International Studies Quarterly (2024) **68**, sqae045

The Changing Face of Nuclear Proliferation

Jeffrey M. Kaplow
William & Mary, USA

A rich literature has identified a number of important drivers of nuclear proliferation. Most of this work, however, treats the determinants of proliferation as constant over the entire nuclear age—the factors leading to proliferation are assumed to be the same in 2010 as they were in 1945. But there are reasons to suspect that the drivers of proliferation have changed over this time: nuclear technology is easier to come by, the global strategic environment has shifted, and the nuclear nonproliferation regime has come into being. To examine changes in the dynamics of nuclear proliferation, I adapt a cross-validation technique frequently used in the machine learning literature. I create a rolling window of training data with which statistical models of proliferation are built, and I then test the predictive power of these models against data from other time periods. The result of this analysis is a temporal map of how the determinants of proliferation have changed over time. My findings suggest that the underlying dynamics of nuclear proliferation have indeed shifted, with important implications both for the literature on nuclear proliferation and for policymakers interested in limiting the future spread of nuclear weapons.

Existe abundante literatura que se ha encargado de identificar varios de los factores importantes que impulsan la proliferación nuclear. Sin embargo, la mayor parte de estos trabajos tratan los determinantes de la proliferación como si fueran constantes a lo largo de toda la era nuclear, es decir, se supone que los factores que conducen a la proliferación son los mismos en 2010 que en 1945. Sin embargo, existen razones para sospechar que los impulsores de la proliferación nuclear han cambiado a lo largo de todo este tiempo: la tecnología nuclear es más fácil de conseguir, el entorno estratégico mundial ha cambiado y, además, ha surgido el régimen de no proliferación nuclear. Con el fin de estudiar estos cambios en la dinámica de la proliferación nuclear, adaptamos una técnica de validación cruzada utilizada con frecuencia en la literatura de aprendizaje automático. Creamos una función de ventana con datos de entrenamiento que nos sirve para construir modelos estadísticos de proliferación. A continuación, probamos el poder predictivo de estos modelos con datos de otros períodos de tiempo. El resultado de este análisis es un mapa temporal de cómo han cambiado los determinantes de la proliferación a lo largo del tiempo. Nuestras conclusiones sugieren que la dinámica subyacente de la proliferación nuclear ha cambiado. Esto conlleva importantes implicaciones tanto para la literatura sobre la proliferación nuclear como para los responsables políticos interesados en limitar la propagación futura de las armas nucleares.

Une littérature abondante a identifié nombre de facteurs importants dans la prolifération nucléaire. Néanmoins, la majorité de ces travaux traite les déterminants de la prolifération comme des constantes dans l'ensemble de l'ère nucléaire ; les facteurs déclenchant la prolifération seraient les mêmes en 2010 qu'en 1945. Or, certains faits laissent à penser que les facteurs de prolifération ont évolué au cours de cette période : la technologie nucléaire est devenue plus facile d'accès, l'environnement stratégique mondial s'est modifié et le régime de non-prolifération nucléaire a fait son apparition. Pour analyser les changements au sein de la dynamique de prolifération nucléaire, j'adapte une technique de validation croisée couramment employée dans la littérature sur l'apprentissage automatique. Je crée une fenêtre dynamique de données d'apprentissage à partir de laquelle construire des modèles statistiques de prolifération, puis j'évalue leur pouvoir prévisionnel par rapport aux données d'autres intervalles. Cette analyse produit une représentation temporelle de l'évolution des déterminants de la prolifération. D'après mes résultats, la dynamique de prolifération nucléaire a bel et bien évolué ; les implications sont importantes tant pour la littérature relative à la prolifération nucléaire que pour les décideurs souhaitant limiter la diffusion des armes nucléaires à l'avenir.

The US intelligence community periodically seeks to identify the countries at greatest risk of pursuing nuclear weapons in the near future. According to now-declassified National Intelligence Estimates, the list of worrisome states in the early 1960s included China—which tested its first nuclear weapon in 1964—along with India, Israel, Sweden, West Germany, Italy, Japan, and Canada (Central Intelligence Agency 1963, 1964).[1] Although no such estimate from recent years is publicly available, unclassified assessments look quite different from the usual suspects of the 1960s. Today's lists of the states most at risk of seeking nuclear weapons in the near future generally include Iran, of course, but often also mention Saudi Arabia, Turkey, South Korea, Egypt, and Japan.[2]

Jeffrey M. Kaplow is an Associate Professor of Government at William & Mary and the Director of NukeLab at William & Mary's Global Research Institute. He is the author of *Signing Away the Bomb: The Surprising Success of the Nuclear Nonproliferation Regime*, published by Cambridge University Press.

Author's Note: Data and replication files are available on the ISQ Dataverse page at https://dataverse.harvard.edu/dataverse/isq or on the author's website at https://jkaplow.net/research. This project was made possible by the support of the Bridging the Gap Project, the Smith Richardson Foundation, the Stanton Foundation, the Reves Center at William & Mary, and William & Mary's Summer Research Fund. I thank Mark Bell, Rebecca Gibbons, Jooeun Kim, Dan Maliniak, Reid Pauly, Paul Poast, Todd Robinson, Maria Rost Rublee, Todd Sechser, Steven Lloyd Wilson, and two anonymous reviewers for helpful comments, along with participants in the William & Mary Government Department Colloquium and the Lansing Lee/Bankard Seminar at the University of Virginia. Research assistance was provided by Zackary Johnson, Julia Qiao, Emilie Smetak, and the student researchers of NukeLab at William & Mary's Global Research Institute. This article is dedicated to the memory of Troy Cullen, whose work with NukeLab made a substantial contribution to this project and whose presence is sorely missed.

[1]Of these, US intelligence analysts believed only India, and "perhaps Israel and Sweden," would actually develop weapons over the next 10 years (Central Intelligence Agency 1964).

[2]There are many such lists produced by think tanks and foreign policy analysts, with fairly wide agreement about the candidate countries. For representative examples, see Allison (2010), Spector (2016), and Brewer (2020).

Kaplow, Jeffrey M. (2024) The Changing Face of Nuclear Proliferation. *International Studies Quarterly*, https://doi.org/10.1093/isq/sqae045

Why does today's list of potential proliferant states bear so little resemblance to those published in the 1960s? The circumstances of some individual countries have clearly changed. Italy, Sweden, Canada, and a reunified Germany face far different security threats today than at the height of the Cold War. And the states on today's list probably would not have been seen as having sufficient wealth or resources to plausibly acquire weapons in the 1960s. But these lists also differ for a broader reason: our understanding of what drives states to seek nuclear weapons has fundamentally changed. The spread of both civilian and sensitive nuclear technology has made indigenous capabilities and technical resources less essential to a successful nuclear effort. At the same time, the global strategic environment has shifted since the beginning of the Cold War, with a corresponding change in the credibility of nuclear alliances and security commitments. The first nuclear weapons also were developed outside the constraining effect of the set of international treaties, agreements, and norms that make up today's nuclear nonproliferation regime.

The extensive academic literature on nuclear proliferation and nuclear restraint, while it has generated many important insights, does not provide much help in understanding the changing face of nuclear proliferation. Research employing quantitative methods has seen particular growth, but this work makes very strong assumptions about the extent to which the drivers of proliferation have remained static over time. By pooling all available cases for a single analysis, existing research, in effect, treats the fundamental dynamics of nuclear proliferation as unchanged since the dawn of the nuclear age.

In this article, I propose a theory of change in the drivers of nuclear proliferation over time. I argue that supply-side factors—the underlying capabilities of states and their ability to access nuclear technologies—have become less important over time, and I generate competing hypotheses about the changing influence of demand-side factors and international institutions. I then test these hypotheses using multiple subsamples of proliferation data delineated by time period, focusing on the change in the predictive power of each of these variables over time. Adapting a cross-validation technique frequently used in the machine learning literature, I create a rolling window of training data with which statistical models of proliferation are built, and I then test the predictive power of these models against data from other time periods. The result of this analysis is a temporal map of how the determinants of proliferation have changed over time. My findings suggest that we should treat with caution the literature's recent emphasis on supply-side drivers of nuclear proliferation. Latent industrial capacity and other supply-side factors were undoubtedly important for the first nuclear states, but the lessons from those nuclear programs may be less relevant for the states at greatest risk of seeking nuclear weapons today.

This article makes several contributions to the nuclear proliferation and international relations literatures. First, I make a theoretical contribution in laying out a theory of change in nuclear proliferation. Shifts in the underlying dynamics of nuclear supply, international threat perceptions, and institutional effectiveness suggest hypotheses about changes over time in the factors that are likely to lead to nuclear pursuit. Advancing a theory of change is essential for our understanding of dynamic phenomena like nuclear proliferation; static theories risk missing essential drivers of international behavior that shift over time. Second, for researchers studying nuclear proliferation, my empirical findings provide additional leverage for understanding the drivers of nuclear proliferation today and in the immediate future. My analysis shows that a central conclusion of past studies of nuclear proliferation—that supply-side factors are the most important variables governing nuclear pursuit—relies heavily on the earliest nuclear weapons programs and may be less relevant to understanding nuclear proliferation today.

Third, the article offers a critique of temporal pooling, the widespread practice of assuming a common data-generating mechanism over time. This is an extremely strong assumption, particularly in issue areas—such as nuclear proliferation—where changes in technology play an important role in state decision-making. Finally, the article presents a flexible empirical approach that can be used to investigate temporal variation in a variety of substantive areas. While by no means a replacement for traditional regression models, the rolling-window cross-validation (RWCV) procedure I describe here offers a novel way to understand the common factors influencing state decision-making across different time periods.

The central message of this article—that scholars should carefully consider how their results vary over time—applies broadly to social scientific studies in international relations and beyond. Many of our findings implicitly assume that the preferences of actors or the constraints they face are static, but often these factors are shifting over time in a way that can bias our conclusions. It seems worthwhile, then, to revisit accepted findings with an eye toward whether the underlying dynamics of an issue have changed—whether the empirical lessons we have learned apply as well to today's challenges as they did to the earlier periods we have studied.

Below I apply this idea to the issue of nuclear proliferation, proceeding in four parts. First, I describe the literature on the drivers of nuclear proliferation, highlighting the potential dangers of reasoning from past cases. Second, I propose a theory of change in nuclear proliferation, generating hypotheses about the importance of various drivers of proliferation over time. Third, I test these hypotheses using a RWCV approach. Finally, I conclude with implications for other scholarly work.

The Drivers of Nuclear Proliferation

The literature on nuclear proliferation has identified a number of factors that may affect state decisions to seek nuclear weapons.[3] These drivers of proliferation and nuclear restraint can be divided into three broad categories: nuclear capabilities (supply-side factors), nuclear motives (demand-side factors), and international institutions.

Nuclear Capabilities

Industrial and economic capacity have long been seen as the most important pre-requisite for nuclear weapons pursuit. A full-scale nuclear weapons program requires financial resources, access to particular raw materials, specialized industrial facilities, and advanced engineering and scientific expertise. States lacking these resources, scholars have argued, represent little risk of proliferation, no matter how much they desire a weapon.

Existing work has largely found support for the importance of supply-side factors in driving proliferation. A body of research has identified significant links between economic capacity and proliferation behavior (Singh and Way 2004; Fuhrmann and Berejikian 2012; Horowitz and

[3]For a comprehensive review, see Sagan (2011).

Narang 2014; Way and Weeks 2014), and an even stronger association between industrial or latent nuclear capability and weapons programs (Singh and Way 2004; Jo and Gartzke 2007; Fuhrmann 2009; Bleek and Lorber 2014; Way and Weeks 2014; Miller 2014b; Fuhrmann and Tkach 2015). Gartzke and Kroenig (2009) conclude that supply-side factors—to include state capability—"are among the most important determinants of nuclear proliferation."

Advocates of supply-side theories of nuclear proliferation also have focused on the provision of nuclear assistance to would-be proliferants. By lowering the barriers to a successful nuclear weapons effort, nuclear assistance may make states more likely to take the initial step and launch a nuclear weapons program. Studies that have examined the role of nuclear assistance in encouraging proliferation have found that sensitive assistance, bilateral civilian assistance, and multilateral civilian assistance are strongly associated with the decisions of states to seek nuclear weapons (Fuhrmann 2009; Kroenig 2009; Brown and Kaplow 2014).[4]

Nuclear Motives

The most basic driver of nuclear weapons pursuit is a state's concern about its own security (Sagan 1996). States justify the expense and risk associated with seeking weapons largely in terms of the security benefits they believe will follow from nuclear acquisition. Countries that have been more frequent participants in international conflict—or that expect to be involved in conflict in the future—will find the cost of nuclear weapons development easier to justify. It is no surprise, then, that a number of studies have shown an association between nuclear weapons pursuit and a state's conflict behavior (Singh and Way 2004; Fuhrmann 2009; Brown and Kaplow 2014; Miller 2014b; Fuhrmann and Horowitz 2015; Fuhrmann and Lupu 2016), the proliferation decisions of neighbors or rivals (Singh and Way 2004; Fuhrmann 2009; Fuhrmann and Berejikian 2012; Fuhrmann and Horowitz 2015; Fuhrmann and Lupu 2016), and the presence of an alliance or security guarantee from a nuclear-armed patron (Bleek and Lorber 2014; Reiter 2014; Fuhrmann and Horowitz 2015; Gerzhoy 2015).[5]

A number of other factors may affect a state's desire for nuclear weapons. Several scholars have found a relationship between the type of regime or the psychology or experiences of its leader and the decision to pursue nuclear weapons (Hymans 2006; Way and Weeks 2014; Fuhrmann and Horowitz 2015). The openness of a particular regime to the international economy, too, may moderate the desire to proliferate, as may norms of nonproliferation behavior (Solingen 2007; Rublee 2009).

Institutions

Since the Treaty on the Non-Proliferation of Nuclear Weapons (NPT) came into force in 1970, only four states have acquired nuclear weapons, and no state has acquired nuclear weapons while a member of the NPT.[6] For some, this is strong evidence of the effectiveness of the NPT in constraining state behavior. The NPT has been hailed as one of the most successful security treaties in history (Cirincione 2008), chiefly because it is seen as a key factor in keeping the number of nuclear weapons states far below the dire predictions of the 1960s.[7]

A number of studies have examined the possible constraining role of the NPT in at least a cursory way. Quantitative analyses have found that whether a state is a member of the NPT in a given year is often significantly associated with a decreased likelihood of proliferation (Jo and Gartzke 2007; Fuhrmann 2009; Bleek and Lorber 2014; Brown and Kaplow 2014). As these authors acknowledge, however, the significant negative correlation between NPT membership and nuclear weapons programs does not tell us much about the independent role of the NPT, because there is likely to be a powerful selection effect with regard to treaty membership. If states are more likely to join the NPT when they have no intention of proliferating, as we might expect, then this correlation may have little to do with the constraining power of the treaty. Studies that seek to avoid this selection problem have come to mixed conclusions about the effectiveness of nonproliferation agreements (Coe and Vaynman 2015; Fuhrmann and Lupu 2016; Smith and Spaniel 2021).[8]

Temporal Pooling in Studies of Proliferation

This body of proliferation research uses past cases of weapons pursuit or restraint to offer—at least implicitly but sometimes quite openly—lessons for academics and policymakers examining today's nuclear proliferation challenges. Scholarship on proliferation promises more than a history lesson; it is useful and important because it advances our understanding of an international security issue of continuing relevance.

When we reason from past cases, we assume that the underlying dynamics of proliferation in the past are similar to those at play today. Studies whose research designs employ detailed case studies or historical analysis commonly make this assumption explicit. In justifying the decision to focus on a particular set of cases or historical examples, scholars defend the external validity of their analysis and explain why their findings are relevant beyond the set of cases that have been chosen.

Studies that adopt a quantitative approach, however, frequently skip this step; it may seem unnecessary for authors to justify the scope of an analysis that employs the full universe of cases. However, empirical models constructed from all previous cases make very strong assumptions about the extent to which the drivers of proliferation have remained unchanged over time. Quantitative approaches to nuclear proliferation typically begin with a dataset that pools nearly all country years since 1938 or 1945—including states with and without nuclear weapons programs—into a single dataset. Regression models that analyze pooled datasets generally assume a common data-generating process over the full dataset; these static parameter models assume that the effect of a given explanatory variable on the likelihood of an outcome is the same for observations in 1945 as for observations in 2010.

One risk, then, is that this approach will obscure important temporal variation for factors of interest. Regression coefficients that appear statistically and substantively

[4]For contrary views, see Montgomery (2013), Miller (2017), and Gibbons (2020). Kemp (2014) argues that foreign assistance is not necessary for states to develop substantial gas centrifuge uranium enrichment programs.

[5]None of these findings are universal, however. For alternative results on nuclear rivalry, see Jo and Gartzke (2007), and on alliances, see Fuhrmann (2009) and Brown and Kaplow (2014).

[6]One state with nuclear weapons, South Africa, later gave them up and joined the NPT. North Korea acquired nuclear weapons only after withdrawing from the NPT.

[7]Most famously, President Kennedy warned that dozens of states might join the nuclear club by the 1970s (Kennedy 1963).

[8]For a review and discussion of the challenges of empirical work in this area, see Kaplow (2022a).

significant based on the full pooled dataset may fail to reach significance within time-limited subsamples of the full dataset. This may be a particular problem in studies of nuclear proliferation, where cases of nuclear pursuit are quite rare and unevenly distributed over time. Cases from early in the nuclear age, such as the United States and the Soviet Union, can have a disproportionate influence on the results of the analysis. It may be that the dynamics of proliferation were simply different for the earliest nuclear states. Perhaps proliferation is not what it used to be.

A Theory of Change in Nuclear Proliferation

How might the drivers of nuclear proliferation have changed over time? I examine this question for each of the three broad categories of proliferation theory described above: nuclear capabilities, nuclear motives, and international institutions.

Nuclear Capabilities

We might expect that supply-side considerations had a greater impact on state decisions to pursue nuclear weapons early in the nuclear age than in more recent years, for several reasons. First, knowledge of the nuclear fuel cycle, weapons design, and weaponization has become much more widespread over time.[9] While scientists in the US and Soviet nuclear programs were forced to work out most of the details of nuclear reactions on their own (or with help from espionage), the physics of nuclear weapons are now broadly understood and openly published; this includes once closely protected weapons design details, such as materials used in boosted weapons and thermonuclear bombs, neutron initiators, and explosive lenses.[10] Technical barriers to uranium enrichment and nuclear reactor design that required tremendous resources to overcome in the first nuclear programs are now routinely covered in nuclear engineering courses worldwide. Easier access to nuclear knowledge reduces the cost to states of pursuing a nuclear weapon; prospective proliferant states need not budget for a Manhattan Project when a good portion of the initial legwork has been done for them. Kemp (2014), for example, argues persuasively that the spread of gas centrifuge uranium enrichment technology, in particular, changed the nuclear proliferation landscape by putting the production of fissile material for a nuclear weapon within reach of states with minimal technical and human resources, even without foreign assistance.

Second, would-be nuclear states now have several options for obtaining the knowledge and technology necessary for nuclear weapons development. While the early weapons states once held nuclear secrets closely, the spread of global nuclear power has opened up multiple sources of nuclear assistance. States may, of course, still receive weapons-related help from their nuclear allies (Kroenig 2009). But they may also benefit from civilian nuclear cooperation agreements or multilateral nuclear aid from the International Atomic Energy Agency (IAEA) (Fuhrmann 2009; Brown and Kaplow 2014). Most providers of civilian nuclear assistance are careful to avoid aiding in weapons efforts, but nuclear technology is inherently dual use. Even seemingly benign forms of aid—general physics instruction, for example, or training in agricultural or medical applications of nuclear technology—can help to build expertise that translates to weapons work. As the A.Q. Khan network has shown, even sensitive technology is frequently available on the open market (Fitzpatrick 2007), and the number of potential suppliers has been increasing over time (Braun and Chyba 2004).[11] Each of these forms of assistance reduces the ultimate cost of a weapons effort, encouraging states that might otherwise have been resource-constrained to push forward with nuclear development. Perversely, the sheer availability of nuclear assistance also makes each individual cooperation agreement or nuclear aid project less important to the decision to begin a nuclear weapons effort.

Third, and partly because of the widespread availability of nuclear knowledge and foreign assistance, almost no state would now begin a nuclear weapons program from scratch. Most would-be proliferants can draw upon indigenous resources such as uranium deposits or a pool of experienced scientists and engineers. Many states have achieved a kind of nuclear latency—the underlying capacity to produce nuclear weapons—by virtue of their civilian nuclear power programs or academic research efforts (Sagan 2010; Mehta and Whitlark 2017). These existing facilities and resources provide a useful base from which to launch a broader weapons program and again alter the calculus for the decision to pursue weapons.

Finally, one way in which supply-side factors encourage nuclear proliferation is by reassuring states about their chances of proliferation success. Given the potential negative consequences of nuclear weapons development, states will pursue weapons only if they feel they have a good chance of reaching the finish line. The more difficult nuclear weapons development seems, the less likely states are to start a program. Early nuclear weapons programs appeared less likely to end in success. Prior to the Manhattan Project, of course, it was not clear that a weapon *could* be developed successfully. Nobel Prize-winning physicist Hans Bethe, who later made significant contributions to the US nuclear effort, recalled that he at first "considered … an atomic bomb so remote that I completely refused to have anything to do with it. … I thought we would never succeed in any practical way" (Rhodes 1986, 415).

Even after the US program and other early efforts demonstrated a weapon was possible, the task still appeared quite difficult. The first nuclear weapons states devoted tremendous resources to their nuclear development efforts. At its peak, for example, the US nuclear weapons program during World War II employed about 125,000 people, with some 500,000 working for the Manhattan Project at some point during the war; the Soviet program employed a similar number (Wellerstein 2013). Later nuclear programs, however, have had the examples of China, Israel, India, South Africa, Pakistan, and North Korea to draw from. States evaluating their own prospects for success today are likely to see weapons as more attainable, regardless of their own capabilities, than did potential proliferants in the 1950s and 1960s.

[9]That certain technologies were no longer secret or restricted does not mean that a country would necessarily be capable of using those technologies in practice. Several authors have pointed to the importance of tacit knowledge—knowledge that comes from human experience rather than impersonal instructions—in the realm of nuclear weapons (MacKenzie and Spinardi 1995; Montgomery 2005). For example, despite acquiring detailed information about US nuclear weapons work via espionage, Soviet weapons designers were forced to reinvent solutions to a number of problems that the United States had already successfully addressed (Holloway 1994).

[10]On the secrecy surrounding early nuclear weapons efforts and its eventual erosion, see Wellerstein (2021).

[11]That nuclear technology is dual use does not necessarily mean that civilian technology will be helpful in a country's nuclear weapons efforts, or even that sensitive nuclear assistance will speed up a country's weapons program. On the difficulty of translating foreign assistance to nuclear weapons progress, see Montgomery (2013).

Pushing back against these trends are the increasing efforts over the years by the international community to restrict trade in sensitive technology, particularly enrichment and reprocessing. The aftermath of India's 1974 nuclear test galvanized collective steps to limit nuclear supply, first in the form of the London Club of nuclear supplier states and later as the Nuclear Suppliers Group.[12] The United States in the 1970s expanded its own use of diplomatic and economic pressure to stop the spread of nuclear technology, both informally—as in the case of efforts to stop West German provision of fuel cycle technology to Brazil (Gray 2012)—and through threatened and imposed economic sanctions (Speier, Chow, and Starr 2001; Reynolds and Wan 2012; Miller 2014b). Despite these efforts, however, the overall trend in nuclear supply suggests that would-be proliferants today have greater access to nuclear technology than those at the dawn of the nuclear age.

Together, then, these arguments suggest the following hypothesis:

Supply hypothesis: The effect of supply-side factors on nuclear proliferation has decreased over time.

Nuclear Motives

The motivation for nuclear proliferation ultimately reflects a state's perception of the usefulness of nuclear weapons for its own security. If leaders believe that acquiring nuclear weapons will translate into better odds of regime survival, they will be more likely to seek weapons. If leaders see nuclear weapons as neutral or as diminishing national security—for example, by inviting preemptive attack—they will be less likely to seek weapons.

It may be that nuclear weapons were viewed as more essential for state security early in the nuclear age, or at least during the Cold War. In the immediate aftermath of World War II—and without an established nuclear taboo (Tannenwald 2007)—states saw nuclear weapons as potentially useful in conflict. And under the shadow of nuclear threats from the United States or the Soviet Union, at least some states believed the acquisition of a nuclear deterrent of their own was the only way to guarantee their security. With the end of the Cold War, however, many came to view the utility of nuclear weapons as significantly diminished. Nuclear deterrence, in particular, receded as a security imperative for many states, to the extent that most nuclear weapons states were willing to pare down their own arsenals considerably in the 1990s (Norris and Kristensen 2010). This logic suggests the following hypothesis:

Demand hypothesis I: The effect of demand-side factors on nuclear proliferation has decreased over time.

If the end of the Cold War brought a reduced risk of a great-power nuclear exchange, it also introduced a loosening of the entrenched alliance system from earlier in the nuclear age. The US and Soviet spheres of influence had formed the basis for many security commitments during the Cold War, some of which had been effective in limiting demand for nuclear weapons among key allies. The extended deterrence promise offered by the US nuclear umbrella, for example, seemed more credible in the context of superpower rivalry. States that would have been satisfied by alliance commitments in the past may now see a need to develop their own nuclear deterrent.

A related argument is that the end of Cold War constraints on US foreign policy enhanced the security threat to potential US adversaries, such as North Korea, Iran, Iraq, and Libya. These states may have felt the threat of US aggression more acutely as the United States' foreign policy preoccupation with the Soviet Union receded. For potential US adversaries, then, demand-side factors may have increased in importance over time.

More generally, nuclear weapons development seems like a sensible response to the general increase in the number of nuclear states and nuclear aspirants over time. As more states have investigated nuclear weapons efforts, and still more have built up a level of indigenous capability that could speed the path to the bomb if they decide to seek it, others may decide to follow suit. This may be because of particular concern about the programs or capabilities of neighbors and potential rivals—as posited by nuclear domino theories (Allison 2005; Miller 2014a)—or just a general response to the changing threat environment.

Changes in counterproliferation policy over time also may have affected the importance of demand-side factors. The increasing reliance on economic sanctions in response to suspected nuclear weapons efforts may lead states to recalculate the net benefit of nuclear pursuit.[13] Similarly, if the prospects of military action against nuclear facilities are perceived to be increasing over time for particular states—perhaps due to the example of Israeli strikes on the Osirak reactor in Iraq and Syria's Al-Kibar nuclear facility—demand-side factors may be more likely now to prompt nuclear restraint.[14]

These arguments suggest the following hypothesis:

Demand hypothesis II: The effect of demand-side factors on nuclear proliferation has increased over time.

International Institutions

Formal international institutions focusing on the spread of nuclear technology did not exist before the IAEA was formed in 1957 and did not actively seek to restrict nuclear weapons proliferation until the NPT entered into force in 1970. It may seem obvious, then, that the effect of international institutions on state proliferation decisions would be greater now than at the beginning of the nuclear age.

A number of analysts, however, see the NPT in particular as a weak institution that has little effect on state behavior.[15] Qualitative analyses of specific cases of nuclear proliferation and restraint have largely seen NPT membership as coincident with, rather than a cause of, the decision to forgo nuclear weapons (Reiss 1995; Campbell, Einhorn, and Reiss 2004). Even analysts who acknowledge a past constraining effect see the NPT and the broader nonproliferation regime in more recent years as teetering on the brink of collapse. Horovitz (2015) has documented a long history of analysts pointing to various flaws in the regime as evidence of its weakened condition and future irrelevance. These flaws include nuclear pursuit outside of the regime by Israel, India, and Pakistan (Asculai 2004; Fahmy 2006); violations from within by Iraq, Iran, and—until its withdrawal from the

[12]On the origins and early work of the Nuclear Suppliers Group, see Anstey (2018).

[13]This effect is in addition to the supply-side impact of sanctions designed to limit nuclear supply, discussed above.

[14]On military attacks against nuclear programs generally, see Fuhrmann and Kreps (2010). On Israel's strike against Iraq's Osirak reactor, see Braut-Hegghammer (2011, 2016). On Israel's strike against Syria's Al-Kibar facility, see Follath and Stark (2009).

[15]Some theorists in the realist tradition see international institutions like the NPT as largely epiphenomenal (Mearsheimer 1994).

treaty in 2003—North Korea (Huntley 2006; Allison 2010; Grand 2010); and a distinct lack of progress on nuclear disarmament (Kmentt 2013; Dhanapala and Duarte 2015). For these scholars, analysts, and practitioners, the NPT seemed to have lost whatever power it held in the early days of the nonproliferation regime, suggesting the following hypothesis:

Institutions hypothesis I: The effect of membership in the nuclear nonproliferation regime on nuclear proliferation has decreased over time.

It may be, on the other hand, that the regime has increased its constraining power over states since its inception. What was once a few isolated treaties has evolved over time into a full regime complex, with a dense network of dozens of agreements, conventions, and informal institutions governing multiple dimensions of nuclear security and proliferation (Carcelli et al. 2014; Kaplow 2022a). The NPT and related institutions help to change the calculus for proliferant states in two ways.

First, the nonproliferation regime has increased the cost over time for states that pursue nuclear weapons. One way that the NPT constrains state behavior is by providing information and a focusing mechanism for enforcement and punishment. As more states join the NPT and other aspects of the regime, this function of the treaty becomes more effective. Since nearly all nonnuclear states are now members of the NPT, states that seek to proliferate in the future will also be in violation of their international commitments, drawing enhanced scrutiny from the international community and increasing the chance of multilateral sanctions or other collective punishment.

Improvements over time in international safeguards have made the information-provision function of the NPT much more effective. Nuclear safeguards have expanded in their focus over time from merely verifying the correctness of state declarations to a state-level approach, seeking to ensure the completeness of declarations and the lack of covert nuclear work (Carlson et al. 1999; Pellaud 2000; Goldschmidt 2001; Tape and Pilat 2008). At the same time, new technologies have strengthened the hand of international inspectors, as verification and monitoring mechanisms began to incorporate commercial satellite imagery, remote monitoring tools, and environmental sampling techniques (Donohue 1998; Chitumbo, Robb, and Hilliard 2002; Schanfein 2008). Since 1997, states have been encouraged to sign an Additional Protocol to their safeguards agreements, allowing international inspectors broader access to nuclear facilities in some NPT member states (Hirsch 2004; Findlay 2007).

Second, the seeming effectiveness of the nuclear nonproliferation regime over time helps to reduce the motivation for nuclear pursuit among its members (Kaplow 2022a, 2022b).[16] State parties to the NPT might initially worry that their commitment to forgo weapons development will not be reciprocated. Over time, however, the treaty has demonstrated its effectiveness. No state has successfully acquired weapons while a member of the NPT, and those states that have been thought to be seeking weapons within the treaty—Iran, for example—have seen a coordinated response from the international community. The regime's track record has helped to build confidence in its members that they need not seek weapons of their own.

This logic suggests the following hypothesis:

Institutions hypothesis II: The effect of membership in the nuclear nonproliferation regime on nuclear proliferation has increased over time.

A Test of the Changing Drivers of Nuclear Proliferation

If the drivers of nuclear proliferation are changing over time, how would we know? As I describe above, most quantitative studies of nuclear proliferation assume the same data-generating process throughout the timeframe of the analysis. These static parameter models are unable to identify changes in the effects of particular factors over time.

Researchers can allow the effects of variables of interest to change over time in a traditional regression framework by interacting temporal dummy variables with key explanatory variables, but results from this approach can be cumbersome to interpret and it is rarely employed in practice. More commonly, scholars will disaggregate their data by time period,[17] but this approach calls for the analyst to identify sometimes arbitrary breakpoints. Methods expressly designed to identify variation in effects over time, such as change point analysis, are better suited for data with significant structural breaks than those that exhibit gradual temporal changes.[18] General time-varying effects models can require more data for valid inference and have not been widely adopted in the international relations literature.[19] None of these methodological approaches have been applied to the case of nuclear weapons proliferation.

To test my hypotheses, I relax the static parameter assumption in two ways. First, as an initial step in understanding temporal variation in the drivers of proliferation, I build empirical models using subsamples within the data that are delineated by time period. Then, as a more robust approach, I use a RWCV technique to identify changes over time in proliferation dynamics.

An Empirical Model of Nuclear Proliferation

For the analyses that follow, I construct empirical models of nuclear proliferation that include variables representing each of the hypotheses described above. While I use different subsamples of data and multiple methods in my models, in each case I employ time series data in which the unit of observation is the country year. The data run from 1950 to 2010. As my dependent variable, I use a binary measure of nuclear pursuit, updating nuclear weapons program data drawn from Jo and Gartzke (2007).[20] I consider states to have nuclear weapons programs only for the period prior to nuclear acquisition; observations after a state has nuclear weapons are dropped from the dataset.

In my models, I represent supply-side drivers of proliferation with three variables. First, I capture economic resources overall with a measure of the state's real Gross Domestic Product (GDP) (K. S. Gleditsch 2002). Second, I represent a country's technical capability with Jo and Gartzke's

[16] The NPT and associated institutions may also have been important in instantiating over time an international norm of nonproliferation (Rublee 2009).

[17] See, for example, Farber and Gowa (1997) and Gowa (1999).

[18] On change point analysis in political science, see Western and Kleykamp (2004) and Spirling (2007). Park (2010), employs a Poisson state space model that allows parameters to vary more frequently than in most change point methods, but this approach can be difficult to implement in a hypothesis-testing framework.

[19] For a general description of this class of models, see Hastie and Tibshirani (1993). A recent application in international relations is Anderson, Mitchell, and Schilling (2016).

[20] To Jo and Gartzke's (2007) coding of nuclear program dates, I add Libya (1970–2003) and Syria (2002–2007), and I adjust the end dates of programs in Iraq (1991) and Iran (2003).

(2007) seven-point composite measure of latent nuclear capacity.[21] Finally, I include a measure of multilateral nuclear assistance—a count of fuel cycle-related projects provided through the Technical Cooperation (TC) program administered by the IAEA (Brown and Kaplow 2014). This measure, however, is limited by time period, offering little empirical leverage before 1971. An alternative measure of civilian nuclear assistance from Fuhrmann (2009) yields similar results.[22]

Three variables capture demand-side drivers of proliferation. First, I include a measure of conflict history, employing a dichotomous variable that is set to 1 if a state has experienced an armed interstate conflict in the previous 5 years (N. P. Gleditsch et al. 2002), and 0 otherwise. Second, I capture nuclear rivalry with a dummy variable set to 1 if a state's rival had a nuclear weapons program in that year (Thompson and Dreyer 2011). Third, to capture the role of alliances or security guarantees in predicting proliferation, I follow common practice in the nuclear proliferation literature and include a binary variable set to 1 if a state has a defense pact with a nuclear weapons state in a given year (Gibler and Sarkees 2004). While these are commonly used measures of demand-side drivers of nuclear proliferation, it is worth noting that they may also capture some elements of supply-side theory. For example, the presence of a nuclear rival may lead states to seek weapons to protect themselves against a nuclear threat (a demand-side consideration), but those rival nuclear states may have incentives to restrict nuclear trade to their adversaries (a supply-side consideration).

To identify the association between weapons pursuit and membership in international institutions, I employ two variables. The first is a dummy variable set to 1 if a state is a member of the NPT in that year. Second, I include a measure of a state's embeddedness in the many institutions that make up the nuclear nonproliferation regime. This is simply the share of eligible agreements that a state has joined—that is, the number of nonproliferation treaties of which a state is a member, divided by the number of nonproliferation treaties the state is eligible to join (Carcelli et al. 2014; Kaplow 2022a).

A Split-Sample Approach

As an initial attempt to understand how the drivers of nuclear proliferation change over time, and to facilitate comparisons with existing work, I first split my data into three subsamples delineated by year: 1950–1969,1970–1989, and 1990–2010.[23] For each subsample, I build a separate model of nuclear pursuit using the variables and data structure described above, employing penalized likelihood logistic regression to address problems of rare-event bias and separation in the data (Firth 1993; Zorn 2005). To address temporal autocorrelation in the dependent variable, I include in each model a cubic polynomial of the number of years since a state last sought nuclear weapons (Carter and Signorino 2010). Because my data include country years in which a state is pursuing nuclear weapons, I also employ a cubic polynomial representing the number of years that have passed in the course of a state's nuclear program (Brown and Kaplow 2014). I report robust standard errors clustered by country.

If the factors associated with nuclear weapons programs have indeed shifted over time, we would expect to see a corresponding change in regression results across the subsamples. Table 1 presents those results. Model 1 is a regression based on the full time span of the data, while models 2–4 show results from the three data subsamples.[24] The change in statistical results over time is quite clear and suggests that the common practice of analyzing a single, pooled dataset may be obscuring interesting temporal variation in factors of interest.

The regression on the full data sample indicates a significant role for supply-side, demand-side, and institutional factors in driving proliferation. While the coefficient on real GDP did not reach statistical significance in the full sample, measures of nuclear capacity and multilateral nuclear assistance were strongly associated with nuclear weapons programs. Among variables representing nuclear motivation, conflict history and nuclear rivalry correlate with weapons development. These results also provide some limited support for institutional explanations for nuclear restraint. NPT member states are much less likely to seek nuclear weapons in the full data sample.

We should interpret the results for the temporal subsamples with some caution because our reduced statistical power in each time period may lead us to conclude there is no significant effect where one in fact exists. Analysis of the data subsamples, however, at least suggests that the importance of various drivers of proliferation may change over time. Nuclear capability seems to decline in importance over time, consistent with the supply hypothesis advanced above. In the earliest data subset, nuclear capacity is significantly associated with nuclear weapons programs, but this relationship disappears after 1970. The measure of multilateral nuclear assistance is not available in the earliest data subsample, but it is significantly associated with nuclear weapons programs between 1970 and 1989. The coefficient on this variable does not reach statistical significance, however, for the 1990–2010 time period.

The relationship over time between demand-side factors and proliferation is less clear. Of my three measures of nuclear motivation, the presence of a rival nuclear weapons program is the only one that reaches significance in the earliest cohort. Between 1970 and 1989, interstate conflict is associated with proliferation. In the most recent data sample, interstate conflict again falls out of significance, while defense pacts with nuclear powers take on more importance. Counter to expectations, having a defense pact with a nuclear state in this time period is associated with a greater likelihood of a nuclear weapons program. This echoes other mixed empirical findings in the literature on nuclear alliances (Bleek and Lorber 2014; Reiter 2014).

My measure of institutional drivers is not available in model 2, but NPT membership is strongly associated with nuclear restraint in the 1970–1989 time period. In the post-Cold War period, NPT membership is associated with an increased likelihood of nuclear pursuit. This result is not too

[21]Because this measure is unavailable after 2001, in the models below I simply extend each country's nuclear capacity from that year through 2010. This seems like a reasonable approximation for most countries because the components of Jo and Gartzke's (2007) index—including items such as known uranium deposits—are largely static. Dropping this variable from my models, however, or dropping all observations after 2001, does not affect my results.

[22]The Fuhrmann (2009) data run only through 2000, while the Brown and Kaplow (2014) data are available through 2010.

[23]This is equivalent to interacting the variables of interest with time dummies for each era being tested. I employ the split-sample approach for ease of interpretation.

[24]To allow for regression analysis across the full time span in model 1, the NPT membership variable is coded as 0 for all states prior to the entry into force of the NPT in 1970, the regime embeddedness variable is coded as 0 for all states prior to the creation of the IAEA in 1957, and the fuel cycle-related TC variable is coded as 0 for all states prior to the first fuel-cycle TC projects in 1971. Omitting these variables from model 1 yields similar results for the other factors in the analysis.

Table 1. Regression analysis of the drivers of proliferation by time period

		Model 1 *1950–2010*	*Model 2* *1950–1969*	*Model 3* *1970–1989*	*Model 4* *1990–2010*
Supply-side factors	Real GDP	−0.132 (0.086)	−0.134 (0.186)	0.073 (0.049)	0.053 (0.035)
	Nuclear capacity	**0.327** **(0.102)**	**0.676** **(0.095)**	0.166 (0.091)	0.063 (0.082)
	Fuel cycle-related IAEA TC	**0.186** **(0.047)**		**0.267** **(0.070)**	0.156 (0.092)
Demand-side factors	Interstate conflict (previous 5 years)	**1.700** **(0.429)**	0.737 (0.400)	**1.809** **(0.608)**	−0.065 (0.809)
	Rival with nuclear weapons program	**0.864** **(0.348)**	**1.749** **(0.593)**	0.451 (0.536)	1.240 (0.835)
	Defense pact with nuclear state	0.467 (0.426)	−0.013 (0.517)	−0.074 (0.566)	**0.835** **(0.352)**
Institutions	NPT membership	**−0.763** **(0.370)**		**−1.218** **(0.479)**	**2.767** **(0.584)**
	Regime embeddedness	−0.247 (0.899)		1.461 (1.393)	**−5.701** **(1.049)**
	Constant	**−4.179** **(0.571)**	**−6.206** **(0.204)**	**−4.057** **(0.475)**	**−3.571** **(0.370)**
	N	8,537	1,976	2,802	3,759

Notes: Penalized likelihood logistic regression coefficients with robust standard errors, clustered on country, reported in parentheses. Bold values are statistically significant ($p < 0.05$). Cubic polynomials of the years without a nuclear weapons program and the years since a program began are included in all models but not shown.

surprising when one considers that the states with nuclear weapons programs during this period—Iran, Iraq, Libya, North Korea, and Syria—were all NPT members.[25] The variable representing embeddedness in the broader nonproliferation regime is strongly associated with nuclear restraint in the post-1990 time period.

Changes in Predictive Power over Time

This analysis is perhaps more useful in demonstrating the risks of temporal pooling than it is in judging the changing drivers of proliferation over time. Finding a lack of statistical significance for a particular variable is not the same as finding that it has no effect (Rainey 2014). While disaggregating the full dataset makes it possible to see changes in the data-generating process over time, it also reduces statistical power in analysis of each subsample. This makes it more difficult to find a significant association even if one exists. This problem is exacerbated by the fact that nuclear weapons programs are relatively rare events. Splitting the sample by time period puts even more pressure on a small number of programs to support wide inferences about the effects of the drivers of proliferation. Another potential problem with this analysis is the arbitrary division of the sample into three equal parts. It may be that pooling in these three time periods—just like pooling across the whole dataset—obscures important variation in the factors of interest.

To at least partially address these issues, I turn to an analysis of the out-of-sample predictive validity of my variables of interest. I examine the contribution made by each category of factors to overall predictions of proliferation, asking, in effect, how the predictive power of the drivers of proliferation changes over time. Among the benefits of shifting away from a traditional regression framework for analysis is that prediction is, fundamentally, a way of capturing the substantive rather than statistical significance of empirical findings. A number of scholars in recent years have offered pointed critiques of the emphasis in scientific publishing on statistical significance (Gelman and Stern 2006; Simmons, Nelson, and Simonsohn 2011; Nuzzo 2014). A focus on predictive validity offers a promising alternative (Ward, Greenhill, and Bakke 2010). Predicting out of sample also avoids some of the model specification issues associated with regression models (Bell 2016), and gives us more confidence that our results depict relationships between variables that are actually present in our data and not just statistical noise.

To evaluate the predictive accuracy of the drivers of proliferation over time, I use a support vector machine (SVM), a type of statistical learning model. SVMs work by maximizing the separation in multidimensional space between two types of outcomes, in this case the presence or absence of a nuclear weapons program. Out-of-sample observations are then classified by which side of that boundary they fall on. Statistical learning techniques are commonly used in

[25] Of course, neither the positive nor negative association between NPT membership and nuclear weapons programs represents a convincing causal claim; states may well be more likely to join the NPT if they do not intend to seek weapons (Fuhrmann and Lupu 2016).

computer science and statistics and are increasingly being employed in political science. This approach is particularly helpful in dealing with data in which the effects of particular variables are likely to be highly conditional (Beck, King, and Zeng 2000). Because nuclear proliferation is a relatively rare event, factors that lead to nuclear pursuit may have little effect on state decision-making in most cases, but a dramatic effect on the observations most at risk of seeking weapons. The generalized linear models frequently used to analyze nuclear proliferation have trouble capturing these complex relationships in the data.

To examine the predictive power of each set of explanatory variables, I introduce a RWCV technique. RWCV allows for an analysis that avoids the temporal pooling issue of the preceding analysis and helps to identify common drivers of nuclear proliferation across different time periods. While RWCV offers the benefit of additional flexibility—it can capture similarities across time in a much more nuanced way than using large temporal subsets of the overall dataset—its results are best interpreted as highlighting the general shape of time-varying effects, rather than identifying the most important driver of proliferation in a particular year. The result of the RWCV analysis is a kind of temporal map, providing a new way for analysts to investigate high-level changes in international phenomena of interest.

To implement the RWCV analysis, I divide the full dataset into overlapping 5-year slices: 1950–1954, 1951–1955, etc. Next, I use a single 5-year subsample as training data to build a model of nuclear proliferation, holding back all the other 5-year subsamples to test the model's out-of-sample predictive validity.[26] That is, I first build a model based on data from 1950 to 1954. Then, I evaluate the ability of that model to predict proliferation from 1951 to 1955, 1952 to 1956, and so on, through all the 5-year time periods up to 2006–2010. I then build a model using data from the next 5-year period, 1951–1955, testing that model's performance on all the other 5-year subsamples of data. I repeat this process until every 5-year subsample has served as the training data for the model.

An important complicating factor in my analysis is the relative rarity of nuclear weapons programs in my data.[27] Statistical models built to maximize predictive accuracy naturally focus on the more frequent case in the data—that is, they tend to explain cases of nuclear restraint more than they do cases of nuclear pursuit. I address this problem with a synthetic minority oversampling technique that adjusts the training data used to build the statistical model (Chawla et al. 2002). It both increases the number of cases of nuclear weapons programs in the training data—generating synthetic cases using a nearest-neighbor method—and decreases the cases of nuclear restraint in the training data through systematic undersampling. As a result, we are left with a more balanced sample of cases in the training data, which leads to a more accurate prediction of the rare event. It is important to note that this procedure does not change the out-of-sample testing data that has been set aside to judge the predictive success of the model. The test data always reflects unchanged observations from the original dataset.

The same issue complicates the task of measuring predictive validity. In the presence of rare-event data, statistical models can be highly accurate overall without helping to predict the cases of most interest. A model that always predicts that states will not proliferate is not particularly useful for our purposes, but it would still accurately predict nearly all cases in my dataset. As an alternative to overall accuracy, I use a metric that is more sensitive to changes in the model's ability to predict cases of proliferation, rather than just nuclear restraint. The positive predictive value (PPV) of the model, also known as its precision, is defined as the number of true positives returned by the model divided by the number of positive predictions. Put another way, PPV is the share of "yes" predictions made by the model that turn out to be correct.[28] Higher values for PPV reflect better prediction of proliferation events.[29]

Figure 1 shows the performance of the full model over time. The horizontal axis shows the first year of each 5-year training window used to build the model, and the vertical axis shows the first year of the 5-year testing window used to evaluate the model's predictive validity. Darker squares represent more accurate predictions; PPVs for this model ranged from 0 to 1, with a mean of 0.19.

The squares along the diagonal of the figure represent in-sample prediction. We would expect the model to have an easier time with such predictions; models trained on a particular 5-year window should more accurately predict proliferation in the same time period. More interesting is the apparent temporal variation in the predictive power of the model as it moves away from the diagonal of the chart. If there were no change in the importance of particular drivers of proliferation over time, we would expect to see roughly the same level of prediction regardless of the training or testing cohort. Instead, we see that predictive power is strongly dependent on the time period covered by the training and testing data. Proliferation in the 1990s and 2000s, for example, is poorly predicted by the 1950s and 1960s (in the upper-left corner of the chart). Models trained in the mid-1970s and later, however, are generally quite accurate in predicting proliferation in the 1990s and 2000s. This result is consistent with a change in the data-generating process over time.

To understand how the predictive power of individual variables has changed over time, I repeated the RWCV procedure above using only supply-side, demand-side, and institutional variables, respectively. The results are shown in figure 2. Here, each chart shows the variation over time in the predictive power of a particular set of variables. All three models illustrate difficulties in applying the lessons of the early years of the nuclear age to more recent decades. In the first chart, models of nuclear supply trained using data from the 1950s through the mid-1970s are largely unable to successfully predict post-Cold War nuclear proliferation. This suggests a significant change in the importance of supply-side factors over time. In the second chart, highlighting the role of nuclear motivation, models trained on the earliest time periods have some predictive capacity against more recent decades, and models built with data from 1960 onward fared well in predicting proliferation through 1990.

[26]The choice of a 5-year rolling window is arbitrary; results appear similar with window size from 6 to 10 years. Windows smaller than 5 years suffer from problems with model convergence.

[27]This issue, known as class imbalance, has prompted a significant literature in computer science. See Sun, Wong, and Kamel (2009).

[28]Other metrics that are suitable for evaluating the prediction of rare events—such as the area under the receiver operating characteristic curve or the F_1 score (Swets 1988; Joshi 2002)—yield similar results.

[29]Predictive performance for models that suffer from class imbalance can also be improved by shifting the arbitrary threshold for what the model considers to be a positive prediction. For this analysis, because PPV is undefined if there are no positive predictions, I set the threshold for a "yes" prediction to the lower of (a) 0.5 or (b) the level necessary to generate at least one positive prediction for a given testing window. The general patterns in the result, however, are visible regardless of the threshold chosen. On threshold adjustment strategies for improving predictive analysis, see Yu et al. (2015), Zou et al. (2016), and Esposito et al. (2021).

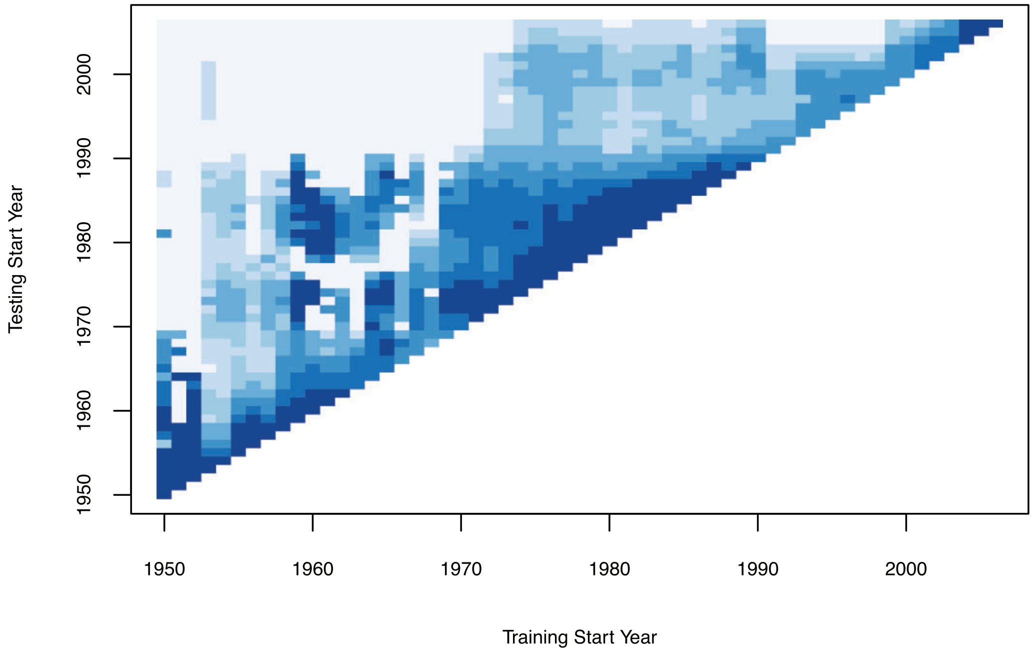

Figure 1. Predictive validity (full model)

The final chart illustrates results from a model of membership in the nuclear nonproliferation regime.[30] Here, it appears that models trained on the earliest and latest years of the regime—before the mid-1960s and after 1990—are stronger predictors of recent proliferation. Like the demand-side model, the institutions model suggests a different data-generating process for nuclear proliferation from the mid-1960s through the end of the Cold War. This makes some intuitive sense—as the NPT has grown to boast near-universal membership in recent years, the effectiveness of the signal of regime membership may have weakened.

We can summarize the findings of the RWCV technique by averaging the results for each 5-year test period across all of the training periods. This procedure yields the mean PPV for each 5-year test period across the full dataset, allowing us to identify which drivers of proliferation were the best predictors in each time period.[31] The results are illustrated in figure 3. The vertical axis shows the average PPV for each test period, starting in the year depicted on the horizontal axis.

The lines on the chart reflect the general trend shown in figure 2, with all three models illustrating more predictive validity in the 1950s, 1970s, and 1980s. The solid line in figure 3, representing supply-side drivers of proliferation, shows the decline in predictive power of this set of factors. A model based only on supply-side variables yielded a PPV of about 0.21 when attempting to predict nuclear weapons programs in the 1950s; the same factors had a PPV of only about 0.06 percentage points when predicting proliferation after 1990. Demand-side factors followed a similar pattern but were most predictive between 1970 and 1990. Institutional factors steadily increased in their ability to predict proliferation between the mid-1960s through the 1980s, with a steep decline in predictive power in the 1990s. For test periods since 2000, however, institutional models have been the best predictor of nuclear weapons programs, lending some support to the idea that the importance of institutional drivers of nuclear restraint has increased despite NPT naysayers.

While not definitive, the findings from these predictive models support the idea that the underlying dynamics of nuclear proliferation have changed over time. My findings suggest that supply-side, demand-side, and institutional factors have declined in predictive importance since the early 1990s, but that the predictive power of institutional factors has rebounded more recently. These temporal patterns in the ability of particular theories to account for nuclear weapons pursuit reveal an often-overlooked aspect of nuclear proliferation. As the international security environment shifts, so to it appears does the nuclear calculus of state leaders.

Conclusion

This paper highlights a potential problem with empirical analyses of nuclear proliferation. Most research that seeks to identify the causes of proliferation—and particularly work taking a quantitative approach to empirical testing—assumes that the effects of particular drivers of proliferation remain constant over time. This is a strong assumption. I argue, instead, that the importance of particular drivers of proliferation has shifted since the dawn of the nuclear age. My theory suggests that supply-side factors—including state capacity and the ability to obtain nuclear assistance—figure less prominently in state decision-making today than in early weapons programs. The trajectory for demand-side drivers of proliferation and for the effect of international institu-

[30]Data are not available for this model before 1957, the year the IAEA was founded. The IAEA is widely seen as the first institution within the nuclear nonproliferation regime.

[31]Put another way, this approach averages the values across each row of the charts shown in figure 2. While figure 2 shows predictive power only above the diagonal (that is, training data are only used to predict test data for future time periods), I calculate the mean PPV using all possible training data. So, for example, the average PPV for the 1970–1974 period includes the PPV for training data from 1950 to 1970, as well as training data from 1970 to 2010. By including training data both before and after the test period, I ensure that all test periods include the same number of training datasets, equally distributed across time.

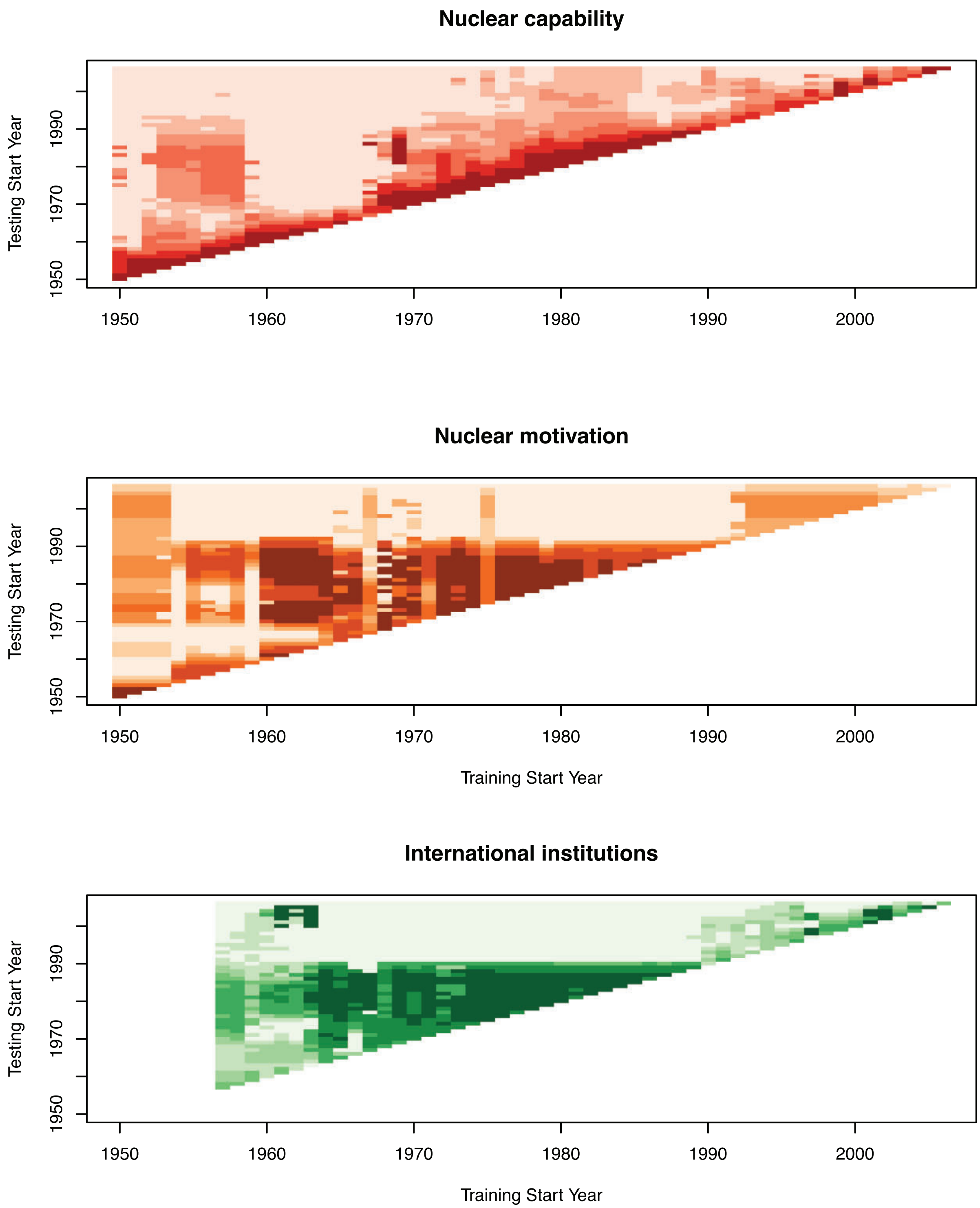

Figure 2. Predictive validity (individual factors)

tions is less clear, and I propose competing hypotheses in each of these areas that lay out the case for a stronger or weaker role over time.

An analysis of the out-of-sample predictive power of each set of variables suggests that the effects of the various drivers of proliferation are not static over time, as existing work assumes, but rather that the data-generating process for nuclear proliferation has changed. Consistent with expectations, I find that supply-side factors have declined in importance; demand-side factors, too, appear less capable over time of predicting nuclear pursuit. The evidence for institutional factors is more mixed, as the predictive power of these variables has recently seemed to reverse a decline.

These findings are important for the ability of researchers to contribute to today's nonproliferation policy debates. This analysis suggests that scholars of nuclear proliferation should be careful about applying results from the entirety of the nuclear age to address con-

Figure 3. Average predictive validity for each 5-year test period

temporary issues in nonproliferation policy. We would, of course, be cautious in applying lessons from the Manhattan Project to our analysis of Iran's contemporary nuclear program; we should recognize that some of our empirical results may stem from the same type of comparison.

This article focuses on the changing nuclear proliferation landscape, but its approach is broadly applicable across international relations and the social sciences. Pooling observations over time is a common mode of quantitative inquiry, and scholars rarely relax the static parameter assumption that follows from this approach. My results suggest that temporal pooling risks biasing our results. Scholars should look for empirical strategies that allow the effects of their variables of interest to change over time, examining whether their findings hold up across different time periods within the data. The RWCV analysis used here is one option for researchers seeking to understand the time-varying effects of their factors of interest.

References

Allison, Graham. 2005. "The Cascade of Proliferation." *Atlantic Monthly* 296 (3): 36–7.

———. 2010. "Nuclear Disorder: Surveying Atomic Threats." *Foreign Affairs* 89 (1): 74–85.

Anderson, Christopher C., Sara McLaughlin Mitchell, and Emily U. Schilling. 2016. "Kantian Dynamics Revisited: Time-Varying Analyses of Dyadic IGO-Conflict Relationships." *International Interactions* 42 (4): 644–76.

Anstey, Isabelle. 2018. "Negotiating Nuclear Control: The Zangger Committee and the Nuclear Suppliers' Group in the 1970s." *The International History Review* 40 (5): 975–95.

Asculai, Ephraim. 2004. *Rethinking the Nuclear Non-Proliferation Regime.* Tel Aviv: Jaffee Center for Strategic Studies, Tel Aviv University.

Beck, Nathaniel, Gary King, and Langche Zeng. 2000. "Improving Quantitative Studies of International Conflict: A Conjecture." *American Political Science Review* 94 (1): 21–35.

Bell, Mark S. 2016. "Examining Explanations for Nuclear Proliferation." *International Studies Quarterly* 60 (3): 520–9.

Bleek, Philipp C., and Eric B. Lorber. 2014. "Security Guarantees and Allied Nuclear Proliferation." *Journal of Conflict Resolution* 58 (3): 429–54.

Braun, Chaim, and Christopher F. Chyba. 2004. "Proliferation Rings: New Challenges to the Nuclear Nonproliferation Regime." *International Security* 29 (2): 5–49.

Braut-Hegghammer, Målfrid. 2011. "Revisiting Osirak: Preventive Attacks and Nuclear Proliferation Risks." *International Security* 36 (1): 101–32.

———. 2016. *Unclear Physics: Why Iraq and Libya Failed to Build Nuclear Weapons.* Ithaca, NY: Cornell University Press.

Brewer, Eric. 2020. "*Toward a More Proliferated World?*" Washington, DC: Center for a New American Security and Center for Strategic and International Studies.

Brown, Robert L., and Jeffrey M. Kaplow. 2014. "Talking Peace, Making Weapons: IAEA Technical Cooperation and Nuclear Proliferation." *Journal of Conflict Resolution* 58 (3): 402–28.

Campbell, Kurt M., Robert J. Einhorn, and Mitchell Reiss. 2004. *The Nuclear Tipping Point: Why States Reconsider Their Nuclear Choices.* Washington, DC: Brookings Institution.

Carcelli, Shannon, Erik Gartzke, Rebecca Davis Gibbons, and Jeffrey M. Kaplow. 2014. "The Nuclear Regime Complex: A New Dataset." *Paper Prepared for the 2014 International Studies Association Annual Convention, Toronto, Canada.*

Carlson, John, Victor Bragin, John Bardsley, and John Hill. 1999. "Nuclear Safeguards as an Evolutionary System." *The Nonproliferation Review* 6 (2): 109–17.

Carter, David B., and Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18 (3): 271–92.

Central Intelligence Agency. 1963. *Likelihood and Consequences of a Proliferation of Nuclear Weapons Systems.* National Intelligence Estimate. Wilson Center Digital Archive.

———. 1964. *Prospects for a Proliferation of Nuclear Weapons over the Next Decade.* National Intelligence Estimate. National Security Archive Electronic Briefing Book 401.

Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. "SMOTE: Synthetic Minority Over-Sampling Technique." *Journal of Artificial Intelligence Research* 16: 321–57.

Chitumbo, K., S. Robb, and J. Hilliard. 2002. "Use of Commercial Satellite Imagery in Strengthening IAEA Safeguards." In *Commercial Satellite Imagery: A Tactic in Nuclear Weapon Deterrence*, eds. Bhupendra Jasani and Gotthard Stein, 23–48. Chichester: Praxis Publishing Ltd.

Cirincione, Joseph. 2008. *Bomb Scare: the History and Future of Nuclear Weapons*. New York, NY: Columbia University Press.

Coe, Andrew J., and Jane Vaynman. 2015. "Collusion and the Nuclear Nonproliferation Regime." *The Journal of Politics* 77 (4): 983–97.

Dhanapala, Jayantha, and Sergio Duarte. 2015. "Is There a Future for the NPT?" *Arms Control Today* 45 (6): 8–10.

Donohue, David L. 1998. "Strengthening IAEA Safeguards through Environmental Sampling and Analysis." *Journal of Alloys and Compounds* 271–3: 11–8.

Esposito, Carmen et al. 2021. "GHOST: Adjusting the Decision Threshold to Handle Imbalanced Data in Machine Learning." *Journal of Chemical Information and Modeling* 61 (6): 2623–40.

Fahmy, Nabil. 2006. "An Assessment of International Nuclear Nonproliferation Efforts after 60 Years." *The Nonproliferation Review* 13 (1): 81–7.

Farber, Henry S., and Joanne Gowa. 1997. "Common Interests or Common Polities? Reinterpreting the Democratic Peace." *The Journal of Politics* 59 (2): 393–417.

Findlay, Trevor. 2007. "Looking Back: The Additional Protocol." *Arms Control Today* 37 (9): 47–51.

Firth, David. 1993. "Bias Reduction of Maximum Likelihood Estimates." *Biometrika* 80 (1): 27–38.

Fitzpatrick, Mark. 2007. *Nuclear Black Markets: Pakistan, A.Q. Khan and the Rise of Proliferation Networks*. London: International Institute for Strategic Studies.

Follath, Erich, and Holgar Stark. 2009. "The Story of 'Operation Orchard': How Israel Destroyed Syria's Al Kibar Nuclear Reactor." *Spiegel Online*.

Fuhrmann, Matthew. 2009. "Spreading Temptation: Proliferation and Peaceful Nuclear Cooperation Agreements." *International Security* 34 (1): 7–41.

Fuhrmann, Matthew, and Jeffrey D. Berejikian. 2012. "Disaggregating Noncompliance: Abstention versus Predation in the Nuclear Nonproliferation Treaty." *Journal of Conflict Resolution* 56 (3): 355–81.

Fuhrmann, Matthew, and Michael C. Horowitz. 2015. "When Leaders Matter: Rebel Experience and Nuclear Proliferation." *The Journal of Politics* 77 (1): 72–87.

Fuhrmann, Matthew, and Sarah E. Kreps. 2010. "Targeting Nuclear Programs in War and Peace: A Quantitative Empirical Analysis, 1941—2000." *Journal of Conflict Resolution* 54 (6): 831–59.

Fuhrmann, Matthew, and Yonatan Lupu. 2016. "Do Arms Control Treaties Work? Assessing the Effectiveness of the Nuclear Nonproliferation Treaty." *International Studies Quarterly* 60 (3): 530–9.

Fuhrmann, Matthew, and Benjamin Tkach. 2015. "Almost Nuclear: Introducing the Nuclear Latency Dataset." *Conflict Management and Peace Science* 32 (4): 443–61.

Gartzke, Erik, and Matthew Kroenig. 2009. "A Strategic Approach to Nuclear Proliferation." *Journal of Conflict Resolution* 53 (2): 151–60.

Gelman, Andrew, and Hal Stern. 2006. "The Difference between 'Significant' and 'Not Significant' is Not Itself Statistically Significant." *The American Statistician* 60 (4): 328–31.

Gerzhoy, Gene. 2015. "Alliance Coercion and Nuclear Restraint: How the United States Thwarted West Germany's Nuclear Ambitions." *International Security* 39 (4): 91–129.

Gibbons, Rebecca Davis. 2020. "Supply to Deny: The Benefits of Nuclear Assistance for Nuclear Nonproliferation." *Journal of Global Security Studies* 5 (2): 282–98.

Gibler, Douglas M., and Meredith Reid Sarkees. 2004. "Measuring Alliances: The Correlates of War Formal Interstate Alliance Dataset, 1816–2000." *Journal of Peace Research* 41 (2): 211–22.

Gleditsch, Kristian Skrede. 2002. "Expanded Trade and GDP Data." *Journal of Conflict Resolution* 46 (5): 712–24.

Gleditsch, Nils Petter et al. 2002. "Armed Conflict 1946–2001: A New Dataset." *Journal of Peace Research* 39 (5): 615–37.

Goldschmidt, Pierre. 2001. "Strengthened Safeguards: Meeting Present and Future Challenges." *IAEA Bulletin* 43 (4): 6–11.

Gowa, Joanne. 1999. *Ballots and Bullets: The Elusive Democratic Peace*. Princeton, NJ: Princeton University Press.

Grand, Camille. 2010. "The Non-Proliferation Treaty in an Era of Proliferation Crises." In *Nuclear Weapons after the 2010 NPT Review Conference*, edited by Jean Pascal Zanders, 13–26. European Union Institute for Security Studies (EUISS).

Gray, William Glenn. 2012. "Commercial Liberties and Nuclear Anxieties: the US-German Feud over Brazil, 1975–7." *The International History Review* 34 (3): 449–74.

Hastie, Trevor, and Robert Tibshirani. 1993. "Varying-Coefficient Models." *Journal of the Royal Statistical Society* 55 (4): 757–96.

Hirsch, Theodore. 2004. "The IAEA Additional Protocol: What It Is and Why It Matters." *The Nonproliferation Review* 11 (3): 140–66.

Holloway, David. 1994. *Stalin and the Bomb*. New Haven, CT: Yale University Press.

Horovitz, Liviu. 2015. "Beyond Pessimism: Why the Treaty on the Non-Proliferation of Nuclear Weapons Will Not Collapse." *Journal of Strategic Studies* 38 (1–2): 126–58.

Horowitz, Michael C., and Neil Narang. 2014. "Poor Man's Atomic Bomb? Exploring the Relationship between 'Weapons of Mass Destruction'." *Journal of Conflict Resolution* 58 (3): 509–35.

Huntley, Wade L. 2006. "Rebels without a Cause: North Korea, Iran and the NPT." *International Affairs* 82 (4): 723–42.

Hymans, Jacques E. C. 2006. *The Psychology of Nuclear Proliferation: Identity, Emotions, and Foreign Policy*. Cambridge: Cambridge University Press.

Jo, Dong-Joon, and Erik Gartzke. 2007. "Determinants of Nuclear Weapons Proliferation." *Journal of Conflict Resolution* 51 (1): 167–94.

Joshi, M. V. 2002. "On Evaluating Performance of Classifiers for Rare Classes." *2002 IEEE International Conference on Data Mining*, 641–4.

Kaplow, Jeffrey M. 2022a. *Signing away the Bomb: The Surprising Success of the Nuclear Nonproliferation Regime*. Cambridge: Cambridge University Press.

———. 2022b. "State Compliance and the Track Record of International Security Institutions: Evidence from the Nuclear Nonproliferation Regime." *Journal of Global Security Studies* 7 (1).

Kemp, R. Scott. 2014. "The Nonproliferation Emperor Has no Clothes." *International Security* 38 (4): 39–78.

Kennedy, John F. 1963. "News Conference 52." Presented at the US State Department, Washington, DC. John F. Kennedy Library.

Kmentt, Alexander. 2013. "How Divergent Views on Nuclear Disarmament Threaten the NPT." *Arms Control Today* 43 (10): 8–13.

Kroenig, Matthew. 2009. "Importing the Bomb: Sensitive Nuclear Assistance and Nuclear Proliferation." *Journal of Conflict Resolution* 53 (2): 161–80.

MacKenzie, Donald, and Graham Spinardi. 1995. "Tacit Knowledge, Weapons Design, and the Uninvention of Nuclear Weapons." *American Journal of Sociology* 101 (1): 44–99.

Mearsheimer, John J. 1994. "The False Promise of International Institutions." *International Security* 19 (3): 5–49.

Mehta, Rupal N., and Rachel Elizabeth Whitlark. 2017. "The Benefits and Burdens of Nuclear Latency." *International Studies Quarterly* 61 (3): 517–28.

Miller, Nicholas L. 2014a. "Nuclear Dominoes: A Self-Defeating Prophecy?" *Security Studies* 23 (1): 33–73.

———. 2014b. "The Secret Success of Nonproliferation Sanctions." *International Organization* 68 (4): 913–44.

———. 2017. "Why Nuclear Energy Programs Rarely Lead to Proliferation." *International Security* 42 (2): 40–77.

Montgomery, Alexander H. 2005. "Ringing in Proliferation: How to Dismantle an Atomic Bomb Network." *International Security* 30 (2): 153–87.

———. 2013. "Stop Helping Me: When Nuclear Assistance Impedes Nuclear Programs." In *The Nuclear Renaissance and International Security*, edited by Adam N. Stulberg and Matthew Fuhrmann, 177–202. Palo Alto, CA: Stanford University Press.

Norris, Robbert S., and Hans M. Kristensen. 2010. "Global Nuclear Weapons Inventories, 1945–2010." *Bulletin of the Atomic Scientists* 66 (4): 77–83.

Nuzzo, Regina. 2014. "Scientific Method: Statistical Errors." *Nature News* 506 (7487): 150–2.

Park, Jong Hee. 2010. "Structural Change in U.S. Presidents' Use of Force." *American Journal of Political Science* 54 (3): 766–82.

Pellaud, Bruno. 2000. "The Strengthened Safeguards System: Objectives, Challenges and Expectations." In *Tightening the Reins*, edited by Erwin Häckel and Gotthard Stein, 89–98. Berlin: Springer.

Rainey, Carlisle. 2014. "Arguing for a Negligible Effect." *American Journal of Political Science* 58 (4): 1083–91.

Reiss, Mitchell. 1995. *Bridled Ambition: Why Countries Constrain Their Nuclear Capabilities*. Washington, DC: Woodrow Wilson Center Press.

Reiter, Dan. 2014. "Security Commitments and Nuclear Proliferation." *Foreign Policy Analysis* 10 (1): 61–80.

Reynolds, Celia L., and Wilfred T. Wan. 2012. "Empirical Trends in Sanctions and Positive Inducements in Nonproliferation." In *Sanctions, Statecraft,* *and Nuclear Proliferation*, edited by Etel Solingen, 56–122. Cambridge: Cambridge University Press.

Rhodes, Richard. 1986. *The Making of the Atomic Bomb*. New York, NY: Simon & Schuster.

Rublee, Maria Rost. 2009. *Nonproliferation Norms: Why States Choose Nuclear Restraint*. Athens, GA: University of Georgia Press.

Sagan, Scott D. 1996. "Why Do States Build Nuclear Weapons? Three Models in Search of a Bomb." *International Security* 21 (3): 54–86.

———. 2010. "Nuclear Latency and Nuclear Proliferation." In *Forecasting Nuclear Proliferation in the 21st Century*, edited by William C. Potter and Gaukar Mukhatzhanova, 80–101. Palo Alto, CA: Stanford University Press.

———. 2011. "The Causes of Nuclear Weapons Proliferation." *Annual Review of Political Science* 14 (1): 225–44.

Schanfein, Mark. 2008. "International Atomic Energy Agency Unattended Monitoring Systems." In *Nuclear Safeguards, Security, and Nonproliferation*, edited by James E. Doyle, 113–33. Burlington, MA: Butterworth-Heinemann.

Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science* 22 (11): 1359–66.

Singh, Sonali, and Christopher R. Way. 2004. "The Correlates of Nuclear Proliferation: A Quantitative Test." *Journal of Conflict Resolution* 48 (6): 859–85.

Smith, Bradley C., and William Spaniel. 2021. "Do Nonproliferation Agreements Constrain?" *Journal of Peace Research* 58 (6): 1163–77.

Solingen, Etel. 2007. *Nuclear Logics: Contrasting Paths in East Asia and the Middle East*. Princeton, NJ: Princeton University Press.

Spector, Leonard S. 2016. "A Proliferation Plateau May Offer Unique Opportunities." *Arms Control Today* 46 (3): 14–9.

Speier, Richard H., Brian G. Chow, and S. Rae Starr. 2001. *Nonproliferation Sanctions*. Santa Monica, CA: RAND Corporation.

Spirling, Arthur. 2007. "Bayesian Approaches for Limited Dependent Variable Change Point Problems." *Political Analysis* 15 (4): 387–405.

Sun, Yanmin, Andrew K. C. Wong, and Mohamed S. Kamel. 2009. "Classification of Imbalanced Data: A Review." *International Journal of Pattern Recognition and Artificial Intelligence* 23 (4): 687–719.

Swets, John A. 1988. "Measuring the Accuracy of Diagnostic Systems." *Science* 240 (4857): 1285–93.

Tannenwald, Nina. 2007. *The Nuclear Taboo: The United States and the Non-Use of Nuclear Weapons since 1945*. Cambridge: Cambridge University Press.

Tape, James, and Joseph F. Pilat. 2008. "Nuclear Safeguards and the Security of Nuclear Materials." In *Nuclear Safeguards, Security, and Nonproliferation*, edited by James E. Doyle, 17–30. Burlington, MA: Butterworth-Heinemann.

Thompson, William R., and David R. Dreyer. 2011. *Handbook of International Rivalries: 1494–2010*. Washington, DC: Congressional Quarterly Press.

Ward, Michael D., Brian D. Greenhill, and Kristin M. Bakke. 2010. "The Perils of Policy by *P*-Value: Predicting Civil Conflicts." *Journal of Peace Research* 47 (4): 363–75.

Way, Christopher, and Jessica L. P. Weeks. 2014. "Making It Personal: Regime Type and Nuclear Proliferation." *American Journal of Political Science* 58 (3): 705–19.

Wellerstein, Alex. 2013. "How Many People Worked on the Manhattan Project?" *Restricted Data: The Nuclear Secrecy Blog*. Accessed March 11, 2022, https://blog.nuclearsecrecy.com/2013/11/01/many-people-worked-manhattan-project/.

———. 2021. *Restricted Data*. Chicago, IL: University of Chicago Press.

Western, Bruce, and Meredith Kleykamp. 2004. "A Bayesian Change Point Model for Historical Time Series Analysis." *Political Analysis* 12 (4): 354–74.

Yu, Hualong et al. 2015. "Support Vector Machine-Based Optimized Decision Threshold Adjustment Strategy for Classifying Imbalanced Data." *Knowledge-Based Systems* 76: 67–78.

Zorn, Christopher. 2005. "A Solution to Separation in Binary Response Models." *Political Analysis* 13 (2): 157–70.

Zou, Quan et al. 2016. "Finding the Best Classification Threshold in Imbalanced Classification." *Big Data Research* 5: 2–8.

Kaplow, Jeffrey M. (2024) The Changing Face of Nuclear Proliferation. *International Studies Quarterly*, https://doi.org/10.1093/isq/sqae045