection effect hypotheses to test for bias.

Effect	Hypothesis	Treatment Group		Comparison Group		Self-Selection
		L	CL	L	CL	
Candidate Selection Effect	Leaders with trait T_1 are more likely to initiate conflict than leaders with trait T_0 only when counterfactual leaders have the same value of *T*.	1	1	0	0	No
Leader Selection Effect	Leaders with trait T_1 are more likely to initiate conflict than leaders with trait T_0 only when counterfactual leaders have the opposite value of *T*.	1	0	0	1	No
Candidate Self-Selection Effect	Leaders with trait T_1 are more likely to initiate conflict than leaders with trait T_0 only when counterfactual leaders have the same value of *T* and following self-selection events	1	1	0	0	Yes
Leader Self-Selection Effect	Leaders with trait T_1 are more likely to initiate conflict than leaders with trait T_0 only when counterfactual leaders have the opposite value of *T* and following self-selection events	1	0	0	1	Yes
Trait Effect	Leaders with trait T_1 are more likely to initiate conflict than leaders with trait T_0 regardless of the counterfactual leader trait T and self-selection.	-	-	-	-	-

Testing these hypotheses against leader traits allows us to indirectly address the relationship between leader traits, the international environment, and conflict initiation. The chosen traits arise from competing and complementary theories of how leader traits affect their likelihood of initiating conflict in four categories: military experience, age, sex, and ideology.

The purpose of testing the selection effect hypotheses first allows us to test for the presence of bias without even needing to consider any of the more specific reasons for why we

[146] For example, T=1 would be an ex-military leader and T=0 would be non-ex-military. Or, in the case of a trait that is continuous like age, T=1 would be an older leader and T=0 would be a younger leader.

might see them. Conversely, if none of the selection effect hypotheses are supported, then there would be nothing to explain. However, if any of the four selection effect hypotheses are supported, then an explanation is required since we know it can only result from some back-door non-causal effect (given the placebo assumption). On its own, this is arguably necessary and sufficient to reject the trait perspective hypothesis given the assumptions required for it to make sense (as explained in the previous chapter). However, without providing any theoretical justification as to why our findings support one of the hypotheses, one could counter that the findings are just random noise, a result of bias in data collection, or some other factor I have not considered.

To further strengthen why each selection hypothesis is plausible evidence to support the environmental perspective that leader traits do not matter, I more specifically discuss the underlying logic for each hypothesis and why we should expect to see them in our findings. Accordingly, I also provide additional supporting hypotheses for each trait using a variety of other variables that are consistent with the central prediction of the selection effect hypothesis. For instance, the candidate selection effect hypothesis is most likely explained by demographics causing candidates to have the same trait in environments where conflict is more likely.

Even though many demographic variables are measured imprecisely and are not appropriate controls, we can still hypothesise that they should predict candidates' traits in the expected direction since even imprecise variables should have at least some positive correlation. However, if there is a negative correlation, it is more difficult to find support for the environmental perspective hypothesis over the trait perspective hypothesis. Note that I do not provide supporting hypotheses for a trait without sufficient theoretical justification. For example, as there is no evidence or reason to believe that female leaders should be more likely to initiate conflict than male leaders because they are more likely to self-select, I do not discuss or postulate reasons for why the candidate self-selection or leader self-selection effect hypotheses could be supported in that respective section.

Note that I develop two types of hypotheses. The first are the selection effect hypotheses that predict conflict initiation for the period *after* leaders are selected. The other hypotheses predict the leader and counterfactual leader's traits at the time of the selection event date and are primarily used to support the corresponding selection effect hypothesis (i.e., that the selection effect is not random noise). Thus, to help distinguish between the various hypotheses and avoid confusion when discussing them together, I refer to the first group as the "selection effect hypotheses" and the latter group as "supporting hypotheses" when specifically discussing and testing their predictions in isolation.

5.1 Military Experience

There is already considerable empirical and theoretical support for the idea that either the candidate, leader or self-selection effect hypotheses could explain why ex-military leaders appear more likely to initiate conflict. Therefore, this is an important trait variable to re-examine in this thesis to disentangle the opposing arguments and evidence. Table 5-2 below summarises the selection effects, the treatment and comparison groups, and the supporting hypotheses for military experience.

Table 5-2: Summary of leader military experience selection effect hypotheses.

Selection Effect Hypotheses	Treatment Group		Comparison Group		Supporting Hypotheses
	Leader Military Experience	Counterfactual Leader Military Experience	Leader Military Experience	Counterfactual Leader Military Experience	
Candidate Selection Effect	Yes	Yes	No	No	FIR, Demographics
Leader Selection Effect	Yes	No	No	Yes	FIR, Affinity Bias
Candidate Self-Selection Effect	Yes	Yes	No	No	Attempt
Leader Self-Selection Effect	Yes	No	No	Yes	Success

5.1.1 Military Experience – Leader Selection Effect Explanations

Why would ex-military leaders be more likely to initiate conflict only when the military experience of the counterfactual leader is not the same? The most likely reason is that people might be more likely to choose ex-military over non-ex-military candidates in more conflict-prone[147] states. Intuitively, this makes sense. If people wanted a leader to handle a seemingly hostile international environment, who better to select than one with prior military experience? A large body of research suggests this intuition may be correct. Known more generally as "veteran bonus theory," several studies find that voters see veterans as more competent in defence and security and that veterans are generally more likely to win elections.[148] In U.S. Senate races between 1982 and 2016, for instance, there was a slight veteran electoral bonus and this bonus for

[147] I explain how to measure this more specifically in the Methodology Chapter. But for context here, this is simply the total number of conflicts a state has been involved in during the ten years prior to the selection event. This is used primarily as a proxy for the unobservable environmental risk.

[148] Jeremy M. Teigen, 'Military Experience in Elections and Perceptions of Issue Competence: An Experimental Study with Television Ads', *Armed Forces & Society* 39, no. 3 (2013): 415–33.

Democratic candidates in open races was comparable to an incumbent advantage.[149] In parliamentary elections in Croatia, candidates with military combat experience also had an electoral advantage in areas that were more exposed to war violence but were penalised by voters in areas that were not.[150] In another study on Croatian elections, ex-military voters were likelier to cast ballots for the nationalist party, but only if they did not show signs of PTSD.[151] This also appears to support the military conservatism theory that exposure to combat decreases veterans' hawkish beliefs or behaviour.

Real-world anecdotes also support this potential first image reversed (FIR) leader selection effect explanation. During the Vietnam War, former U.S. President Bill Clinton was vehemently opposed to the war and draft, yet in 1969 he expressed interest in joining the Arkansas Reserve Officer Training Corps (ROTC). Why? In a 1992 letter, Clinton apologised to the ROTC recruitment officer for ultimately pulling out, stating, "I decided to accept the draft in spite of my beliefs for one reason: to maintain my political viability within the system."[152] Clinton's concerns here perhaps echo the sentiment of many aspiring politicians in states that often engage in international conflict – it is highly beneficial for one's political career to have previously served in the military. Indeed, in the U.S. during the Cold War, prior military experience was almost a requirement to be a U.S. President, as Hubert Humphrey learned the hard way. During the 1960 U.S. Democratic presidential primaries, Robert Kennedy accused Humphrey of being a "draft-

[149] David K. Richardson, 'The Electoral Impact of Military Experience: Evidence From US Senate Elections (1982–2016)', *Armed Forces & Society*, 2021, 0095327X211038032.
[150] Josip Glaurdić and Christophe Lesschaeve, 'Warrior Candidates: Do Voters Value Combat Experience in Postwar Elections?', *Political Research Quarterly*, 2021, 10659129211040524.
[151] Christophe Lesschaeve, 'Voting after Violence: How Combat Experiences and Postwar Trauma Affect Veteran and Voter Party Choices in Croatia's 2003 Postwar Elections', *Armed Forces & Society* 46, no. 2 (2020): 259–80.
[152] 'The 1992 Campaign; A Letter by Clinton on His Draft Deferment: "A War I Opposed and Despised"', *The New York Times*, 13 February 1992, sec. U.S., https://www.nytimes.com/1992/02/13/us/1992-campaign-letter-clinton-his-draft-deferment-war-opposed-despised.html.

dodger" for not fighting in the Second World War.[153] Humphrey subsequently lost the 1960 bid to John F. Kennedy and the 1968 U.S. presidential election to Richard Nixon, both of whom were Second World War veterans. This is not to say that military experience guarantees being elected – there are plenty of examples where it did not – but it does seem to confer an advantage, especially in conflict-prone states where the leader is expected to have a strong command of their military. Indeed, as Jeremy Teigen noted concerning the tendency for Americans to choose ex-military Presidents:

> "With the centrality of potential military conflicts seemingly continuous, it is natural that voters of the [Cold War] era might be more than those in other eras to scrutinise candidates' ability to lead a military. As the importance of the president's role as commander in chief has grown, so too has the potential for voters to use biographical military information to judge candidates."[154]

Furthermore, we can see another hint of this possible FIR leader selection effect for veterans in the difference between ex-military leaders in the United States and Japan (Figure 5-1 and Figure 5-2). Before the end of the Second World War, the United States had no president with military experience in over 30 years. In the 50 years following 1945, however, every U.S. President was a veteran until Bill Clinton in 1994. Is it just a coincidence that there is a correlation between U.S. ex-military leaders and a time when Americans were worried about total or nuclear war with the Soviet Union? Conversely, Japan is an illustrative case for the opposite effect. Before the Second World War ended, 50% of Japanese leaders had military experience, but only 16% did afterwards.

[153] Arnold A. Offner, *Hubert Humphrey: The Conscience of the Country* (Yale University Press, 2018), 149.
[154] Jeremy Teigen, *Why Veterans Run: Military Service in American Presidential Elections, 1789-2016* (Temple University Press, 2018), 42.

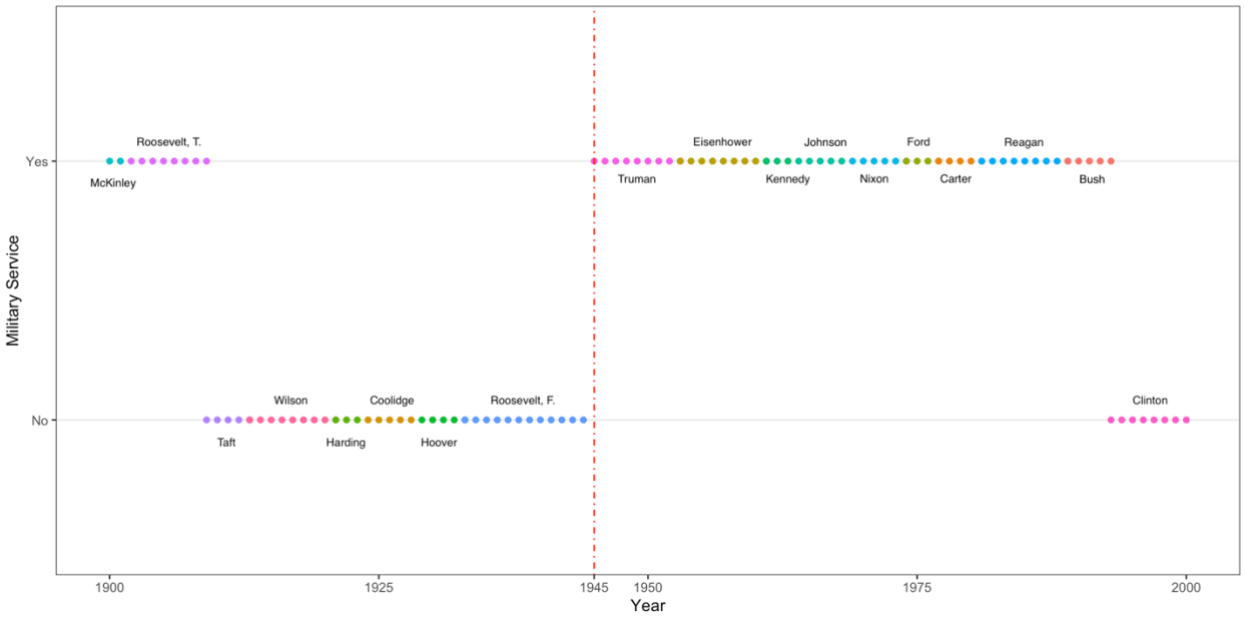


Figure 5-1: Ex-military leaders in the United States 1900-2000

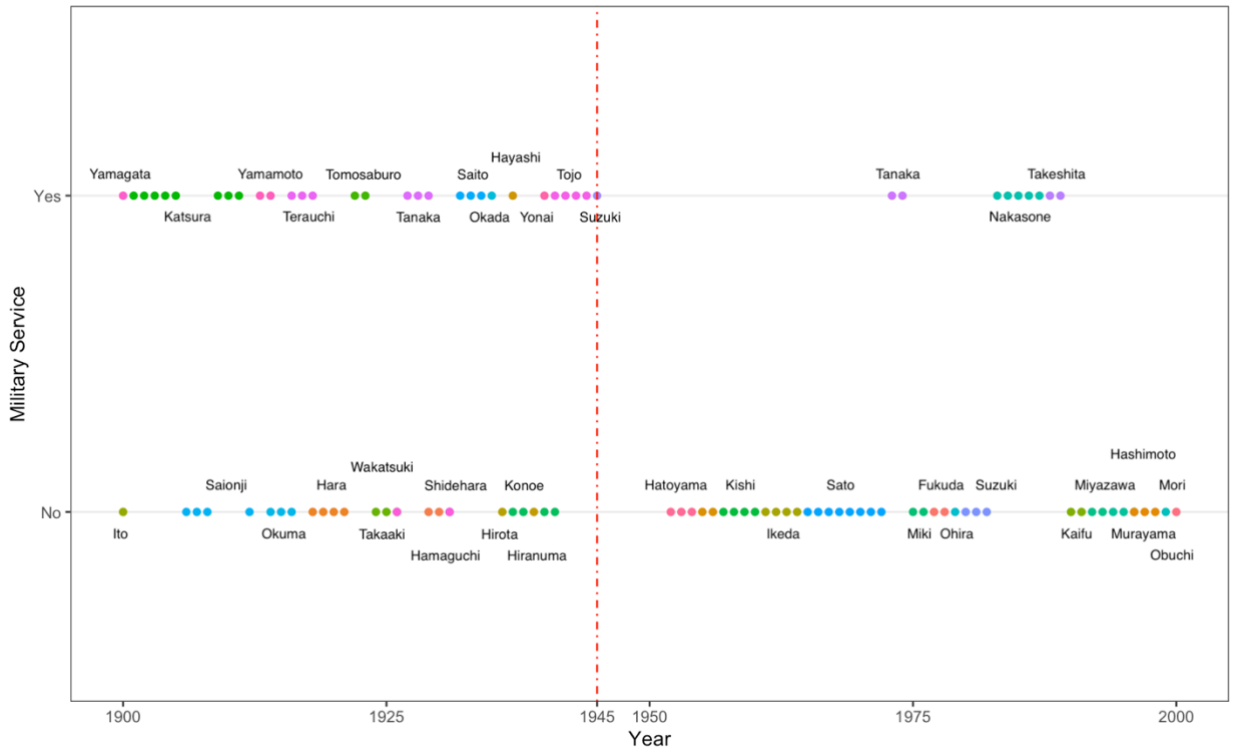


Figure 5-2: Ex-military leaders in Japan 1900-2000

This suggests that deciding whether to select ex-military leaders may be linked to international conflict. After its defeat in the Second World War, Japan lost its great power status and subsequently denounced using the military to settle international disputes. Known as the "Yoshida

Doctrine", the Japanese constitution proclaimed that "the Japanese people forever renounce war as a sovereign right of the nation" and that military forces "will never be maintained."[155] Subsequently, they did not have any ex-military Prime Ministers for thirty years. On the other hand, the United States became a superpower almost overnight and subsequently engaged in a Cold War against the Soviet Union for ideological and economic supremacy. Consequently, veteran U.S. Presidents were then the norm throughout this period. Furthermore, the difference in the number of ex-military leaders in these two countries does not appear to be due to a surplus of veterans. If it were, we would expect Japan to have a similar number of ex-military leaders, if not more, than the U.S. post-1945. While these two cases are anecdotes, they are perhaps the clearest evidence of why the effect of ex-military leaders on conflict might entirely be due to a leader selection effect. Thus, the FIR leader selection hypothesis for military experience is formulated as follows:

Military Experience - FIR Leader Selection Effect Hypotheses:

1. Ex-military leaders are more likely to initiate conflict than non-ex-military leaders only when the leader and counterfactual leader do not have the same military experience.
2. Ex-military candidates are more likely to be chosen over non-ex-military candidates in more conflict-prone states.

Alternatively, affinity bias could also explain a leader selection effect. Again, affinity bias is when people tend to vote for candidates with the same traits as themselves, so states with larger shares of veterans in the population should be more likely to vote for ex-military candidates over non-ex-military candidates. States with a larger population of veterans are also those that are probably frequently engaged in conflict due to wars or conscription. Accordingly, an affinity bias in favour of ex-military candidates would likely result in a disproportionate selection of ex-

[155] 'The Constitution of Japan' (Prime Minister of Japan and His Cabinet, 1946).

military leaders in more conflict-prone states. Thus, based on these two predictions, the affinity bias hypothesis is formulated based as follows:

Military Experience – Affinity Bias Leader Selection Hypothesis:

1. Ex-military leaders are more likely to initiate conflict than non-ex-military leaders only when the leader and counterfactual leader do not have the same military experience.
2. Ex-military candidates are more likely to win elections against non-ex-military candidates in states with larger shares of veterans in the population.
3. States with larger shares of veterans in the population are more likely to initiate conflict than states with lower shares of veterans in the population.

5.1.2 Military Experience – Candidate Selection Effect Explanations

A leader selection effect could also be the result of First-Image Reversed selectorate preferences. If people were consistently voting for candidates with military experience over those without it, then we should probably expect this same preference to manifest when choosing candidates first as well as the final leader. Indeed, if veteran bonus theory is right and veterans are more likely to win elections – and even more so in response to the international system – then political parties are likely to take notice and be more likely to nominate them due to their greater electability. Furthermore, if Heads of State were similarly more likely to appoint ex-military over non-ex-military individuals as leaders in the belief that ex-military leaders are better at dealing with international conflict, then they probably would initially only consider ex-military individuals in the first place. For instance, if Winston Churchill had randomly died just after Chamberlain's resignation, King George and Parliament would likely have only considered other ex-military individuals to be the new Prime Minister. Thus, we can formulate the FIR candidate selection hypothesis as follows:

FIR Candidate Selection Effect Hypotheses

1. Ex-military leaders are more likely to initiate conflict than non-ex-military leaders only when the leader and counterfactual leader have the same military experience.
2. Candidates are more likely to be ex-military than non-ex-military in more conflict-prone states.

Although the US-Japan example above hints that demographics might not be relevant, it still could confound any selection for veterans in conflict-prone states more generally. Presumably, states that utilise conscription have many veterans in the population, for obvious reasons. And most states that do use conscription do so because they are engaged in wars or might expect to be in the future. Israel is the most obvious example here, as it has had universal conscription for males and females in some form since its independence in 1949. Accordingly, it is no surprise that 14 of the 17 Israeli leaders between 1949 and 2004 were ex-military.[156] Other countries may not have continuous military conscription like Israel but have implemented it during wars and so have certain years where a significant number of veterans are in the population. This is probably the case for many World War participants, such as the U.S., Germany, Japan, Italy, the U.K., France, Russia, Italy, and to a lesser extent, Canada, Australia, and South Africa. Furthermore, even without conscription, there could be a pro-military culture whereby volunteers might make up a significant portion of the military. In the U.K. during the beginning stages of the First World War, for instance, around 24% of the male population had volunteered and, including those who were eventually conscripted, around 46% of all military-aged males had served in the war.[157] Accordingly, the demographic consequences of war and international conflict can last for decades

[156] According to the LEAD dataset, see: Ellis, Horowitz, and Stam, 'Introducing the LEAD Data Set'.
[157] Jay M. Winter, 'Britain's 'Lost Generation'of the First World War', *Population Studies* 31, no. 3 (1977): 449–66.

and likely reflect the pool of candidates with prior military experience. Most males are between 18 and 20 years old when they either enlist or are conscripted, and leaders enter office around 58 years old, on average[158], so any candidate selection effect for veterans could last 40 years. For example, any aspiring British politician in the 1950s would likely have been part of the cohort of men who served between 1914 and 1918. Thus, in a decade in which the U.K. was the 5th most conflict-prone state,[159] they would have likely had a veteran leader even if they were randomly selected.

The demographic and FIR explanation might work together, however. Even if it was not just a random chance due to demographics, there is still an expectation that political figures have served in the military during wars if eligible, as the above examples of Clinton and Humphrey support. Selectorates in states that frequently engage in conflict might want an ex-military leader to confront interstate threats but can also easily put one into office if there are many veteran candidates.

Conversely, it could also mean that the domestic and FIR explanations could oppose each other. That is, selectorates might not want ex-military leaders in the belief they are too aggressive and likely to start more conflicts but are unlikely to have a choice because demographics result in numerous ex-military candidates. If nearly half of older males were veterans in the U.K. in the 1950s, for example, then all else equal, half of the party leaders were likely veterans, so voters might have had to choose one even if they ultimately did not want to. While veteran bonus theory suggests that veterans are more electable because of positive stereotypes, there is also plenty of evidence that some tend to view veterans with negative stereotypes. They can be seen as a hero,

[158] According to the LEAD dataset, see: Ellis, Horowitz, and Stam, 'Introducing the LEAD Data Set'.
[159] According to the MIDB v5.0 dataset, see: Charles S. Gochman and Zeev Maoz, 'Militarized Interstate Disputes, 1816-1976: Procedures, Patterns, and Insights', *Journal of Conflict Resolution* 28, no. 4 (1984): 585–616.

victim, or villain.[160] Or they can be seen as angry, withdrawn, violent, or mentally ill.[161] While these studies shed light on how people perceive veterans in general, they tell us very little about how people might perceive ex-military leaders in the expectation of foreign policy consequences. In other words, even if people did not want ex-military leaders because of these negative stereotypes, they might have little choice but to select one because wars and conscription tend to result in many ex-military candidates. Thus, based on these two predictions, we can formulate the demographic candidate selection effect hypothesis:

Military Experience – Demographic Candidate Selection Effect Hypotheses

1. Ex-military leaders are more likely to initiate conflict than non-ex-military leaders only when the leader and counterfactual leader have the same military experience.
2. Candidates are more likely to be ex-military as the share of veterans in the population increases.
3. States with larger shares of veterans in the population are more likely to initiate conflict than states with lower shares of veterans in the population.

5.1.3 Military Experience – Self-Selection Explanation

The last possible leader selection explanation is that ex-military leaders are only more likely to initiate conflict following self-selections. This means that it would not be military experience directly causing conflicts to increase, but rather self-selection. However, the candidate self-selection hypothesis predicts that this would only occur when the counterfactual leader is ex-military (e.g., coup attempts in which an ex-military coup leader tries to depose an ex-military

[160] Helen McCartney, 'Hero, Victim or Villain? The Public Image of the British Soldier and Its Implications for Defense Policy', *Defense & Security Analysis* 27, no. 1 (2011): 43–54.
[161] James E. Bordieri and David E. Drehmer, 'Vietnam Veterans: Fighting the Employment War', *Journal of Applied Social Psychology* 14, no. 4 (1984): 341–47.

incumbent). Conversely, the leader self-selection hypothesis predicts that this might only occur when the counterfactual leader is not ex-military (e.g., coup attempts in which an ex-military coup leader overthrows a non-ex-military incumbent). Why either of these could be the case is intuitive since there is perhaps no trait more essential to self-selecting than being in the military, given that one must have access to weapons and must convince the entire military apparatus that their coup is not to be countered. More specifically, considering that ex-military leaders are probably more likely to attempt and successfully self-select, it could be that any self-selection explanation is because of one of these two scenarios. Indeed, according to the LEAD dataset, 68% of leaders who entered office "irregularly" were ex-military, so we already know that ex-military leaders are probably more likely to self-select. Therefore, we can postulate two more hypotheses for why we might see either a candidate or leader self-selection effect:

Military Experience – Candidate Self-Selection Hypothesis:

1. Ex-military leaders are more likely to initiate conflict than non-ex-military leaders only when the leader and counterfactual leader have the same military experience and following self-selection events.
2. Ex-military candidates are more likely to attempt to self-select than non-ex-military candidates.

Military Experience – Leader Self-Selection Hypothesis:

1. Ex-military leaders are more likely to initiate conflict than non-ex-military leaders only when the leader and counterfactual leader do not have the same military experience and following self-selection events.
2. Ex-military candidates are more successful at self-selecting than non-ex-military candidates.

5.2 Non-Combat Military Experience

Considering several studies already suggest that non-combat ex-military leaders are more likely to initiate conflict, these findings might result from one of the selection effects. One important caveat to note, however, is that the finding of non-combat leaders being more conflict-prone in previous research is only in reference to leaders without military experience. More specifically, the authors disaggregate military experience type using two dummy variables, combat and non-combat, and then regress those on conflict initiation.[162] Compared to leaders without military experience, they only find that leaders without combat experience are more likely to initiate conflict, but not those with combat experience. However, the confidence intervals between non-combat and combat leaders overlap, meaning there is no difference between the two. Confusingly, this means that ex-military leaders without combat experience are more likely to initiate conflict than those without military experience but are not more likely to initiate conflict than those with combat experience. Accordingly, we then need to theorise how selection effects might explain why non-combat leaders are more likely than non-ex-military leaders to initiate conflict and why combat leaders are *not* more likely to initiate conflict than non-ex-military leaders. Table 5-3 below summarises the selection effects, the treatment and comparison groups, and the supporting hypotheses for military experience type.

[162] Horowitz and Stam, 'How Prior Military Experience Influences the Future Militarized Behavior of Leaders', 545.

Table 5-3: Summary of leader military experience type selection effect hypotheses.

Selection Effect Hypotheses	Treatment Group		Comparison Group		Supporting Hypotheses
	Leader Military Experience Type	Counterfactual Leader Military Experience Type	Leader Military Experience Type	Counterfactual Leader Military Experience Type	
Candidate Selection Effect	Non-Combat	Non-Combat	None	None	FIR, Demographics
Leader Selection Effect	Non-Combat	None	None	Non-Combat	FIR
Candidate Self-Selection Effect	Yes	Yes	No	No	Attempt
Leader Self-Selection Effect	Yes	No	No	Yes	Success

5.2.1 Non-Combat Military Experience – Leader Selection Effect Explanations

Why would non-combat veterans only be more likely to initiate conflict than non-veterans when the military experience type of the counterfactual leader is not the same? I believe this could likely result from a similar FIR preference for ex-military candidates – that people are likelier to choose non-combat veterans for leader over both non-ex-military and combat veterans in environments where conflict is more likely. But this FIR preference must also necessarily mean that there would be no FIR preference for combat veterans over non-veterans (assuming previous research is right and combat veteran leaders are not more likely to initiate conflict than non-ex-military leaders). However, it seems counterintuitive to suggest that combat veterans would not be more likely to be chosen over non-combat veterans or non-veterans in response to increased conflict. Indeed, in the runup to the 2004 U.S. Presidential election, Democratic Party nominee John Kerry proclaimed, "We are a nation at war…As president, I will wage this war

with the lessons I learned in Vietnam."[163] The incumbent, George W. Bush, was also a Vietnam-era veteran from 1968 to 1974, but never left the U.S. throughout the entire war, which Kerry also stressed to point out in a different speech.[164] Surely, Kerry felt that his combat experience would provide him with some electoral advantage over Bush's lack of combat experience or else why would he point it out? Of course, Kerry lost the election, which is not entirely because he had combat experience. However, this could hint that Kerry was wrong and that non-combat experience could be an electoral bonus, and combat experience an electoral negative. One possible explanation could be the correlation between combat experience and having a military career. According to the LEAD dataset, 71% of combat veterans were military careerists, whereas only 52% of non-combat leaders were, which is probably because those with military careers are more likely to have seen combat due to their long service in the military. There is some evidence among the American electorate that having a military career is an electoral negative, as Samuel Huntington points out; "While American people like their political candidates to be military heroes, they want their military experience to be an interlude in, or a sideline to, an otherwise civilian career."[165] This might also help explain why John McCain – a combat veteran with a 22-year military career – lost the 2008 U.S. to Barack Obama – a non-veteran – despite the U.S. being engaged in two overseas wars. Thus, if selectorates prefer veteran leaders in response to increased international conflict – but only those without military careers – then non-combat veterans would be more likely to be chosen over non-veterans and combat veterans in more conflict-prone states. Thus, we can then formulate the FIR leader selection effect hypothesis as follows:

---

[163] 'Kerry: We Are Here to Make America Stronger', CNN, 30 July 2004, https://edition.cnn.com/2004/allpolitics/07/30/dems.kerry.transcript.2/index.html.
[164] Candy Crowley et al., 'Nader Blasts "Unpatriotic Corporations"', CNN, 24 February 2004, https://edition.cnn.com/2004/allpolitics/02/23/elec04.prez.main/index.html.
[165] Huntington, *The Soldier and the State*, 159.

Non-Combat Experience – FIR Leader Selection Effect Hypotheses:

1. Leaders with non-combat military experience are more likely to initiate conflict than leaders with combat experience and no military experience only when the leader and counterfactual leader do not have the same military experience type.
2. Candidates with non-combat military experience are more likely to be chosen over candidates with combat experience and no military experience in more conflict-prone states.

5.2.2 Non-Combat Military Experience – Candidate Selection Effect Explanations

Suppose we were to find that non-combat ex-military leaders are more likely than non-ex-military leaders to initiate conflict only when counterfactual leaders have the same military experience type. In that case, this is likely a result of the same demographic explanation as military experience in general. Again, states with universal conscription, such as Israel and France, are probably more conflict-prone than states without conscription since the implementation of this policy is at least partly due to past experiences or future expectations of conflict or wars. Likewise, states without conscription but nonetheless have a strong pro-military culture, such as in the United States between the Vietnam War and Iraq/Afghanistan War, might also tend to have more non-combat than combat veterans in the population. Accordingly, there would likely be a large number of non-combat veterans in these states' populations, so they are more likely to have non-combat candidates and leaders, even by sheer chance. Thus, we can formulate the FIR and demographic candidate selection hypothesis:

Non-Combat Experience – FIR Candidate Selection Effect Hypothesis:

1. Leaders with non-combat military experience are more likely to initiate conflict than leaders with combat experience and no military experience only when the leader and counterfactual leader have the same military experience type.
2. Candidates are more likely to be non-combat veterans in more conflict-prone states.

Non-combat – Demographic Candidate Selection Effect Hypothesis:

1. Leaders with non-combat military experience are more likely to initiate conflict than leaders with combat experience and no military experience only when the leader and counterfactual leader have the same military experience type.
2. Candidates are more likely to be non-combat veterans as the veteran population percentage increases.
3. States with higher shares of veterans in the population are more likely to initiate conflict than states with lower shares of veterans in the population.

5.2.3 Non-Combat Military Experience – Self-Selection Explanation

Why would non-combat leaders be more likely to initiate conflict only after self-selecting? One possibility is that self-selections, such as coups, are more likely to occur in smaller states that have not fought in any wars and hence probably have little to no combat veterans among the political elite. Accordingly, if ex-military leaders were only more likely to initiate conflict after self-selecting, it would likely be mostly non-combat leaders who do so. Indeed, of all those leaders who entered "irregularly" in WLF, 41% were non-combat veterans and 37% were combat veterans. Considering that 72% of all ex-military leaders in LEAD are combat veterans, non-combat ex-military leaders are 19% more likely to enter office irregularly than combat ex-military leaders. Therefore, any self-selection effect we see among non-combat

candidates could be because they are more likely to attempt self-selections and successfully do so. Thus, we can formulate two self-selection hypotheses:

Military Experience – Candidate Self-Selection Hypothesis:

1. Leaders with non-combat military experience are more likely to initiate conflict than leaders with combat experience or no military experience only when the leader and counterfactual leader have the same military experience type and following self-selection events.
2. Candidates with non-combat military experience are more likely to attempt to self-select than candidates with combat experience or no military experience.

Military Experience – Leader Self-Selection Hypothesis:

1. Leaders with non-combat military experience are more likely to initiate conflict than leaders with combat experience or no military experience only when the leader and counterfactual leader do not have the same military experience type and following self-selection events.
2. Candidates with non-combat military experience are more successful at self-selecting than candidates with combat experience or no military experience.

5.3 War-Win Experience

Recall that a few studies find some evidence to suggest that leaders who have previously fought in a war that was won are more likely to initiate conflict than those who have not.[166] Like military experience type, however, this finding only refers to those with no military experience. The authors disaggregate war experience type using two dummy variables, war winner and war

[166] Horowitz and Stam, 'How Prior Military Experience Influences the Future Militarized Behavior of Leaders'; Horowitz, Stam, and Ellis, *Why Leaders Fight*.

loser, and then regress those on conflict initiation along with non-combat and combat leaders, but the confidence intervals overlap between war winners and other military experience types. Accordingly, we only need to theorise how selection effects might explain why leaders with war-win experience are likelier than non-ex-military leaders to initiate conflict. Furthermore, I do not consider a self-selection explanation for war win experience since there is no reason to speculate that war winners should be more likely to attempt or be more successful at self-selecting than war losers or those with no military experience. If there were either a candidate or leader self-selection event for war winners, this is most likely because they have military experience rather than war experience. Table 5-4 below summarises the expected selection effects, the treatment and comparison groups, and the supporting supporting hypotheses for war experience type.

Table 5-4: Summary of leader war experience type selection effect hypotheses.

Selection Effect Hypotheses	**Treatment Group**		**Comparison Group**		**Supporting Hypotheses**
	Leader War Experience Type	**Counterfactual Leader War Experience Type**	**Leader War Experience Type**	**Counterfactual Leader War Experience Type**	
Candidate Selection Effect	War-Winner	War-Winner	No Military Experience	No Military Experience	FIR
Leader Selection Effect	War-Winner	No Military Experience	No Military Experience	War-Winner	FIR

1 War-Win Experience – Leader Selection Effect Explanation

Why would leaders with war win experience only be more likely to initiate conflict than leaders with war loss experience and no war experience when the counterfactual leader has the same war experience type? The most likely reason is an FIR preference – that selectorates are likelier to choose war winners over others in environments where conflict is more likely. Considering I used the examples of Kerry and McCain above – both of whom fought in a lost war – this could mean that only combat veterans who have fought in a successful war enjoy an

electoral advantage. Indeed, the U.S. had a World War Two veteran President every year between 1952 and 1992, but never had a Vietnam Veteran President. In fact, a Vietnam veteran lost three straight elections between 2000 and 2008 (Gore, Kerry, and McCain). Admittedly, there is not much evidence beyond the U.S. anecdotes since most countries do not have elections with candidates with war loss and war win experience, let alone elections between a war loser and non-veteran. For example, in the post-war years, Germany had multiple elections with candidates who fought in a lost war, but none with those who have fought in a war that was won, and rarely have the runners-up been non-military. Nevertheless, the U.S. anecdotes are at least intuitive and possibly illustrate a more general pattern across states that war winners are more electable than war losers and those without military experience. Thus, we can formulate the war-winner FIR leader selection effect hypothesis:

War-Winner Experience – FIR Leader Selection Hypotheses:

1. Leaders with war win experience are more likely to initiate conflict than leaders with war loss experience or no war experience only when the leader and counterfactual leader do not have the same war experience type.
2. Candidates with war win experience are more likely to be chosen over candidates with war loss experience and no war experience in more conflict-prone states.

5.3.1 War-Win Experience – Candidate Selection Explanations

Why would war-winner leaders be more likely to initiate conflict than war losers and non-ex-military leaders only when the counterfactual leaders have the same war experience type? On the one hand, a First Image Reversed preference among selectorates does not appear to be a likely explanation. That is, there is little reason to suspect that elites or the mass public would prefer a candidate who has fought in a successful war over one that fought in a failed war.

Perhaps this might matter if they were a leading general, but surely not for ordinary soldiers as most leaders were. Furthermore, given that wars are rare, it is unlikely that leader selection events would have a mix of either candidate, or a candidate that has fought in both a successful and failed war. Indeed, according to the LEAD dataset, only 28 out of 3008 leaders have fought in both a successful and failed war, of which nearly half were French and Greek leaders. Rather, it is more likely that states that have won wars in the past tend to be more conflict-prone in the future because their victory likely reinforced their strength and position vis-à-vis other states. In contrast, states that have lost wars will likely have suffered hefty economic losses, become less powerful, and hence lost the means to be involved in conflict. As a result, conflict-prone states would likely have more candidates who have fought in a victorious war than conflict-averse states. The comparison between Japan and the U.S. above supports this possibility. The U.S. initiated over five times more conflicts in the 14 years after 1945 than the 14 years during the interwar period, whereas Japan initiated 17 in the 14-year interwar period but 0 in the 14 years after 1945[167] Similarly, Germany did not initiate a single conflict between 1945 and 1980.[168] Thus, for war experience we can test the following FIR candidate selection hypothesis for war winners:

War-Winners – FIR Candidate Selection Hypothesis:

1. Leaders with war-win experience are more likely to initiate conflict than leaders with war-loss experience and no war experience only when the leader and counterfactual leader have the same war experience.
2. Candidates are more likely to have war-win experience than war-loss experience and no war experience in more conflict-prone states.

[167] Gochman and Maoz, 'Militarized Interstate Disputes, 1816-1976'.
[168] Ibid.

5.4 Age

Why older leaders appear more likely to initiate conflict than younger leaders is likely explained by either candidate or leader selection effects. I do not postulate a self-selection hypothesis since the effect of age on conflict in previous studies is stronger in democracies and thus little reason to suspect that the self-selection hypothesis explains why older leaders are more likely to initiate conflict than younger leaders. Since age is continuous, what the selection effect hypotheses mean is less clear than binary variables, so to aid understanding it is more helpful to think of only two leader ages: older and younger. Table 5-5 below summarises the expected selection effects, the treatment and comparison groups, and the supporting hypotheses for age.

Table 5-5: Summary of leader age selection effect hypotheses.

Selection Effect Hypotheses	**Treatment Group**		**Comparison Group**		**Supporting Hypotheses**
	Leader Age	**Counterfactual Leader Age**	**Leader Age**	**Counterfactual Leader Age**	
Candidate Selection Effect	Older	Older	Younger	Younger	FIR, Demographics, Affinity Bias
Leader Selection Effect	Older	Younger	Younger	Older	FIR, Affinity Bias

5.4.1 Age – Leader Selection Effect Explanations

Why would older leaders be more likely to initiate conflict only when the counterfactual leaders are not of a similar age? The most intuitive explanation for this potential finding, I believe, is that voters or Heads of State could consistently be choosing an older over younger candidate for leader in environments where conflict is more likely. If this were the case, however, it is more likely a result of people choosing candidates with more political experience than choosing older candidates, given that age and political experience are inextricably linked. Indeed, according to the LEAD dataset, a leader's prior experience and entry age are highly correlated (Figure 5-3).

Consequently, if people choose more experienced candidates over less experienced candidates in response to interstate threats, this would necessarily result in older candidates being chosen over younger candidates. Equally, even if people were unaware of a candidate's political experience, they could use their age as a heuristic for it instead – rightly or wrongly – when choosing among candidates. However, regardless of whether people are directly voting for more politically experienced candidates or voting for older candidates as a heuristic for political experience in case of uncertainty, there is some empirical support in psychology that there could be such a preference for older candidates. Older-looking composite images of Presidential candidates are more likely to win mock elections during wartime.[169] Furthermore, people endorse younger candidates when they desire change and older candidates when they desire stability.[170] This preference for older candidates during times of instability might be driven by a sense that older individuals are better at reasoning through social conflicts.[171] So, if voters use a candidate's age as a proxy for political experience in the absence of other information, they might be following the simple heuristic "If there is a need for stability, then choose an older leader."[172] Since increased international conflict likely brings uncertainty, selectorates in states that frequently engage in conflict might be likelier to choose older leaders over younger ones.

[169] Brian R. Spisak, 'The General Age of Leadership: Older-Looking Presidential Candidates Win Elections during War', *PLoS One* 7, no. 5 (2012).
[170] Brian R. Spisak et al., 'The Age of Exploration and Exploitation: Younger-Looking Leaders Endorsed for Change and Older-Looking Leaders Endorsed for Stability', *The Leadership Quarterly* 25, no. 5 (2014): 805–16.
[171] Igor Grossmann et al., 'Reasoning about Social Conflicts Improves into Old Age', *Proceedings of the National Academy of Sciences* 107, no. 16 (2010): 7246–50.
[172] Spisak et al., 'The Age of Exploration and Exploitation', 807; Steven W. Webster and Andrew W. Pierce, 'Older, Younger, or More Similar? The Use of Age as a Voting Heuristic', *Social Science Quarterly* 100, no. 3 (2019): 635–52.

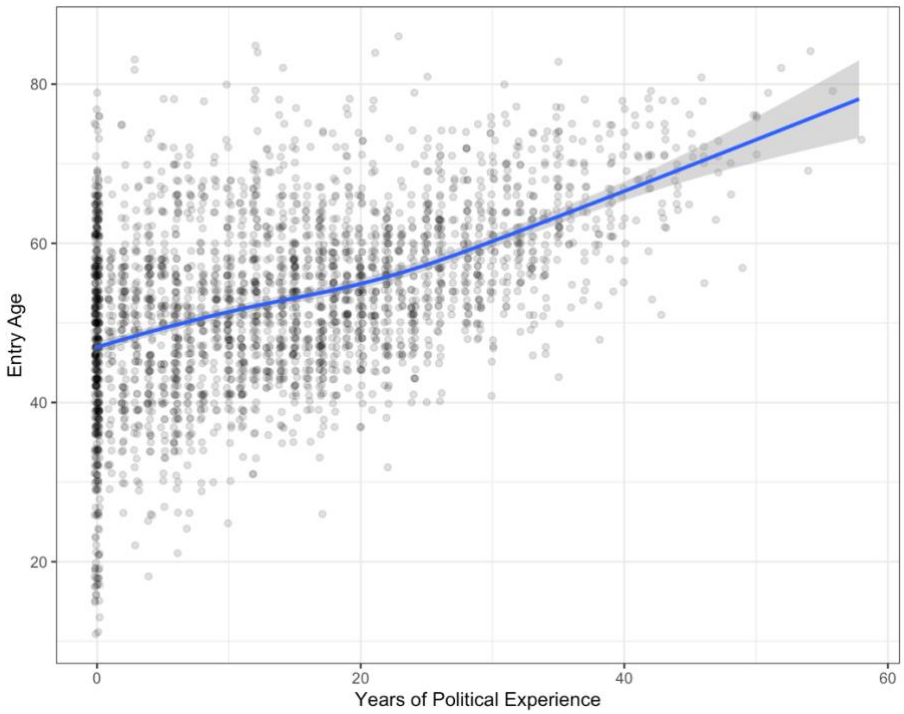


Figure 5-3: Correlation between leaders' years of previous political experience and entry age in LEAD.
Note: Bivariate OLS with 95% confidence intervals.

Beyond scientific studies, several real-world anecdotes also provide evidence for this FIR leader selection effect. For instance, in a 1984 U.S. election debate during the height of the Cold War, 73-year-old Ronald Reagan joked of his opponent, 56-year-old Walter Mondale: "I want you to know that also I will not make age an issue of this campaign. I am not going to exploit, for political purposes, my opponent's youth and inexperience." In a later interview, Mondale admitted that Reagan's quip "…had gotten me there. That was really the end of my campaign that night. I think."[173] Similarly, during the 2008 U.S. Presidential election between 47-year-old Barack Obama and 72-year-old John McCain, Joe Biden warned:

> "It will not be six months before the world tests Barack Obama like they did John Kennedy. The world is looking. We're about to elect a brilliant 47-year-old senator

[173] 'The Reagan/Mondale Debate' (PBS, 24 September 2000), https://www.pbs.org/newshour/spc/debatingourdestiny/doc1984.html.

> president of the United States of America… we're going to have an international crisis, a generated crisis, to test the mettle of this guy."[174]

If a staunch ally of a Presidential candidate publicly expresses the view (perhaps accidentally) that a young and inexperienced leader is likely to face a greater number of interstate threats, then surely this concern might also be held by voters, especially in states that frequently engage in international conflict. This is not to say prior political experience or age trumps policy or other relevant interests of voters. Indeed, John McCain lost the 2008 election to Obama despite the U.S. being engaged in two major overseas conflicts. Rather, when given a choice between an older or younger candidate, we might expect voters or Heads of State to choose the former for leader in environments where conflict is more likely. Thus, we can postulate the FIR leader selection effect hypothesis for age:

Age – FIR Leader Selection Hypotheses:

1. Older leaders are more likely to initiate conflict than younger leaders only when the leader and counterfactual leader are not the same age.
2. Older candidates are more likely to be chosen over younger candidates in more conflict-prone states.

Furthermore, affinity bias could explain why we might see a leader selection effect. For affinity voting, several pieces of evidence suggest that voters are more likely to cast ballots for similar-aged candidates.[175] So, if there are more older people in a population, then any older candidates would likely enjoy more electoral support than in states with lower ages. Additionally, younger people tend to vote at much lower rates than older people across many

[174] John M Broder, 'Biden's Crisis Remarks Reverberate', *The New York Times*, 20 October 2008, https://thecaucus.blogs.nytimes.com/2008/10/20/bidens-crisis-remarks-reverberate.
[175] Semra Sevi, 'Do Young Voters Vote for Young Leaders?', *Electoral Studies* 69 (2021): 102200.

states, with few exceptions.[176] Known as the "life cycle model of political turnout, " younger individuals tend to have fewer resources, interest, and political knowledge than older individuals, and so are simply less likely to attend the polls.[177] All in all, states with higher life expectancies and elderly in the population might be more conflict-prone because they tend to be wealthier and better developed. Indeed, powerful states are more likely to initiate conflict; relative power is correlated with GDP;[178] and GDP is correlated with life expectancy.[179] Consequently, candidates will tend to be older in conflict-prone states since they are more powerful, wealthy, and have higher life expectancies than conflict-averse states. In this web of correlations, then, the relationship between older leaders and past conflict might only arise because there is a higher probability that older individuals make it to the top of the political elite and become viable candidates for leader.

Age – Affinity Bias Leader Selection Hypothesis:

1. Older leaders are more likely to initiate conflict than younger leaders only when the leader and counterfactual leader are not the same age.
2. Older candidates are more likely to be chosen over younger candidates in states with higher life expectancies.
3. States with higher life expectancies are more likely to initiate conflict.

[176] Kaat Smets, 'A Widening Generational Divide? The Age Gap in Voter Turnout through Time and Space', *Journal of Elections, Public Opinion & Parties* 22, no. 4 (2012): 407–30.
[177] Kaat Smets, 'Revisiting the Political Life-Cycle Model: Later Maturation and Turnout Decline among Young Adults', *European Political Science Review* 8, no. 2 (2016): 225–49.
[178] CINC has a spearman correlation coefficient of .823. See: Hyung Min Kim, 'Comparing Measures of National Power', *International Political Science Review* 31, no. 4 (2010): 419.
[179] Rafia Shafi and Samreen Fatima, 'Relationship between GDP, Life Expectancy and Growth Rate of G7 Countries', *International Journal of Sciences* 8, no. 06 (2019): 74–79.

5.4.2 Age – Candidate Selection Explanations

Why would older leaders be more likely to initiate conflict only when the counterfactual leaders are of a similar age? This is probably for the same reasons as affinity bias above, but rather than being directly chosen, these states might have a surplus of older candidates, which are then more likely to be nominated and hence be chosen as leader. Thus, the demographic candidate selection effect hypothesis is formulated as follows:

Age – Demographic Candidate Selection Effect Hypotheses:

1. Older leaders are more likely to initiate conflict than younger leaders only when the leader and counterfactual leader are the same age.
2. Candidates are more likely to be older in states with higher life expectancies.
3. States with higher life expectancies are more likely to initiate conflict.

5.5 Female Leaders

It is more difficult to consider how selection could explain why female leaders appear more likely to initiate conflict than male leaders. On the one hand, all the available evidence suggests we should not expect a leader selection effect – that is, female candidates should not be more likely to be chosen over male candidates for leaders in environments where conflict is more likely. More specifically, we have no reason to suspect there should be a First Image Reversed selectorate preference for female leaders. Indeed, countries that have experienced an international conflict with fatalities have never appointed a female defence minister,[180] and people generally prefer male over female leaders in more hostile international security environments.[181] Likewise, we probably would not expect a self-selection explanation since no

[180] Tiffany D. Barnes and Diana Z. O'Brien, 'Defending the Realm: The Appointment of Female Defense Ministers Worldwide', *American Journal of Political Science* 62, no. 2 (2018): 355–68.
[181] Nam Kyu Kim and Alice J. Kang, 'External Threat Environments and Individual Bias against Female Leaders', *Political Science Research and Methods* 10, no. 1 (2022): 1–17.

female leaders are probably not more likely than male leaders to self-select. Thus, I believe only the candidate selection effect hypothesis is likely (Table 5-6).

Table 5-6: Summary of leader sex selection effect hypotheses.

Selection Effect Hypotheses	Treatment Group		Comparison Group		Supporting Hypotheses
	Leader Sex	Counterfactual Leader Sex	Leader Sex	Counterfactual Leader Sex	
Candidate Selection Effect	Female	Female	Male	Male	Gender Constraints

5.5.1 Female Leaders – Candidate Selection Effect Explanation

Why would female leaders be more likely to initiate conflict only when the counterfactual leaders have the same sex? I believe the likely causal explanation here is due to candidate constraints. Most studies on female political participation and international conflict restrict their analyses to shorter time periods and fewer states, which essentially controls for factors not considered in large-*n* studies on female leaders between 1875 and 2004. Most female leaders only really began to enter office during the last half of the Cold War – when conflicts were generally more frequent than at other periods. Margaret Thatcher, Isabel Peron, Golda Meir, Corazon Aquino, and Indira Gandhi are the more notable female leaders during this period. What is also common among these women is that they were leaders of some of the more powerful and conflict-prone states where gender equality tends to be higher: the United Kingdom, Argentina, Israel, the Philippines, and India, respectively. During this same period, many smaller, weaker and less conflict-prone states also arose due to decolonisation, where women's political participation is lower or even forbidden, and so tend to not have female leaders. Thus, it may be that progressive ideas regarding female political participation are correlated with international conflict due to some unknown common cause that arose during this time. Admittedly, there is little evidence for what this cause may be, but I suspect that since

gender equality also correlates with wealth, GDP, and economic development,[182] more powerful and conflict-prone states might have been quicker to adopt these social changes regarding women in politics because they had the resources and political stability to do so. As a result, it could be that females are less constrained in becoming candidates in more powerful and conflict-prone powerful, and hence environments where conflict is more likely. Thus, we can formalise the gender constraint selection effect hypothesis for why female leaders are more likely to initiate conflict:

Female Leaders – Gender Constraint Candidate Selection Effect Hypotheses:

1. Female leaders are more likely to initiate conflict than male leaders only when the leader and counterfactual leader are the same sex.
2. Candidates are more likely to be female than male in more gender-equal countries.
3. States with higher levels of gender equality are more likely to initiate conflict.

5.6 Right-Wing Ideology

I suggest there is sufficient reason to believe that any of the three selection effect hypotheses might explain why right-wing leaders appear more likely to initiate conflict than left-wing and centre candidates. Table 5-7 below summarises the expected selection effects, the treatment and comparison groups, and the supporting hypotheses for ideology.

[182] Armin Falk and Johannes Hermle, 'Relationship of Gender Differences in Preferences to Economic Development and Gender Equality', *Science* 362, no. 6412 (2018): eaas9899.

Table 5-7: Summary of leader ideology selection effect hypotheses

Selection Effect Hypotheses	Treatment Group		Comparison Group		Supporting Hypotheses
	Leader Ideology	Counterfactual Leader Ideology	Leader Ideology	Counterfactual Leader Ideology	
Candidate Selection Effect	Right-Wing	Right-Wing	Left-Wing, Centre	Left-Wing, Centre	FIR
Leader Selection Effect	Right-Wing	Left-Wing, Centre	Left-Wing, Centre	Right-Wing	FIR
Candidate Self-Selection Effect	Right-Wing	Right-Wing	Left-Wing, Centre	Left-Wing, Centre	Attempt
Leader Self-Selection Effect	Right-Wing	Left-Wing, Centre	Left-Wing, Centre	Right-Wing	Success

5.6.1 Right-Wing Ideology – Leader Selection Effect Explanations

There are several reasons to suspect that selectorates could be more likely to choose right-wing over left-wing candidates for leader in more dangerous environments. Right-wing leaders tend to be hawkish and emphasise the need to confront interstate threats – a stance that would likely appeal to selectorates in places where international conflict is frequent. When people perceive their political environments as increasingly hostile, they shift from a dovish to a hawkish bias.[183] The most famous example of this is the British public's discontent with Neville Chamberlain over his policy of appeasement toward Nazi Germany, which ultimately led to his resignation and the appointment of a more hawkish leader, Winston Churchill. Furthermore, there is evidence that dovish governments are more likely to lose power after offering concessions in international bargaining.[184] In US elections between 1980 and 2004, for instance, Gadarian finds that voters punished Democratic candidates less for being too hawkish on foreign policy than for being too dovish, concluding that "this may indicate that voters prefer a

[183] Sandeep Baliga, David O. Lucca, and Tomas Sjöström, 'Domestic Political Survival and International Conflict: Is Democracy Good for Peace?', *The Review of Economic Studies* 78, no. 2 (2011): 459.
[184] Schultz, 'The Politics of Risking Peace'.

president they perceive as strong (i.e., hawkish) on national security issues rather than one that shares their own ideal position on foreign policy."[185] Dovish incumbents are more likely to be punished at the polls than hawkish incumbents when initiating a high-level militarised interstate dispute.[186] Given this relationship, it is likely that right-wing candidates would also enjoy a similar electoral advantage where conflict is frequent.

In a lengthy literature review on the ideology-foreign policy nexus, Costel concludes that "the overwhelming majority of these studies show that…right-leaning governments support more hawkish foreign policy actions than their left-wing counterparts."[187] In Western European democracies, conservative parties are more likely than liberal parties to support increased military spending and expanding the military presence abroad.[188] In the United Kingdom, Labour Party supporters are almost three times more likely to support reduced military spending than Conservative Party supporters.[189] Likewise, French right-wing supporters were likelier to take a hawkish stance on Algerian independence, the Suez Canal, and the development of nuclear capabilities.[190] Thus, we can formulate the FIR leader selection effect hypotheses for why right-wing leaders are more likely to initiate conflict:

[185] Shana Kushner Gadarian, 'Foreign Policy at the Ballot Box: How Citizens Use Foreign Policy to Judge and Choose Candidates', *The Journal of Politics* 72, no. 4 (2010): 1055.
[186] Osman Sabri Kiratli, 'Where Do Hawks and Doves Fly When Shots Are Fired?', *Party Politics*, 2020, 1–12.
[187] Costel, 'Hawks versus Doves: The Influence of Political Ideology on the Foreign Policy Behavior of Democratic States', 57.
[188] Richard C. Eichenberg, *Public Opinion and National Security in Western Europe: Consensus Lost?* (Springer, 1989).
[189] Hans-Dieter Klingemann, Richard Hofferbert, and Ian Budge, 'A Theory of Democratic Policymaking', in *Parties, Policies, and Democracy*, by Hans-Dieter Klingemann, Richard Hofferbert, and Ian Budge (Boulder, CO: Westview, 1994), 1–19.
[190] Michael T. Koch and Patricia Sullivan, 'Should I Stay or Should I Go Now? Partisanship, Approval, and the Duration of Major Power Democratic Military Interventions', *The Journal of Politics* 72, no. 3 (2010): 619.

Right-Wing Ideology - FIR Leader Selection Effect Hypothesis:

1. Right-wing leaders are more likely to initiate conflict than left-wing and centre-leaders only when the leader and counterfactual leader have the same ideology.
2. Right-wing candidates are more likely to be chosen over centre and left-wing candidates in more conflict-prone states.

5.6.2 Right-wing Ideology: Candidate Selection Effect Explanation

Why would right-wing leaders only be more likely to initiate conflict than left-wing and centre leaders when the counterfactual leaders have the same ideology? The most likely explanation is due to the same FIR preference as above but for right-wing candidates. It is more difficult to consider why demographics could explain a candidate selection effect. While ideology is certainly correlated with demographics within states, why people might adopt right-wing beliefs in the U.S. differs from, say, Finland or Brazil. For example, older age is correlated with right-wing ideology in the U.S., but left-wing ideology in Serbia.[191] The only common predictor for right-wing beliefs across states seems to be resistance to immigration.[192] However, anti-immigration attitudes are probably not correlated globally with any of the traits investigated here. Nevertheless, we can formulate the FIR candidate selection effect hypothesis for why right-wing leaders are more likely to initiate conflict:

[191] Bojan Todosijević, 'Left-Right Ideology: Its Meaning and Effects on Party Preferences in Serbia', *Социолошки Преглед* 50, no. 2 (2016): 161–78.
[192] Ferruh Yılmaz, 'Right-Wing Hegemony and Immigration: How the Populist Far-Right Achieved Hegemony through the Immigration Debate in Europe', *Current Sociology* 60, no. 3 (2012): 368–81; Hans-George Betz, 'The New Politics of Resentment: Radical Right-Wing Populist Parties in Western Europe', *Comparative Politics*, 1993, 413–27; Jeroen Van der Waal, Willem De Koster, and Peter Achterberg, 'Ethnic Segregation and Radical Right-Wing Voting in Dutch Cities', *Urban Affairs Review* 49, no. 5 (2013): 748–77.

Right-Wing Ideology – FIR Candidate Selection Effect Hypothesis:

1. Right-wing leaders are more likely to initiate conflict than left-wing and centre leaders only when the leader and counterfactual leader have the same ideology.
2. Candidates are more likely to be right-wing in more conflict-prone states.

5.6.3 Right-Wing Ideology – Leader Self-Selection Explanations

The other potential explanation is that right-wing leaders are more likely to attempt to self-select, as well as be successful. Given the evidence above considering right-wing ideology being correlated with hawkishness, we can test the right-wing ideology self-selection hypothesis:

Right-Wing Ideology – Candidate Self-Selection Hypothesis:

3. Right-wing leaders are more likely to initiate conflict than left-wing and centre leaders only when the leader and counterfactual leader have the same ideology and following self-selection events.
4. Right-wing candidates are more likely to attempt to self-select than left-wing and centre candidates.

Right-Wing Ideology – Leader Self-Selection Hypothesis:

3. Right-wing leaders are only more likely to initiate conflict than left-wing and centre leaders only when the leader and counterfactual leader do not have the same ideology and following self-selection events.
4. Right-wing candidates are more successful at self-selecting than left-wing and centre candidates.

5.7 Holistic Measures

For the two holistic measures, hawkishness and leader risk, we would expect to see the same broad hypotheses above since they combine all these traits. Since they are latent variables, however, there is no specific theoretical evidence to suspect any selection effects beyond those we have reviewed above for each trait. Both measures are already positively correlated with the high-risk traits above, so the same causal logic applies, and a lengthy review is not necessary for why either could be a product of selection effects. Thus, instead of testing every single hypothesis above for each trait that is correlated with hawkishness, we can deduce that if any of the five main hypotheses were supported for why hawkish leaders are more likely to initiate conflict, then it is likely due to the same selection effects we see for the other traits, and so we would only need to re-test that specific hypothesis when investigating the hawkishness score. Since hawkishness is a continuous variable like age, it is easier to understand the hypotheses by considering only two types, Hawk and Dove. Table 5-8 below summarises the expected selection effects, the treatment and comparison groups, and the supporting hypotheses for ideology.

Table 5-8: Summary of leader hawkishness selection effect hypotheses.

Selection Effect Hypotheses	**Treatment Group**		**Comparison Group**		**Supporting Hypotheses**
	Leader Ideology	**Counterfactual Leader Ideology**	**Leader Ideology**	**Counterfactual Leader Ideology**	
Candidate Selection Effect	Hawk	Hawk	Dove	Dove	See other traits
Leader Selection Effect	Hawk	Dove	Dove	Hawk	See other traits
Candidate Self-Selection Effect	Hawk	Hawk	Dove	Dove	See other traits
Leader Self-Selection Effect	Hawk	Dove	Dove	Hawk	See other traits

6 Methodology

To what extent do a leader's traits influence their decision to start conflicts? While previous research suggests several traits are influential, the key question that remains is whether these relationships are possibly biased due to selection effects. Accordingly, the methodology I use is principally designed to test for the selection effect hypotheses using quantitative methods.

In this chapter, I first explain how I operationalise selection events and identify counterfactual leaders within them. I thoroughly review each selection mechanism (i.e., elections, appointments, successions, self-selections) and the criteria needed to identify the counterfactual leader in each case. Second, I describe how traits of both leaders and counterfactual leaders are constructed, as well as conflict initiations and other variables used when testing the supporting hypotheses. Lastly, I describe the general pattern for how data was analysed in the findings chapter.

6.1 Key Terms of Reference

To reiterate, I use the term candidate to describe both leaders and counterfactual leaders in selection events. I realise that describing a leader of a failed coup as a "candidate" is jarring, yet there is no other word that better describes both leaders and counterfactual leaders. Another option might be "potential leaders", but I already distinguish between leaders and counterfactual leaders, and so it would likely create more confusion than clarity. I use the term "individual" when referring to a unique person regardless of how many times they were a candidate. For example, Winston Churchill is the leader once (1951-1955), the counterfactual leader once (1945-1950), and a candidate twice (1945 and 1951 elections),

6.2 Scope

I limit the scope of my analysis to the years between 1945 and 2004 primarily due to data limitations and research efficiency. While expanding the scope would strengthen external validity, there is substantially less biographical information on leaders and even more so on counterfactual leaders before 1945. The time spent researching these cases would be inefficient as most would result in little to no information. Furthermore, the MID dataset (where my outcome variable described below is taken from) ends in 2015 and so would drastically increase the number of individuals needed to collect data for a small 17-year time period that I do not anticipate to significantly alter my findings. In other words, there is little reason to suspect fundamental differences in how and why leaders are selected between 2005 and 2015 compared to 1945 to 2004, so the scope is sufficiently large enough to make reasonable generalisations dependent on the findings.

6.3 Units of Analysis - Leader Selection Events

The units of analysis are leader selection events which I define as any actual or potential leader transition. I include actual leader transitions to ensure that cases in which a new leader is selected are included. I also include potential leader transitions to ensure that cases when an incumbent is re-elected or a leader thwarts a coup attempt are included as well, since selection has occurred even though no leader transition occurred.

Leader selection events must determine the position of the leader of the state as defined previously – the individual who is ultimately responsible for the foreign policy of their state. Since we are ultimately concerned with whether selection explains why leaders with high-risk traits are more likely to initiate conflict, then it only makes sense to investigate only the traits of individuals who are the ultimate decision-making authority of their state. Following standard practice in IR, leaders are operationalised using the LEAD dataset, and I use that individual's position to identify and code the corresponding leader selection event. For example, LEAD

considers the leader of France to be the Prime Minister before 1959 but the President after since the 1958 constitution granted more power to the President. Therefore, before 1959, selection events in France are only included if they are general elections or appointments that determine the Prime Minister, whereas after 1959, selection events are only included if they are elections or successions that determine the President.

6.4 The Counterfactual Leader

The counterfactual leader is the individual who was the next closest individual to being leader in the selection event. Considering the key research objective is to assess whether the effect of leader traits on conflict depends on the trait of the counterfactual treatment (i.e., the opposite trait of the counterfactual leader), it is necessary to more thoroughly explain how I identify and operationalise both leader selection events and the counterfactual leader.

6.5 Operationalising Leader Selection Events and Counterfactual Leaders

While there is wide variation in how states select leaders, selection events are any of four selection mechanisms: elections, appointments, succession, and self-selection. In most cases, leaders are selected in only one, but there may be some ambiguity in certain cases due to complexities and variation across states. The importance of operationalising leader selection events into these four different selection mechanisms is primarily to identify how leaders are selected into office, who can be considered the counterfactual leader, and whether a leader selection effect is even at least a plausible explanation. For example, the First-Image Reversed (FIR) hypotheses depend on the selectorate choosing candidates with high-risk traits over those with low-risk traits. Therefore, to test the FIR hypotheses, we must exclude cases such as when U.S. Vice President Lyndon Johnson automatically became President following the death of John F. Kennedy. An FIR explanation would not make sense in this case, although it would make sense in the subsequent 1964 U.S. general election when Johnson defeated Barry Goldwater.

Therefore, it is important to categorise leader selection events more specifically to ensure any supported hypotheses are internally valid. A review of how I operationalise selection mechanisms, sub-mechanisms, and counterfactual leaders follows.

6.5.1 Elections

The first selection mechanism is elections. Because elections vary in how votes determine who becomes leader, I further disaggregate elections into several sub-mechanisms: presidential, parliamentary, party leadership, and normative.

6.5.1.1 Presidential elections

Presidential elections are the clearest sub-mechanism to infer whether voters choose candidates with certain traits over others. Voters are given a clear choice between two or more candidates and there is little confusion that whoever wins the electoral or popular vote will become leader. Even if people vote based on party loyalty rather than the traits of the specific nominees, they are still directly voting for the future leader of their state. Thus, even though we cannot observe any FIR selectorate preference for say, a veteran over a non-veteran in more dangerous times, we can still observe a baseline preference (i.e., if veterans are more likely to win elections against non-veterans than vice versa), and so it follows that if veterans are not more likely to win elections in more conflict-prone states then at least we have more evidence that this is probably not explained by an FIR preference.

In Presidential elections determining the counterfactual leader is rather straightforward; whoever wins the second-most individual votes in simple or absolute majority systems (e.g., Chile pre-1989), or second-most electoral votes (e.g., United States). Both of these vote determination types occur in most Presidential elections since they tend to be a two-candidate contest in practice, and there is little reason why we might expect anyone else to be the

counterfactual leader besides the runner-up. Although in some types of Presidential elections, the counterfactual leader is more difficult to determine, such as in the Lay de Lemas system in Uruguay, since the candidate with the most votes in the party with the most votes become President. In these cases, we could consider the counterfactual leader to either be the candidate with the second-most votes (even if their party lost) or the second-place candidate of the party that won. A strong case could be made for either, although I use the former when determining the counterfactual leader since they received the second-most votes in total.

6.5.1.2 Parliamentary Elections

The second sub-mechanism is Parliamentary elections, where voters choose leaders more indirectly than in Presidential elections. In most parliamentary elections, votes are tallied at the local level, determining which party gains a majority and which party leader or nominee becomes Prime Minister. This is, of course, what usually happens as parliamentary elections are varied in how votes formally determine who becomes Prime Minister. While voters can cast ballots based on who they want to become Prime Minister, they could also cast ballots due to purely local interests for their MP, and so who becomes state leader is not necessarily due to voters' preferences for leader. In other words, it is more difficult in parliamentary elections to infer whether voters choose candidates as opposed to party or local interests. Nevertheless, I include parliamentary elections when testing FIR hypotheses to improve external validity, as there will likely be fewer elections between candidates with two different traits. If the evidence supports one of the FIR hypotheses, we can always conduct robustness checks by excluding parliamentary elections.

The counterfactual leader in parliamentary elections is observable yet slightly more difficult to determine than in Presidential elections. This is primarily because who wins the most seats does not necessarily guarantee the Premiership, in theory at least. In Westminster-type elections, such as in the U.K., the leader of the party who wins enough seats to gain a majority is

expected to be appointed as Prime Minister by the Head of State. While the Head of State can pick anybody they want in theory, their choice is almost always entirely constrained by norms of appointing leaders based on their party's relative share of seats. Therefore, it is reasonable to assume that, in most cases, the counterfactual leader in parliamentary elections is the leader or nominee of the party who wins the second-most seats since they were the next closest to becoming Prime Minister. In more complicated cases like Germany, where the Chancellor is technically elected after the general election in the Bundestag via secret ballots, the counterfactual leader is still the leader of the runner-up party since the Chancellor is always the leader of the party with the most seats. In other parliamentary systems where coalitions are the norm, such as in the Netherlands and Denmark, the coalition parties propose an individual as Prime Minister following negotiations after the election, which the Head of State may or may not approve and might even choose someone who is not a party leader. In these cases, it is even more difficult to determine the counterfactual leader since who wins the most seats often does not become Prime Minister, but not always, and we can only observe the counterfactual leader if we could observe who the coalition parties "almost" chose. While this makes determining the counterfactual leader more difficult in non-Westminster type parliamentary elections, I believe the country-specific norms of selecting the Head of Government is the most appropriate way to determine the counterfactual leader. If the selected leader is from the same party as the winning party, then the counterfactual leader is the leader of the winning party instead of the leader of the runner-up party.

6.5.1.3 Leadership Elections

The third election sub-type is leadership elections. These mostly occur in Parliamentary democracies following the resignation or unexpected death of the incumbent Prime Minister. The incumbent party remains in power but holds an election within their party to determine the successor. In these cases, leaders are selected either by formal leadership elections where votes

are tallied, or informally where party elites do not necessarily tally votes but otherwise agree on the replacement. Accordingly, the counterfactual leader in leadership elections is the individual with the second-most votes in a formal electoral contest or the individual explicitly stated as the alternative choice in more informal leadership contests. By explicitly stated, I mean that there has to be some scholarly evidence or multiple statements made by those involved in the informal leadership election that the individual was the next choice.

6.5.1.4 Government Elections

The fourth election type is government elections. In these cases, the central government either formally votes, or informally agree, for who they want as leader. Like leadership elections, these almost always occur following the resignation or death of the incumbent (e.g., Ireland). Determining the counterfactual is simple – the runner-up – but difficult to observe since many government elections occur behind closed doors. Therefore, we can only determine the counterfactual leader in these cases if they have been identified in government documents or academic sources. In Ireland, for example, Prime Ministers are ultimately appointed by the President but are first nominated during Dali debates, of which the Ireland government has transcripts for as far back as 1919.[193]

6.5.1.5 Normative Elections

The last election sub-type is normative elections, whereby norms rather than votes essentially determine winners. For example, the leader of Switzerland in LEAD is considered to be the President of the Federal Council, who is elected yearly. However, the Federal Council are expected to vote for the most senior member who has never previously served as Federal

[193] 'Debates', Houses of the Oireachtas, accessed 15 December 2021, https://www.oireachtas.ie/en/debates/.

President, who is almost always the Vice President.[194] Thus, even though Switzerland's leaders are technically "elected" by the Federal Council, who will eventually win is rarely in doubt. In these cases, the counterfactual leader is the individual who is second-in-line based on previous norms. So, for Switzerland, the counterfactual leader is almost always the new Vice President.

The different types of elections are meant to explain more precisely who can be considered the counterfactual leader and why. However, I would like to address several issues with operationalising elections and counterfactual leaders. First is the problem of elections that may not be free and fair. While this is undoubtedly a concern, I initially make no distinction as to whether an election is free-and-fair when determining the counterfactual leader for several reasons. First, most rigged elections have more to do with voting restrictions rather than allowing opposition candidates to become leader. Elections in South Africa during Apartheid, for instance, are not considered free and fair in NELDA , although the main opposition candidates in these elections were not restricted from becoming Prime Minister had they won. Second, while there might be some rigged elections where the incumbent regime only allows certain opposition candidates to run in the belief they are unelectable, this does not appear to be the case more generally in existing election datasets or my research.[195] Usually, incumbents seem to ban opposition parties outright rather than specific candidates, so the runners-up can still be considered counterfactual leaders. In other words, excluding elections that are not free and fair would control away the very system potentially producing a disproportionate number of candidates with certain backgrounds. Therefore, I do not initially require elections to be free and

[194] The Federal Council, 'History of the Federal Presidency', The Federal Council: Portal of the Swiss Government, accessed 17 April 2023, https://www.admin.ch/gov/en/start/bundespraesidium/bundespraesidenten-und-schweizer-geschichte.html.

[195] I use the National Elections Across Democracy and Autocracy (NELDA) dataset to determine elections. See, Susan D. Hyde and Nikolay Marinov, 'Which Elections Can Be Lost?', *Political Analysis* 20, no. 2 (2012): 191–210.

fair, although I use this requirement when testing First-Image Reversed hypotheses, as will be explained in more detail later in this chapter.

The second issue is what to do with unopposed elections. Ultimately, I decided to include unopposed elections as selection events even though there is no counterfactual leader for several reasons. First, it was not always clear whether the election was unopposed or there was just insufficient information to determine the runner-up (mostly for elections in smaller states farther back in time). In elections that were clearly unopposed (such as winning 99% of the vote against the runner-up – blank ballots – I decided to include them and code them as being unopposed since they were unopposed might be also be correlated with conflict in some way.

6.5.2 Appointments

The second selection mechanism is appointments. These largely occur in either semi-Presidential or constitutional monarchies where the Head of State (President or Monarch) chooses the individual they want to serve as Head of Government. Like leadership elections, political appointments tend to follow the resignation or death of the incumbent leader, and the individual appointed is leader until the next general election is held. In most cases, determining whether a leader has been selected by appointment is clear and easy to observe – the Head of State formally appoints someone. However, in parliamentary systems there is often ambiguity in determining whether a leader is appointed by the Head of State or elected following a general election. Again, Heads of State may choose whoever they like as leader, in theory, but in practice their choice is often highly restricted by norms or legislative frameworks, and so they almost always choose the leader or nominee of the party with the most seats following an election. Accordingly, I consider Prime Ministers as being selected following a general election rather than being appointed, since the Head of State nearly always appoints the winner, except in select countries as described previously.

Alternatively, an individual may become leader by being hand-picked by their predecessor. These cases are rare and usually only occur in states with charismatic leaders or dictatorships (e.g., Francois Duvalier appointing his son, Jean-Claude Duvalier, as successor). Though, this means that there is some conceptual overlap with monarchical succession, which I more thoroughly explain later. But for brevity, I distinguish the two based on whether succession is formally stated to be monarchical. If a predecessor appoints their own family as leader, but there is no formal monarchical succession mechanism, then this is better grouped within the appointed selection event type.

The last way leaders can also be appointed is by a foreign Head of State. Though, some might argue these individuals are not "leaders" since there is a debate as to whether states that have leaders installed by foreign powers should be considered independent. Nevertheless, I include these types here because leader datasets include many such cases. In Archigos, for instance, U.S. General Douglas Macarthur is considered leader of Japan between 1945 and 1951, although the Correlates of War does not consider Japan an independent state during this period. Thus, depending on how leaders are defined and operationalised in other leader trait studies, it is necessary to include foreign-installed leaders to get the full picture when investigating selection and trait explanations.

Across all appointment sub-mechanisms, the counterfactual leader is the individual the Head of State "almost" chose but ultimately did not. More specifically, the counterfactual leader can either be the first person asked by the Head of State to lead a government but 1) simply declined or 2) eventually could not get enough support to form a government and so stepped down before they officially become leader (i.e., India 1979), or 3) the individual the Head of State "almost" chose but did not (e.g., Rab Butler in 1957). For the last criterion, the counterfactual leader is difficult to observe since we cannot know who was almost chosen unless the Head of State explicitly states who else they had in mind or there is a general consensus among experts in the literature. For instance, sources consistently agree that Rab Butler was the alternative

candidate to Alexander Douglas-Home to succeed Harold Macmillan as UK Prime Minister,[196] so it is reasonable to infer that Butler is the counterfactual leader in this case.

6.5.3 Successions

The third selection mechanism is successions. These are selection events in which a leader comes into power based on being next in line of some formal or informal pre-determined order. However, two sub-types of successions, constitutional and hereditary, require different criteria when determining the counterfactual leader.

6.5.3.1.1 Constitutional Successions

Constitutional successions are when leaders enter office as the second-in-line based on formal legislative rules outlying the order of succession. The counterfactual leader is relatively straightforward in constitutional succession since the orders of succession are written into legal frameworks, so we can easily deduce who was next-in-line in each case. In the U.S., for instance, the Vice President assumes the Presidency upon the resignation or death of the President, and if the Vice President is not available, then the Speaker of the House of Representatives instead becomes President. So, when Vice President Lyndon Johnson became U.S. President in 1963 following the assassination of John F. Kennedy, the counterfactual leader would have been Speaker of the House, John W. McCormack.

The second succession sub-type is hereditary succession. This is perhaps the oldest leader selection type and can be based on a formal or informal framework, where the next closest relative becomes leader. Although, hereditary successions are varied with regard to the specific type of inheritance rules. Agnatic primogeniture - where the eldest son is the heir – has traditionally been the most common type of hereditary selection. However, states in the Arabian

[196] Anthony Howard, *RAB: The Life of RA Butler* (A&C Black, 1987), 246–47.

Gulf use a mix of succession rules where the heir can be the eldest son, or oldest surviving son of a predecessor.[197] Nevertheless, identifying the counterfactual leader in monarchical succession – the heir apparent – is straightforward since these individuals are usually explicitly stated to be the new heir following a succession. However, in monarchical successions with vague orders of succession that do not specify who the new heir apparent should be, identifying the counterfactual leader is more difficult. Nonetheless, in these cases, the counterfactual leader is the individual who would most likely be leader if the current leader died. Who the "most likely" individual is subjective but is determined by scholarly or media consensus stating who is likely the new heir apparent.

As noted previously, there is some conceptual overlap between succession and appointments by a Head of State since many dictators end up appointing one of their children as successor (e.g., Duvalier dynasty in Haiti). The main difference, however, is that who becomes leader in succession is determined by formal legislation or informal customs, whereas in appointments by a predecessor, a choice is made without any corresponding framework. Thus, determining whether a leader in these cases enters office through appointment or succession is based on whether there were any legislative or hereditary frameworks implemented. If a dictator declares an individual the successor simply due to nepotism, then this is categorised within the appointment selection event type rather than succession.

6.5.4 Self-Selections

The last selection mechanism is self-selection. These most often include coups and revolutions, and even cases where an ambitious individual consolidates power and takes over unopposed, such as when Joseph Stalin became leader of the Soviet Union in 1923. Since there

[197] John E. Peterson, 'The Nature of Succession in the Gulf', *The Middle East Journal*, 2001, 580–601.

is often no formalised leader selection process or statement of who is state leader in these cases, the leader in LEAD and Archigos is determined based on who had the most *de facto* political control.

Determining the counterfactual leader in self-selections is not as clear-cut as the other selection event types. On the one hand, we could consider the deposed leader as the counterfactual leader since they would have remained leader had the coup never occurred or failed. On the other hand, some coup leaders depose the incumbent but then appoint another individual to become leader, so it could be argued that the coup leader's alternative choice is the counterfactual leader. While there are good arguments for both options, I believe the best way to determine the counterfactual leader in self-selections is the following rule; if a coup leader installs themselves as leader, then the counterfactual leader is best considered their deposed predecessor. However, if the coup leader appoints another individual as leader based on any of the *other* selection event types, such as allowing the second-in-line to become leader or un-annulling election results, then the counterfactual leader is the third-in-line and runner-up in those cases, respectively.

Self-selections also include failed coups since these are potential leader transitions. In most of these cases, the counterfactual leader is the leader of the failed coup since it is clear they attempted to seize power for themselves. In some cases, however, the coup leader's primary aim is explicitly stated to install someone else as leader (usually a prior leader) and so the counterfactual leader is not the coup leader, but rather the individual who would have been installed had the coup been successful. I do not consider auto-coups since it is not clear there was a potential leader transition or who the counterfactual leader could be.

6.6 Sources

Elections were primarily determined using the National Elections Across Democracy and Autocracy (NELDA) dataset.[198] If there is evidence an election occurred, but is not in NELDA, then I included the election and used secondary sources to determine the counterfactual leader.

To my knowledge, there is no existing database of appointments. Fortunately, this only occurs once each time a leader is selected and so were generally easy to observe. However, in many cases there were few sources and it was not uncommon for the literature to use the term "appointed" and "elected" to describe the same leader transition. Furthermore, in cases farther back in time and in smaller states, the only sources were the news media. In short, finding information when leaders were appointed was very difficult. Nevertheless, most sources I used to determine appointments and the counterfactual leaders within them were from area-specific political science journal articles, although these tended to be limited to Europe, India, Israel and Japan.

Monarchical successions were similarly difficult to find information as appointments, although slightly less so. The heir-apparent is often explicitly stated due to norms or can be inferred through previous successions. However, constitutional successions were fairly easy to identify since nearly all cases could be determined due to being written into their constitutions.

Coups and coup attempts were taken directly from Cline Center Coup d'État Project Dataset, with only slight deviations if the coup targeted someone other than the leader as defined (i.e., a coup targeted the Presidency, but the leader is the Prime Minister).[199]

[198] Hyde and Marinov, 'Which Elections Can Be Lost?'
[199] Buddy Peyton et al., 'Cline Center Coup D'état Project Dataset', *Cline Center for Advanced Social Research* 2 (2020): 0.

6.7 Leader and Counterfactual Leader Traits

Traits are both the outcome and explanatory variables and main focus of this dissertation, so it is necessary to first explain how I operationalise the traits of both selected leaders and counterfactual leaders. Again, since I am interested in a possible selection effect for those traits that are associated with the initiation of international conflict, I limit the operationalisation of traits to only those so far discussed in this dissertation: age, military experience, war experience, gender, ideology, and leader hawkishness. However, it should be mentioned that I do not include Horowitz and Stam's leader risk score when testing my hypotheses in subsequent analyses. The leader risk score is constructed using international conflict initiations and hence correlated with my outcome already, so any statistical techniques would necessarily result in a positive association regardless.

Each leader and counterfactual leader trait are individually coded using the following strategy: First, I use the LEAD dataset to identify the relevant traits of each individual. However, I also re-checked each trait to confirm that the LEAD dataset is correct. If the LEAD dataset is inconsistent with a bibliographic source (e.g., Saddam Hussein having military experience when no evidence suggests he did), I use my coding instead of LEAD. For those individuals that are not in LEAD, I then use a variety of bibliographic sources to identify each trait of interest.

However, one potential source of bias is that there could be a substantial difference in the available information about leaders and counterfactual leaders between stronger and weaker states. For instance, we can probably expect that there will be significantly more biographical information on candidates in great powers like the United States than in smaller states like Cape Verde, simply because the political elites in the former have had more influence throughout history than those of the latter. As a result, I may incorrectly code candidates in weaker states as having no military experience because of a lack of information, which might distort any association. To remedy this possible source of bias, I include a four-level confidence scale for

every individual based on the amount of biographical information available. A confidence level of four indicates the individual is coded in LEAD, three indicates there is substantial information to indicate an individual does or does not have a given trait, whereas a confidence level of two indicates there is a reasonable amount of information to determine two or three traits, one indicates very little information beyond one or two, and zero indicates almost no information to identify any. These confidence scores aim to check for bias due to missing information.

6.7.1 Age

Age is measured simply as the birth year subtracted from the leader selection term year. A more precise measurement of age would include the day and month, but this is not considered due to difficulties in finding information on a significant number of individuals. This limitation is not an issue, however, as this would only result in a maximum mismeasurement of 364 days (e.g., selection term begins on January 1st, but the individual was born on December 31st), which is unlikely to affect the results in any meaningful way. In other words, when considering why older or younger candidates might be chosen, it is unlikely the selectorate would care about or even notice a one-year age difference between them.

6.7.2 Military Experience

Prior military experience is determined using the same criteria as Horowitz and Stam.[200] An individual is determined to have prior military experience if they have previously served in the formal military forces of any country (e.g., Israeli candidates serving in the Turkish or British militaries before 1949). Individuals who were conscripted into the military but chose to undergo

[200] Horowitz and Stam, 'How Prior Military Experience Influences the Future Militarized Behavior of Leaders'.

civil service, as is often common in Austria, are not considered to have prior military experience since that experience is not theorised to have the same socialisation effect as the military.

6.7.3 Military Experience Type

I also disaggregate military experience into two dummy variables depending on whether the individual has seen combat. Combat veterans are coded 1 if they have deployed to the location of any recognised international or internal war during their military service, and 0 if otherwise. Following previous research, I consider a recognised war as any conflict listed in the UCDP/PRIO Armed Conflict Dataset (v22.1).[201] Following Horowitz and Stam, I also do not code for whether an individual deployed to a war zone actively faced imminent death in combat since it is too difficult to determine among a larger number of cases.[202] Conversely, non-combat veterans are coded 1 if they have served in the military but never served in a war zone (e.g., Ronald Reagan) and 0 if they had no military experience. Lastly, non-ex-military leaders are those that have never served in the military.

6.7.4 War Win/Loss Experience

I further disaggregate combat experience into two binary variables depending on if the individual has seen combat in a war that was won or lost in the Dyadic MID Dataset (v3.0).[203] War-Win is a binary variable indicating 1 if they have fought in a war that was won and 0 if otherwise. War-loss is construed the same except for fighting in a lost war. There are obvious

[201] Shawn Davies, Therese Pettersson, and Magnus Öberg, 'Organized Violence 1989–2021 and Drone Warfare', *Journal of Peace Research* 59, no. 4 (2022): 593–610; Nils Petter Gleditsch et al., 'Armed Conflict 1946-2001: A New Dataset', *Journal of Peace Research* 39, no. 5 (2002): 615–37.
[202] Horowitz and Stam, 'How Prior Military Experience Influences the Future Militarized Behavior of Leaders', online appendix.
[203] Zeev Maoz et al., 'The Dyadic Militarized Interstate Disputes (MIDs) Dataset Version 3.0: Logic, Characteristics, and Comparisons to Alternative Datasets', *Journal of Conflict Resolution* 63, no. 3 (2019): 811–35.

problems with operationalising combat experience and war experience into two dummy variables, as it probably cannot capture the complexity of each experience. Nonetheless, the main benefit is to test whether selection effects explain the findings of previous research, and so constructing them using the same methods is suitable for this purpose.

6.7.5 Right-Wing Ideology

An individual's ideology is taken from the Heads of Government (HoG) dataset and is a trichotomous variable of left, centre, and right. Since the HoG dataset is limited to a select number of leaders, I do not attempt to insert my version of their political ideology if the individual is not included in HoG. While it would be ideal to have a more nuanced view of ideology than a left-centre-right trichotomy across all individuals, determining the ideological leaning of each individual is extremely research-intensive and beyond the aims and scope of this research.

6.7.6 Hawkishness

Hawkishness is determined for each individual using a simplified version of Carter and Smith's original measure.[204] Since the hawkishness score uses several measures that do not apply to many counterfactual leaders (e.g., irregular entry), I recalculate each individual's hawkishness score based only on a select number of indicators (all military-related variables, age and sex). This is a slight limitation but is necessary to ensure that all individuals' hawkishness scores are measured similarly without an imbalance of variables distorting the measure. Furthermore, since military-related variables largely influence the hawkishness score, my simplified measure should result in the same initial positive association with conflict initiation using the original measure.

[204] Carter and Smith, 'A Framework for Measuring Leaders' Willingness to Use Force'.

6.7.7 Conflict Initiation

Since we are ultimately interested in the extent to which traits cause leaders to start conflicts, the outcome in the selection effect hypotheses tests is the initiation of an international conflict. To reiterate, my definition of international conflict is synonymous with militarised interstate disputes, so they are constructed using the Militarised Interstate Disputes Dataset (v5.0).[205] However, as this variable is only included in the monadic analysis where each observation is the leader-spell-term-year, a more detailed explanation is needed regarding how it is more specifically constructed for use in the monadic dataset.

Since the fundamental purpose is to compare conflicts between leaders after selection events, ensuring that conflicts are included in the appropriate leader following a selection event is necessary. For example, even though Ronald Reagan was U.S. President between 1981 and 1989, a 1987 conflict initiated by the U.S. must only be included in the observations of the 1984 U.S. election, and not the 1980 election. Furthermore, since leaders often enter office months after they were selected in the selection event, it is necessary to ensure the selected leader initiated the conflict. For example, Ronald Reagan was elected on November 4th 1980 but did not enter office until January 20th 1981. Therefore, we need to ensure that any conflicts initiated by the U.S. between these two dates are included in the observations of Jimmy Carter, his predecessor.

To abide by these conditions, I create a leader spell and leader term for each observation in the monadic data. A leader spell is the entire time a leader was in office consecutively. A leader term is the period in between selection events within each leader spell. The year is self-explanatory. For example, Tage Erlander continuously led Sweden between 1946 and 1969. In his case, there is only one leader spell since he never left office, and he has eight separate terms

[205] Gochman and Maoz, 'Militarized Interstate Disputes, 1816-1976'.

because his party won eight straight elections. A different example is Belgium's Gaston Eyskens. He has three different leader spells since he entered office three separate times, but only one term in each spell as he was never re-elected. The start date for each leader spell is the date they first entered office and the end date is their last day in office. If a leader spell has only one term (e.g., Esykens above), then the start date and end are the same as the spell. For leaders selected in more than one selection event during the spell, the end date of the first term is the date of the selection event. The start date of the second term is the day after the selection event. To use Reagan again, Ronald Reagan has one spell from 21/01/1981 to 20/01/1989. Reagan has two terms. The first one starts on 21/01/1981 since he was elected on 04/11/1980 (the selection event date). Since he won re-election on 06/11/1984 – the date of the U.S. general election – the end date of the first term is the election date, and the start date of the second term is the day after, 07/11/1984. Since he only could serve two terms, the spell end date and the end date are the same (20/01/1989).

With the above rules in mind, international conflicts are constructed as a dichotomous variable indicating 1 if there was a conflict initiation in a given year during the leader term year and 0 if no conflict initiation occurred. For leaders who initiated more than one conflict in the same year during the same term, the observation for the leader is repeated for however many total conflicts was initiated. So, to use Reagan as an example again, he initiated five conflicts in 1983, so he has five observations coded as conflict initiation = 1 for that term year. He initiated only one in 1984 before the November election, so he has only one observation coded as conflict initiation = 1. However, he did not initiate any conflicts between November and December 1984, so he has one observation coded as conflict initiation = 0.

Following previous research, I only include conflicts if the state is the "originator" and "initiator" of the conflict, which is determined using the same criteria in the MIDB (v5.0) dataset.[206]

6.8 Supporting Hypotheses Explanatory Variables

Since I also test the outcome of candidates' traits and outcomes of selection events between two candidates given several explanatory variables, it is necessary to explain how each is constructed.

6.8.1 Past International Conflict Involvement

Several of my hypotheses predict certain traits should be more likely among candidates in "conflict-prone states." To test this, I use the total number of past conflict involvements within the ten years prior to the selection event. Like the outcome, conflict involvements are taken from the MIDB (v5.0) Dataset and are all conflicts in which the state participated.[207] The ten-year cut-off is ultimately arbitrary but is the most appropriate since international conflicts are rare and there is a trade-off between variation and temporal correlation. Conflicts within one year are probably more causally related to future conflict, but using this would include too many zeroes and underweight the likelihood of future conflict where it was extremely likely, such as the U.S. in 1940, which had zero conflict involvements within the previous year. On the other hand, a twenty-year cut-off would greatly increase variation in the number of conflict involvements, but many would be too distant in the past and overweight the probability of conflict where it was lower, such as in Iraq in 2004, which had thirty-nine conflicts within the

[206] This is determined by the variables *orig* = 1 and *sidea* = 1. See, Ibid.
[207] Ibid.

previous twenty years but only one in the previous ten years.[208] Thus, a ten-year cut-off is perhaps the most suitable compromise between increased variation and temporal correlation and is a reasonable proxy for the probability of conflict given the environment when leaders' enter office.

I also make no distinction as to whether a state was an "originator" (primary participant of the conflict) or "joiner", as doing so would necessarily leave out many important conflicts that can have huge implications on the future conflict behaviour of a state as well as the selection of leaders. If we were to exclude joiners, then the U.S. involvement in the Korean War, for example, would not be included since they joined on the side of the primary defender, South Korea.[209] This particular conflict, just a few years after the Second World War, fuelled the rivalry between the U.S. and China/North Korea and likely increased the U.S.'s probability of future conflict with these two countries in the future. Additionally, the Korean War might have also shaped the perceptions of US voters as to how they viewed their immediate security environment, possibly shifting their preferences for a leader they believe is better able to handle these two nuclear ambitious foes (and eventually capable in the case of China).

6.8.2 Veteran Population Proportion

The key demographic explanatory variable for the military experience traits is the number of veterans in the population. However, no census data exists across states on the number of veterans in the population, except in the United States. Nevertheless, the NMC Dataset (v6.0) provides the total population and military personnel for each year from 1816 to 2016.[210] Therefore, I use the NMC data to construct an estimated veteran population percentage since it

[208] In 2004, Iraq had initiated 39 conflicts in the past 20 years, but only 1 in the past 10 years. Most of these occurred during the Iran-Iraq war in the 1980s.
[209] Dispute #51 in the MIDB dataset.
[210] J. David Singer, Stuart Bremer, and John Stuckey, 'Capability Distribution, Uncertainty, and Major Power War, 1820-1965', *Peace, War, and Numbers* 19 (1972): 48.

is the only known data with this information across states. While our first instinct might be to divide the military personnel by the total population in a given year, this would not tell us how many are in the population when leaders are selected. If we used this method, for example, the estimated veteran population percentage for the United States in 1980 would be <1%., but we know from census data that 18% of all American adults were veterans in 1980 (37% of all adult males).[211] Therefore, I construct an estimated veteran percentage using a 40-year rolling average of the total military personnel divided by the total population. This measure is not entirely precise, but the general trend within countries and general differences between countries are at least reasonably accurate, given what we know about their past involvement in wars or conscription (Figure 6-1). Thus, the estimated veteran population percentage is accurate enough that positive or negative associations are sufficient evidence of the hypotheses rather than specific effect sizes.

[211] Gretchen Livingston, 'Profile of U.S. Veterans Is Changing Dramatically as Their Ranks Decline', *Pew Research Center* (blog), 11 November 2016, https://www.pewresearch.org/fact-tank/2016/11/11/profile-of-u-s-veterans-is-changing-dramatically-as-their-ranks-decline/.

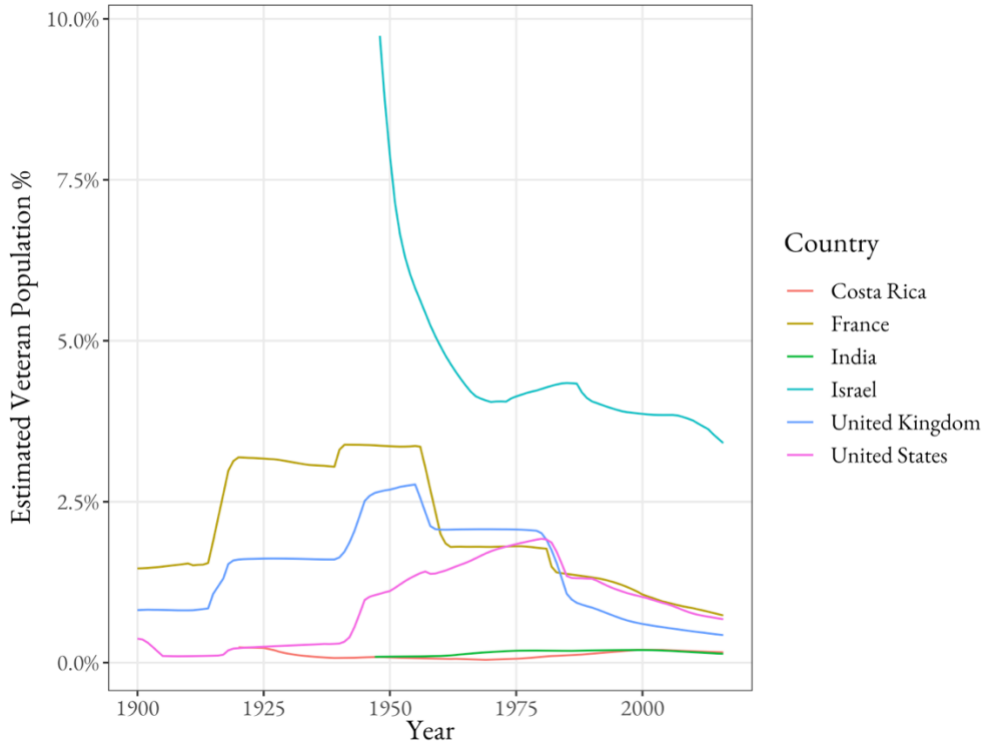


Figure 6-1: Example of the estimated veteran population percentage among selected countries.

6.8.3 Gender Equality

For testing the candidate selection hypotheses for female leaders, I use the relative level of gender equality of a state at the time of the selection event. This is taken directly from the historical gender equality index.[212] However, since a gender equality score of 70 would be considered on the lower end in 2003 but at the higher end in 1960, I standardize each score within the same year to account for relative differences in time. In this way, Saudi Arabia in 2000 will not be considered to have the same level of gender equality as Canada in 1950.

[212] Selin Dilli, Sarah G. Carmichael, and Auke Rijpma, 'Introducing the Historical Gender Equality Index', *Feminist Economics* 25, no. 1 (2019): 31–57.

6.8.4 Life Expectancy

Since the primary explanatory variable for the demographic hypotheses of age is life expectancy, I use the adult life expectancy for a state in a given year from the United Nations.[213]

6.8.5 Incumbency

One potential confounder for the age-related hypothesis is incumbency since incumbents are probably more likely to be older than challengers, so it is necessary to ensure that the candidate's age, rather than being an incumbent, is responsible for why older candidates might be elected over younger candidates. Incumbent is a binary variable indicating 1 if the candidate was incumbent in the election and 0 if otherwise.

6.8.6 Past War Win / War Loss

Past war-win and war-loss are binary variables of whether a state has won or lost a war, respectively, in the forty years before the selection event. The forty-year period accounts for the likelihood of a candidate with war experience being selected since leaders tend to enter office at 57 years, on average. These two variables are constructed using the Dyadic MID dataset (v4.02) and the same criteria for war-win and war-loss above.[214]

---

[213] 'World Population Prospects - Population Division - United Nations', United Nations, accessed 8 November 2022, https://population.un.org/wpp/Download/Standard/Population/.
[214] More specifically, I only include MIDs that are coded as a war (*hihost* = 5) and the outcome is either a victory (*outcome* = 1) or loss (*outcome* = 2). Gochman and Maoz, 'Militarized Interstate Disputes, 1816-1976'.

6.9 Data Analysis

6.10 Testing the Selection Effect Hypotheses

To test each selection effect hypothesis, I use quantitative methods following the counterfactual leader strategy similar to that described in Chapter Four. For each trait I follow the general pattern when testing the relevant hypotheses: First, I find the population treatment effect by regressing the leader trait on conflict initiation among the entire population of leaders. This ensures that the trait perspective hypothesis is initially supported in my data before I test for bias due to selection.

Second, I find the sample treatment effect by regressing the leader trait on conflict initiation only among observations where the counterfactual leader trait is observed. The primary purpose of this step is to ensure that the exclusion of missing counterfactual leaders does not distort inferences in relation to selection effects based on later steps. For example, if military experience was null or non-significant after excluding missing counterfactual leaders, then we would likely expect to find a null candidate or leader selection effects as well.

Third, I find the placebo treatment effect by regressing the counterfactual leader's trait on conflict initiation. This step confirms that there is evidence of bias due to selection as long as the placebo assumption is valid. For example, if the effect of ex-military counterfactual leaders is the same as that for ex-military leaders, then this is initial evidence of bias due to selection effects.

Fourth, I subset the data to only include non-self-selection events and find the candidate selection effect. This is the effect of the leader trait only when the counterfactual leader trait is the same, which specifically tests the candidate selection effect hypothesis. Next, I then find the leader selection effect which is the effect of the leader trait only when the counterfactual leader trait is the opposite, and specifically test the leader selection effect hypothesis. I then repeat these two steps, but only among self-selection events if hypothesised and sufficient observations

exist. A comparison between these candidate and leader selection effects reveals whether we have evidence to support any of the selection effect hypotheses.

As the outcome variable in each case is conflict initiation (as constructed above), I use a logistic regression for each step. Furthermore, since observations are clustered within individual leaders, I use clustered standard errors by the individual leader. To help visualise the selection effects, I also provide a predicted probability plot derived from the models.

The above is the general regression strategy for most of the traits analysed. However, I use slight deviations depending on the nature of the trait. For instance, since age and hawkishness are continuous variables, steps four and five are collapsed into one model with an interaction term. The selection effects are interpreted using values two standard deviations away from the mean for age, and the minimum and maximum values for hawkishness. Furthermore, I make slight adjustments due to other issues that arise during the analysis, which I explain in more detail in the respective section.

6.10.1 Testing the Placebo Assumption

Recall that the key assumption supporting any selection hypothesis is that the counterfactual leader trait must have a non-causal effect on conflict initiation. The most likely way this assumption can be violated is if the counterfactual leader was the previous leader. More specifically, if a given trait causes leaders to be more conflict-prone, then it could be that a previous leader with this trait started more conflicts, which increased the probability of conflict when the new leader took over. To ensure that the placebo assumption is not violated, I more specifically test the assumption for each trait if there is evidence for one or more selection effects. To do this, I regress an interaction term between the trait and whether the counterfactual leader was the previous leader. If this assumption is satisfied, then the effect of the counterfactual leader trait on conflict will not depend on whether they were the previous

leader. If there is a significant interaction term, then this assumption is violated and is evidence against the respective selection effect hypothesis, and evidence for the trait hypothesis.

The other way the assumption can be violated is if the counterfactual leader was given a prominent role in government immediately following the selection event. While this is a possibility, I do not test for this. First, it is very difficult to identify the post-selection position of many counterfactual leaders, let alone when they officially started said position. Second, even if we could identify their subsequent position and official start date, it is even more difficult to individually assess whether each counterfactual leader's position in government was meaningful enough to causally impact a leader's foreign policy. Lastly, it is far more likely that the first scenario above would violate the assumption since the counterfactual leader was the leader. Therefore, if the test for the first scenario validates the assumption, then it is unlikely that the traits of a runner-up-turned-minister could still have a causal effect on conflict.

6.10.2 Testing the Supporting Hypothesis

Following the selection effect and placebo assumption tests, I test whether there is evidence for the corresponding supporting hypotheses. Again, the supporting hypotheses are those that more specifically predict why we might see a candidate or selection effect, and so are used to strengthen the respective selection effect hypotheses. To do this, I use the following general method: first, I test whether any explanatory variables are positively associated with the trait among candidates. As the outcome is the trait, I convert the LSD dataset into long format where each observation is the candidate in the selection event. Depending on the nature of the trait (i.e., binary or continuous), I either use a logistic or OLS regression with robust standard errors clustered within individuals to ensure the same individuals are not driving the results. I also test these same variables using robust standard errors within countries to determine if a concentration of candidates with the trait within countries drives the results.

Next, I subset the LSD dataset to only include elections and appointments where the two candidates had opposing traits. This step aims to test any hypotheses that predict high-risk candidates will be chosen over low-risk candidates. However, there is a trade-off between internal and external validity when considering how best to test these hypotheses. On the one hand, there is likely to be a low number of appointments and elections where the winner and runner-up had opposing traits, so it is worth including as many cases as possible to ensure external validity. On the other hand, any choice might not reflect the selectorate preferences in autocracies or rigged elections, which could jeopardise internal validity. For instance, if elections are rigged then obviously any voter preferences are not revealed beyond what the incumbent regime has allowed. Therefore, I try to strike a reasonable balance between these two problems by limiting cases based on the following conditions:

1. For appointments, the polity score must be above 0 and must not have been determined by a foreign Head of State.
2. For elections, the election must be considered free and fair in the NELDA dataset. To determine this criterion, an election must meet four conditions from NELDA; 1) opposition parties were allowed; 2) there was a choice of candidates on the ballot; 3) there was a clear prior indication an incumbent would give up power; and 4) the results were not cancelled.[215] More restrictive criteria will be used for robustness checks if there is any evidence for an FIR hypothesis, such as excluding elections where opposition members were harassed, or whether a foreign state interfered in the election. However, I do not consider these in my initial tests due to their highly subjective nature.
3. For elections that are not included in NELDA, the polity score must be above 0 and the runner-up must not have been the leader of the party with the most seats (for cases

[215] In the NELDA dataset, these are variables NELDA3, NELDA5, NELDA7, and NELDA 32, respectively.

when the selectorate's true choice in the election was not observed due to failing to gain a majority or coalitionary bargaining).

The above is the general approach I use for each trait, although I may deviate slightly to account for things not previously considered. Data was collated and coded using Excel and analysed using R statistical software.

7 Findings

To what extent do the personal traits of leaders influence their decision to start international conflicts? While previous research suggests traits matter greatly, I theorised earlier in the thesis that selection likely explains these findings. In short, "high-risk" leaders are not risky, just selected into high-risk environments, and so the relationship between traits and conflict initiation is likely spurious. However, given the problems with unobservable variables, we cannot get a precise statistical estimate of leader traits on conflict while controlling for selection. Therefore, the point of my data analysis is to assess the totality of evidence and reduce our uncertainty about the question of whether traits matter given selection. Thus, the questions this chapter seeks to address are not "What is the point estimate of military experience on conflict" but rather more general questions about likely pathways of causation that quantitative analysis can help to answer. Are effects in the expected direction as the hypotheses? Do relationships weaken or get stronger as predicted? How close or far apart are the confidence intervals? I seek to answer these types of questions using the available data and my new dataset.

The organisation of this chapter is as follows: First, I provide an overview of the data and summary statistics within the data. Second, I then investigate each high-risk trait independently, one by one, to assess whether the results support the central argument that selection effects likely explain why leaders with the high-risk value of the trait appear more likely to initiate conflict than those with the low-risk value. To do this, I first test the selection effect hypotheses to investigate whether the effect of each trait is potentially biased due to candidate, leader, or self-selection effects. If the evidence supports one of the selection effect hypotheses, I test the placebo assumption to ensure these findings are internally valid. Lastly, I test the supporting hypotheses to determine whether the causal logic of the selection effect hypothesis is plausible (i.e., that the selection effect reveals bias and not random noise). Ultimately, I find evidence for the candidate selection effect hypothesis for military experience, non-combat experience, war-

win experience, age, and hawkishness. Furthermore, the candidate selection effect we see for all the military-related traits is likely explained by demographics – in each case, the estimated share of veterans in the population strongly predicts whether a candidate will be ex-military, non-combat, and a war winner, and the estimated share of veterans in the population when leaders enter office predicts conflict initiations during their term. I find a similar candidate selection effect for age, although the supporting demographic hypothesis is weaker since states with increased life expectancies are not more likely to initiate conflict. However, I find no evidence for either candidate or leader selection effect hypotheses for female leaders and right-wing leaders, although the findings for these two traits are limited due to small sample sizes. Lastly, I find no evidence among any of the traits that selectorates are choosing leaders with high-risk traits over those with low-risk traits as past conflict increases, contrary to the FIR supporting hypotheses.

7.1 A Note on Datasets Used

First, I must note that I use different datasets (based on my original data collection sample) separately depending on the variables of interest in subsequent analyses. First is the Leader Selection Dataset (LSD) dataset, in which every observation is a single leader selection event. The LSD data is used to provide descriptive statistics about the distributions of leader and counterfactual leader traits and selection events and when testing supporting hypotheses as to whether candidates with certain traits are more likely to be chosen. Second is the LSD – Monadic dataset, in which every observation is a single year within the leader selection term. I use the LSD – Monadic dataset only when testing the selection effect hypotheses. The third dataset is the LSD – candidate dataset, which is essentially the LSD dataset converted into a long format whereby every observation is the individual candidate in the selection event. I use the LSD – candidate dataset primarily when testing the supporting hypotheses for candidate selection effects since the outcome is the candidates' traits.

7.2 LSD Summary Statistics

Since these are observational data and there are likely numerous ways in which bias can potentially manifest and distort the findings, it is necessary to first provide an overview of these data to identify any issues or potential problems that may arise when testing the various hypotheses in subsequent analysis.

There are a total of 2,680 leader selection events. Of these, I could not determine the counterfactual leader in 971 cases (36.2%) due to inability to identify the individual in the literature. Across each of the selection mechanisms, missing counterfactual leaders are primarily a problem in the appointment selection mechanism (Table 7-1). However, this is not unexpected considering that Heads of State rarely explicitly acknowledge the individual they almost chose but did not. Most appointments in which the counterfactual leader could be determined were identifiable because they were the first to be asked by the Head of State but declined, or there was a general agreement in the literature of who the Head of State also considered. For elections, the majority of missing counterfactual leaders is due to either unopposed elections or unique cases that are too difficult to determine.[216] The large number of missing counterfactual leaders in appointments and successions should not distort results, however, since there is little reason to believe that these two selection mechanisms would primarily drive the effect of any of the traits. In any case, this potential missing information bias will be directly checked in subsequent analyses since I include the coefficient for all leaders and the coefficient for only those leaders where the counterfactual leader trait is observed. If there is a significant difference between the two, then we know that the missing counterfactual leader is likely causing this bias, and adjustments may need to be considered.

[216] For instance, Recep Erdogan's party was elected in 2002, but even though he was party leader he was not allowed to become Prime Minister as he was technically banned from politics. He eventually become Prime Minister in 2003 after a province by-election where he became MP and hence Prime Minister.

Table 7-1: Number of observed and unobserved counterfactual leaders in selection terms by selection mechanism

		Selection Mechanism					
		Elected	Appointed	Succession	Self-Selected	Unknown	Total
Counterfactual Leaders	Observed	1129 (76%)	48 (12%)	117 (57%)	413 (59%)	14 (19%)	1709 (64%)
	Unobserved	360 (24%)	351 (88%)	89 (43%)	83 (28%)	24 (32%)	972 (36%)

Table 7-2 below lists the number of observations and percentages [217] for each selection mechanism across regime types. Unsurprisingly, the majority of elections occur in democracies, and the majority of self-selections occur in autocracies. Part of the inconsistency in some cases (e.g., self-selection occurring in a democracy) is because many dates in the Polity IV dataset where a democracy to autocracy transition (and vice versa) occurs on either the day before or after the selection date, but is too difficult to ascertain whether the selection followed or preceded the change.

Table 7-2: Number of observations and percentage of each selection mechanism by regime type

		Selection Mechanism				
		Elected	Appointed	Succession	Self-Selected	Unknown
Regime Type	Autocracy	274 (20%)	93 (26%)	52 (26%)	287 (59%)	14 (19%)
	Anocracy	253 (18%)	99 (28%)	42 (21%)	138 (28%)	24 (32%)
	Democracy	844 (62%)	162 (46%)	107 (53%)	61 (13%)	37 (49%)

[217] Based on Polity IV data. Democracy is coded if the polity score is 5 or greater and autocracy is coded if the score is less than -5. Anocracies are between -4 and 4.

7.3 Leaders and Counterfactual Leaders: Summary Statistics

There are 1,738 leader "spells" (e.g., Ronald Reagan 1981-1989) and 2,680 leader selection "terms" (e.g., Ronald Reagan 1981-1984 and Ronald Reagan 1984-1989). There are 1,511 individual leaders, as 40% (n=611) had multiple selection terms. The median number of leader selection terms for each leader spell is one with a median length of 2.09 years in office. The top-five leaders with the most selection terms are Alfredo Matiauda (12), Kim Il-Sung (9), Souvanna Phouma (9), Josip Tito (8), and Tage Erlander (8). However, several leaders (Il-Sung, Phouma, and Tito) have multiple terms due to unopposed elections or failed coups. When excluding unopposed elections or failed coups, the top-five leaders with the most selection terms are Tage Erlander (8), Alfredo Matiauda (7), Robert Menzies (7), Lyden Pindling (6), and Dawda Jawara (6).

For counterfactual leaders, there are a total of 1,708, of which 1,336 were unique individuals since 22% (n=373) were counterfactual leaders in more than one selection event. Contrary to leaders, individuals with more than one appearance as a counterfactual leader are mostly in democracies since those who attempt failed coups tend to be jailed or killed afterwards. Accordingly, the top-five individuals with the most counterfactual leader appearances were runners-up multiple times in parliamentary elections: Menachim Begin (6), Shimon Peres (6), Lim Siang (6), Carl Romme (5), and Joop Den Uyl (5).

7.3.1 Geographic Statistics

Leader selection events are relatively balanced geographically (Figure 7-1).[218] The plural majority are in Europe (26%) due to the higher number of parliamentary democracies, and hence the relative frequency of elections and leader turnover occur compared to the other selection

[218] The percentage of leader selection events by region: Europe (26%), Americas (24%), Asia (24%), Africa (23%), Oceania (4%)

mechanisms. The average number of leader selection events per country is 14.57 (standard deviation = 10.28), with Switzerland having the most (n=60) due to their yearly "elections" and five countries only having one leader selection event.[219] Therefore, the results are unlikely to be biased due to spatial autocorrelation.

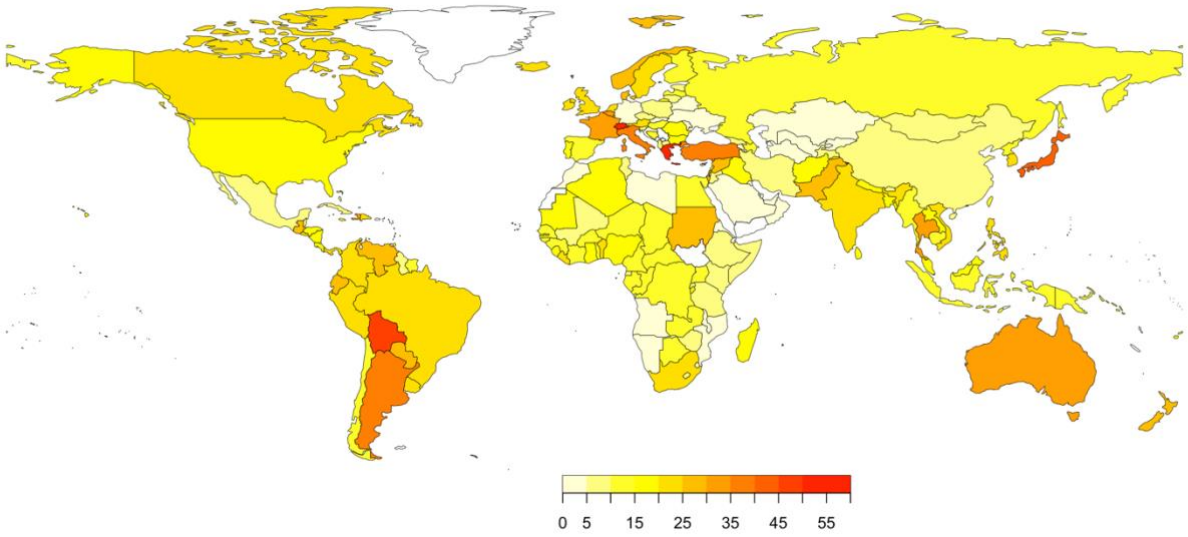


Figure 7-1: Heatmap of leader selection events by country

The distribution of leader selection events and mechanisms is relatively balanced over time, though there is an increase in elections during the 1990s, which is expected due to the sudden rise in democracies after the fall of the U.S.S.R. (Figure 7-2). Conversely, self-selections decrease over time for presumably the same reason. Thus, bias due to temporal autocorrelation is not anticipated.

[219] Brunei, Oman, Tibet, Eritrea, East Timor. The former three have one because they are monarchical systems and had one leader change during the entire 60-year period. Eritrea achieved independence in 1993 and has had one leader since, whereas East Timor achieved independence in 2002 and has had only one leader due to the short time period.

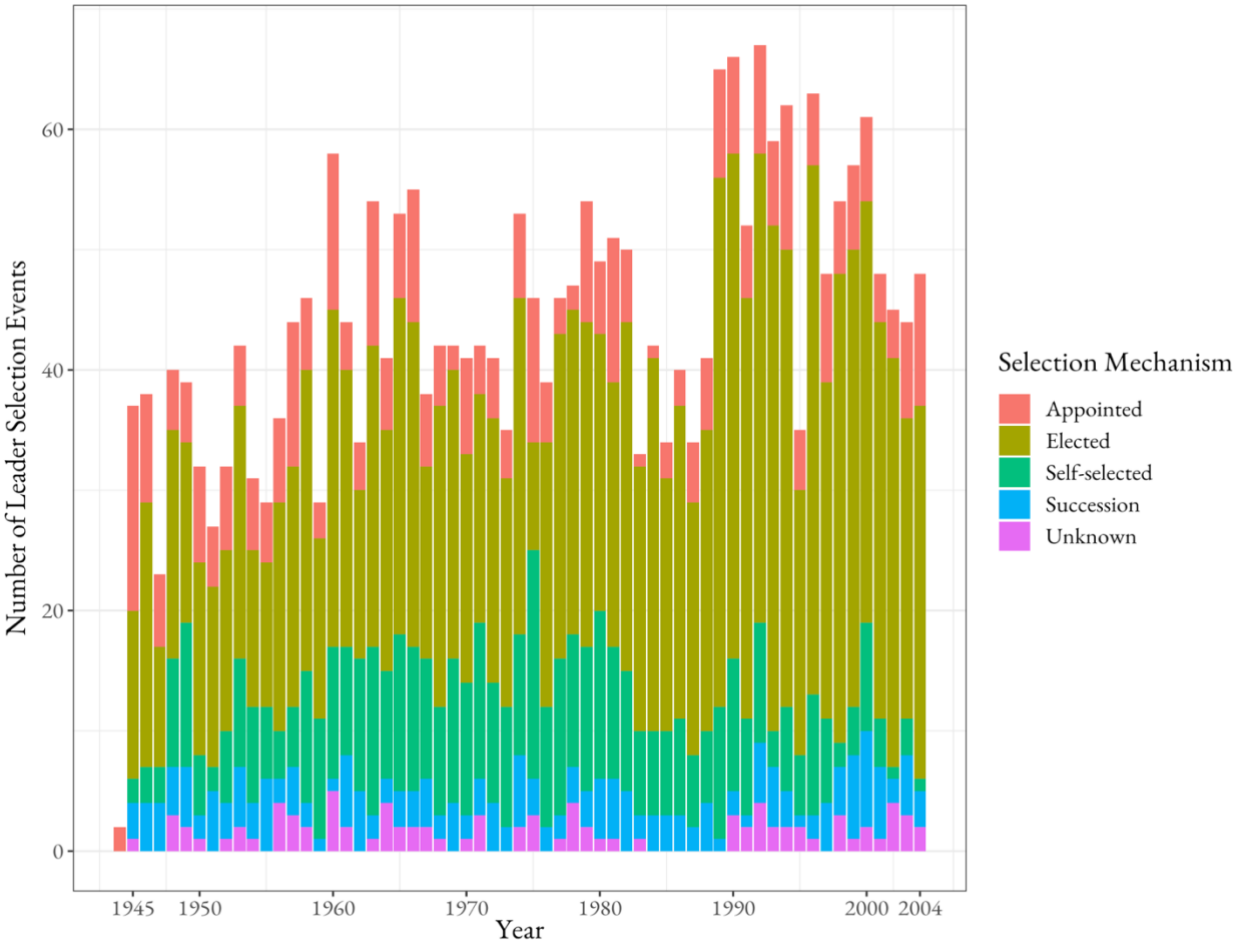


Figure 7-2: Number of selection events and mechanisms by year

7.3.2 Selection Mechanism Statistics

Figure 7-3 below breaks down the percentage of leader selection events by mechanisms and the sub-mechanisms (e.g., parliamentary or presidential). The majority of elections were parliamentary (43%), followed by presidential (38%), central government elections (16%), party leadership elections (3%), and 2 were an unknown election type.[220] For parliamentary elections, 15% (n=98) were those in which the party leader/candidate with the most seats did not become leader due to not gaining a majority, coalitionary bargaining, failing to form a government, or the Head of State ultimately appointed someone else. There were no presidential elections in which

[220] For instance, some sources indicate that a leader was "elected", but there was insufficient information to ascertain the type of election.

the individual with the most votes (electoral when used) did not become leader. For appointments, the Head of State accounted for 52% of cases, followed by the central government (35%), a foreign Head of State (9%), the predecessor (5%), and the rest were too difficult to determine due to lack of information. For self-selections, 53% were failed coup attempts, 41% were successful coups, whereas the rest entered office through consolidation of power, war, or revolutions. For successions, 49% were constitutional, followed by normative (35%), hereditary (15%), and rotation (1%).

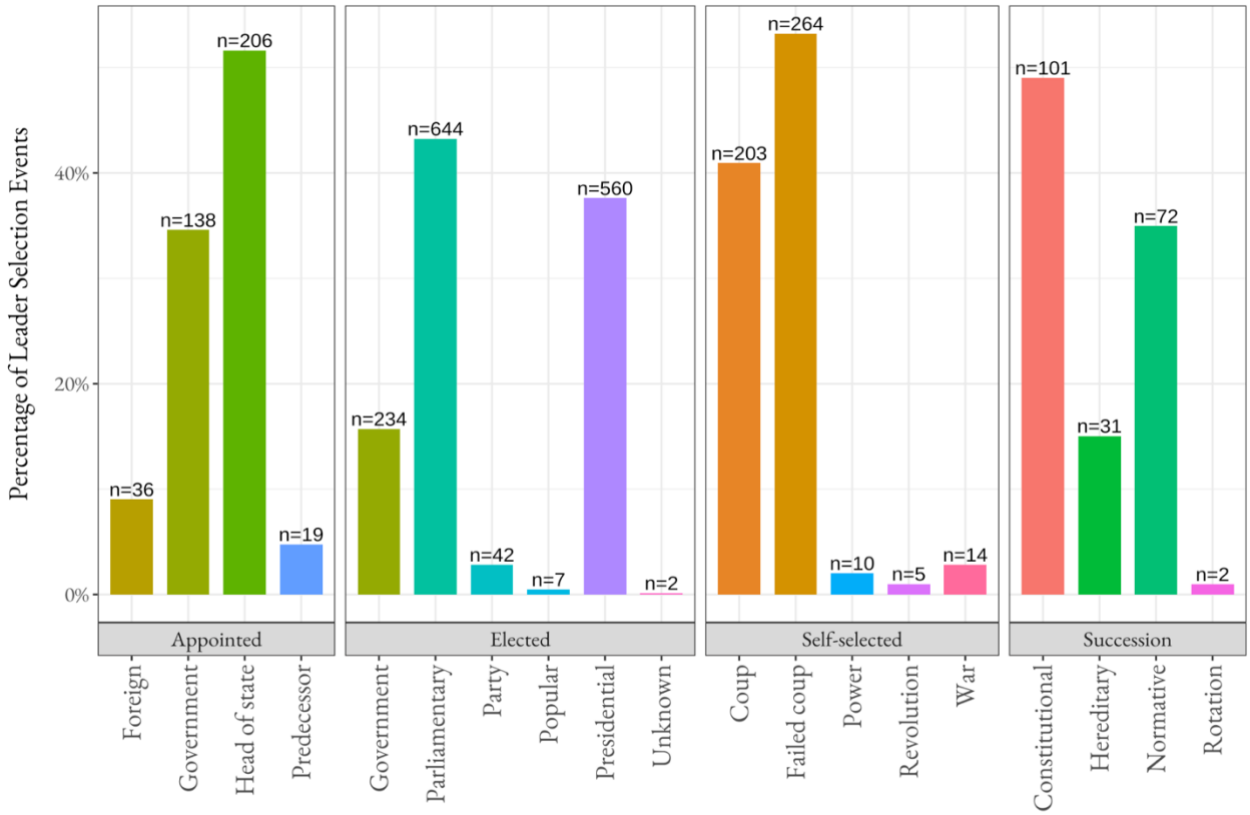


Figure 7-3: Percentage of selection events by selection mechanism

7.3.3 Conflict Initiation Statistics

There are a total of 1,507 conflict initiations across all leader selection terms. While the monadic analyses use a binary measure of conflict initiation, the distribution of the total conflict initiations per leader selection term is heavily right-skewed, consistent with other research using

militarised interstate dispute initiation as the outcome (Figure 7-4). Among all leader selection terms, 77% did not result in a single conflict initiation. The top-five leaders with the most conflict initiations were Ayatollah Khomeini (61), Saddam Hussein (42), Mao Zedong (36), Leonid Brezhnev (29) and Nikita Khrushchev (26). 46 countries did not initiate a conflict between 1945 and 2004. Thus, clustered standard errors within individual leaders are required when regressing the leader and counterfactual leader traits on conflict, as anticipated.

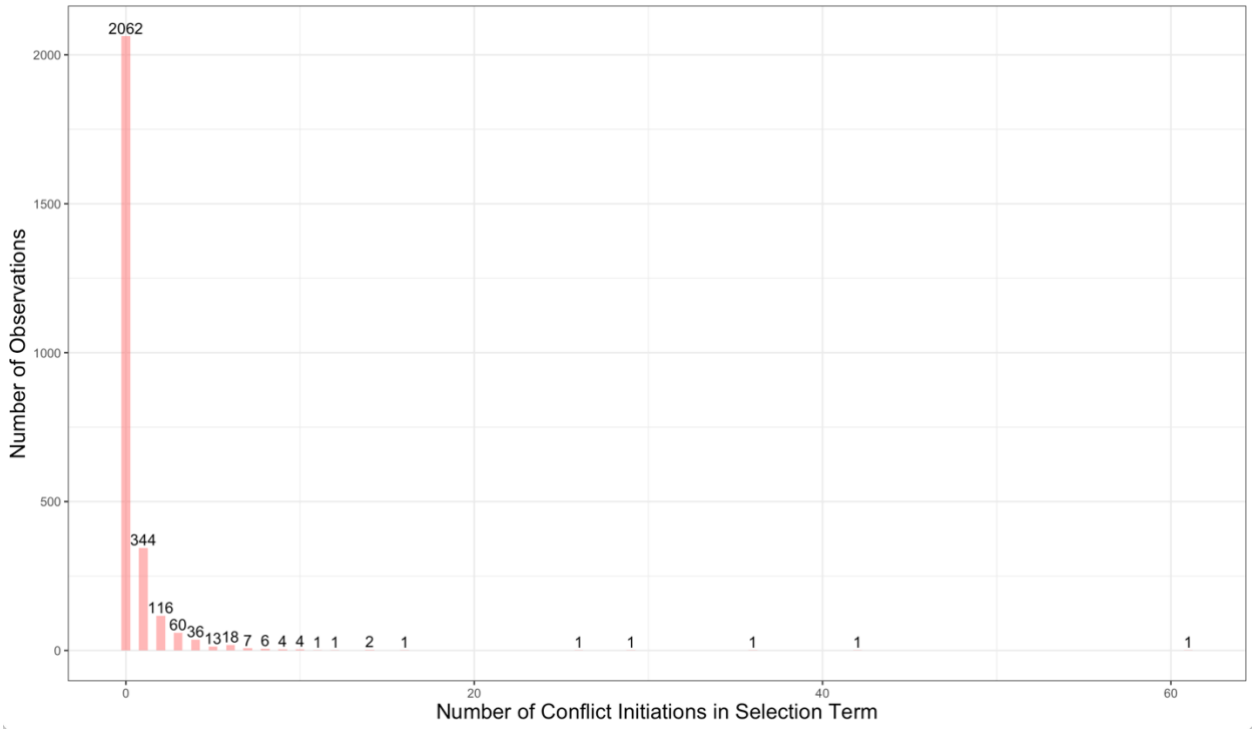


Figure 7-4: Histogram of conflict initiations per leader selection term.

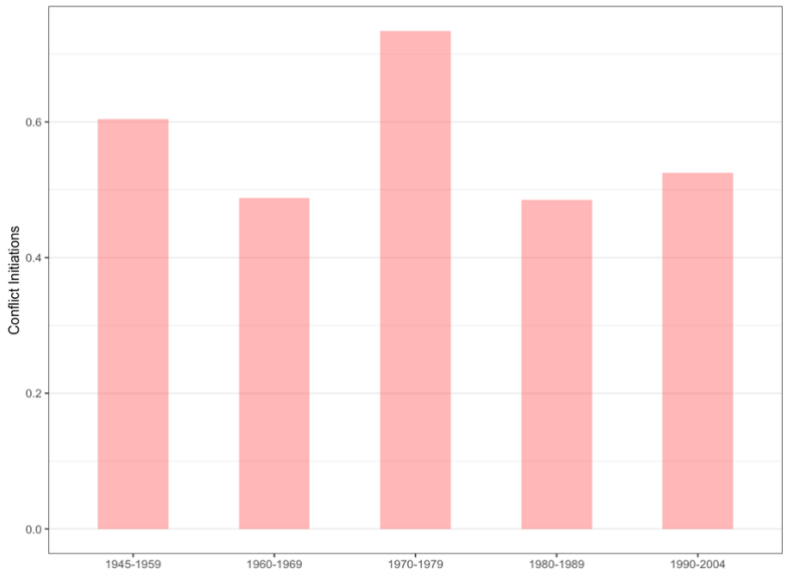


Figure 7-5: Average number of conflict initiations per leader selection term by decade.

While the number of conflict initiations slightly increases over time, the average number of conflict initiations per leader selection term remains relatively the same (Figure 7-5). Therefore, there is no reason to suspect temporal confounding.

When regressing the selection mechanisms on conflict initiation in the monadic data, there are significant differences in the probability of conflict initiation (Figure 7-6). Leaders who have self-selected into office are the most likely to initiate conflict, whereas leaders who have entered office through succession are the least likely.[221] Because self-selections are much more likely to initiate conflict than other mechanisms and could be correlated with traits such as age and military experience, testing for the self-selection effect hypotheses in later analyses is important. Why successions are less likely to initiate conflict than elections and appointments, however, could be because those who enter office through constitutional successions are usually temporary figures until the next election or appointment, and so might have limited powers or more constraints to their decision-making capabilities. In any case, there is no theoretical reason

[221] There is no significant difference in conflict initiations between constitutional and hereditary selection sub-types. The constitutional logit coefficient is 0.07 and 95% confidence interval is between 0.04 and 0.12. The hereditary logit coefficient is 0.05 and 95% confidence interval is between 0.03, and 0.10.

why we would expect successions to confound leader traits or selection effects in later analyses to the same degree of self-selection, and so there is no reason to adjust for this particular selection mechanism when testing the various hypotheses.

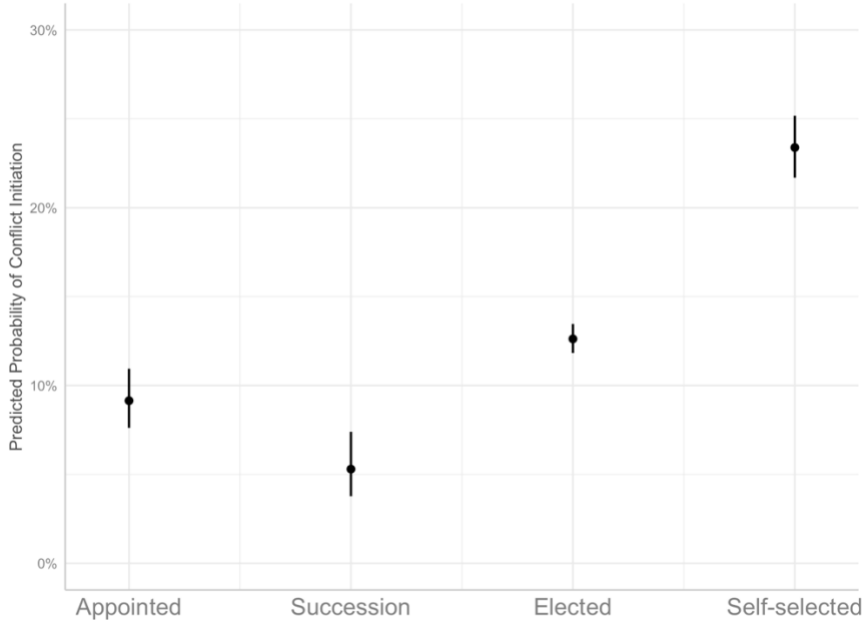


Figure 7-6: Predicted probability of conflict initiation by selection mechanism.

7.4 A Note on Information Bias

As noted in the methodology section, a key concern is bias due to missing information on individuals in more conflict-averse states. Of the 2,773 unique individuals in the LSD – Candidate dataset, 68% (n=1543) had extensive information from the LEAD dataset or in the form of biographies or other detailed life histories, 14% (n=324) had a moderate amount of sources such as brief biographies in news or journal articles to determine all traits, 7% (n=167) had very limited sources where it was only possible to determine two or three traits, and 3% (n=67) had little to no sources of information where it was only possible to determine one or two traits at most. Thus, potentially miscoded traits due to a lack of information are only a possible concern for a relatively small number of individuals, and so it is not expected to significantly alter the findings.

In the next sections, I begin the main analysis for each trait. To reiterate from the Methodology chapter, I follow the same general strategy for each. First, I provide more specific summary statistics for the trait of interest to identify any potential problems or sources of bias. I then test the selection effect hypotheses using the LSD – Monadic dataset. If any selection effect hypotheses are supported, I test the placebo assumption using the same data. Lastly, I then test the supporting hypotheses. To test the supporting hypotheses for any potential candidate selection effects, I use the LSD – Candidate data, where the outcome is the trait of every leader and counterfactual leader. To test the supporting hypotheses for potential leader selection effects, I use a subset of the LSD data that only includes elections and appointments where selectorates chose between the leader and counterfactual leader.

7.5 Prior Military Experience

7.5.1 Military Experience: Summary Statistics

In total, 32% of leaders and 29% of counterfactual leaders had prior military experience. Only 7% (n=118) of the 1,709 counterfactual leaders lacked information to discern whether they had been in the military, most of which were in Caribbean states. The number and proportion of ex-military candidates generally decline over time, which is expected due to the elapse of time since the Second World War, when a large proportion of many countries' populations fought, except for a brief increase in Africa during the 1970s (Figure 7-7).

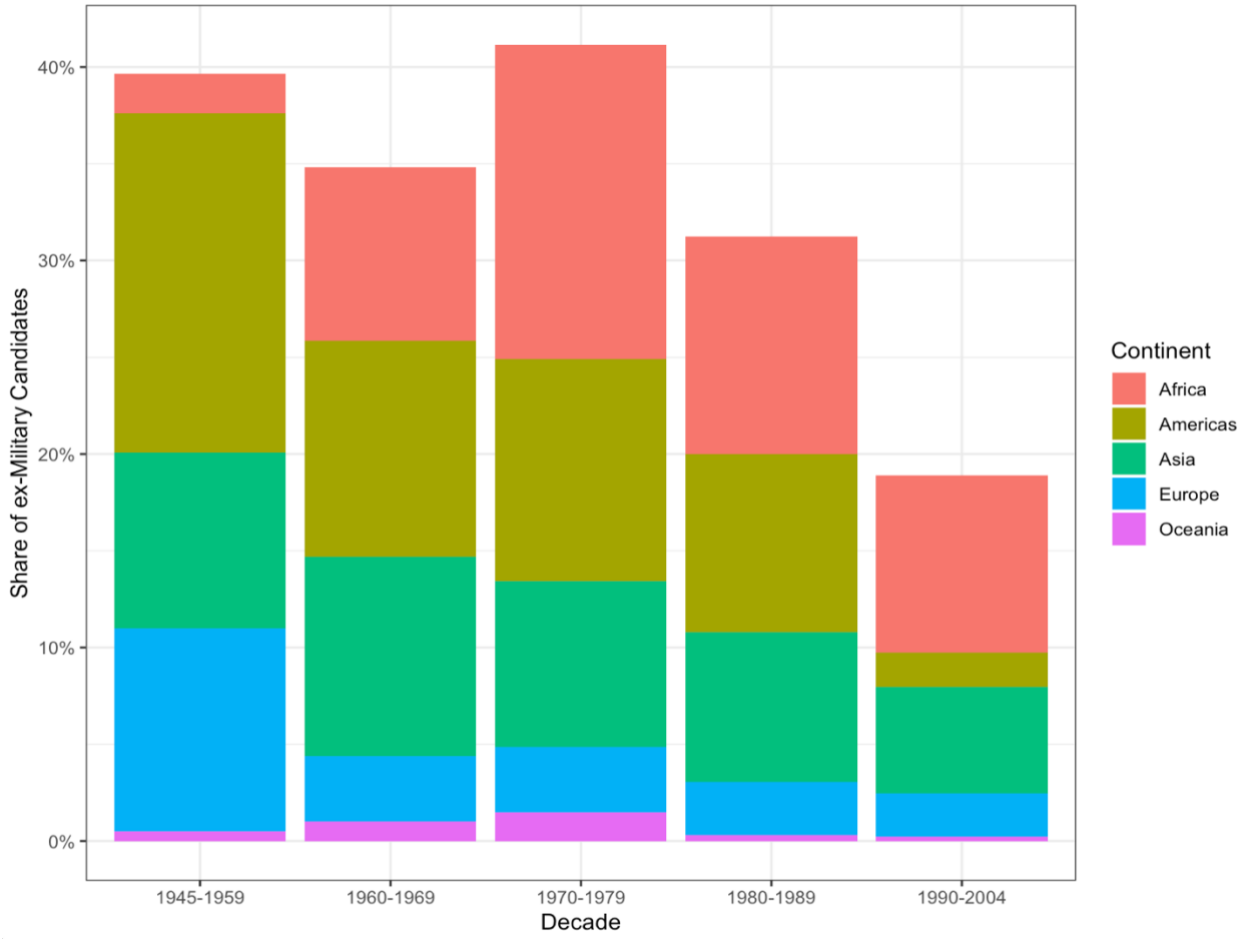


Figure 7-7: Percentage of leaders and counterfactual leaders with prior military experience by decade and continent

Most ex-military leaders and counterfactual leaders appeared in either self-selections or elections (43% each). Among self-selections, veterans were 88% of leaders who self-selected but were counterfactual leaders in 75% of failed coups. For self-selections in countries with at least ten leader selection terms, the highest percentage of ex-military candidates were in El Salvador (93%), Mauritania (93%), Iraq (91%), Burkina Faso (83%), and Ghana (82%). For elections, veterans were the winners in 13% and the runners-up in 9%. Additionally, in electoral contests where the top-two candidates are a veteran and non-veteran, veterans are more likely to win than lose (58%), a finding consistent with veteran bonus theory. The highest percentage of ex-military electoral candidates in countries with at least ten selection terms were in the United States (87%), France (77%), Israel (77%), Paraguay (53%), and Guatemala (53%).

7.5.2 Military Experience: Selection Effect Hypotheses Tests

To reiterate, I use the counterfactual leader method to test whether we see evidence for the selection effect hypotheses. All three selection effects have been theorised to potentially explain the positive relationship between ex-military leaders and conflict initiation. Table 7-3 below provides the hypotheses for reference.

Table 7-3: Summary of military experience selection effect hypotheses

Selection Effect Hypotheses	**Treatment Group**		**Comparison Group**		**Supporting Hypotheses**
	Leader Military Experience	**Counterfactual Leader Military Experience**	**Leader Military Experience**	**Counterfactual Leader Military Experience**	
Candidate Selection Effect	Yes	Yes	No	No	FIR, Demographics
Leader Selection Effect	Yes	No	No	Yes	FIR, Affinity Bias
Candidate Self-Selection Effect	Yes	Yes	No	No	Attempt
Leader Self-Selection Effect	Yes	No	No	Yes	Success

7.5.3 Military Experience: Main Findings

Table 7-4: Military experience selection effect regression results.

				Self-Selected = No		Self-Selected = Yes	
	Population Treatment Effect	Sample Treatment Effect	Placebo Treatment Effect	Candidate Selection Effect	Leader Selection Effect	Candidate Selection Effect	Leader Selection Effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Military Experience (Leader)	0.69*** (0.35, 1.02)	0.65** (0.23, 1.07)		1.55*** (1.07, 2.04)	-0.16 (-0.89, 0.56)	0.97 (-0.11, 2.05)	-1.56* (-2.82, -0.30)
Military Experience (Counterfactual Leader)			1.25*** (0.85, 1.64)				
Constant	-2.10*** (-2.34, -1.86)	-1.88*** (-2.19, -1.57)	-2.11*** (-2.29, -1.93)	-2.21*** (-2.44, -1.98)	-1.81*** (-2.43, -1.20)	-1.72*** (-2.63, -0.82)	-0.33 (-1.53, 0.86)
Observations	10,680	6,557	6,565	3,630	1,127	949	851
Log Likelihood	-4,233.85	-2,884.82	-2,769.28	-1,314.39	-434.06	-537.76	-428.43
Akaike Inf. Crit.	8,471.70	5,773.64	5,542.56	2,632.77	872.12	1,079.52	860.86

Note: Robust standard errors clustered by individual leader *p<.05; **p<.01; ***p<0.001

Table 7-4 above shows the regression results for military experience. For each model, logistic regression with robust clustered standard errors by the individual leader is used. Moving from left to right, we can see that the population treatment effect (model 1) and sample treatment effect (model 2) of prior military experience on conflict initiation are not significantly different, confirming that the exclusion of missing counterfactual leaders should not bias the relationship between leaders' military experience and conflict initiation in subsequent models. In model 3, the placebo treatment effect (the effect of counterfactual leaders' military experience on conflict initiation) is also positive and significant, and while the effect size is twice as large as the sample treatment effect, the confidence intervals of these two measures overlap, meaning the difference between the treatments is not statistically significant (i.e., sample and placebo treatments). This suggests a selection effect for leaders with military experience in environments where conflict is more likely.

Models 4 and 5 test each selection effect hypothesis more specifically for leaders among non-self-selections. Model 4 shows that the probability of conflict is nearly three and a half times greater when both candidates are veterans compared to when both are non-veterans, which is consistent with the candidate selection effect hypothesis. Conversely, model 5 shows that when leaders and counterfactual leaders have opposing traits, there is no significant

difference in conflict initiations, contrary to the leader selection effect hypothesis. Among self-selected leaders, however, we see the opposite findings. Model 6 shows no candidate self-selection effect, and Model 7 shows the opposite leader self-selection effect. Thus, no evidence exists for the candidate or leader self-selection effect hypothesis.

Figure 7-8 below better visualises the candidate and leader selection effects in the above tables in full. Starting with non-self-selected leaders (left facet), we can see the effect of military experience on conflict initiation is entirely driven by a candidate selection effect. Since there is no significant difference in the effect of military experience when the military experience of the counterfactual is not the same, there is no evidence for the leader selection effect hypothesis. Among self-selected leaders, we see a similar candidate selection effect though the confidence intervals are wider and overlap due to the smaller sample size; there are 3,179 observations of non-veterans in the candidate selection group, but only 284 observations of non-veterans in the candidate selection group. Conversely, the leader self-selection effect is in the opposite direction – conflict is significantly more likely to occur when non-veterans remove veterans from power than the opposite. In essence, the results thus far support the candidate selection effect hypothesis.

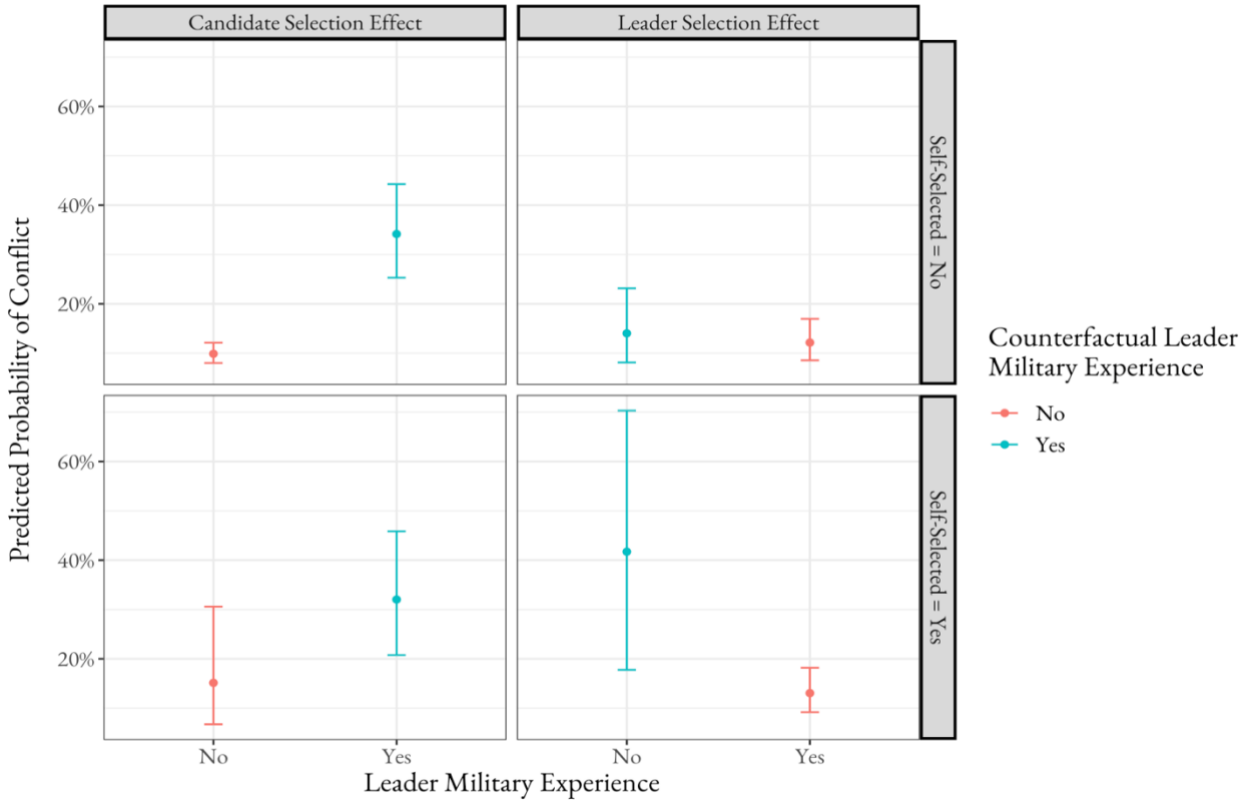


Figure 7-8: Predicted probability of conflict by leader military experience selection effects

7.5.4 Military Experience: Placebo Assumption Test

Recall that the key assumption for the selection effect models above is that a counterfactual leader's military experience cannot be associated with conflict initiations except through selection. That is, there is still a possibility that the candidate selection effect we see above is simply because ex-military counterfactual leaders were previous leaders, and so caused more conflicts which created a dangerous environment for the future leader. To test for this, I include an interaction term between counterfactual leaders' military experience and a binary variable indicating whether the counterfactual leader was the previous leader in the same data used in model 4. The results visualised in Figure 7-9 confirm that the counterfactual leader's military experience on conflict initiation does not depend on being a previous leader. In other words, the counterfactual leader's military experience is a valid placebo and is likely not causing

the candidate selection effect we see among non-self-selected leaders in Figure 7-8 above. Thus, we have further evidence to support the candidate selection effect hypothesis.

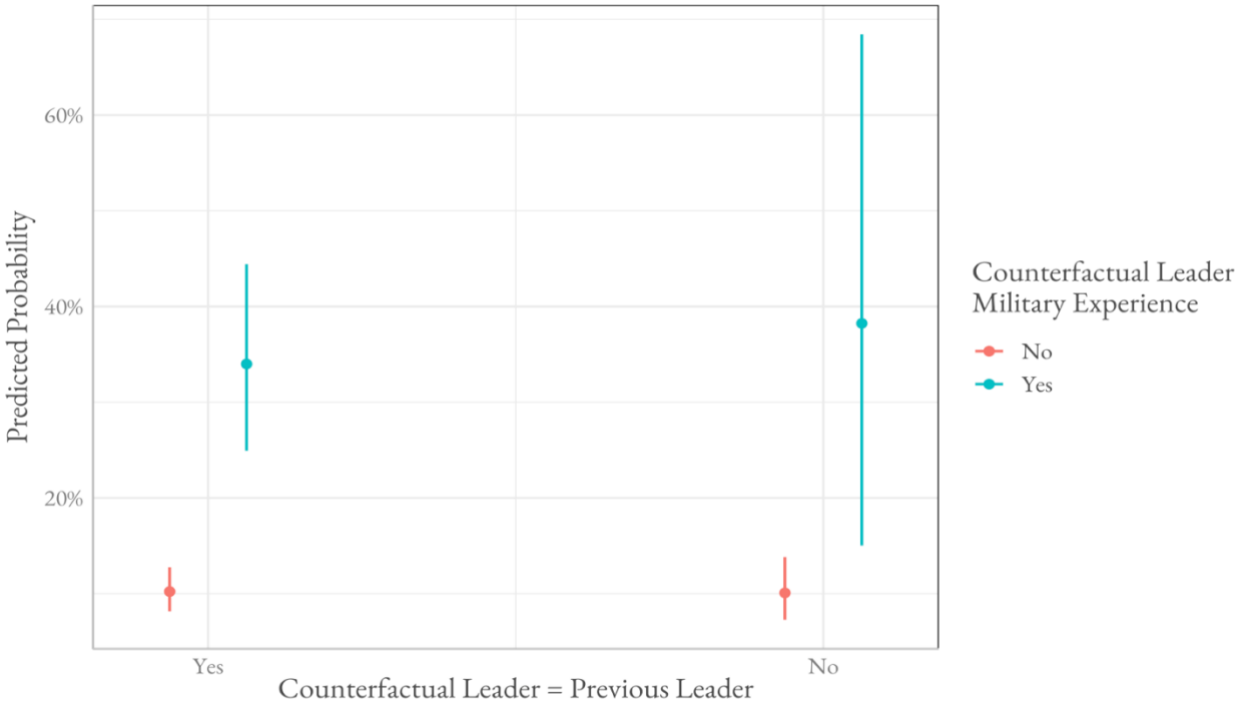


Figure 7-9: Effect of counterfactual leader's military experience as an interaction with whether they were the previous leader before the selection event.
Note: Same data used in Model 4 of Table 7-4.

7.5.5 Military Experience: Supporting Hypotheses Tests

While the above findings so far support the candidate selection effect hypothesis, we still do not know exactly why ex-military leaders are more likely to initiate conflict than non-ex-military leaders only when the counterfactual leader has the same military experience. Again, this could be due to random noise or unexpected factors. Therefore, I more specifically test whether the demographic or FIR supporting hypotheses explain this candidate selection effect. To do this, I use the LSD - Candidate dataset, where each observation is the candidate in the leader selection event (i.e., the leader and counterfactual leader). The outcome is a binary indicator of whether each candidate has prior military experience. The two main explanatory variables are

the share of veterans in the population and past conflict. Second, I subset the LSD dataset to include only selection events in which the outcome results from the selectorate's choice (i.e., free and fair elections and appointments). Lastly, I only include self-selections to investigate whether ex-military candidates are more successful at overthrowing non-ex-military incumbents than the reverse.

Table 7-5: Candidate military experience regression results.

	Ex-Military Candidate				Ex-Military Candidate Chosen		Ex-Military Self-Select Winner
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Past Conflict	0.04***			0.03	-0.02		
	(0.02, 0.06)			(-0.005, 0.07)	(-0.05, 0.02)		
Est. Veteran Pop %		0.25***		0.27*		-0.12	
		(0.12, 0.38)		(0.01, 0.53)		(-0.41, 0.17)	
Self-Selection Attempt			2.40***	2.54***			
			(2.13, 2.66)	(2.22, 2.86)			
Constant	-0.95***	-0.97***	-1.04***	-1.41***	0.34	0.38	0.46***
	(-1.09, -0.81)	(-1.13, -0.81)	(-1.17, -0.90)	(-1.68, -1.14)	(-0.10, 0.79)	(-0.15, 0.90)	(0.14, 0.78)
Clustered SE	Individual	Individual	Individual	Country	Country	Country	Country
Observations	4,070	4,070	4,070	4,070	195	188	194
Log Likelihood	-2,504.08	-2,512.71	-2,311.86	-2,249.05	-133.16	-128.42	-129.44
Akaike Inf. Crit.	5,012.16	5,029.41	4,627.73	4,506.10	270.32	260.83	260.87
Note:							*p<.05; **p<.01; ***p<0.001

Table 7-5 above shows the regression results for military experience. Models 1 through 3 show that a candidate's military experience is positively associated in isolation with past conflict, the estimated veteran proportion, and self-selection attempt, respectively when standard errors are clustered by the individual. However, Model 4 shows that when clustered standard errors by country are used, past conflict is no longer significant and the effect of the estimated share of veterans in the population decreases and the relationship weakens. This means that while both predict whether candidates are ex-military, this is largely because ex-military candidates tend to be concentrated within the same countries. In other words, there is more support for the demographic hypothesis than the affinity bias hypothesis. To support the affinity bias

hypothesis, we would expect the effect to be stronger when country-level clustered errors are included, since this would mean the relationship is stronger within countries than between them.

Models 5 and 6 more specifically test whether selectorates are more likely to choose ex-military candidates over non-ex-military candidates conditional on past conflict and the estimated share of veterans in the population. Both models show, however, that ex-military candidates are less likely than non-ex-military candidates to be chosen as both variables increase, although these relationships are not statistically significant. Consequently, no other evidence supports the leader selection effect hypothesis. Lastly, model 7 shows that ex-military candidates are significantly more successful at overthrowing non-ex-military leaders than the reverse. According to the model, in self-selection attempts that are only between ex-military and non-ex-military candidates, an ex-military candidate is 50% more successful at overthrowing a non-ex-military leader than the reverse. However, this is puzzling given that the leader self-selection effect in Table 7-4 above is negative and significant. This suggests that ex-military coup leaders are significantly more successful at overthrowing non-military leaders than when non-military coup leaders try to overthrow ex-military leaders. Afterwards, however, these ex-military leaders are significantly *less likely* to initiate conflict than the non-veteran coup leaders who successfully overthrow military leaders. This finding, however, is explained by Ayatollah Khomeini's 1979 self-selection (incumbent Mohammed Reza is ex-military). When Khomeini is excluded from model 7 in the selection table above, the effect of non-military leaders on conflict decreases by 31%, and the relationship is no longer significant (coef = -1.00, p = .10). Nevertheless, considering the leader self-selection effect is negative, the fact that veterans are more successful at overthrowing non-veterans than the reverse does not explain why ex-military leaders are more likely to initiate conflict.

Returning to the demographic hypothesis, Figure 7-10 and Figure 7-11 below show that both predictions of the demographic candidate selection effect hypothesis are supported; the estimated veteran proportion is positive and significantly associated with the selection of ex-

military candidates and conflict initiation. In essence, there is stronger evidence of the candidate selection effect hypothesis since the corresponding demographic hypothesis is also supported.

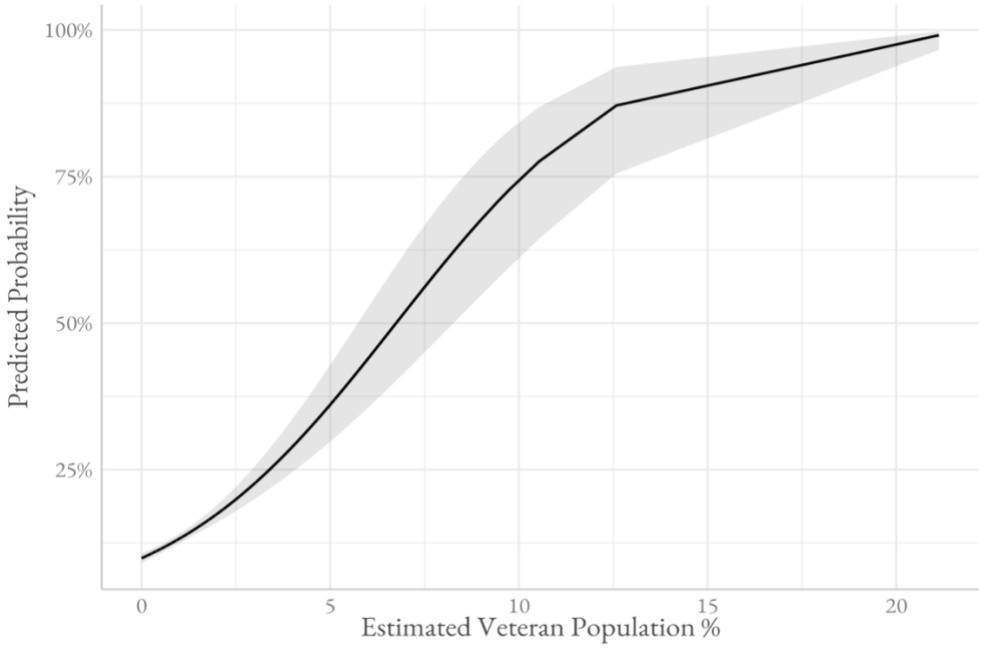


Figure 7-10: Predicted probability of conflict initiation by estimated veteran population percentage.

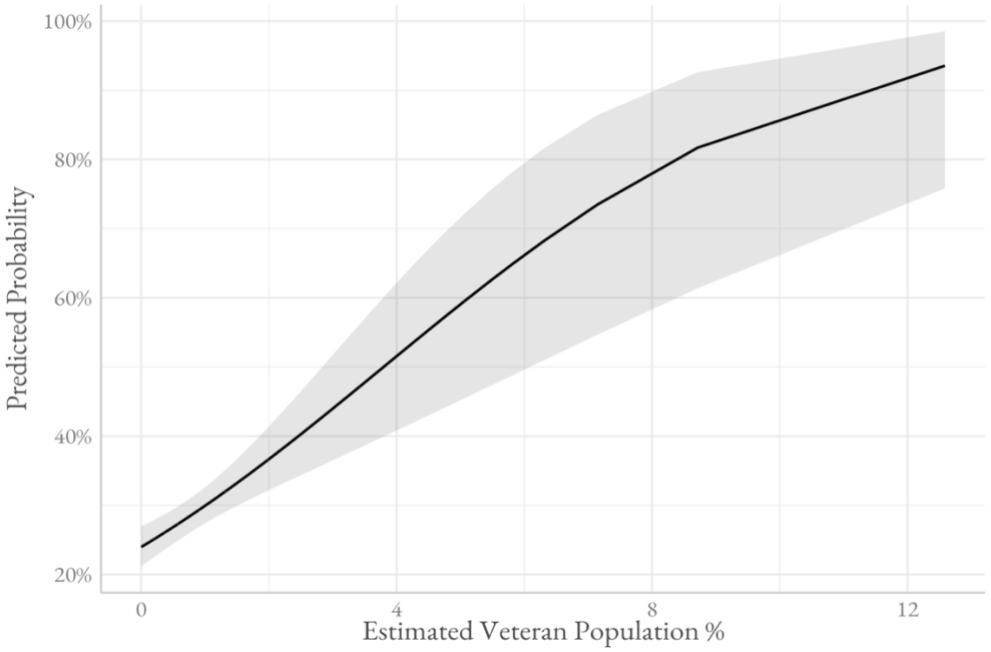


Figure 7-11: Predicted probability of ex-military candidate by estimated population percentage.

7.5.6 Military Experience: Conclusion

The findings above support the environmental perspective argument: ex-military leaders are likely not more likely than non-ex-military leaders to initiate conflict, just more likely to be selected into environments where conflict is already more likely. The evidence also suggests that this is because of demographics: conflict-prone states tend to have more veterans in the population and so tend to have more ex-military candidates and hence ex-military leaders than conflict-averse states.

7.6 Military Experience Type

While the above findings investigate military experience in general, it is important to investigate if we see the same selection effects for non-combat military experience, given that several findings suggest only leaders with this military experience type are more likely to initiate conflict than non-veterans.

7.6.1 Military Experience Type: Summary Statistics

In total, 10% of leaders and counterfactual leaders were non-combat veterans, whereas 22% of leaders and 15% of counterfactual leaders were combat veterans. However, 12% (n=208) of the 1,709 counterfactual leaders lacked sufficient biographical information to discern whether they had combat experience. Of these, 41% were counterfactual leaders of failed coups, with very few biographical sources beyond a brief excerpt in older newspapers. The plural majority of combat veterans appeared in elections (49%), whereas the plural majority of non-combat veterans appeared in self-selections (45%). Among self-selections, combat veterans were 54% of leaders and non-combat veterans were 32%. For counterfactual leaders in self-selections, combat veterans were 25% of counterfactual leaders in failed coups, and non-combat veterans were 34%.

In elections, combat veterans were the winners in 13% and the runners-up in 19%, whereas non-combat veterans were the winners in 5% and the runners-up in 7%. Additionally, in electoral contests between candidates with different combat experience, combat veterans won 55% of elections against non-veterans and 68% against non-combat veterans. In contrast, non-combat veterans won 65% of elections against non-veterans. These findings are consistent with veteran bonus theory.

Across all selection mechanisms, the highest percentage of combat veteran candidates among countries with at least ten leader selection terms were in Egypt (93%), Mauritania (70%), France (67%), Congo – Kinshasha (63%) and Burkina Faso (61%). Conversely, the highest percentage of non-combat veteran candidates were in Cambodia (70%), Sudan (70%), Burundi (42%), Iraq (37%) and Syria (35%).

7.6.2 Military Experience Type: Selection Effect Hypotheses Tests

Recall that either of the three selection effects hypotheses was predicted to explain a positive relationship between non-combat ex-military leaders and conflict initiation. To test whether we see evidence for any of the three hypotheses, I use the standard regression strategy outlined in the Methodology Chapter and used in section 7.5.2. above. Table 7-6 below lists the selection effect hypotheses for reference.

Table 7-6: Military experience type selection effect hypotheses.

Selection Effect Hypotheses	Treatment Group		Comparison Group		Supporting Hypotheses
	Leader Military Experience Type	Counterfactual Leader Military Experience Type	Leader Military Experience Type	Counterfactual Leader Military Experience Type	
Candidate Selection Effect	Non-Combat	Non-Combat	None	None	FIR, Demographics
Leader Selection Effect	Non-Combat	None	None	Non-Combat	FIR
Candidate Self-Selection Effect	Yes	Yes	No	No	Attempt
Leader Self-Selection Effect	Yes	No	No	Yes	Success

7.6.3 Military Experience Type: Main Findings

Table 7-7: Military experience type selection effect regression results

				Self-Selected = No		Self-Selected = Yes	
	Population Treatment Effect	Sample Treatment Effect	Placebo Treatment Effect	Candidate Selection Effect	Leader Selection Effect	Candidate Selection Effect	Leader Selection Effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Military, Combat (Leader)	0.68*** (0.32, 1.03)	0.63** (0.16, 1.10)		1.26*** (0.69, 1.83)	0.15 (-0.63, 0.94)	1.29 (-0.15, 2.73)	-1.45* (-2.82, -0.07)
Military, Non-Combat (Leader)	0.72** (0.20, 1.24)	0.65 (-0.09, 1.40)		1.64* (0.16, 3.13)	0.51 (-0.35, 1.36)	1.31 (-1.84, 4.46)	-2.26** (-3.70, -0.82)
Military, Combat (Counterfactual Leader)			1.16*** (0.68, 1.65)				
Military, Non-Combat (Counterfactual Leader)			1.57*** (0.86, 2.27)				
Constant	-2.10*** (-2.34, -1.86)	-1.87*** (-2.18, -1.56)	-2.10*** (-2.28, -1.92)	-2.21*** (-2.44, -1.98)	-1.78*** (-2.40, -1.17)	-1.65*** (-2.56, -0.74)	-0.01 (-1.29, 1.28)
Observations	10,668	6,224	6,241	3,485	1,238	568	933
Log Likelihood	-4,231.09	-2,735.35	-2,599.43	-1,211.72	-554.30	-322.51	-470.18
Akaike Inf. Crit.	8,468.18	5,476.71	5,204.86	2,429.44	1,114.59	651.02	946.36

Note: Robust standard errors clustered by individual leader

*p<05; **p<.01; ***p<0.001

Table 7-7 above shows the regression results for military experience type. Each model's outcome is the initiation of international conflict in the leader selection term year. Logistic regression is conducted with robust standard errors clustered by the individual leader. Model 1 shows the population treatment effect of non-combat and combat experience among all leaders, whereas model 2 shows the sample treatment effect of both among leaders where the counterfactual leader's combat experience is observed. While the effect of combat experience is not statistically different between the two models, non-combat experience is slightly over the traditional .05 significance level cut-off (p=.08). Model 3 shows the placebo treatment effect of a counterfactual leader's combat experience on conflict is positive and significant. Like military experience above, both are positive and significant but are around twice as large as the population treatment effect, although the confidence intervals overlap substantially.

Comparing the selection effects of non-combat experience between models 4 and 5 among non-self-selections, we can see that the effect of military experience type on conflict initiation depends on the counterfactual leaders' military experience type, contrary to the trait hypothesis. Model 4 compares the effect of leaders' combat experience when both leaders and counterfactual leaders have the same category of combat experience (i.e., non-military, military–combat, or military-non-combat). The probability of conflict is three and half times greater when both are combat veterans compared to when both are non-veterans, and just over five times greater when both are non-combat veterans compared to when both are non-veterans. Like military experience above, both findings are consistent with the candidate selection effect hypothesis. Conversely, model 5 compares the effect of leaders' combat experience when leaders and counterfactual leaders have different categories of combat experience. While the effect of both combat and non-combat experience is positive, neither is significant, contrary to the leader selection effect hypothesis.

Models 6 and 7 are the same as models 4 and 5 but only among self-selections. Model 6 shows that there is no candidate self-selection effect, but model 7 shows there is a negative

leader self-selection effect. That is, the probability of conflict when non-veteran leaders remove a non-combat veteran from power is five and a half times greater than when non-combat veteran leaders remove non-veterans.

Figure 7-12 below better visualises the candidate and leader selection effects. A similar pattern as military experience among non-self-selected leaders emerges as before; the positive and significant effect of non-combat experience on conflict initiation is likely due to candidate selection effects. That is, non-combat leaders are more likely to initiate conflict than non-ex-military leaders only when the counterfactual leader has the same military experience type. Furthermore, since there is no significant difference in the effect of non-combat experience when the counterfactual leader has the opposite trait, there is therefore no evidence for the leader selection effect hypothesis.

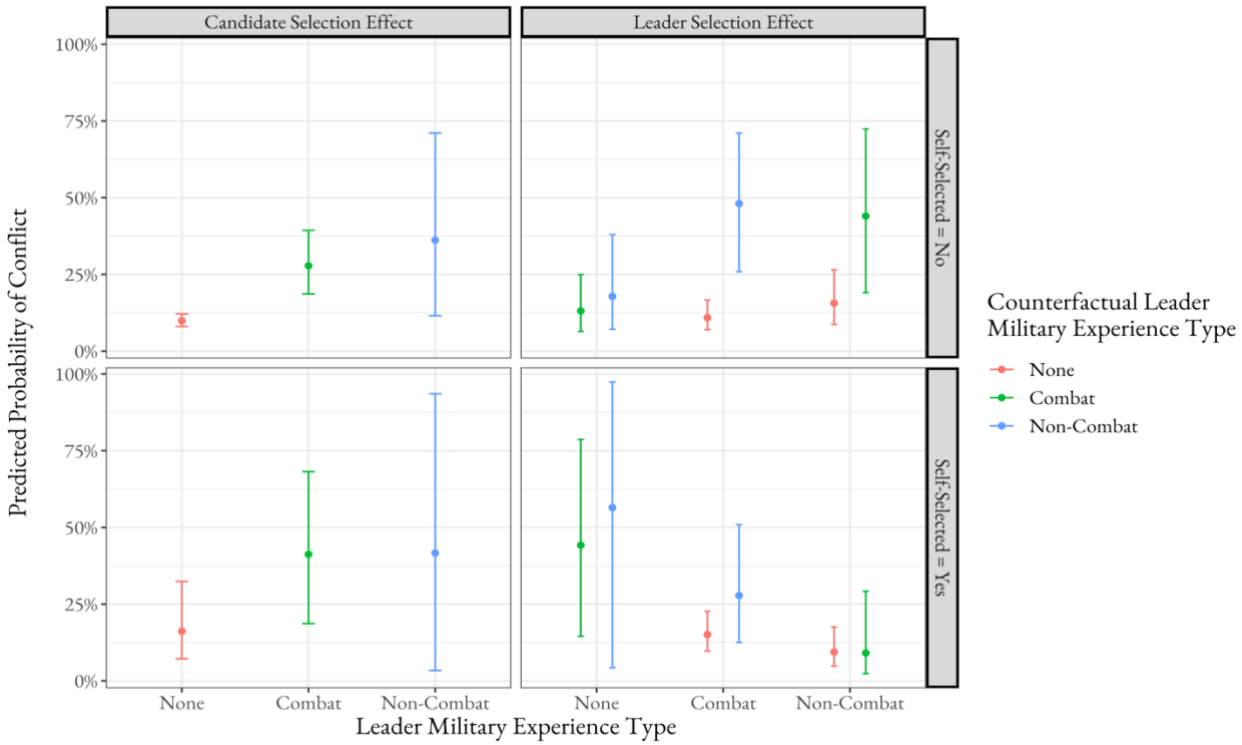


Figure 7-12: Predicted probability of conflict by leader military experience type, selection effect, and self-selection.

7.6.4 Military Experience Type: Placebo Assumption Test

Figure 7-13 below shows that the placebo assumption is validated. The candidate selection effect among non-self-selected leaders is likely not because non-combat counterfactual leaders were previous leaders, creating an environment where conflict was more likely for the future leader. More specifically, the probability of conflict is not significantly different depending on whether non-combat counterfactual leaders were the previous leader. However, this is not the case among self-selected leaders and mirrors the puzzling finding from Figure 7-12.

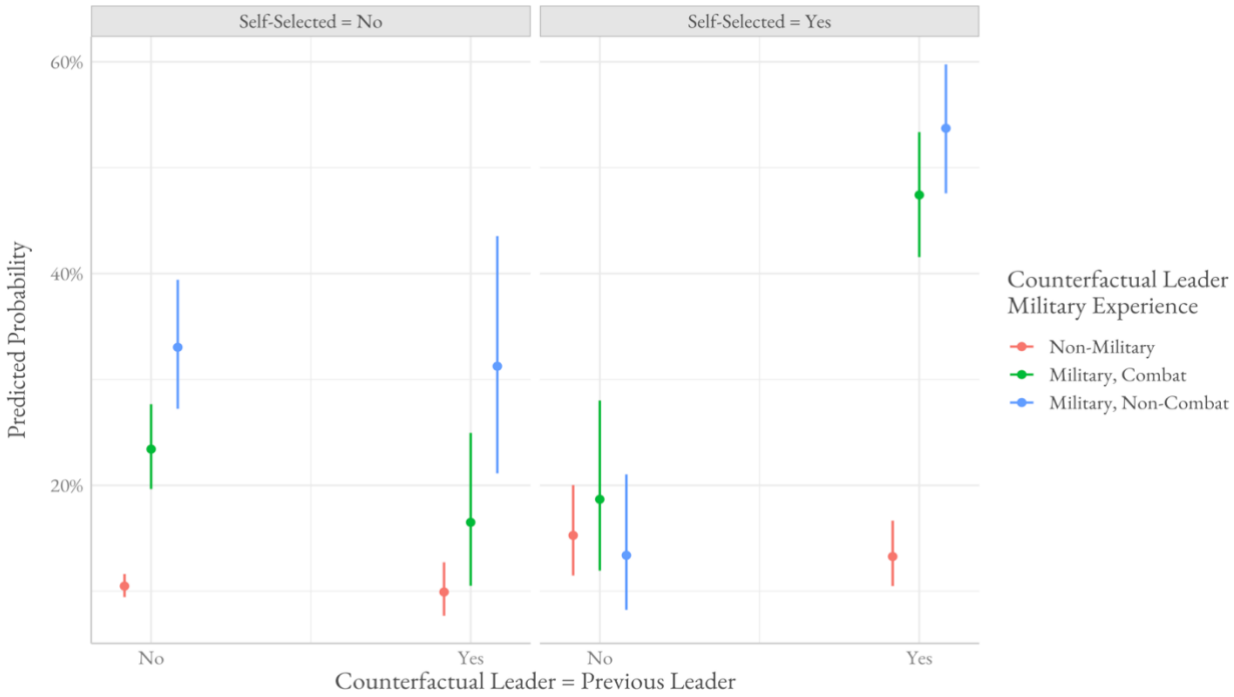


Figure 7-13: Regression results plot of testing the counterfactual leader placebo assumption for military experience.

7.6.5 Military Experience Type: Supporting Hypotheses Tests

To more specifically test whether the candidate selection effect above is explained by the demographic or FIR supporting hypotheses, I conduct the same regression strategy as military experience above, except with slight alterations to account for the three differences in military

experience type. To test the supporting hypotheses for candidate selection effects, the outcome is the candidate military experience type, and so I run two separate regressions for non-combat and combat candidates using non-ex-military candidates as the reference group. To test the supporting hypotheses for leader selection effects, the outcome is whether the non-combat or combat candidate was chosen over non-ex-military candidates, and I use both reference groups of candidates with no-military experience or the opposite experience type (e.g., non-combat chosen over combat).

Table 7-8: Candidate military experience type regression results (non-ex-military reference group).

	Non-Combat Candidate				Non-Combat Candidate Chosen		Non-Combat Wins Self-Selection
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Past Conflict	0.04*** (0.02, 0.06)			0.04 (-0.01, 0.09)	-0.01 (-0.04, 0.01)		
Est. Veteran Pop %		0.21 (-0.09, 0.50)		0.25 (-0.10, 0.60)		-0.21 (-0.46, 0.04)	
Self-Selection Attempt			2.65*** (2.29, 3.01)	2.79*** (2.36, 3.22)			
Constant	-2.13*** (-2.34, -1.92)	-2.11*** (-2.39, -1.84)	-2.35*** (-2.57, -2.12)	-2.73*** (-3.13, -2.33)	0.15 (-0.19, 0.49)	0.26 (-0.12, 0.63)	0.06 (-0.37, 0.49)
Clustered SE	Individual	Individual	Individual	Country	Country	Country	Country
Observations	2,356	2,356	2,356	2,356	203	198	103
Log Likelihood	-867.62	-875.53	-765.04	-746.01	-140.29	-136.12	-71.35
Akaike Inf. Crit.	1,739.25	1,755.06	1,534.08	1,500.02	284.57	276.24	144.70
Note:							*p<.05; **p<.01; ***p<0.001

Table 7-9: Candidate military experience type regression (combat experience reference group).

	Non-Combat Candidate				Non-Combat Candidate Chosen		Non-Combat Wins Self-Selection
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Past Conflict	0.001 (-0.02, 0.02)			0.01 (-0.01, 0.04)	-0.002 (-0.11, 0.11)		
Est. Veteran Pop %		-0.17 (-0.42, 0.09)		-0.18 (-0.40, 0.03)		-0.04 (-0.29, 0.21)	
Self-Selection Attempt			0.40* (0.05, 0.74)	0.35 (-0.07, 0.76)			
Constant	-0.61*** (-0.88, -0.34)	-0.45** (-0.75, -0.16)	-0.72*** (-0.99, -0.45)	-0.63** (-1.04, -0.21)	-0.77 (-1.63, 0.09)	-0.72* (-1.41, 0.04)	-1.03** (-1.81, -0.25)
Clustered SE	Individual	Individual	Individual	Country	Country	Country	Country
Observations	820	820	820	820	64	63	38
Log Likelihood	-532.74	-529.62	-529.64	-526.61	-39.74	-39.36	-21.90
Akaike Inf. Crit.	1,069.47	1,063.24	1,063.28	1,061.21	83.49	82.71	45.80

Note: *p<.05; **p<.01; ***p<0.001

The above tables show the regression results for non-combat ex-military candidates when the reference group is a non-ex-military candidate (Table 7-8) and a candidate with combat experience (Table 7-9). Starting with candidate selection tests, model 1 shows that past conflict is positively and significantly associated with the selection of non-combat candidates only when compared with non-ex-military candidates. Model 2 shows that the estimated veteran population percentage is also positively associated with the selection of non-combat candidates only when compared with non-ex-military candidates. Model 3 includes both explanatory variables together with robust standard errors clustered by country. Regardless of the reference group, past conflict and the estimated veteran percentage are both no longer significantly associated with the selection of non-combat candidates. The results are similar for military experience above; model 1 was largely a result of non-combat candidates being concentrated within the same countries. For non-combat candidates, while the estimated veteran percentage is not significant in model 4, it does, however, gain significance and the effect size increase when past conflicts is excluded from the model ($p = .02$), but we do not see the same for past conflicts when the veteran percentage is excluded ($p = .09$). It appears, then, that self-selection attempt is a suppressor variable, as that the estimated veteran percentage is only somewhat correlated with the selection of non-combat leaders among non-self-selection events (Figure 7-14). Considering we already

know that the estimated veteran population percentage is positively associated with conflict initiation, these findings further support the demographic candidate selection effect hypothesis, at least among non-self-selected leaders.

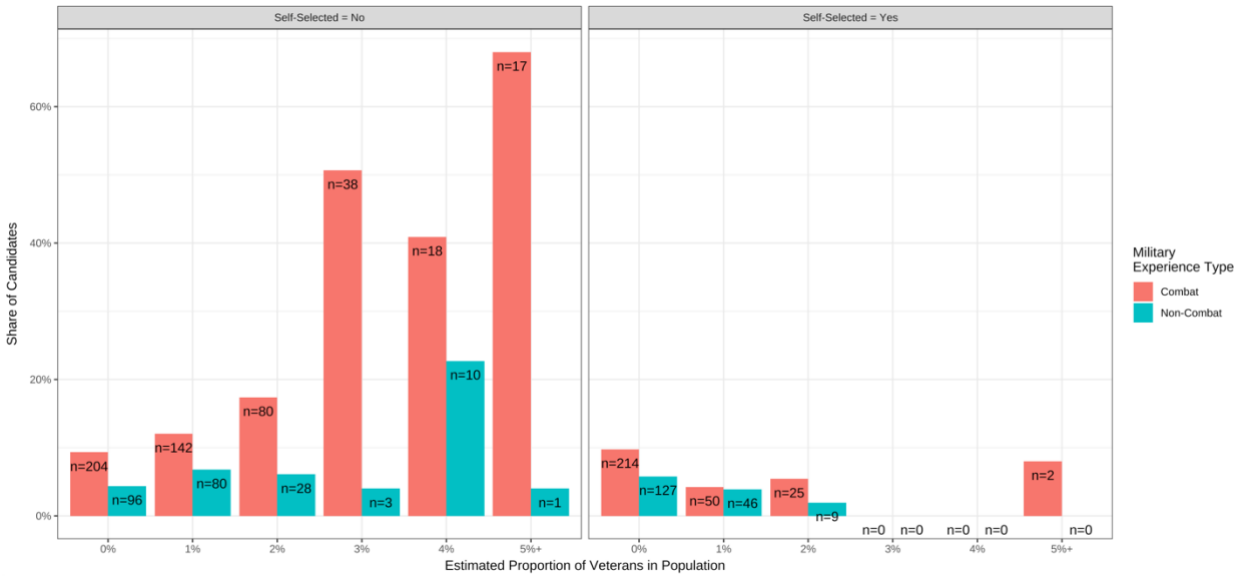


Figure 7-14: Share of combat/non-combat candidates by estimated share of veterans in population and self-selection.

7.6.6 Military Experience Type: Conclusion

The totality of the findings above moderately supports the environmental argument that non-combat leaders are not more likely to initiate conflict than non-ex-military leaders, just more likely to be selected into environments where conflict is already more likely. The evidence suggests that this is because of demographics: conflict-prone states tend to have more veterans in the population and so tend to have more non-combat candidates, and hence non-combat leaders, than conflict-averse states.

7.7 War Win Experience

7.7.1 War Experience Type: Summary Statistics

Among individuals with combat experience, I determined whether they had war win or war loss experience in all but one leader.[222] Table 7-10 below summarises the frequency and proportion of war experience across all leaders and counterfactual leaders.

Table 7-10: Frequency and proportion of leaders and counterfactual leaders by war experience

			Military, Combat			
	No Military Experience	Military, Non-Combat	War Loss	War Draw	War Win	Both War Win and Loss
Leaders	1793 (68%)	260 (12%)	144 (5%)	273 (10%)	156 (6%)	15 (<1%)
Counterfactual Leaders	1125 (75%)	145 (10%)	63 (4%)	88 (6%)	78 (5%)	2 (<1%)

7.7.2 War Experience Type: Selection Effect Hypotheses Tests

I conduct the standard regression strategy to test the selection effect hypotheses for war-winner experience. However, I excluded self-selected leaders from the analysis primarily because the sample sizes of war winners are too small within this selection mechanism, and we did not predict either self-selection effect hypothesis. In any case, the findings in Table 7-11 below are essentially identical when including self-selection despite slightly increased confidence intervals due to reduced observations. Figure 7-15 below summarises the selection effect and supporting

[222] Mohammed Yusuf of Afghanistan.

Figure 7-15: War experience type selection effect hypotheses.

Selection Effect Hypotheses	Treatment Group		Comparison Group		Supporting Hypotheses
	Leader War Experience Type	Counterfactual Leader War Experience Type	Leader War Experience Type	Counterfactual Leader War Experience Type	
Candidate Selection Effect	War-Winner	War-Winner	No Military Experience	No Military Experience	FIR
Leader Selection Effect	War-Winner	No Military Experience	No Military Experience	War-Winner	FIR

7.7.3 Leader War Experience Type: Main Findings

Table 7-11: War experience type selection effect regression results.

	Population Treatment Effect (1)	Sample Treatment Effect (2)	Placebo Treatment Effect (3)	Candidate Selection Effect (4)	Leader Selection Effect (5)
War Loser	0.27 (-0.20, 0.74)	-0.02 (-0.68, 0.64)		1.01 (-0.06, 2.08)	-0.78* (-0.88, 0.86)
War Winner	1.13*** (0.62, 1.63)	1.18*** (0.58, 1.78)		2.21*** (1.21, 3.21)	0.43 (-0.34, 1.20)
(CL) War Loser			0.40 (0.-32, 1.11)		
(CL) War Winner			1.12*** (0.40, 1.81)		
Constant	-1.95*** (-2.14, -1.76)	-1.76*** (-2.02, -1.51)	-1.76*** (-2.00, -1.51)	-1.81*** (-2.08, -1.53)	-1.67*** (-2.20, -1.14)
Observations	10,673	6,232	6,241	5,322	940
Log Likelihood	-4,224.62	-2,714.33	-2,736.34	-2,224.77	-453.45
Akaike Inf. Crit.	8,455.23	5,434.66	5,478.68	4,455.54	912.90

Note: Robust standard errors clustered by individual leader. *p<05; **p<.01; ***p<0.001

Table 7-11 shows the regression results for war experience type. The first two models confirm that the effect of war winners on conflict is positive and significant, and there is no significant difference in the likelihood of conflict initiation between war losers and those without

military experience. Model 3 shows that the placebo effect of war win and war loss experience is to the population treatment effect, suggesting a selection effect for candidates with war-win experience compared to those with no military experience (and no evidence for war losers). Turning to the selection effects, Model 4 shows the probability of conflict for war experience among states where the leader and counterfactual leader had the same war experience type. We can see a large candidate selection effect for war winners, and none for war losers. That is, the probability of conflict is nine times greater when both candidates are war winners than when both candidates have no military experience. Model 5 shows a positive leader selection effect for war winners but is not statistically significant. Interestingly, there is a negative leader selection effect for war losers.

Figure 7-16 below visualises the selection effects from models 4 and 5 of Table 7-11 above. The left facet clearly shows the positive candidate selection effect for war winners compared to non-military leaders. Interpreting the leader selection effects, however, is more difficult than the other military-related variables considering that there are multiple values for which to compare (i.e., no military experience, non-war military experience, war-win experience, war-loss experience, both war-win and war-loss experience). So, to simplify the visualisation, I only show the difference between war-winners and those with no military experience. The candidate selection effect is substantial; when the leader and counterfactual leaders' are both war winners, the probability of conflict is just over twice as likely to occur than when both have no military experience. Therefore, the candidate selection effect hypothesis is supported. However, there is no evidence for the leader selection effect hypothesis since selecting a war-winner candidate over a non-ex-military candidate does not lead to any more conflicts than selecting a non-ex-military candidate over a war-winner candidate. Therefore, there is no evidence to support the leader selection effect hypothesis.

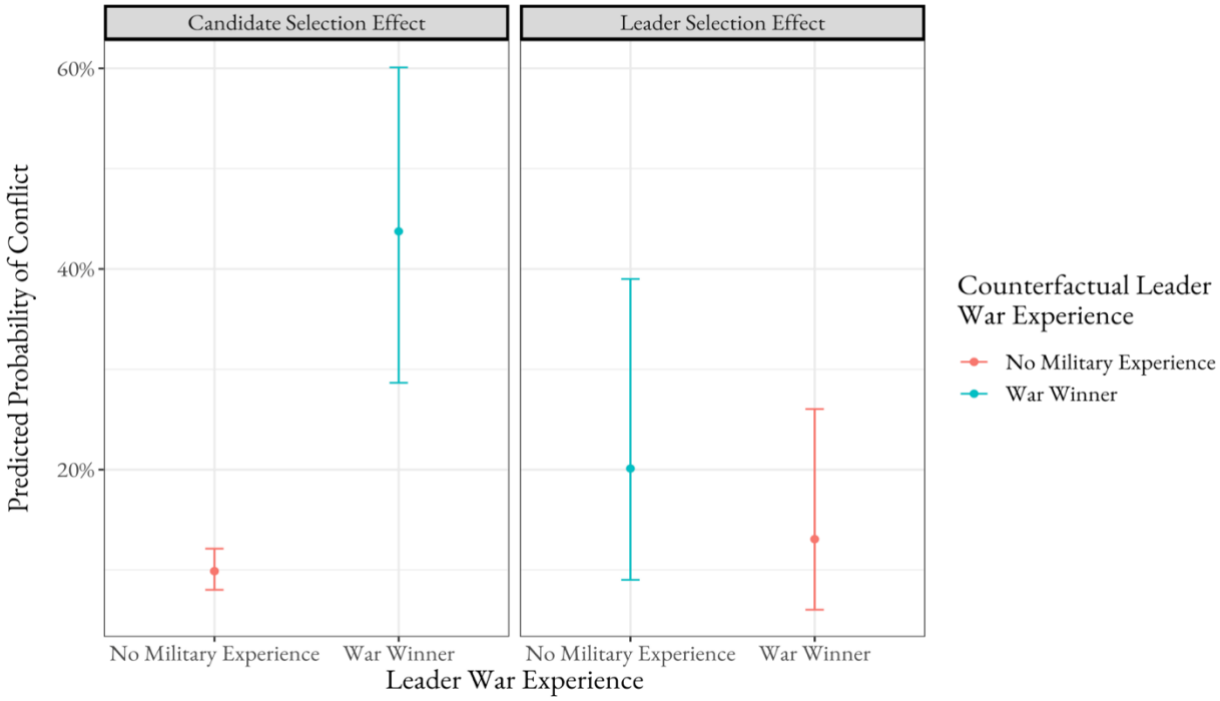


Figure 7-16: Predicted probability of conflict by leader war experience and selection effect (among non-self-selected leaders).

7.7.4 War Experience Type: Placebo Assumption Test

Figure 7-17 below confirms that the candidate selection effect above does not result from counterfactual leaders with war-win experience being previous leaders. More specifically, the predicted probability of conflict for counterfactual leaders with war-win experience is not statistically significant depending on whether they were the previous leader. Thus, this validates the central assumption that the counterfactual leader's war win experience has no causal effect on conflict initiation.

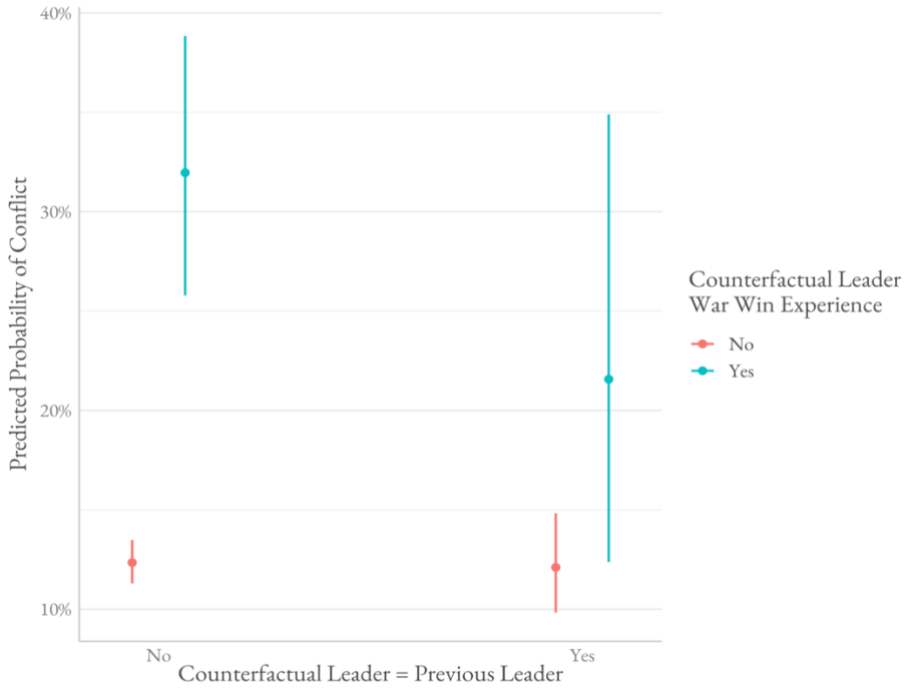


Figure 7-17: War Win experience placebo assumption test

7.7.5 War Experience Type: Supporting Hypotheses Tests

To more specifically test whether the candidate selection effect above is due to any of the supporting hypotheses, I conduct the same statistical analysis as other military-related variables above except, again, I do not include self-selections due to the small sample size of war winners and war losers within that selection mechanism.

Table 7-12: Candidate war experience type regression results.

	Outcome: War Winner Candidate				Outcome: War Winner Candidate Wins Election		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Past Conflict	0.07*** (0.04, 0.10)			0.04 (-0.01, 0.08)	-0.03 (-0.10, 0.04)		
Est. Veteran Pop %		0.82*** (0.57, 1.07)		0.62** (0.23, 1.00)		-0.12 (-0.82, 0.58)	
Past War Win			1.78*** (1.23, 2.33)	0.84 (-0.28, 1.97)			-1.11 (-2.98, 0.76)
Constant	-2.83*** (-3.15, -2.50)	-3.26*** (-3.66, -2.86)	-2.74*** (-3.04, -2.44)	-3.40*** (-3.93, -2.88)	0.02 (-0.75, 0.78)	-0.03 (-1.02, 0.97)	0.10 (-0.61, 0.82)
Clustered SE	Individual	Individual	Individual	Country	Country	Country	Country
Observations	3,030	3,030	3,171	3,030	55	53	55
Log Likelihood	-801.57	-761.11	-823.55	-733.71	-37.28	-36.32	-36.37
Akaike Inf. Crit.	1,607.14	1,526.21	1,651.10	1,475.42	78.56	76.65	76.75
Note:							*p<.01; **p<.05; ***p<0.001

Table 7-13: Regression results – war loss experience supplementary selection tests

	Outcome: War Loss Candidate			
	(1)	(2)	(3)	(4)
Past Conflict	0.03*** (0.01, 0.05)			0.02 (-0.02, 0.05)
Est. Veteran Pop %		0.44*** (0.21, 0.67)		0.34 (-0.03, 0.72)
Past War Loss			1.19*** (0.68, 1.69)	1.00* (0.18, 1.83)
Constant	-2.70*** (-2.99, -2.41)	-2.92*** (-3.28, -2.56)	-2.75*** (-3.07, -2.44)	-3.13*** (-3.74, -2.52)
Clustered SE	Individual	Individual	Individual	Country
Observations	2,991	2,991	3,141	2,991
Log Likelihood	-770.79	-753.12	-804.78	-733.03
Akaike Inf. Crit.	1,545.59	1,510.24	1,613.57	1,474.05
Note:				*p<.05; **p<.01; ***p<0.001

Table 7-12 and Table 7-13 above show the regression results for war winners and war losers, respectively. For war losers, I exclude the FIR hypothesis tests among elections since there are too few observations (n=16), and the results are null due to low variation among the explanatory variables. In model 1, war winners and losers are more likely to be candidates as past conflict increases, and war winners are more so than war losers. Model 2 shows that both are also more likely to be candidates as the estimated share of veterans in the population increases, and again more so for war winners than war losers. Model 3 shows that past war win is positively and significantly associated with the selection of war winner candidates, and the

same for past war loss and war loser candidates. However, model 4 shows that only the estimated veteran population percentage is significant for war winners when country-level clustered errors are included. This is likely because past war wins are rare and the variable is constructed over a 40-year period, so there is likely little variation within countries, dampening the effect when country-level correlation is considered. Furthermore, the estimated veteran population percentage and past wars are correlated (Figure 7-18), which could also weaken the effect since there is more variation in the former. Indeed, on its own, past war is robustly associated with war-winner candidates when including robust country-level clustered standard errors (coef = 1.77, p <.001).

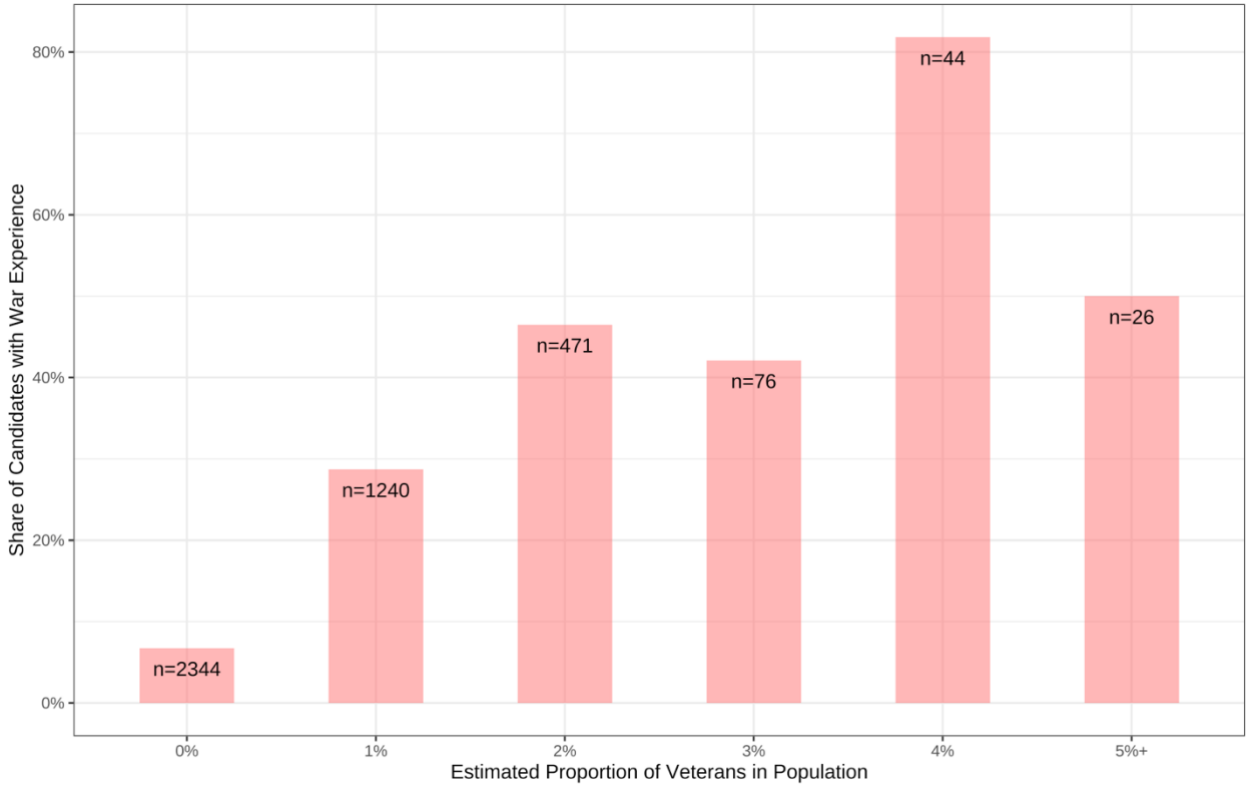


Figure 7-18: Share of candidates with war experience (war or loss) by the estimated proportion of veterans in the population.

Turning to models 5 through 7 in Table 7-12, we can see that war-win candidates are not more likely to win elections over non-military candidates conditional on any of the explanatory

variables. Thus, we have no further evidence to support either the FIR leader selection effect hypothesis or the demographic leader selection effect hypothesis for why war winners appear more likely to initiate conflict.

7.7.6 War Experience Type: Conclusion

The totality of the findings above moderately supports the environmental argument that leaders with war-win experience are not especially conflict-prone and are instead just more likely to be selected into conflict-prone environments. The evidence suggests that this is because of demographics: states that have won past wars are more likely to initiate conflict than conflict-prone states, have a higher percentage of veterans in the population, and so tend to have more ex-military candidates and hence ex-military leaders than conflict-averse states.

7.8 Age

7.8.1 Age: Summary Statistics

The birth year could not be determined in 1% of leaders and 9% of counterfactual leaders. The average age of leaders is 56.3 (sd = 11.5), and the average age of counterfactual leaders is 56.9 (sd = 11.9) and is normally distributed for both. In countries with at least 10 candidates, the five with the oldest candidates are Cyprus (67), Nepal (65), Greece (64), Japan (64), and India (64), whereas the five with the youngest candidates are Macedonia (38), Latvia (44), Slovenia (44), Solomon Islands (45), and Estonia (46). Table 7-14 below provides the average age of all candidates across each selection mechanism.

Table 7-14: Average age of leaders and counterfactual leaders by selection mechanism.

	Elections	Appointments	Successions	Self-Selections
Leaders	58.5	53.9	53.5	54.2
Counterfactual Leaders	57.7	53.1	48.9	56.8

7.8.2 Age: Selection Effect Hypotheses Tests

I conduct the standard regression strategy as above to test the initial selection effect hypotheses for age. While I only hypothesised leader and candidate selection effects, I also disaggregated between non-self-selections and self-selections in case there is still evidence for a self-selection effect despite not predicting so. Additionally, since age is a continuous variable, my results table below is structured slightly differently than the military-related variables above. More specifically, since I cannot subset the data between when leaders' and counterfactual leaders' age is the same or opposite, I instead use an interaction term between the two and then show the candidate and selection effects using the predicted probability between two leaders that are two standard deviations away from the mean (i.e., leaders that are 34 years old and 79-years-old). Table 7-15 below lists the selection effect hypotheses for reference.

Table 7-15: Summary of selection effect hypotheses for age.

Selection Effect Hypotheses	**Treatment Group**		**Comparison Group**		**Supporting Hypotheses**
	Leader Age	**Counterfactual Leader Age**	**Leader Age**	**Counterfactual Leader Age**	
Candidate Selection Effect	Older	Older	Younger	Younger	FIR, Demographics, Affinity Bias
Leader Selection Effect	Older	Younger	Younger	Older	FIR, Affinity Bias

7.8.3 Leader Age: Main Findings

Table 7-16: Age selection effect regression results.

				Selection Effects	
	Population Treatment Effect	Sample Treatment Effect	Placebo Treatment Effect	Non-Self-Selected Leaders	Self-Selected leaders
	(1)	(2)	(3)	(4)	(5)
Age	0.03** (0.01, 0.05)	0.03* (0.01, 0.06)		0.12** (0.05, 0.20)	-0.04 (-0.15, 0.07)
(CL) Age			0.03*** (0.01, 0.05)	0.12** (0.05, 0.20)	-0.05 (-0.17, 0.08)
Age*(CL) Age				-0.002** (-0.003, -0.001)	0.001 (-0.001, 0.003)
Constant	-3.46*** (-4.49, -2.43)	-3.47*** (-4.91, -2.02)	-3.42*** (-4.30, -2.54)	-9.77*** (-14.18, -5.35)	-0.62 (-6.98, 5.75)
Observations	10,596	6,394	6,472	4,805	1,589
Log Likelihood	-4,228.33	-2,796.79	-2,808.45	-1,822.74	-818.59
Akaike Inf. Crit.	8,460.65	5,597.58	5,620.89	3,653.48	1,645.17

Note: Robust standard errors clustered by individual leader. *p<05; **p<.01; ***p<0.001

A comparison between models 1 and 2 confirms that the population treatment effect of age is not biased by missing counterfactual leaders. Model 3 confirms that the counterfactual leader's age has the same effect on conflict initiation as a leader's age, suggesting a selection effect for older leaders in environments where conflict is more likely. Models 4 and 5 more specifically show the selection effects using an interaction term between leaders' and counterfactual leaders' age among non-self-selected and self-selected leaders, respectively.[223] Since the selection effects are difficult to infer from coefficients alone, Figure 7-19 below visualises the results for models 4 and 5. To help interpret whether we have evidence for the candidate or leader selection effect hypotheses, I show the predicted probability of conflict between leaders and counterfactual leaders that are two standard deviations away from the mean age of 56.5 (34 and 79). Accordingly, the candidate selection effect shows the predicted probability of conflict for a 34-year-old leader when the counterfactual leader is also 34 years old, compared to the predicted probability of conflict when both are 79 years old. Conversely, the

[223] Since age is continuous variable, I simply interact the leader and counterfactual age to infer selection effects, as opposed to subsetting each selection effect as with the military-related traits above.

leader selection effect shows the predicted probability of conflict for a 34-year-old leader when the counterfactual leader's age is 79, compared to the predicted probability of conflict when the leader's age is 79 and the counterfactual leader's age is 34. As Figure 7-19 shows, among both non-self-selected and self-selected leaders, the probability of conflict is significantly greater when both leaders and counterfactual leaders are 79 years old than when both are 34. Conversely, there is no leader selection effect for older leaders in either selection mechanism – the predicted probability of conflict is the same regardless of whether an older candidate is selected over a younger candidate, or vice versa. In other words, the results support the candidate selection effect hypothesis that older leaders are only more likely to initiate conflict when the counterfactual leader is of similar age.

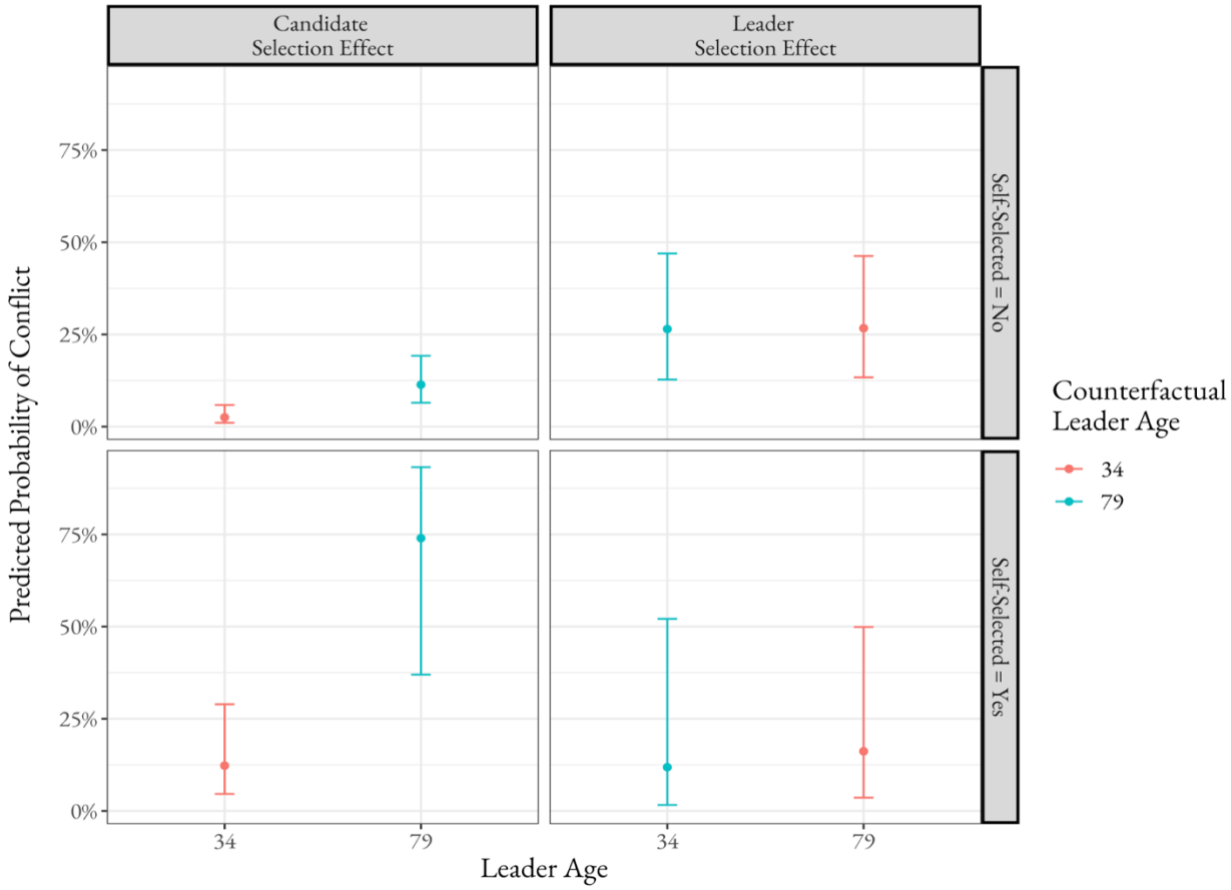


Figure 7-19: Predicted probability of conflict for leader age and selection effects among non-self-selected and self-selected leaders.

7.8.4 Age: Placebo Assumption Tests

Figure 7-20 below confirms that the candidate selection effect above does not result from older counterfactual leaders being previous leaders. The interaction term between counterfactual leaders' age and whether they were previous leaders is insignificant ($p = 0.96$ among non-self-selections, $p = .06$ among self-selections). We can see this because the confidence intervals overlap within each facet, although the assumption is weaker among self-selected leaders since the effect size of being previous leaders for older leaders is much larger than not being the previous leader and only slightly overlaps. Thus, this validates that the candidate selection effect we see for age is not because older counterfactual leaders were also previous leaders, although only among non-self-selected leaders.

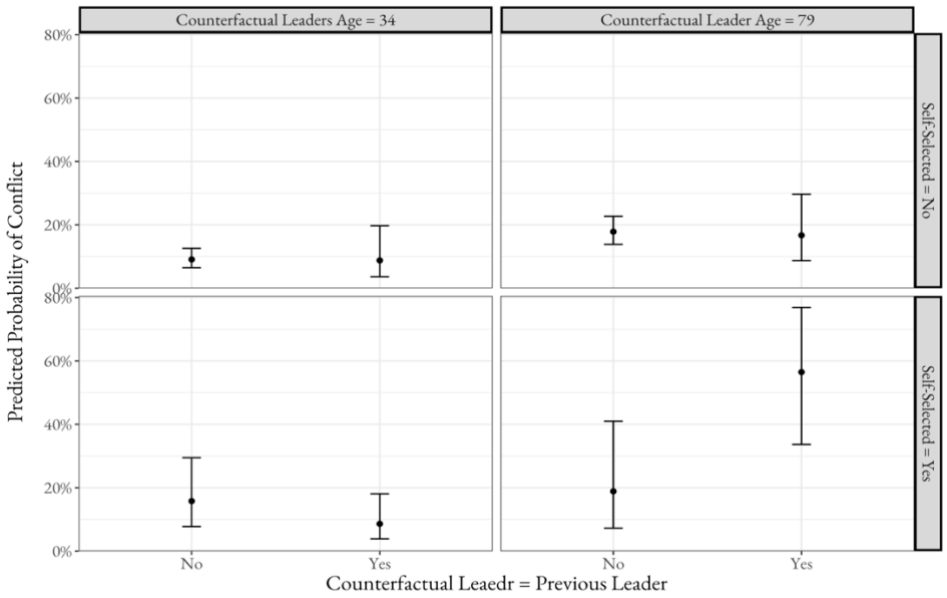


Figure 7-20: Regression results for testing the placebo assumption for age.

7.8.5 Age: Supporting Hypotheses Tests

To more specifically test whether the candidate selection effect above is due to demographics or an FIR selectorate preference, I conduct the same statistical analysis as the other military-related traits above with slight alterations to account for the continuous nature of

age. To test for the FIR and affinity bias supporting hypotheses, I only compare selection events between candidates that were at least ten years apart in age in appointments and elections. This ensures we can reliably infer that the outcome reflects the selectorate's preferences between two candidates with an observable age difference. The outcome is whether the older candidate was chosen over the younger candidate.

Table 7-17: Candidate age regression results.

	Candidate Age				Older Candidate Chosen		Older Candidate Wins Self-Selection
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Past Conflict	0.20***			0.08	0.01		
	(0.13, 0.27)			(-0.01, 0.18)	(-0.03, 0.05)		
Life Expectancy		0.20***		0.04		0.01	
		(0.15, 0.24)		(-0.02, 0.10)		(-0.01, 0.04)	
Self-Select Attempt			-9.31***	-29.15**			
			(-10.53, -8.08)	(-48.33, -9.97)			
Incumbent				2.39***			
				(1.48, 3.30)			
Incumbent (Winner)					0.89***	0.99***	
					(0.43, 1.35)	(0.49, 1.48)	
Incumbent (Loser)					-0.35	-0.36	
					(-0.93, 0.22)	(-0.95, 0.23)	
Constant	53.60***	41.99***	55.19***	51.87***	-0.16	-1.17	-0.53**
	(53.02, 54.18)	(39.12, 44.86)	(54.70, 55.69)	(47.80, 55.93)	(-0.52, 0.19)	(-2.59, 0.25)	(-0.86, -0.20)
Clustered SE	Individual	Individual	Individual	Country	Country	Country	Country
Observations	4,210	3,765	4,210	1,975	408	372	167
Adjusted R^2	0.01	0.05	0.05				
Log Likelihood				-7,271.40	-266.82	-241.32	-76.69
Note:							*p<.01; **p<.05; ***p<0.001

Table 7-17 above shows the supporting hypotheses test results for age. Models 1 and 2 show that candidates are more likely to be older as past conflict and life expectancy increase. Model 3 shows that older candidates are significantly less likely to attempt to self-select, which is evidence against a candidate self-selection effect, though it was not predicted. Model 4 includes all the explanatory variables as covariates with robust standard errors clustered within countries. Life expectancy loses effect size and significance, although past conflict remains weakly positively associated with the selection of older candidates (Model 3). Although, on its own, life expectancy remains robust with country-level robust clustered standard errors (coef = .198, p <

.001). The relationship between candidates' age and life expectancy weakens when country-level standard errors are included because past conflicts and life expectancy are positively correlated (Figure 7-21). However, life expectancy is not positively associated with conflict initiation, contrary to the last prediction of the demographic candidate selection effect hypothesis (coef = 0.003, p = .55). Nevertheless, given that candidates in conflict-prone states are more likely to be older than in conflict-averse states, and that conflict-prone states tend to have higher life expectancies, it is still likely that demographics explains the candidate selection effect.

Models 4 and 5 test whether older candidates are more likely to be elected than younger candidates as past conflicts and life expectancy increases, respectively, controlling for incumbency status for both the winner and runners-up. Model 4 shows that the number of past conflicts is not associated with whether older candidates are chosen, contrary to the FIR leader selection effect hypothesis. Conversely, model 5 shows that older candidates are not more likely to be chosen as the life expectancy increases, contrary to the affinity bias hypothesis.

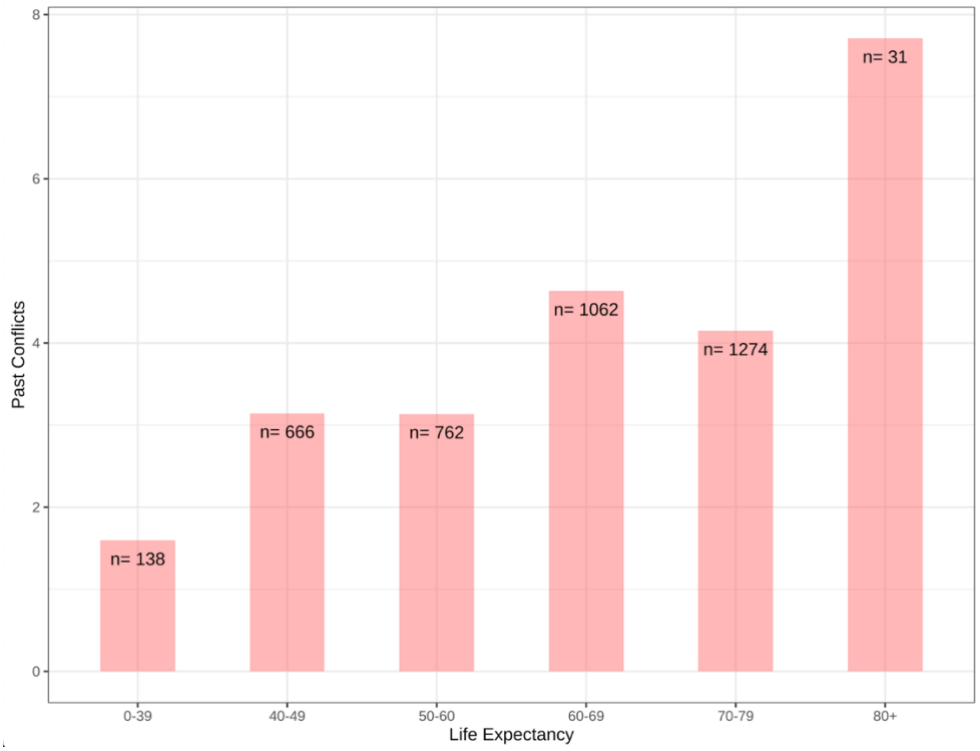


Figure 7-21: Average number of past conflict involvement by life expectancy bins.

7.8.6 Age: Conclusion

The totality of evidence above suggests that older leaders only appear more likely than younger leaders to initiate conflict because conflict-prone states tend to have higher life expectancies and hence older candidates. Meaning that conflict-prone states will tend to select older leaders more than conflict-averse states, even if by chance.

7.9 Biological Sex

7.9.1 Sex: Summary Statistics

Among all leader selection terms, females were 2% of candidates (2.4% of leaders and 2.2% of counterfactual leaders). The proportion of female candidates increases over time as 87% appeared in selection events in the last half of the dataset (after 1977). Geographically, 50% of all female candidates were in Asia, followed by Europe (21%) and the Americas (20%). Most female candidates (68%) appeared in elections and were slightly more likely to win elections when facing male candidates (54% to 46%). Compared to male candidates, female candidates are three years younger on average, more likely to be left-wing (42% to 26%), less likely to be centre (21% to 42%), and slightly more likely to be right-wing (38% to 33%). No female candidates had ever served in the military.

7.9.2 Sex: Selection Effect Hypotheses Tests

To evaluate the effect of a leader's sex on conflict given possible selection effects, I use the same regression method as previous traits, although I do not disaggregate by self-selection

due to the small number of female leader observations within this selection mechanism.[224] Table 7-18 summarises the predicted selection effect hypothesis for reference.

Table 7-18: Summary of selection effect hypotheses for biological sex.

Selection Effect Hypotheses	**Treatment Group**		**Comparison Group**		**Supporting Hypotheses**
	Leader Sex	**Counterfactual Leader Sex**	**Leader Sex**	**Counterfactual Leader Sex**	
Candidate Selection Effect	Female	Female	Male	Male	Gender Constraints

7.9.3 Biological Sex: Main Findings

Table 7-19: Female leader selection effect regression results.

	Population Treatment Effect	Sample Treatment Effect	Placebo Treatment Effect	Candidate Selection Effect	Leader Selection Effect
	(1)	(2)	(3)	(4)	(5)
Female	0.10 (-0.66, 0.86)	0.04 (-0.80, 0.88)		-0.97 (-3.26, 1.31)	-0.04 (-1.15, 1.07)
(CL) Female			0.03 (-0.61, 0.66)		
Constant	-1.83*** (-2.00, -1.66)	-1.66*** (-1.88, -1.45)	-1.66*** (-1.87, -1.45)	-1.67*** (-1.88, -1.45)	-1.45*** (-2.11, -0.78)
Observations	10,724	6,947	6,947	6,696	251
Log Likelihood	-4,314.73	-3,050.19	-3,050.20	-2,927.29	-121.03
Akaike Inf. Crit.	8,633.47	6,104.37	6,104.40	5,858.57	246.06

Note: Robust standard errors clustered by individual leader. *p<05; **p<.01; ***p<0.001

In Table 7-19, models 1 and 2 show that female leaders' population treatment effect and sample treatment effect on conflict are positive, although the sample treatment effect is 60% smaller. In model 3, the placebo effect is similar to the sample treatment effect, suggesting a positive selection effect for female leaders in more conflict-prone environments. However, we

[224] Among non-self-selection terms with a leader and counterfactual leader with the opposing sex, there are only two male leaders, and only one initiated any conflicts (Argentinian Jorge Videla).

can see both the candidate selection (model 4) and leader selection effect (model 5) are negative. The reason both selection effects are negative, but the sample treatment effect is positive, is likely due to Simpson's Paradox – when a trend among data groups reverses when those data are combined primarily due to large differences in sample sizes. Indeed, there are 6,666 observations when both the leader and counterfactual leader are male and only 146 when both are female. Nevertheless, Figure 7-22 below visualises the selection effects. As we can see, there is little to no evidence for either candidate or leader selection effect hypothesis for female leaders since the effect is negative in each case. In other words, selecting a female over a male for leader leads to fewer conflicts than the reverse.

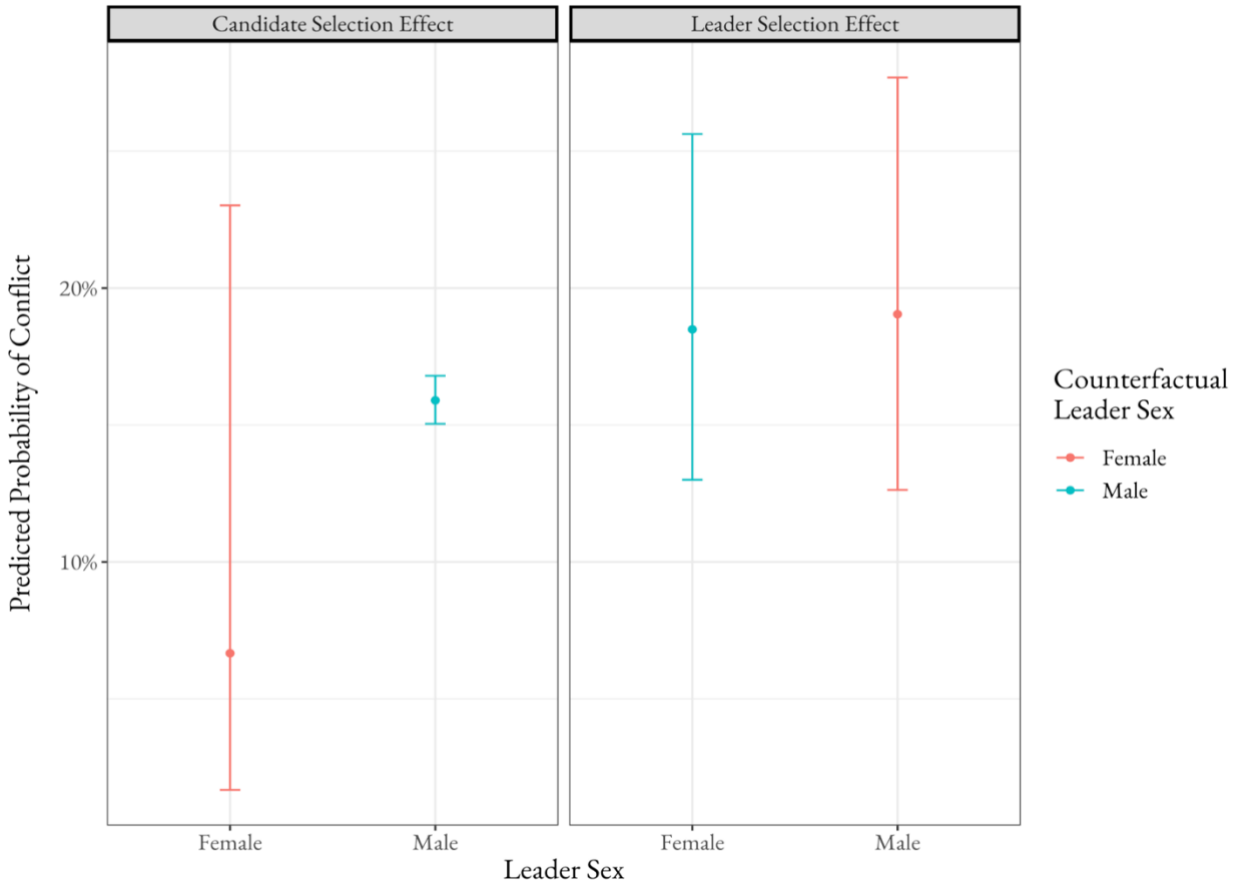


Figure 7-22: Predicted probability of conflict by leader sex and selection effect

Because there is no evidence of a candidate or leader selection effect for female leaders, so there is no reason to check the placebo assumption. Furthermore, the results would likely be non-significant anyway due to the small sample of female leaders.

7.9.4 Sex: Supporting Hypotheses Tests

Although there is no evidence that candidate or leader selection effects explain why female leaders are more likely to initiate conflict, I include their supporting hypotheses tests for transparency.

Table 7-20: Candidate sex regression results.

	Outcome: Female Candidate			Outcome: Female Wins Election	
	(1)	(2)	(3)	(4)	(5)
Past Conflict	0.03* (0.01, 0.05)		0.03 (-0.01, 0.07)	0.05 (-0.07, 0.16)	
Gender Equality		0.09 (-0.50, 0.69)	0.11 (-0.55, 0.77)		-0.05 (-0.60, 0.50)
Constant	-3.88*** (-4.25, -3.52)	-3.54*** (-3.94, -3.14)	-3.69*** (-4.17, -3.21)	-0.19 (-1.23, 0.85)	0.07 (-0.69, 0.83)
Clustered SE	Individual	Individual	Country	Country	Country
Observations	4,356	3,115	3,115	40	31
Log Likelihood	-476.70	-403.80	-400.85	-27.27	-21.45
Akaike Inf. Crit.	957.39	811.61	807.69	58.54	46.91

Note: *p<.01; **p<.05; ***p<0.001

Model 1 shows that females are more likely than males to be candidates as past conflicts increase. However, Model 2 shows that gender equality is unrelated to female candidacy, paradoxically, contrary to the first prediction of the candidate constraints hypothesis. This is presumably due to the higher number of female candidates in South Asia (Pakistan, India, Bangladesh, and Sri Lanka), where gender equality scores tend to be significantly lower.

Furthermore, gender equality is negatively associated with past conflicts and insignificant,[225] contrary to the second prediction of the candidate constraints hypothesis. Model 3 regresses both explanatory variables together with country-level clustered standard errors. While past conflict is no longer significant in this model, when gender equality is excluded, it remains slightly significant even with country-level clustered standard errors ($p = .049$). Thus, we have weak evidence for the FIR hypothesis that conflict-prone states are more likely to have female candidates than conflict-averse states.

Models 4 and 5 test whether female candidates are more likely to be chosen over male candidates in appointments and free and fair elections where the top-two candidates are female and male. Model 4 shows that female candidates are more likely to be chosen over male candidates as past conflict increases, but the relationship is insignificant. Thus, there is no evidence to support an FIR leader selection effect. Model 5 shows another paradoxical finding: female candidates are less likely to be chosen over male candidates as the gender equality score increases, contrary to the demographic hypothesis.

[225] Bivariate OLS model with country-level clustered standard errors. Coefficient = -.14, $p = .20$.

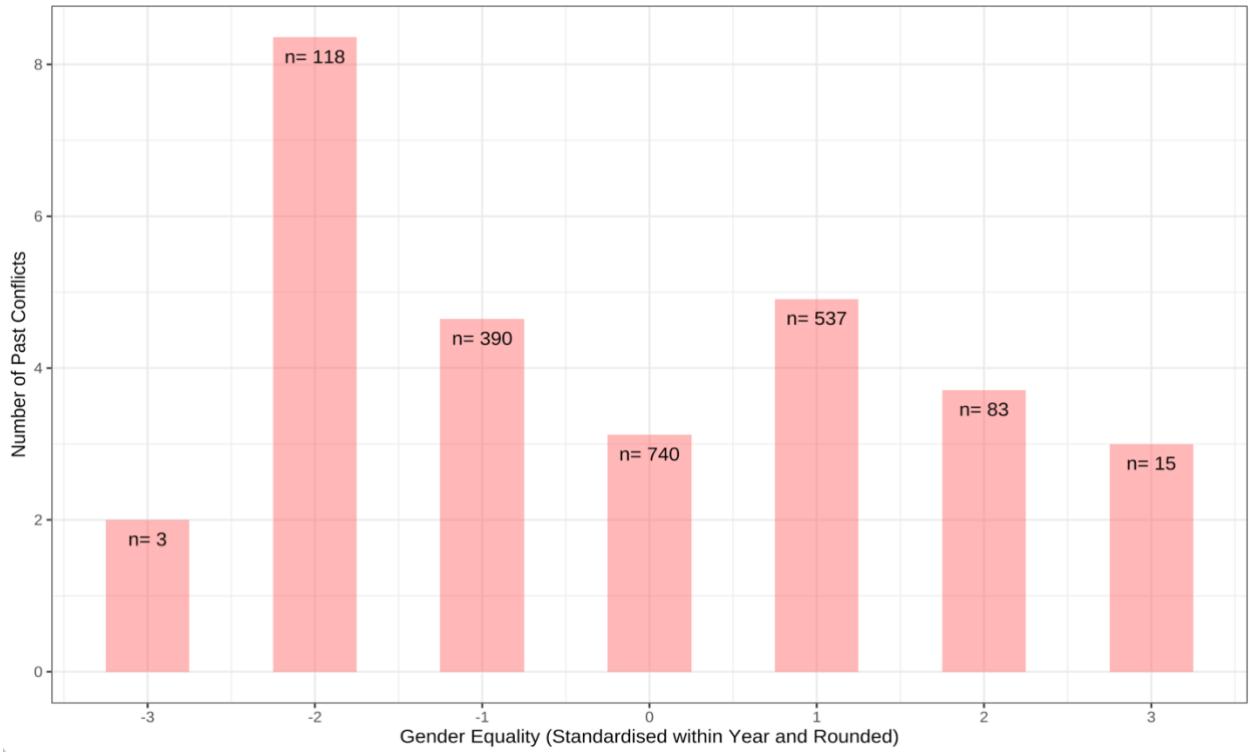


Figure 7-23: Average number of past conflicts by gender equality score.
Note: Gender equality is standardised within year and rounded to the nearest whole number.

7.10 Ideology

7.10.1 Ideology: Summary Statistics

Using the HoG dataset, ideology was identified for 33% (n=897) of leaders and 19% (n=679) of counterfactual leaders. Because ideology likely correlates with other traits, Table 7-21 below lists the cross-trait correlations. As expected, right-wing candidates are more likely to be military veterans, older, and male than left-wing candidates.

Table 7-21: Percentage of candidate ideology by traits.

	N	Veteran	Combat Veteran	War Experience	Age	Female
Left-Wing	256	13%	12%	8%	55.3	4%
Centre	447	10%	7%	6%	56.4	1%
Right-Wing	305	24%	18%	18%	58.4	2%

7.10.2 Ideology: Selection Effect Hypotheses Tests

It is first necessary to note, however, that during analysis several issues with ideology became apparent and will likely bias the results significantly. First, not only are a significant number missing as noted previously, but there is a potential problem of autocorrelation since the overwhelming majority of candidates are missing ideology in Africa (99%) and Asia (95%). Indeed, ideology is missing from all candidates in some of the more notable conflict-prone countries, such as India, Turkey, and Israel. Secondly, there is a substantial lack of variation in candidates' ideology within countries; all of Switzerland's 111 candidates are centre, all but one of Japan's 37 candidates are right-wing, the U.S. has no left-wing candidates (only centre and right-wing), and the U.K. has no centre candidates. Accordingly, country-level random effects are a potential problem but cannot be controlled for, given these two problems. Thus, it is necessary to keep these significant limitations in mind in the subsequent analysis. Furthermore, I do not include self-selection in the regression results again due to the small number of observations.[226]

[226] Among self-selections, there are only a total of 99 observations and 11 conflict initiations where ideology is observed for both candidates.

Table 7-22: Summary of selection effect hypothesis for leader ideology.

Selection Effect Hypotheses	Treatment Group		Comparison Group		Supporting Hypotheses
	Leader Ideology	Counterfactual Leader Ideology	Leader Ideology	Counterfactual Leader Ideology	
Candidate Selection Effect	Right-Wing	Right-Wing	Left-Wing, Centre	Left-Wing, Centre	FIR
Leader Selection Effect	Right-Wing	Left-Wing, Centre	Left-Wing, Centre	Right-Wing	FIR
Candidate Self-Selection Effect	Right-Wing	Right-Wing	Left-Wing, Centre	Left-Wing, Centre	Attempt
Leader Self-Selection Effect	Right-Wing	Left-Wing, Centre	Left-Wing, Centre	Right-Wing	Success

7.10.3 Leader Ideology: Main Findings

Table 7-23: Leader ideology regression results

	Population Treatment Effect (1)	Sample Treatment Effect (2)	Placebo Treatment Effect (3)	Candidate Selection Effect (4)	Leader Selection Effect (5)
Centre	-0.27 (-0.97, 0.43)	-0.32 (-1.20, 0.56)		-0.38 (-1.90, 1.14)	0.44 (-0.66, 1.53)
Right-Wing	0.62* (0.02, 1.22)	0.23 (-0.81, 1.27)		-0.22 (-2.76, 2.32)	0.27 (-0.84, 1.38)
(CL) Centre			NA		
(CL) Right-Wing			0.48 (-0.31, 1.26)		
Constant	-2.53*** (-2.96, -2.10)	-2.26*** (-2.93, -1.59)	-2.48*** (-2.99, -1.98)	-2.92*** (-4.23, -1.60)	-2.18*** (-2.91, -1.45)
Observations	3,104	1,009	1,054	396	613
Log Likelihood	-888.51	-303.34	-314.68	-63.69	-228.82
Akaike Inf. Crit.	1,783.03	612.69	635.36	133.39	463.64

Note: Robust standard errors clustered by individual leader. *p<05; **p<.01; ***p<0.001

Table 7-23 above shows the regression results for leader ideology and selection effects. Model 1 shows that right-wing leaders are more likely than left-wing leaders to initiate conflict

absent any other factors. Model 2 shows that the sample treatment effect is substantially lower than the population treatment effect and is no longer significant, suggesting that counterfactual leaders with missing ideology biases the population treatment effect. This is not unexpected, however, considering that the sample size decreases by 60% when the missing counterfactual leaders are excluded. Model 3 shows the placebo treatment effect, which is positive but insignificant, presumably again due to the small number of observations.

For selection effects, Models 4 and 5 show the results for candidate and leader selection effects, respectively, and Figure 7-24 provides a more detailed visualisation of each. Unfortunately, the candidate selection effects for left and right-wing leaders are largely uninterruptable since there are rarely selection events in which both the leader and counterfactual leader have the same ideology.[227] The lack of counterfactual leaders' ideology also results in too few observations to reliably interpret leader selection effects, although not to the same degree since there are 385 observations between a right-wing and left-wing candidate. In any case, there is no evidence of leader selection effects for right-wing leaders since the confidence intervals overlap with the other ideology. For instance, the model predicts that selecting a right wing over a left-wing candidate leads to a 6% chance of conflict, whereas selecting a left-wing over a right-wing candidate leads to a 10% chance of conflict. Thus, no evidence supports the leader selection effect hypothesis for why right-wing leaders appear more likely to initiate conflict.

[227] There are only 39 observations where both were left-wing and 48 observations where both were right-wing, which is significantly smaller compared to the 308 when both were centre.

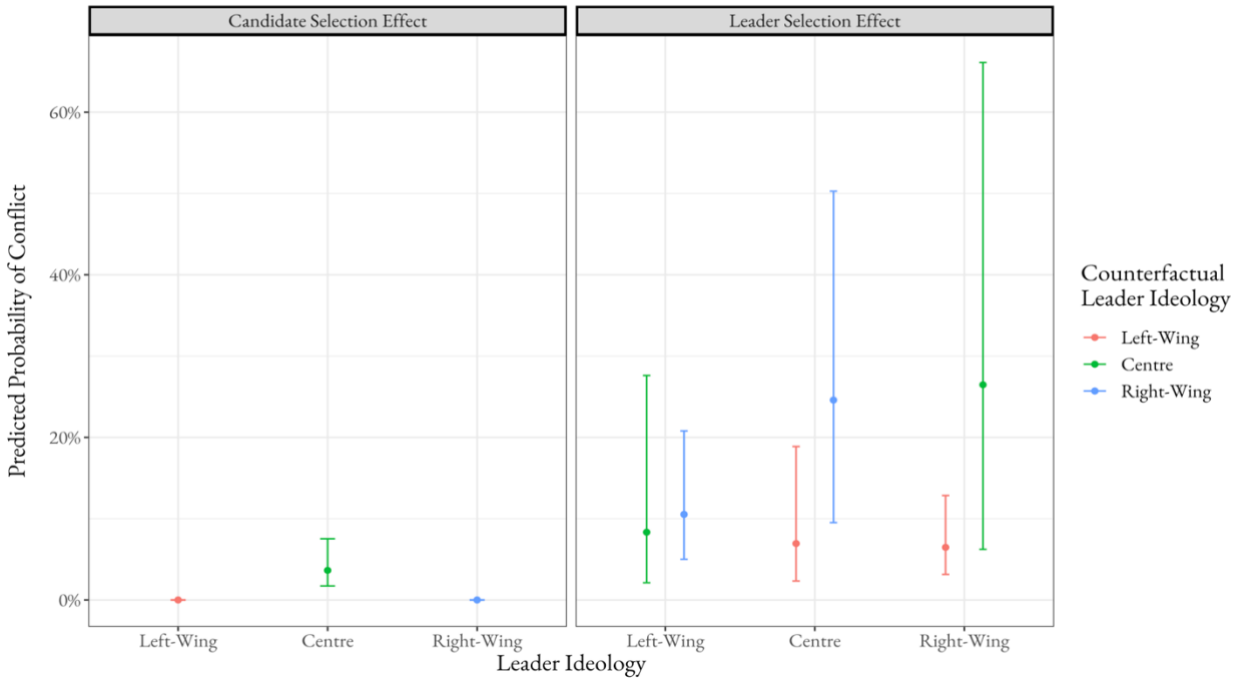


Figure 7-24: Predicted probability of conflict by leader ideology and selection effect

7.10.4 Ideology: Supporting Hypotheses Tests

Although there is no evidence for either candidate or leader selection effects above, it is still necessary to investigate whether we see evidence of either effect in the supplementary analysis. Like the other traits above, I use the LSD – Candidate dataset. However, due to the few free-and-fair elections between candidates of different ideologies, I only test whether right-wing candidates are more likely to be chosen when the runner is left-wing or centre. Table 7-24 below shows the results.

Table 7-24: Candidate ideology regression results.

	Outcome: Right-Wing Candidate		Outcome: Right-Wing Candidate Wins Election
	(1)	(2)	(3)
Past Conflict	0.08** (0.03, 0.13)	0.08 (-0.01, 0.16)	-0.02 (-0.06, 0.03)
Constant	-1.00*** (-1.28, -0.71)	-1.00*** (-1.55, -0.44)	-0.25 (-1.02, 0.52)
Clustered SE	Individual	Country	Country
Observations	1,214	1,214	104
Log Likelihood	-739.90	-739.90	-70.27
Akaike Inf. Crit.	1,483.80	1,483.80	144.53
Note:			*p<.01; **p<.05; ***p<0.001

Model 1 shows that past conflict is associated with the selection of right-wing candidates, although the relationship is no longer significant in model 2 when country-level clustered standard errors are included. Nonetheless, considering that previous research does not adjust for country-level correlation, this adds some support to the candidate selection effect hypothesis for why right-wing leaders appear more likely to initiate conflict. That is, candidates are more likely to be right-wing than left-wing in more conflict-prone states, but this is not because right-wing candidates are more likely to be selected candidates as past conflict increases. Model 3 shows that increased past conflicts are not associated with whether a right-wing candidate wins an election over a left-wing or centre candidate, contrary to the FIR leader selection effect hypothesis.

7.11 Hawkishness

7.11.1 Hawkishness: Summary Statistics

Figure 7-25 shows the distribution of hawkishness scores among candidates. Most candidates have low hawkishness scores because the military-related variables increase the hawkishness score, and most of the candidates in the lowest bin are not ex-military. The

candidates with the highest possible hawkishness scores (1.47) are Kleber Dadjo of Togo and Muhammad Zia-ul-Haq of Pakistan.

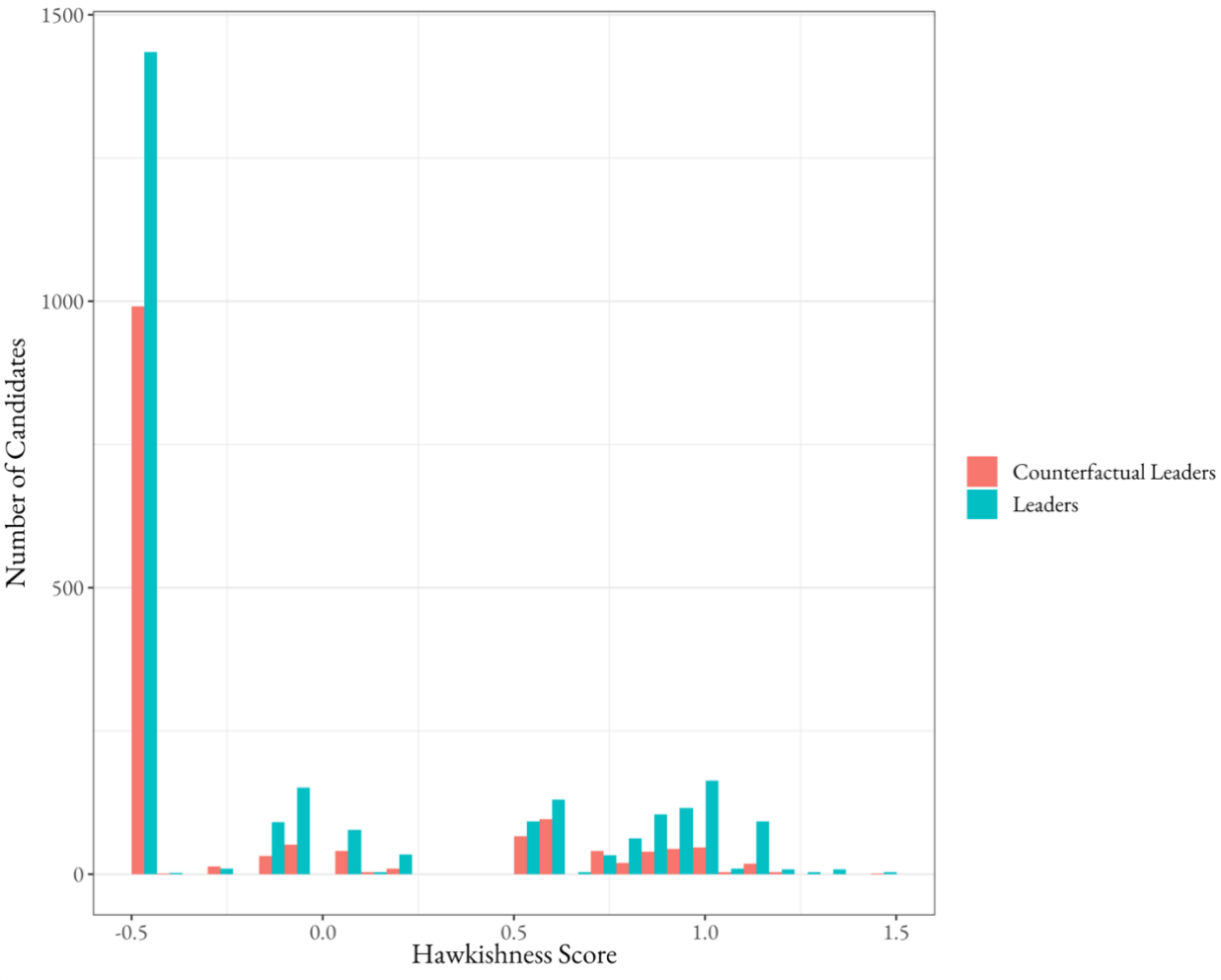


Figure 7-25: Histogram of hawkishness score by leaders and counterfactual leaders

Table 7-25 shows the average hawkishness score by each value of the other high-risk traits. Interestingly, the differences in average hawkishness between the values of some traits suggest the opposite of previous research. Combat leaders tend to be more hawkish than non-combat leaders, war losers more than war-winners, and males more than females. However, military experience is the primary trait driving higher hawkishness scores, so any candidate selection effect is likely to be explained by the same supporting hypotheses for military experience above.

Table 7-25: Hawkishness score among candidates for each trait.

	Military Experience		Military Experience Type		War Experience		Age		Sex		Ideology		
	No	Yes	Non-Combat	Combat	Loser	Winner	<57	>=57	Male	Female	Left-Wing	Centre	Right-Wing
Hawkishness (average)	-.38	.81	.72	.87	.89	.84	-.07	.01	-0.2	-.43	-0.16	-0.32	-0.02

7.11.2 Hawkishness: Selection Effect Hypotheses Tests

To test the initial selection effect hypotheses for hawkishness, I conduct the regression strategy as above for age since hawkishness is a continuous variable. Again, I only test the initial selection effect hypothesis since any findings supporting any of them are likely due to the same supporting hypotheses for the other traits above since hawkishness is constructed using those traits. Table 7-26 below lists the summary of hawkishness selection effect hypotheses for reference.

Table 7-26: Hawkishness summary of selection effect hypotheses.

Selection Effect Hypotheses	**Treatment Group**		**Comparison Group**		**Supporting Hypotheses**
	Leader Ideology	**Counterfactual Leader Ideology**	**Leader Ideology**	**Counterfactual Leader Ideology**	
Candidate Selection Effect	Hawk	Hawk	Dove	Dove	See other traits
Leader Selection Effect	Hawk	Dove	Dove	Hawk	See other traits
Candidate Self-Selection Effect	Hawk	Hawk	Dove	Dove	See other traits
Leader Self-Selection Effect	Hawk	Dove	Dove	Hawk	See other traits

7.11.3 Leader Hawkishness: Main Findings

Table 7-27: Hawkishness selection effect regression results.

				Selection Effects	
	Population Treatment Effect	Sample Treatment Effect	Placebo Treatment Effect	Non-Self-Selected Leaders	Self-Selected Leaders
	(1)	(2)	(3)	(4)	(5)
Hawkishness	0.63*** (0.42, 0.84)	0.57*** (0.30, 0.84)		0.48** (0.14, 0.81)	-0.16 (-0.92, 0.60)
(CL) Hawkishness			0.99*** (0.64, 1.34)	0.60** (0.23, 0.96)	0.99* (0.02, 1.96)
Hawkishness *(CL) Hawkishness				0.43 (-0.15, 1.01)	-0.26 (-1.30, 0.77)
Constant	-1.93*** (-2.10, -1.77)	-1.72*** (-1.94, -1.51)	-1.64*** (-1.84, -1.45)	-1.82*** (-2.02, -1.63)	-1.28*** (-1.97, -0.58)
Observations	10,573	6,202	6,296	4,701	1,501
Log Likelihood	-4,113.44	-2,655.04	-2,656.88	-1,748.92	-804.25
Akaike Inf. Crit.	8,230.87	5,314.08	5,317.75	3,505.84	1,616.50

Note: Robust standards errors clustered by individual leader. *p<.01; **p<.05; ***p<0.001

Figure 7-26: Predicted probability of conflict by leader hawkishness and selection effect

Above, Table 7-27 shows the full regression results for hawkishness. Model 1 shows the population treatment effect of hawkishness score, confirming that leaders with higher hawkishness scores are more likely to initiate conflict than leaders with lower hawkishness scores. Model 2 shows the sample treatment effect of hawkishness is not sensitive to the exclusion of missing counterfactual leaders. Model 3 shows that the placebo treatment effect is positive, significant, and not statistically different from the sample treatment effect since the confidence intervals overlap. This initially confirms a positive selection effect for more hawkish leaders in environments where conflict is more likely given the placebo assumption.

Turning to the selection effects, model 4 shows that the interaction term between leaders' and counterfactual leaders' hawkishness score is not significant among non-self-selected leaders. However, given the selection effects are difficult to infer from coefficients alone, Figure 7-27 below visualises the results for models 4 and 5. To better visualise and interpret whether we have evidence for either the candidate and leader selection effects, I show the predicted

probability of conflict between leaders and counterfactual leaders with minimum and maximum hawkishness scores (-0.46 and 1.47, respectively), which I call "doves" and "hawks" for simplicity. Accordingly, the candidate selection effect shows the predicted probability of conflict when the leader and counterfactual leader are both hawks, compared to when the leader and counterfactual leader are both doves. Conversely, the leader selection effect shows the predicted probability of conflict for a dove when the counterfactual leader is a hawk, compared to the predicted probability of conflict when the leader is a hawk and the counterfactual leader is a dove. As the top row shows, among non-self-selected leaders, the predicted probability of conflict is significantly greater when both leaders and counterfactual leaders are hawks than when both are doves. Conversely, there is no leader selection effect in either selection mechanism – the probability of conflict is the same regardless of whether a hawk is selected over a dove, or vice versa. In other words, the results strongly support the candidate selection effect hypothesis that hawkish leaders are only more likely to initiate conflict when the counterfactual leaders have similar hawkishness scores.

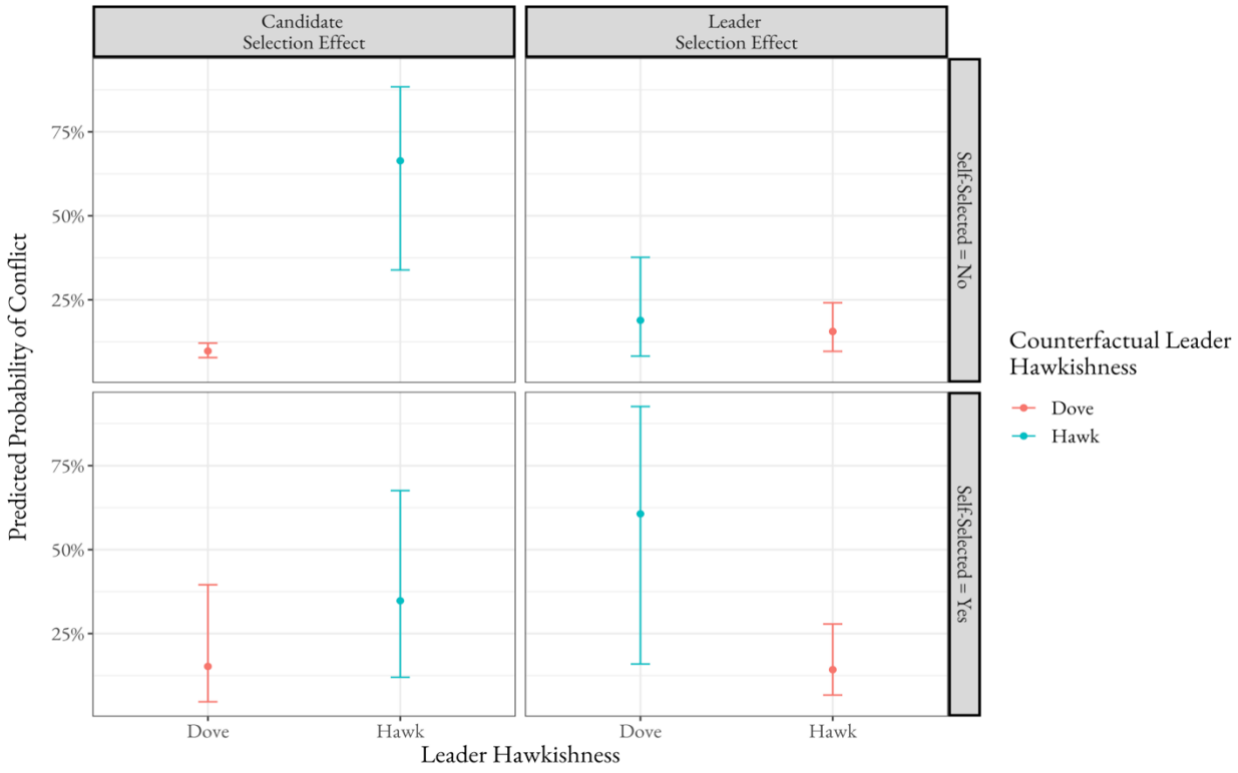


Figure 7-27: Predicted probability plot of leader hawkishness, counterfactual leader hawkishness, and selection effects.

7.12 Conclusion

Overall, the results consistently show that selection likely explains why leaders with high-risk traits appear more likely to initiate conflict than those with low-risk traits. More specifically, leaders with high-risk traits only appear more likely to initiate conflict because they are more likely to be candidates in environments where conflict is more likely. Only the results among female and right-wing leaders do not support any selection effect hypotheses, although this is more difficult to infer given the small number of observations. Therefore, the totality of evidence favours the environmental perspective hypothesis that all of the military-related variables, age and hawkishness, do not explain why leaders start conflicts. If we could control for the environmental risk, I suspect that none of the traits would be positively associated with conflict initiation. Thus, what matters is the environment leaders with high-risk traits are selected into, not their traits.

8 Discussion and Conclusion

"To what extent do the personal traits of leaders influence their decision to start international conflicts?" This question I set out to answer from the beginning. While previous research on this question suggests they have a significant influence, my findings suggest otherwise. Rather than influencing conflicts, my results suggest that leaders with these traits are just selected into environments where conflict is more likely. Selection likely explains why leaders with perceived high-risk traits, such as military experience and old age, appear to be the most conflict-prone. In short, the personal traits of leaders likely have little to no influence on their decision to start international conflicts.

I have come to this conclusion based on two important findings for International Relations. First, the counterfactual leaders' traits are also positively associated with conflict for all traits investigated. This "placebo" effect is not significantly different to the "treatment" effect (i.e., the effect of the leader's trait on conflict). In other words, the traits of counterfactual leaders have the same effect on conflict as the traits of leaders. How can this be when counterfactual leaders likely have no meaningful influence on the leader's decision-making? I argue that selection can only explain this. If the effect of leader traits on conflict were not biased by selection, we would expect the counterfactual leader's trait to have no effect. Considering we see the same effect, there is considerable reason to be sceptical that the effect of the counterfactual leader's traits on conflict is due to selection, but not the leader's traits.

The findings further suggest that the positive placebo effect is due to candidate selection effects. More specifically, leaders with high-risk traits are more likely to initiate conflict than those with low-risk traits *only* when the counterfactual leader has the same traits. Put differently, if both candidates are veterans, whichever leader gains power will result in more conflicts than if both were non-veterans. However, electing a veteran over a non-veteran does not lead to more conflicts than the opposite. I argue these two differences suggest that ex-military leaders only

appear more likely to initiate conflict than their counterparts with no military experience because veterans are more likely to be candidates in environments with a greater probability of conflict.

My findings further suggest that the candidate selection effects are likely due to demographics rather than a First-Image Reversed preference, or random noise. For all the military-related variables, the number of past conflict involvements and the estimated veteran population percentage predict the military experience of both leaders and counterfactual leaders. Likewise, past conflict involvements and the estimated veteran population percentage are significantly associated. These two findings suggest that ex-military candidates are more likely than non-ex-military candidates to be selected into conflict-prone states because these states tend to have a higher percentage of veterans in the population. However, when adjusting for within-country correlation, only the estimated veteran percentage remains robust when predicting the military experience of leaders and counterfactual leaders. For age, the findings are similar except the demographic variable is life expectancy. That is, conflict-prone states are more likely than conflict-averse states to have a larger pool of older candidates from which to choose their leaders. For biological sex and ideology, I do not find evidence of selection effects, although this is due to the small number of cases where leaders and counterfactuals have the same sex or ideology. However, considering the placebo effect for both biological sex and ideology is positive, this does not entirely rule out selection as an explanation.

Table 8-1: Summary of hypotheses supported.

Trait	High-Risk Trait Type	Observed Selection Effect	Hypotheses Supported
Military Experience	Ex-Military	Candidate	Demographics
Military Experience Type	Non-Combat Ex-Military	Candidate	None
War Experience Type	War-Win Ex-Military	Candidate	Demographics
Age	Older	Candidate	First-Image Reversed Affinity Bias
Sex	Female	None	None
Ideology	Right-Wing	None	None
Hawkishness Score	Hawkish	Candidate	NA

While my findings and core argument suggest that leader traits such as age and military experience have no causal effect on leaders' decision to start conflicts (and sex and ideology to a lesser although more uncertain extent), I do not argue that leader traits *in general* do not, or will never, have a direct causal effect on some international political outcome. While I did not investigate personality constructs in this dissertation primarily due to the difficulties mentioned in Chapter Two, my ontological position is that personality probably explains many international political outcomes better than some of the more widely investigated structural and institutional factors, such as military power and regime type.[228] However, this position is difficult to empirically support until measures of personality can be more accurately and consistently applied across a wide range of leaders. Furthermore, there is also a chicken-and-egg dynamic between non-biological traits and personality (e.g., does the military make people more hawkish or are hawkish people more likely to join the military?), which adds more difficulty to testing this position even if we could accurately measure personality across leaders. Nevertheless, ignoring

[228] My position is similar to that expressed in Nye (2020). While Nye investigates morals more specifically, I believe that these are probably one of the larger components of the causal effect of personality on foreign police decision making. See: Joseph S. Nye, *Do Morals Matter?: Presidents and Foreign Policy from FDR to Trump* (Oxford University Press, USA, 2020).

selection effects and environment causes for the moment, my core argument is that any leader-level causes of conflict stem from extreme personalities rather than individual traits.

8.1 Anticipating Objections: Counterfactual Leaders

Perhaps the most likely objection to my research is how I operationalise the concept of counterfactual leaders. Indeed, the fundamental assumption of my research strategy is that these individuals were the "most likely" person to be selected leaders in a counterfactual world given a minimal rewrite of history. There are two points to discuss here. First, it could be argued that many of these individuals are not the most-likely counterfactual leaders, and instead could be someone else. Recall the scenario between Walter Mondale and Gary Hart in the 1984 U.S. election and Democratic primary from Chapter Four. In my data, Walter Mondale is the counterfactual leader following the 1984 U.S. election. However, considering he lost in a landslide, it could be argued that had Gary Hart won the much more closely contested 1984 Democratic Party primaries, he could have been more likely to beat Reagan than Mondale. In other words, a minimal rewrite of history would mean that Hart should be considered the counterfactual leader – not Reagan.

I fully expect that my determination of counterfactual leaders could be argued differently for several or more cases. However, the Hart and Mondale example only reinforces my findings for candidate selection effects. Both Mondale and Hart were all military veterans, and so were Reagan and his Vice-President at the time George H.W. Bush. So, the U.S. would likely have had an ex-military leader in 1984, no matter how minimally we rewrite history. In short, I suspect that even if we could observe the true counterfactual leaders in each selection event, my findings regarding candidate selection effects would remain robust.

The second point is that there are two types of counterfactual leaders – the selected counterfactual leader and the random-death counterfactual leader. In the first case, these are the individuals I focus on in this thesis – those who were not selected leader but were next closest to. In the second case, I did not focus on these individuals but are nonetheless part of the same

core idea – those who are the "most likely" leader in a counterfactual world, except only if the current leader were to die randomly. To yet again use Reagan as an example, the counterfactual selected leader for Reagan's first term is Jimmy Carter since he was the incumbent and lost the 1980 election. The random-death counterfactual leader for Reagan's first term is George H.W. Bush since he was the Vice-President and would have inherited the Presidency if Reagan were to have randomly died during his term. In the initial stages of formulating my research strategy I originally planned to include the random-death counterfactual leader since they satisfy the "minimal-rewrite rule" more so than the selected counterfactual leader. That is, we only need to assume that the leader had a worse diet and died of a heart attack or was travelling in a car at the wrong place in time. However, I ultimately decided not to include random-death counterfactual leaders for several reasons. First, using these individuals only allows testing for the candidate selection effects since they were not selected against in elections or coup attempts. Indeed, George H.W. Bush was selected *for* Vice-President. Second, the placebo assumption is considerably more likely to be violated since those who would inherit the position tend to be Vice-Presidents or Deputy Prime Ministers – individuals who have a meaningful effect on foreign policy. Lastly, there is considerably more uncertainty as to which counterfactual leader would likely be in states with no constitutional succession mechanism. For instance, the U.K. does not have any legal framework for determining the Prime Minister following the death of the incumbent. Thus, with these three problems, I decided that their inclusion would not greatly benefit the analysis or help to answer the research question.

In Chapter Two I framed the research question as fitting within an environmental-trait framework although there is a related conceptual framework I did not previously discuss; ultimate/proximate causation. In essence, ultimate causation is the "why" some outcome occurred, whereas proximate causation is the "how" and this framework is mostly used in biology to help frame the difference between evolutionary explanations for animal behaviour (i.e., why), such as reproductive fitness, and functional reasons (i.e., how), such as cognitive

processes.[229] Nevertheless, the ultimate-proximate framework could also be used similarly for leader-level analyses in international relations. For instance, if male leaders were more likely to start conflicts and wars than female leaders, one ultimate cause could be that males are more risk-prone than females *in general* as a result of evolution through natural selection (e.g., males who were risk-prone left more descendants and hence genetic material favouring risk-proneness than males who were risk-averse) and so male leaders then would be more risk prone than female leaders, on average. The proximate cause for this ultimate explanation would then be the endocrinological processes that predispose females to be more risk averse than males. Conversely, another potential ultimate cause of this disparity could be socialisation, where male leaders might have learned to be more risk-prone than female leaders as a result of gender norms. The proximate cause for this ultimate explanation, then, would be the social transmission of cultural norms, or also any social incentives/constraints of how leaders are expected to act given their gender.

The ultimate, or "why", and proximate, or "how", of leader behaviour also naturally falls within a temporal spectrum; ultimate causes will be further back in time than their proximate causes from the outcome. Indeed, the ultimate-proximate distinction is implicitly framed in any Directed Acyclic Graph (DAG) – since the standard practice for creating DAG's is to temporally order all explanatory variables from the outcome variable by left-to-right or top-to-bottom, then parent explanatory variables are more ultimate causes whereas descendant variables of these are more proximate. Accordingly, all the DAGs I implemented previously in this dissertation implicitly use the ultimate/proximate distinction since I temporally order them from left-to-right and also separate the same variables into time periods. If ex-military leaders were more likely to initiate conflict, for example, and people were voting for ex-military over non-ex-military leaders

[229] Thomas C. Scott-Phillips, Thomas E. Dickins, and Stuart A. West, 'Evolutionary Theory and the Ultimate–Proximate Distinction in the Human Behavioral Sciences', *Perspectives on Psychological Science* 6, no. 1 (2011): 38–47; André Ariew, 'Ernst Mayr's' Ultimate/Proximate'distinction Reconsidered and Reconstructed', *Biology and Philosophy* 18 (2003): 553–65.

in response to a hostile international environment (i.e., FIR leader selection effect), then the why, or ultimate causal explanation, is that ex-military leaders are more likely than non-ex-military leaders to initiate conflict because people choose them in dangerous times. Accordingly, the how, or proximate causes, would then be what happens after selectorate's preferences for leader change in response to the international system and the moment before the conflict initiation. In other words, the environment to conflict initiation causal path is the ultimate explanation and any intermediate causal paths within the environment-conflict initiation causal path are proximate explanations.

8.2 The Problem with Leader Risk

Although I did not investigate the leader risk score in my analysis due to autocorrelation, my findings also explain why "high-risk" leaders appear conflict-prone in the findings of *Why Leaders Fight* (WLF). In Horowitz and Stam's model to construct the risk score, a peace-year cubic spline is included (the total number of years since the leader's state last initiated a conflict, cubed), but this is exogenous to leaders and strongly correlates with conflict initiation. According to their model, the peace-year cubic spline has a variable importance factor nearly four times higher than the next closest leader attribute, non-combat military experience, and nearly twenty times that of most other leader attribute variables.[230] As a result, this necessarily models in systemic-level factors and artificially elevates the risk score for leaders in states that frequently engage in conflict. This is evident when plotting a leader's risk percentile with the system's risk percentile (Figure 8-1). Both are positively and linearly correlated because the large effect of past conflict elevates the leader risk score to match that with the systems risk scores.

[230] According to Horowitz and Stam's leader risk model, the variable importance factor for "years since conflict initiation" variable is 24. For military non-combat experience, it is 7, and between 0 and 2 for most other attributes.

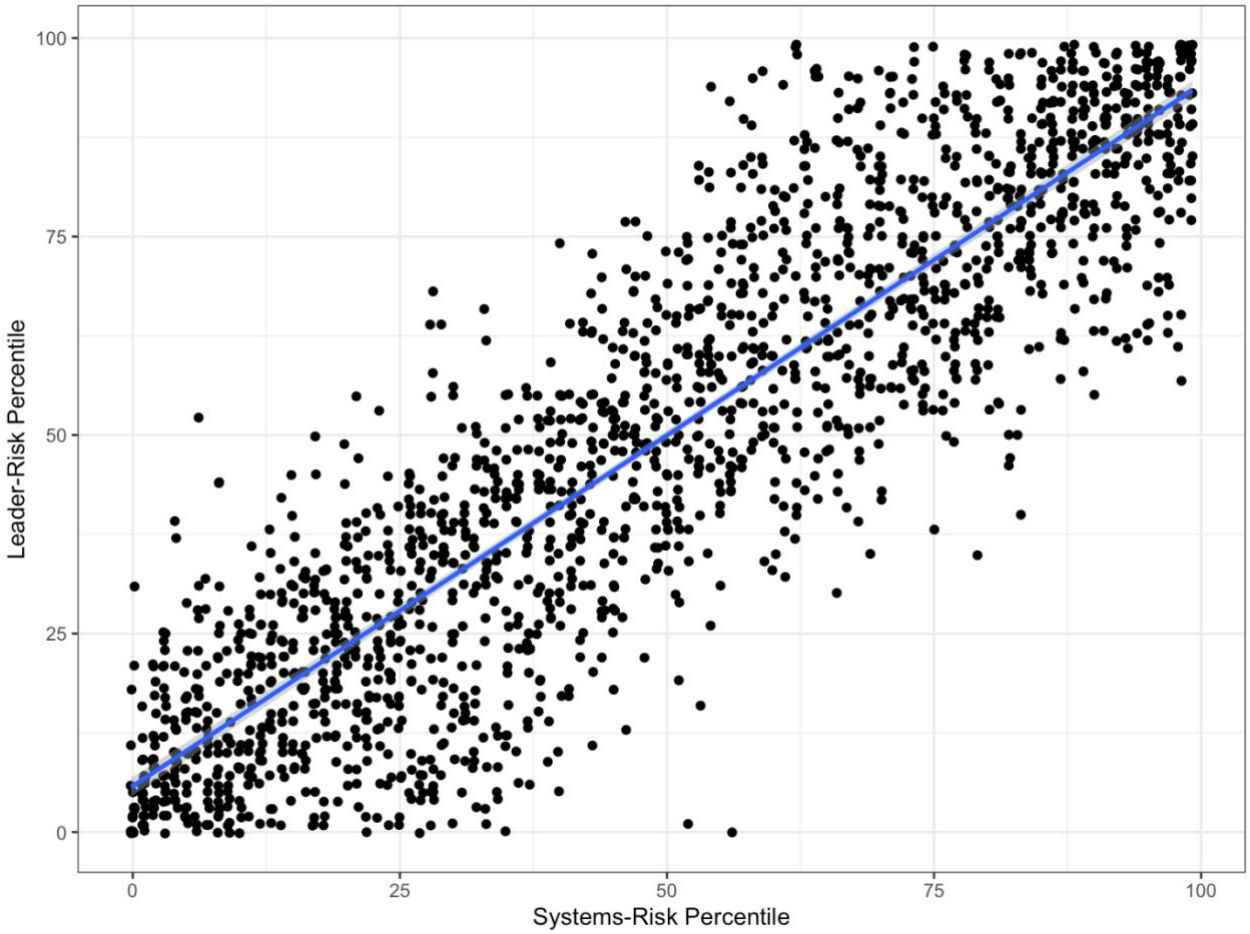


Figure 8-1: Correlation between leader risk and systems risk percentiles

It is not surprising, then, that Horowitz and Stam found that the leader risk score predicts some of the most conflict-prone leaders. The most important systems-level factor inflates risk scores for leaders where conflict is more frequent, hence the correlation with systems risk. When reconstructing Horowitz *et al.'s* leader risk score without the peace-year spline, we see very different results of how well leader risk compares to systems risk in predicting the most conflict-prone leaders (Table 8-2).[231]

[231] Top 2% of leaders with the most conflict initiations.

Table 8-2: Comparison of predictive value between original and revised leader risk percentiles

	Percentage of cases where leader risk performs _____ compared to systems risk		
	Worse	Same	Better
Leader risk (original)	43%	2%	55%
Leader risk (without peace-year spline)	71%	5%	24%

The revised leader risk performs much worse when the peace-year spline is excluded. The inclusion of systems-level variables introduced the selection effect into their models and made it appear as though high-risk leaders were considerably more likely to initiate conflict than low-risk leaders. In other words, what made leaders "high-risk" in Horowitz and Stam's findings is largely due to being selected into high-risk environments, not their backgrounds or attributes. While the risk score does appear to better explain 24% of the most conflict-prone leaders, since my findings have shown evidence for selection effects, it could be that the variation explained by the traits is just accounting for the missing variation explained by the systems risk.

8.3 Issues Regarding International Conflicts as the Outcome

While this research has solely investigated how selection confounds the relationship between leader traits and conflict initiation, there are also several other issues with using Militarized Interstate Dispute (MID) data discovered during data analysis that I did not anticipate.

8.3.1 Measurement Bias: Militarized Interstate Disputes as Outcome Variable

One significant issue that has become apparent is that many international conflicts do not appear to be the product of leader decision-making. I originally went into this study assuming that leaders' decision-making caused every conflict since that is what previous research had assumed, and international conflicts are defined in a way that suggests leaders are responsible. But when looking more closely at how international conflicts are operationalised in the MID dataset, it has become apparent this assumption is questionable.

To understand why, first consider how international conflicts are defined in the MID dataset - "an overt action taken by the official military forces of representatives of a state."[232] Furthermore, this overt action by military personnel must be an "explicit, non-routine, and governmentally authorised action" against the target's states military forces, or in the case of civilians, any seizures, attacks or pursuit across international boundaries.[233] However, nowhere is the requirement that the decision to use force must result from a leader's decision-making. While leaders probably do indeed order their military forces to carry out the conflict in many cases, a cursory review of the case narratives in the MID dataset, however, suggests many conflicts do not appear to be a result of the current leader's decision-making. The clearest example of which is the number of conflicts that are nothing more than brief confrontations between naval vessels and fishing boats. Consider the case narratives of three conflicts:

1) "Canadian coast guard boats seized a Portuguese fishing boat suspected of fishing in Canadian Waters."[234]

2) "The U.S. Coast Guard seized a Cuban fishing boat because it was taking lobsters illegally."[235]

---

[232] Gochman and Maoz, 'Militarized Interstate Disputes, 1816-1976', Coding Manual, 1.
[233] Ibid., coding manual, 2.
[234] MID 4550. Gibler, *International Conflicts, 1816-2010*, 2:53–54.
[235] MID 2958. Ibid., 2:7–8.

3) "A South Korean naval patrol intercepted a Chinese [fishing] vessel and attempted to seize it. The Chinese instead rammed the patrol, a fight ensued, and the fishing boat sank, killing three of the fisherman."[236]

It is questionable that the leader for these three would be primarily responsible for initiating these three conflicts. Indeed, the leaders in each case were probably unaware these events occurred until the other state protested. Fishing disputes like these are not rare and appear numerously throughout the MID dataset. Since they are not separately categorised, however, researchers have little choice but to include them in their analyses unless each case is individually identified and removed.

Border skirmishes are another type of conflict frequent throughout the MID dataset yet do not appear to result from the current leader's decision-making. While these are more likely to involve a leader's decision-making than fishing disputes, many seem to be just the result of rogue soldiers engaging foreign soldiers on their own initiative. For example, in 1986, Soviet and Chinese troops briefly exchanged fire across the border, but a later investigation determined the clash was accidental, and both sides apologised.[237] Given that the border clash was unintentional (or at least claimed to be), it is questionable to infer that some aspect of the leaders at the time, Mikhael Gorbachev and Deng Xiaoping, were causally linked.

Another seemingly irrelevant type of conflict frequently mentioned throughout the MID dataset is isolated aerial engagements. Again, consider the three following case narratives:

1) "An American civilian propeller plane crashed after being confronted by Egyptian jets for crossing into Egyptian airspace."[238] (Egypt initiated conflict)

2) "East German officials seized American soldiers after they were forced to land their helicopter in East Germany due to a storm."[239] (U.S. initiated conflict)

---

[236] MID 4493. Ibid., 2:802.
[237] MID 2718. Ibid., 2:1001.
[238] MID 2906. Ibid., 2:507–8.
[239] MID 2854. Ibid., 2:206.

3) "Two American fighters buzzed a Libyan civilian airliner over Greece"[240] (U.S. initiated conflict)

Again, we can see the problem with using these types of disputes as evidence for some aspect of a leader's decision-making. However, the more dominant structural theories of why states – not leaders – initiate conflict are not susceptible to this problem as they make no assumptions about national-level decision-making. Even if rogue soldiers start a border skirmish on their accord, structural explanations are still valid because the system could still have caused the event, regardless of the individual who caused it. All that matters is that a hypothesised structural variable, say relative power differences, caused a chain of events resulting in a border skirmish. The trait perspective, on the hand, is highly dependent on the assumption that conflicts result from a leader's decision-making since a leader's traits can obviously only have a causal effect through the leaders.

While there is no doubt that leaders are probably responsible for many conflicts, leaders do not – and cannot – micromanage the tactical minutiae of how their military forces operate. Accordingly, leaders will delegate some decision-making authority to their subordinates. While a leader is almost certainly involved in the decision-making process that leads to launching a cruise missile into Sudan (e.g., Bill Clinton), leaders probably have better things to do with their time than order sailors to seize fishing boats. Many leaders do not even have full control over their armed forces in the way the U.S. President does. Though, even the U.S. President does not have the sole authority to declare war, at least formally.

8.3.2 The Problem of Predecessors

Another complication is that conflict initiations could be due to a predecessor's foreign policy, not the current leader. Consider a 1983 dispute under Ronald Reagan in which the U.S.

[240] MID 2978.

sent an aircraft carrier into Libyan waters as a "show of force."[241] It is fairly well known that Reagan was no friend to Libya during his Presidency, so it is not unreasonable to assume this dispute directly resulted from his decision-making. However, in 1979 the Carter Administration gave the go-ahead to naval commanders to move within three miles off the Libyan coast, which was unsurprisingly protested by Libya since they claimed their territorial borders extended up to twelve miles.[242] Thus, it is possible that Reagan had nothing to do with this dispute since the authorisation to go into Libyan territorial waters preceded Reagan. The carrier Captain could have followed orders by a superior to take the ship somewhere, and simply took the shortest route through Libya's claimed maritime boundaries because he had the authorisation to do so. Thus, we cannot just assume that a "show of force" like this is related to a leader's decision-making, especially when the policy granting that action predates the current leader.

8.4 Are International Conflicts Really Conflicts?

Despite the uncertainty as to whether leaders are responsible for initiating conflicts, another issue is whether conflicts should even be included in the first place. Again, the MID dataset uses several conditions for case inclusion, and I have no objections to how they operationalise international conflicts. Whether or not conflicts actually meet these conditions, however, is another story. The MID codes for disputes are based on newspapers, which usually just say *what* happened, but rarely give enough context to determine whether it meets their case inclusion criteria, let alone whether the leader was involved. Consider another "show of force" in the MID dataset under Ronald Reagan, this time against Yemen in 1982.[243] There is no case narrative for this dispute, but a deeper investigation indicates this is a biannual U.S. exercise,

[241] MID 3065.
[242] Gibler, *International Conflicts, 1816-2010*, 2:496.
[243] MID 3613.

Operation Bright Star, that was first signed into action by Jimmy Carter.[244] Since this is a routine exercise that predated his time in office, this should not be considered a Reagan-initiated conflict. How many more conflicts are similarly the product of predecessors' decision-making remains uncertain.

8.5 Endogeneity in Conflicts

Another issue with using international conflicts as the outcome is that for a conflict to occur, the target state must protest the actions of the initiator state.[245] This means that endogeneity is built into measuring conflicts since weak states are probably less likely to protest the actions of powerful initiators than vice versa. Conversely, conflicts started by weak states are probably more likely to be included since other states have little to fear in protesting their actions. This means many instances where leaders of powerful states initiated a conflict are not included in my analysis since the target state was weaker and did not protest. The problem of missing conflicts is evident when comparing the MID dataset to a similar one of the U.S. Use of Force (USUOF).[246] The USUOF dataset only includes uses of military force if there is evidence the U.S. President explicitly decided to use force and does not exclude them based on whether the target state protested. Accordingly, many irrelevant events are included and relevant events are excluded in the MID dataset – 51% of disputes are not considered U.S. uses of force, and only 36% of disputes cover all U.S. uses of force between 1870 and 1992.[247] This difference also appears to affect the relationships between leader traits and international conflicts. When using the USUOF instead of the MID data, Wharton finds that younger American Presidents are more likely to use force than older Presidents, and those with military careers are more likely to use

[244] Jason Brownlee, *Democracy Prevention: The Politics of the US-Egyptian Alliance* (Cambridge University Press, 2012), 39.
[245] Gochman and Maoz, 'Militarized Interstate Disputes, 1816-1976'.
[246] Benjamin O. Fordham and Christopher C. Sarver, 'Militarized Interstate Disputes and United States Uses of Force', *International Studies Quarterly* 45, no. 3 (2001): 455–66.
[247] Ibid., 463–64.

force than those without (contrary to the findings in WLF that military careerists are less likely to initiate conflict than veterans without military careers).[248] Thus, my findings are also sensitive to how conflict is measured.

High-level officials may order their government forces to carry out a militarised action based on policies created during a previous administration. In these cases, it is more difficult to understand how to interpret the decision-making role of a current leader.

8.6 Future Research

Throughout this thesis, several promising avenues of future research have become apparent. One potential option is re-evaluating whether a leader's political ideology influences conflict. Given the small number of leaders with ideology available in the only dataset which attempts to identify leaders' ideology – the Heads of Government dataset – there is a large research gap to be filled by identifying the ideology of significantly more leaders. Furthermore, this potential avenue for future research could also be improved by including the ideology of counterfactual leaders to ensure that any findings are not biased due to either candidate or leader selection effects using the same methods I utilised in this dissertation.

Future research should also investigate whether the MID dataset is biased due to excluding events in which the target state did not protest the militarised actions of the initiator state. Again, weak states are probably less likely to protest disputes started by powerful states, and so it is likely that there is a disproportionate number of dispute dyads between powerful states. If this were the case, then research on why equal dyads appear more likely to result in conflict than unequal dyads could be biased. Instead of disputes being more likely to occur among equal dyads, it could be that disputes are only more likely to be *protested* among equal dyads.

[248] Wharton, 'Three Essays on Decisions to Use Military Force', 85.

Lastly, the overall research design and methodology of this dissertation could be used across the social sciences to help answer other research questions investigating the extent to which leader traits are causal variables. For instance, recent articles and academic studies argue that countries led by female leaders had more favourable COVID-19 outcomes, such as fatality rates, than countries led by male leaders.[249] Using my counterfactual leader strategy, however, I suspect many of the studies showing an association between female leaders and these outcomes will also show a candidate selection effect. That is, if selection events for each leader prior to COVID-19 were included, we should see the same effect between female-runners up and lower COVID fatality outcomes, for example. In other words, female runners-up also "did better" at fighting COVID than male runners-up. Of course, this nonsensical relationship would only be because states with better COVID-19 outcomes probably had greater resources to deal with pandemics prior to March 2020 and were also more likely to select female political candidates than states with worse COVID-19 outcomes.[250] Again, using my counterfactual leader strategy for such a study is more efficient than attempting to control for an endless number of potential country-level confounders since the trait of the runners-up is the only variable needed to test for and confirm bias due to selection.

[249] Pedro-José Martínez-Córdoba, Bernardino Benito, and Isabel-María García-Sánchez, 'Efficiency in the Governance of the Covid-19 Pandemic: Political and Territorial Factors', *Globalization and Health* 17 (2021): 1–13.Luca Coscieme et al., 'Women in Power: Female Leadership and Public Health Outcomes during the COVID-19 Pandemic', *MedRxiv*, 2020, 2020–07; Jon Henley, Eleanor Ainge Roy, and Eleanor de Jong, 'Are Female Leaders More Successful at Managing the Coronavirus Crisis?', *The Guardian*, 25 April 2020, sec. World news, https://www.theguardian.com/world/2020/apr/25/why-do-female-leaders-seem-to-be-more-successful-at-managing-the-coronavirus-crisis.

[250] Windsor et al have a similar argument about how "cultural values" helped mitigate COVID-19 fatality rates and confounds the negative association between female leaders and the COVID-19 fatality rate association, although they also argue that female leaders amplify the beneficial effect of these cultural values. See: Leah C. Windsor et al., 'Gender in the Time of COVID-19: Evaluating National Leadership and COVID-19 Fatalities', *PloS One* 15, no. 12 (2020): e0244531.

8.7 Conclusion

In conclusion, a leader's traits, such as military experience and age, likely do not influence their decision to start international conflicts. I have shown that these relationships are likely biased due to selection effects. Not only are these relationships likely spurious due to selection effects, but there are important questions about whether leaders are even deciding to use force in many conflicts. In other words, any study exploring the effect of leader traits using militarised interstate disputes could be flawed if fishing disputes drive their findings. Leaders may matter for IR, but what matters more is the environments they are selected into.

Bibliography

Almond, Gabriel. *The American People and Foreign Policy*. 8th ed. United States of America: Fredereick A. Praeger, 1960.

Archer, John. 'Testosterone and Human Aggression: An Evaluation of the Challenge Hypothesis'. *Neuroscience & Biobehavioral Reviews* 30, no. 3 (2006): 319–45.

Arena, Philip, and Glenn Palmer. 'Politics or the Economy? Domestic Correlates of Dispute Involvement in Developed Democracies'. *International Studies Quarterly* 53, no. 4 (2009): 955–75.

Bachman, Jerald G., Peter Freedman-Doan, David R. Segal, and Patrick M. O'Malley. 'Distinctive Military Attitudes among US Enlistees, 1976-1997: Self-Selection versus Socialization'. *Armed Forces & Society* 26, no. 4 (2000): 561–85.

Bak, Daehee, and Glenn Palmer. 'Testing the Biden Hypotheses: Leader Tenure, Age, and International Conflict'. *Foreign Policy Analysis* 6, no. 3 (2010): 257–73.

Bakker, Edwin. 'Characteristics of Jihadi Terrorists in Europe (2001-2009)'. *Jihadi Terrorism and the Radicalisation Challenge. European and American Experiences. Furnham, Ashgate*, 2011, 141.

Baliga, Sandeep, David O. Lucca, and Tomas Sjöström. 'Domestic Political Survival and International Conflict: Is Democracy Good for Peace?' *The Review of Economic Studies* 78, no. 2 (2011): 458–86.

Barceló, Joan. 'Are Western-Educated Leaders Less Prone to Initiate Militarized Disputes?' *British Journal of Political Science*, 2018, 1–32.

Barnes, Tiffany D., and Diana Z. O'Brien. 'Defending the Realm: The Appointment of Female Defense Ministers Worldwide'. *American Journal of Political Science* 62, no. 2 (2018): 355–68.

Barnhart, Joslyn N., Robert F. Trager, Elizabeth N. Saunders, and Allan Dafoe. 'The Suffragist Peace'. *International Organization* 74, no. 4 (2020): 633–70.

Baturo, Alexander, and Johan A. Elkink. 'What Countries Select More Experienced Leaders? The PolEx Measure of Political Experience'. *British Journal of Political Science*, 2021, 1–10.

Bearce, David H., and Thomas R. Cook. 'The First Image Reversed: IGO Signals and Mass Political Attitudes'. *The Review of International Organizations* 13, no. 4 (2018): 595–619.

Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 'Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable'. *American Journal of Political Science* 42, no. 4 (1998): 1260–88.

Bertoli, Andrew, Allan Dafoe, and Robert F. Trager. 'Is There a War Party? Party Change, the Left–Right Divide, and International Conflict'. *Journal of Conflict Resolution* 63, no. 4 (2019): 950–75.

Betts, Richard K. *Soldiers, Statesmen, and Cold War Crises*. Cambridge, Massachusetts: Harvard University Press, 1979.

Betz, Hans-George. 'The New Politics of Resentment: Radical Right-Wing Populist Parties in Western Europe'. *Comparative Politics*, 1993, 413–27.

Book, Angela S., Katherine B. Starzyk, and Vernon L. Quinsey. 'The Relationship between Testosterone and Aggression: A Meta-Analysis'. *Aggression and Violent Behavior* 6, no. 6 (2001): 579–99.

Bordieri, James E., and David E. Drehmer. 'Vietnam Veterans: Fighting the Employment War'. *Journal of Applied Social Psychology* 14, no. 4 (1984): 341–47.

Brambor, Thomas, and Johannes Lindvall. 'The Ideology of Heads of Government, 1870–2012'. *European Political Science* 17, no. 2 (2018): 211–22.

Brewer, Paul R., Kimberly Gross, Sean Aday, and Lars Willnat. 'International Trust and Public Opinion about World Affairs'. *American Journal of Political Science* 48, no. 1 (2004): 93–109.

Broder, John M. 'Biden's Crisis Remarks Reverberate'. *The New York Times*, 20 October 2008. https://thecaucus.blogs.nytimes.com/2008/10/20/bidens-crisis-remarks-reverberate.

Brownlee, Jason. *Democracy Prevention: The Politics of the US-Egyptian Alliance*. Cambridge University Press, 2012.

Bueno de Mesquita, Bruce. *The Logic of Political Survival*. Cambridge, Massachusetts: MIT Press, 2003.
Calin, Costel, and Brandon Prins. 'The Sources of Presidential Foreign Policy Decision Making: Executive Experience and Militarized Interstate Conflicts'. *International Journal of Peace Studies* 20, no. 1 (2015): 17–34.
Caprioli, Mary. 'Gender Equality and State Aggression: The Impact of Domestic Gender Equality on State First Use of Force'. *International Interactions* 29, no. 3 (2003): 195–214.
Caprioli, Mary, and Peter F. Trumbore. 'Human Rights Rogues in Interstate Disputes, 1980–2001'. *Journal of Peace Research* 43, no. 2 (2006): 131–48.
Carnes, Nicholas, and Noam Lupu. 'What Good Is a College Degree? Education and Leader Quality Reconsidered'. *The Journal of Politics* 78, no. 1 (2016): 35–49.
Carter, Jeff, and Timothy Nordstrom. 'Term Limits, Leader Preferences, and Interstate Conflict'. *International Studies Quarterly* 61, no. 3 (2017): 721–35.
Carter, Jeff, and Charles E. Smith. 'A Framework for Measuring Leaders' Willingness to Use Force'. *American Political Science Review* 114, no. 4 (2020): 1352–58.
Chaudoin, Stephen. 'Promises or Policies? An Experimental Analysis of International Agreements and Audience Reactions'. *International Organization*, 2014, 235–56.
Chiozza, Giacomo, and Hein Erich Goemans. *Leaders and International Conflict*. Cambridge University Press, 2011.
CNN. 'Kerry: We Are Here to Make America Stronger', 30 July 2004. https://edition.cnn.com/2004/allpolitics/07/30/dems.kerry.transcript.2/index.html.
Collier, Paul, Anke Hoeffler, and Dominic Rohner. 'Beyond Greed and Grievance: Feasibility and Civil War'. *Oxford Economic Papers* 61, no. 1 (2009): 1–27.
Cooke, Rachel. 'David Grossman: "I Cannot Afford the Luxury of Despair"'. *The Observer*, 28 August 2010, sec. Books. https://www.theguardian.com/books/2010/aug/29/david-grossman-israel-hezbollah-interview.
Coscieme, Luca, Lorenzo Fioramonti, Lars F. Mortensen, Kate E. Pickett, Ida Kubiszewski, Hunter Lovins, Jacqueline Mcglade, Kristin Vala Ragnarsdottir, Debra Roberts, and Robert Costanza. 'Women in Power: Female Leadership and Public Health Outcomes during the COVID-19 Pandemic'. *MedRxiv*, 2020, 2020–07.
Costel, Calin. 'Hawks versus Doves: The Influence of Political Ideology on the Foreign Policy Behavior of Democratic States'. University of Tennessee, Knoxville, 2010.
Council, The Federal. 'History of the Federal Presidency'. The Federal Council: Portal of the Swiss Government. Accessed 17 April 2023. https://www.admin.ch/gov/en/start/bundespraesidium/bundespraesidenten-und-schweizer-geschichte.html.
Cox, Gary W. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge University Press, 1997.
Crisman-Cox, Casey, and Michael Gibilisco. 'Audience Costs and the Dynamics of War and Peace'. *American Journal of Political Science* 62, no. 3 (2018): 566–80.
Crowley, Candy, Dan Lothian, Mike Roselli, and Kelly Wallace. 'Nader Blasts "Unpatriotic Corporations"'. CNN, 24 February 2004. https://edition.cnn.com/2004/allpolitics/02/23/elec04.prez.main/index.html.
Davies, Shawn, Therese Pettersson, and Magnus Öberg. 'Organized Violence 1989–2021 and Drone Warfare'. *Journal of Peace Research* 59, no. 4 (2022): 593–610.
De Mesquita, Bruce Bueno, James D. Morrow, Randolph M. Siverson, and Alastair Smith. 'Testing Novel Implications from the Selectorate Theory of War'. *World Politics* 56, no. 3 (2004): 363–88.
Derpanopoulos, George, Erica Frantz, Barbara Geddes, and Joseph Wright. 'Are Coups Good for Democracy?' *Research & Politics* 3, no. 1 (2016): 2053168016630837.

Dilli, Selin, Sarah G. Carmichael, and Auke Rijpma. 'Introducing the Historical Gender Equality Index'. *Feminist Economics* 25, no. 1 (2019): 31–57.
Drezner, Daniel W. 'The Realist Tradition in American Public Opinion'. *Perspectives on Politics* 6, no. 1 (2008): 51–70.
Dube, Oeindrila, and S. P. Harish. 'Queens'. *Journal of Political Economy* 128, no. 7 (2020): 2579–2652.
Dyson, Stephen Benedict, and Thomas Briggs. 'Leaders and Foreign Policy: Surveying the Evidence'. In *Oxford Research Encyclopedia of Politics*, 2017.
Eichenberg, Richard C. *Gender, War, and World Order: A Study of Public Opinion*. Cornell University Press, 2019.
———. *Public Opinion and National Security in Western Europe: Consensus Lost?* Springer, 1989.
Eisner, Manuel. 'Long-Term Historical Trends in Violent Crime'. *Crime and Justice* 30 (2003): 83–142.
Ellis, Cali Mortenson, Michael C. Horowitz, and Allan C. Stam. 'Introducing the LEAD Data Set'. *International Interactions* 41, no. 4 (2015): 718–41.
Endicott, Travis W. 'Combat Experience and the Foreign Policy Positions of Veterans'. *Social Science Quarterly* 101, no. 4 (2020): 1413–29.
Falk, Armin, and Johannes Hermle. 'Relationship of Gender Differences in Preferences to Economic Development and Gender Equality'. *Science* 362, no. 6412 (2018): eaas9899.
Fearon, James D. 'Domestic Political Audiences and the Escalation of International Disputes'. *American Political Science Review*, 1994, 577–92.
Fordham, Benjamin O., and Christopher C. Sarver. 'Militarized Interstate Disputes and United States Uses of Force'. *International Studies Quarterly* 45, no. 3 (2001): 455–66.
Fukuyama, Francis. 'Women and the Evolution of World Politics'. *Foreign Affairs*, 1998, 24–40.
Gadarian, Shana Kushner. 'Foreign Policy at the Ballot Box: How Citizens Use Foreign Policy to Judge and Choose Candidates'. *The Journal of Politics* 72, no. 4 (2010): 1046–62.
Gelpi, Christopher, and Peter D. Feaver. 'Speak Softly and Carry a Big Stick? Veterans in the Political Elite and the American Use of Force'. *American Political Science Review* 96, no. 4 (2002): 779–93.
Gibler, Douglas M. *International Conflicts, 1816-2010: Militarized Interstate Dispute Narratives*. Vol. 2. Rowman & Littlefield, 2018.
Gift, Thomas, and Daniel Krcmaric. 'Who Democratizes? Western-Educated Leaders and Regime Transitions'. *Journal of Conflict Resolution* 61, no. 3 (2017): 671–701.
Gildea, Ross James. 'Psychology and Aggregation in International Relations'. *European Journal of International Relations* 25th Anniversary Special Issue (2020): 1–18.
Glaurdić, Josip, and Christophe Lesschaeve. 'Warrior Candidates: Do Voters Value Combat Experience in Postwar Elections?' *Political Research Quarterly*, 2021, 10659129211040524.
Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and H\a avard Strand. 'Armed Conflict 1946-2001: A New Dataset'. *Journal of Peace Research* 39, no. 5 (2002): 615–37.
Gochman, Charles S., and Zeev Maoz. 'Militarized Interstate Disputes, 1816-1976: Procedures, Patterns, and Insights'. *Journal of Conflict Resolution* 28, no. 4 (1984): 585–616.
Goemans, Henk E., Kristian Skrede Gleditsch, and Giacomo Chiozza. 'Introducing Archigos: A Dataset of Political Leaders'. *Journal of Peace Research* 46, no. 2 (2009): 269–83.
Goldstein, Joshua S. *War and Gender: How Gender Shapes the War System and Vice Versa*. Cambridge University Press, 2003.
Gourevitch, Peter. 'The Second Image Reversed: The International Sources of Domestic Politics'. *International Organization* 32, no. 4 (1978): 881–912.
Grossman, Guy, Devorah Manekin, and Dan Miodownik. 'The Political Legacies of Combat: Attitudes toward War and Peace among Israeli Ex-Combatants'. *International Organization* 69, no. 4 (2015): 981–1009.

Grossmann, Igor, Jinkyung Na, Michael EW Varnum, Denise C. Park, Shinobu Kitayama, and Richard E. Nisbett. 'Reasoning about Social Conflicts Improves into Old Age'. *Proceedings of the National Academy of Sciences* 107, no. 16 (2010): 7246–50.

Harvey, Frank P. *Explaining the Iraq War: Counterfactual Theory, Logic and Evidence*. Cambridge University Press, 2011.

Heaney, Katie. 'Gender Inequalities and International Military Aggression: The Role of Feminism in Achieving Peace'. *Res Publica-Journal of Undergraduate Research* 14, no. 1 (2009): 9.

Henley, Jon, Eleanor Ainge Roy, and Eleanor de Jong. 'Are Female Leaders More Successful at Managing the Coronavirus Crisis?' *The Guardian*, 25 April 2020, sec. World news. https://www.theguardian.com/world/2020/apr/25/why-do-female-leaders-seem-to-be-more-successful-at-managing-the-coronavirus-crisis.

Hermann, Margaret G., and Charles F. Hermann. 'Who Makes Foreign Policy Decisions and How: An Empirical Inquiry'. *International Studies Quarterly* 33, no. 4 (1989): 361–87.

Herre, Bastian. 'Identifying Ideologues: A Global Dataset on Political Leaders, 1945–2020'. *British Journal of Political Science*, 2022, 1–9.

Holsti, Ole R. 'Public Opinion and Foreign Policy: Challenges to the Almond-Lippmann Consensus'. *International Studies Quarterly* 36, no. 4 (1992): 439–66.

Holsti, Ole R., and James N. Rosenau. 'The Domestic and Foreign Policy Beliefs of American Leaders'. *Journal of Conflict Resolution* 32, no. 2 (1988): 248–94.

Horowitz, Michael C., and Matthew Fuhrmann. 'Studying Leaders and Military Conflict: Conceptual Framework and Research Agenda'. *Journal of Conflict Resolution* 62, no. 10 (2018): 2072–86.

Horowitz, Michael C., and Allan C. Stam. 'How Prior Military Experience Influences the Future Militarized Behavior of Leaders'. *International Organization* 68, no. 3 (2014): 527–59.

Horowitz, Michael C., Allan C. Stam, and Cali M. Ellis. *Why Leaders Fight*. Cambridge University Press, 2015.

Horowitz, Michael, Rose McDermott, and Allan C. Stam. 'Leader Age, Regime Type, and Violent International Relations'. *Journal of Conflict Resolution* 49, no. 5 (2005): 661–85.

Houses of the Oireachtas. 'Debates'. Accessed 15 December 2021. https://www.oireachtas.ie/en/debates/.

Howard, Anthony. *RAB: The Life of RA Butler*. A&C Black, 1987.

Huntington, Samuel P. *The Soldier and the State: The Theory and Politics of Civil–Military Relations*. Cambridge, Massachusetts: Harvard University Press, 1957.

Hyde, Susan D., and Nikolay Marinov. 'Which Elections Can Be Lost?' *Political Analysis* 20, no. 2 (2012): 191–210.

Igelström, Erik, Peter Craig, Jim Lewsey, John Lynch, Anna Pearce, and Srinivasa Vittal Katikireddi. 'Causal Inference and Effect Estimation Using Observational Data'. *J Epidemiol Community Health* 76, no. 11 (2022): 960–66.

Ikenberry, G. John, Michael Mastanduno, and William C. Wohlforth. 'Unipolarity, State Behavior, and Systemic Consequences'. *World Politics* 61, no. 1 (2009): 1–27.

Imbens, Guido W., and Donald B. Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences*. New York: Cambridge University Press, 2015.

Ish-Shalom, Piki. 'The Triptych of Realism, Elitism, and Conservatism'. *International Studies Review* 8, no. 3 (2006): 441–68.

Jaquette, Jane S. 'States Make War'. *Foreign Affairs* 78, no. 1 (1999): 128.

Jervis, Robert. 'Do Leaders Matter and How Would We Know?' *Security Studies* 22, no. 2 (2013): 153–79.

———. *How Statesmen Think*. Princeton University Press, 2017.

Johnson, Dominic DP. 'Survival of the Disciplines: Is International Relations Fit for the New Millennium?' *Millennium* 43, no. 2 (2015): 749–63.

Johnson, Dominic DP, and Bradley A. Thayer. 'The Evolution of Offensive Realism: Survival under Anarchy from the Pleistocene to the Present'. *Politics and the Life Sciences* 35, no. 1 (2016): 1–26.
Johnson, Dominic DP, and Dominic Tierney. 'Bad World: The Negativity Bias in International Politics'. *International Security* 43, no. 3 (2018): 96–140.
Johnson, Dominic DP, and Monica Duffy Toft. 'Grounds for War: The Evolution of Territorial Conflict'. *International Security* 38, no. 3 (2013): 7–38.
Jones, Benjamin F., and Benjamin A. Olken. 'Do Leaders Matter? National Leadership and Growth since World War II'. *The Quarterly Journal of Economics* 120, no. 3 (2005): 835–64.
Jones, Daniel M., Stuart A. Bremer, and J. David Singer. 'Militarized Interstate Disputes, 1816–1992: Rationale, Coding Rules, and Empirical Patterns'. *Conflict Management and Peace Science* 15, no. 2 (1996): 163–213.
Kaarbo, Juliet. 'Prime Minister Leadership Style and the Role of Parliament in Security Policy'. *The British Journal of Politics and International Relations* 20, no. 1 (2018): 35–51.
Kadera, Kelly, and Gerald Sorokin. 'Measuring National Power'. *International Interactions* 30, no. 3 (2004): 211–30.
Kaufmann, Chaim. 'Threat Inflation and the Failure of the Marketplace of Ideas: The Selling of the Iraq War'. *International Security* 29, no. 1 (2004): 5–48.
Kertzer, Joshua D. *Resolve in International Politics*. Vol. 2. Princeton University Press, 2016.
Kertzer, Joshua D., and Ryan Brutger. 'Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory'. *American Journal of Political Science* 60, no. 1 (2016): 234–49.
Kertzer, Joshua D., and Kathleen M. McGraw. 'Folk Realism: Testing the Microfoundations of Realism in Ordinary Citizens'. *International Studies Quarterly* 56, no. 2 (2012): 245–58.
Kertzer, Joshua D., and Dustin Tingley. 'Political Psychology in International Relations: Beyond the Paradigms'. *Annual Review of Political Science* 21 (2018): 319–39.
Kim, Hyung Min. 'Comparing Measures of National Power'. *International Political Science Review* 31, no. 4 (2010): 405–27.
Kim, Nam Kyu, and Alice J. Kang. 'External Threat Environments and Individual Bias against Female Leaders'. *Political Science Research and Methods* 10, no. 1 (2022): 1–17.
Kiratli, Osman Sabri. 'Where Do Hawks and Doves Fly When Shots Are Fired?' *Party Politics*, 2020, 1–12.
Klingemann, Hans-Dieter, Richard Hofferbert, and Ian Budge. 'A Theory of Democratic Policymaking'. In *Parties, Policies, and Democracy*, by Hans-Dieter Klingemann, Richard Hofferbert, and Ian Budge. Boulder, CO: Westview, 1994.
Klingler, Jonathan D., and J. Tyson Chatagnier. 'Are You Doing Your Part? Veterans' Political Attitudes and Heinlein's Conception of Citizenship'. *Armed Forces & Society* 40, no. 4 (2014): 673–95.
Koch, Michael T., and Sarah A. Fulton. 'In the Defense of Women: Gender, Office Holding, and National Security Policy in Established Democracies'. *The Journal of Politics* 73, no. 1 (2011): 1–16.
Koch, Michael T., and Patricia Sullivan. 'Should I Stay or Should I Go Now? Partisanship, Approval, and the Duration of Major Power Democratic Military Interventions'. *The Journal of Politics* 72, no. 3 (2010): 616–29.
Krcmaric, Daniel, Stephen C. Nelson, and Andrew Roberts. 'Studying Leaders and Elites: The Personal Biography Approach'. *Annual Review of Political Science* 23 (2020): 133–51.
Kreps, Sarah E., and Geoffrey PR Wallace. 'International Law, Military Effectiveness, and Public Support for Drone Strikes'. *Journal of Peace Research* 53, no. 6 (2016): 830–44.
Kugler, Jacek, and Marina Arbetman. 'Choosing among Measures of Power: A Review of the Empirical Record'. *Power in World Politics*, 1989, 49–77.

Kurizaki, Shuhei, and Taehee Whang. 'Detecting Audience Costs in International Disputes'. *International Organization* 69, no. 4 (2015): 949–80.
Lai, Brian, and Dan Slater. 'Institutions of the Offensive: Domestic Sources of Dispute Initiation in Authoritarian Regimes, 1950–1992'. *American Journal of Political Science* 50, no. 1 (2006): 113–26.
Lebow, Richard Ned. 'What's so Different about a Counterfactual?' *World Politics* 52, no. 4 (2000): 550–85.
Lemke, Douglas, and Suzanne Werner. 'Power Parity, Commitment to Change, and War'. *International Studies Quarterly* 40, no. 2 (1996): 235–60.
Lesschaeve, Christophe. 'Voting after Violence: How Combat Experiences and Postwar Trauma Affect Veteran and Voter Party Choices in Croatia's 2003 Postwar Elections'. *Armed Forces & Society* 46, no. 2 (2020): 259–80.
Lieberman, Evan S. 'Nested Analysis as a Mixed-Method Strategy for Comparative Research'. *American Political Science Review* 99, no. 3 (2005): 435–52.
Lippmann, Walter. *Essays in the Public Philosophy*. Transaction Publishers, 1955.
Livingston, Gretchen. 'Profile of U.S. Veterans Is Changing Dramatically as Their Ranks Decline'. *Pew Research Center* (blog), 11 November 2016. https://www.pewresearch.org/fact-tank/2016/11/11/profile-of-u-s-veterans-is-changing-dramatically-as-their-ranks-decline/.
Lobell, Steven E., Norrin M. Ripsman, Jeffrey W. Taliaferro, and Jeffrey W. Taliaferro. *Neoclassical Realism, the State, and Foreign Policy*. Cambridge University Press, 2009.
Lopez, Anthony C. 'The Evolutionary Psychology of War: Offense and Defense in the Adapted Mind'. *Evolutionary Psychology* 15, no. 4 (2017): 1474704917742720.
Maoz, Zeev, Paul L. Johnson, Jasper Kaplan, Fiona Ogunkoya, and Aaron P. Shreve. 'The Dyadic Militarized Interstate Disputes (MIDs) Dataset Version 3.0: Logic, Characteristics, and Comparisons to Alternative Datasets'. *Journal of Conflict Resolution* 63, no. 3 (2019): 811–35.
Martill, Benjamin, and Adrian Rogstad. 'The End of Consensus? Folk Theory and the Politics of Foreign Policy in the Brexit Referendum'. *Global Affairs* 5, no. 4–5 (2019): 347–67.
Martin, Mike. *Why We Fight*. Oxford University Press, 2018.
Martínez-Córdoba, Pedro-José, Bernardino Benito, and Isabel-María García-Sánchez. 'Efficiency in the Governance of the Covid-19 Pandemic: Political and Territorial Factors'. *Globalization and Health* 17 (2021): 1–13.
McCartney, Helen. 'Hero, Victim or Villain? The Public Image of the British Soldier and Its Implications for Defense Policy'. *Defense & Security Analysis* 27, no. 1 (2011): 43–54.
Mearsheimer, John J. *The Tragedy of Great Power Politics*. WW Norton & Company, 2001.
Mikulaschek, Christoph. 'Cueing Foreign Elite Consensus or Divisions: The Effect of Unanimity in International Organizations on Public Opinion'. In *Political Economy of International Organization Conference*, 2018.
Offner, Arnold A. *Hubert Humphrey: The Conscience of the Country*. Yale University Press, 2018.
Palmer, Glenn, Tamar London, and Patrick Regan. 'What's Stopping You?: The Sources of Political Constraints on International Conflict Behavior in Parliamentary Democracies'. *International Interactions* 30, no. 1 (2004): 1–24.
Park, So Yeon (Ellen). 'Why Do Leaders Matter?: Exploring Leaders' Risk Propensities and Nuclear Proliferation'. PhD Dissertation, University of California, 2021.
Peterson, John E. 'The Nature of Succession in the Gulf'. *The Middle East Journal*, 2001, 580–601.
Peyton, Buddy, Joseph Bajjalieh, Dan Shalmon, Michael Martin, and Jonathan Bonaguro. 'Cline Center Coup D'état Project Dataset'. *Cline Center for Advanced Social Research* 2 (2020): 0.
Pomeroy, Caleb. 'The Psychology of Power: The First Image Reversed and International Security', 2021.

Post, Abigail S., and Paromita Sen. 'Why Can'ta Woman Be More like a Man? Female Leaders in Crisis Bargaining'. *International Interactions* 46, no. 1 (2020): 1–27.
Potter, Philip BK. 'Does Experience Matter? American Presidential Experience, Age, and International Conflict'. *Journal of Conflict Resolution* 51, no. 3 (2007): 351–78.
Powell, Jonathan, and Karina Mukazhanova-Powell. 'Demonstrating Credentials? Female Executives, Women's Status, and the Use of Force'. *Journal of Women, Politics & Policy* 40, no. 2 (2019): 241–62.
Rathbun, Brian C. 'It Takes All Types: Social Psychology, Trust, and the International Relations Paradigm in Our Minds'. *International Theory* 1, no. 3 (2009): 345–80.
Rauch, Carsten. 'Challenging the Power Consensus: GDP, CINC, and Power Transition Theory'. *Security Studies* 26, no. 4 (2017): 642–64.
Renshon, Jonathan, and Stanley A. Renshon. 'The Theory and Practice of Foreign Policy Decision Making'. *Political Psychology* 29, no. 4 (2008): 509–36.
Richardson, David K. 'The Electoral Impact of Military Experience: Evidence From US Senate Elections (1982–2016)'. *Armed Forces & Society*, 2021, 0095327X211038032.
Rohrer, Samuel R. 'What Makes a Prime Minister Great?: A Leadership Trait Analysis of the Effectiveness of British Prime Ministers from 1902 to 2004'. *Research & Politics* 1, no. 3 (2014): 2053168014558970.
Rosen, Stephen Peter. *War and Human Nature*. Princeton, NJ: Princeton University Press, 2005.
Sapolsky, Robert M. *Behave: The Biology of Humans at Our Best and Worst*. Penguin, 2017.
Schramm, Madison, and Alexandra Stark. 'Peacemakers or Iron Ladies? A Cross-National Study of Gender and International Conflict'. *Security Studies* 29, no. 3 (2020): 515–48.
Schultz, Kenneth A. 'The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?' *International Organization* 59, no. 1 (2005): 1–38.
Sechser, Todd S. 'Are Soldiers Less War-Prone than Statesmen?' *Journal of Conflict Resolution* 48, no. 5 (2004): 746–74.
Sevi, Semra. 'Do Young Voters Vote for Young Leaders?' *Electoral Studies* 69 (2021): 102200.
Shafi, Rafia, and Samreen Fatima. 'Relationship between GDP, Life Expectancy and Growth Rate of G7 Countries'. *International Journal of Sciences* 8, no. 06 (2019): 74–79.
Sheafer, Tamir, and Shira Dvir-Gvirsman. 'The Spoiler Effect: Framing Attitudes and Expectations toward Peace'. *Journal of Peace Research* 47, no. 2 (2010): 205–15.
Sievers, Christian. 'Johnson: Putins Krieg Hat Europa Gestärkt'. ZDF, 28 June 2022. https://www.zdf.de/nachrichten/politik/boris-johnson-g7-gipfel-gemeinschaft-100.html.
Sieweke, Jost, and Simone Santoni. 'Natural Experiments in Leadership Research: An Introduction, Review, and Guidelines'. *The Leadership Quarterly* 31, no. 1 (2020): 101338.
Singer, J. David. 'Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816–1985'. *International Interactions* 14, no. 2 (1988): 115–32.
Singer, J. David, Stuart Bremer, and John Stuckey. 'Capability Distribution, Uncertainty, and Major Power War, 1820-1965'. *Peace, War, and Numbers* 19 (1972): 48.
Smets, Kaat. 'A Widening Generational Divide? The Age Gap in Voter Turnout through Time and Space'. *Journal of Elections, Public Opinion & Parties* 22, no. 4 (2012): 407–30.
———. 'Revisiting the Political Life-Cycle Model: Later Maturation and Turnout Decline among Young Adults'. *European Political Science Review* 8, no. 2 (2016): 225–49.
Smith, Brent L., and Kathryn D. Morgan. 'Terrorists Right and Left: Empirical Issues in Profiling American Terrorists'. *Studies in Conflict & Terrorism* 17, no. 1 (1994): 39–57.
Spisak, Brian R. 'The General Age of Leadership: Older-Looking Presidential Candidates Win Elections during War'. *PLoS One* 7, no. 5 (2012).
Spisak, Brian R., Allen E. Grabo, Richard D. Arvey, and Mark van Vugt. 'The Age of Exploration and Exploitation: Younger-Looking Leaders Endorsed for Change and

Older-Looking Leaders Endorsed for Stability'. *The Leadership Quarterly* 25, no. 5 (2014): 805–16.
Stadelmann, David, Marco Portmann, and Reiner Eichenberger. 'Military Careers of Politicians Matter for National Security Policy'. *Journal of Economic Behavior & Organization* 116 (2015): 142–56.
Teigen, Jeremy. *Why Veterans Run: Military Service in American Presidential Elections, 1789-2016*. Temple University Press, 2018.
Teigen, Jeremy M. 'Military Experience in Elections and Perceptions of Issue Competence: An Experimental Study with Television Ads'. *Armed Forces & Society* 39, no. 3 (2013): 415–33.
Tetlock, Philip E., and Aaron Belkin. 'Counterfactual Thought Experiments in World Politics'. *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*, 1996, 1.
Thayer, Bradley A. *Darwin and International Relations: On the Evolutionary Origins of War and Ethnic Conflict*. University Press of Kentucky, 2009.
'The Constitution of Japan'. Prime Minister of Japan and His Cabinet, 1946.
The New York Times. 'The 1992 Campaign; A Letter by Clinton on His Draft Deferment: "A War I Opposed and Despised"'. 13 February 1992, sec. U.S. https://www.nytimes.com/1992/02/13/us/1992-campaign-letter-clinton-his-draft-deferment-war-opposed-despised.html.
'The Reagan/Mondale Debate'. PBS, 24 September 2000. https://www.pbs.org/newshour/spc/debatingourdestiny/doc1984.html.
Todosijević, Bojan. 'Left-Right Ideology: Its Meaning and Effects on Party Preferences in Serbia'. *Социолошки Преглед* 50, no. 2 (2016): 161–78.
Tomz, Michael. 'Domestic Audience Costs in International Relations: An Experimental Approach'. *International Organization*, 2007, 821–40.
Toronto, Nathan. 'Military Recruitment Data Set'. *Ft. Leavenworth, KS: Foreign*, 2007.
Trager, Robert F., and Lynn Vavreck. 'The Political Costs of Crisis Bargaining: Presidential Rhetoric and the Role of Party'. *American Journal of Political Science* 55, no. 3 (2011): 526–45.
TRIP Journal Article Database Release (Version 3.3). 'Teaching, Research, and International Policy Project'. Accessed 22 March 2023. https://trip.wm.edu/data/dashboard/journal-article-database.
United Nations. 'World Population Prospects - Population Division - United Nations'. Accessed 8 November 2022. https://population.un.org/wpp/Download/Standard/Population/.
Van der Waal, Jeroen, Willem De Koster, and Peter Achterberg. 'Ethnic Segregation and Radical Right-Wing Voting in Dutch Cities'. *Urban Affairs Review* 49, no. 5 (2013): 748–77.
Van Evera, Stephen. 'Foreward'. In *American Foreign Policy and the Politics of Fear: Threat Inflation since 9/11*, edited by Trevor A. Thrall and Jane K. Cramer. New York: Routledge, 2009.
Wallace, Geoffrey PR. 'International Law and Public Attitudes toward Torture: An Experimental Study'. *International Organization*, 2013, 105–40.
Wallenfeldt, Jeff. 'Edward Frederick Lindley Wood, 1st Earl of Halifax'. Brittanica, 9 March 2023. https://www.britannica.com/biography/Edward-Frederick-Lindley-Wood-1st-earl-of-Halifax.
Waltz, Kenneth N. *Man, the State, and War: A Theoritical Anylsis*. New York: Columbia University Press, 1959.
Webster, Steven W., and Andrew W. Pierce. 'Older, Younger, or More Similar? The Use of Age as a Voting Heuristic'. *Social Science Quarterly* 100, no. 3 (2019): 635–52.
Weeks, Jessica L. 'Strongmen and Straw Men: Authoritarian Regimes and the Initiation of International Conflict'. *American Political Science Review*, 2012, 326–47.
Wharton, Jaron. 'Three Essays on Decisions to Use Military Force'. PhD Dissertation, Duke University, 2018.

Wilson, Margo, and Martin Daly. 'Competitiveness, Risk Taking, and Violence: The Young Male Syndrome'. *Ethology and Sociobiology* 6, no. 1 (1985): 59–73.
Wilson, Timothy D. *Strangers to Ourselves*. Harvard University Press, 2004.
Windsor, Leah C., Gina Yannitell Reinhardt, Alistair J. Windsor, Robert Ostergard, Susan Allen, Courtney Burns, Jarod Giger, and Reed Wood. 'Gender in the Time of COVID-19: Evaluating National Leadership and COVID-19 Fatalities'. *PloS One* 15, no. 12 (2020): e0244531.
Winter, Jay M. 'Britain's 'Lost Generation'of the First World War'. *Population Studies* 31, no. 3 (1977): 449–66.
Wolford, Scott. 'The Turnover Trap: New Leaders, Reputation, and International Conflict'. *American Journal of Political Science* 51, no. 4 (2007): 772–88.
Wurm, Sebastian Chávez. 'Arming the Revolution. The Shining Path and Its Struggle for Resources (1980–1993)'. *Jahrbuch Für Geschichte Lateinamerikas* 49, no. 1 (2012): 261–84.
Yılmaz, Ferruh. 'Right-Wing Hegemony and Immigration: How the Populist Far-Right Achieved Hegemony through the Immigration Debate in Europe'. *Current Sociology* 60, no. 3 (2012): 368–81.