Modern Warfare: The role of state-sponsored cyber attacks in international conflict

Adam Kaplan*

April 2019

Abstract

The technological revolution did not stop short of international conflict. In 2010 the world was shocked to find out that a cyber weapon developed by nation states, Stuxnet, was used against another nation state, Iran. In this paper, I provide an overview of the research done on international cyber conflict, clarify the basic properties/assumptions made about cyber weapons, and for the first time (to my knowledge) include a quantitative analysis of past and current theories, assumptions about cyber weapons and their role in asymmetric warfare. I find that geographic contiguity has a positive effect on state-sponsored cyber attack onset and cyber weapons are expensive. Further, my results suggest that this high barrier of entry makes cyber weapons unsuitable for asymmetric warfare for non-espionage purposes. They also suggest that cyber weapons do not necessarily replace traditional weapons nor are always used in addition to them.

1 Introduction

Almost all of our life is in the "cloud," and this trend is only growing. It is not only individuals, but also companies and governments that are digitizing everything from national security secrets to missile launch panels. All this digitization is met with euphoria by the consumers. However, older attacks such as Stuxnet, which destroyed nuclear centrifuges in

*Undergraduate student at University of Rochester

Iran (Karnouskos, 2011), and more recent attacks such as the Russian interference in the 2016 U.S. elections show that computers are not a safe place for all this information. On top of it all, currently, policymakers are unsure how to legislate the digital world. As figure 1 shows, the number of cyber attacks per year is increasing, making this problem ever more important to tackle.

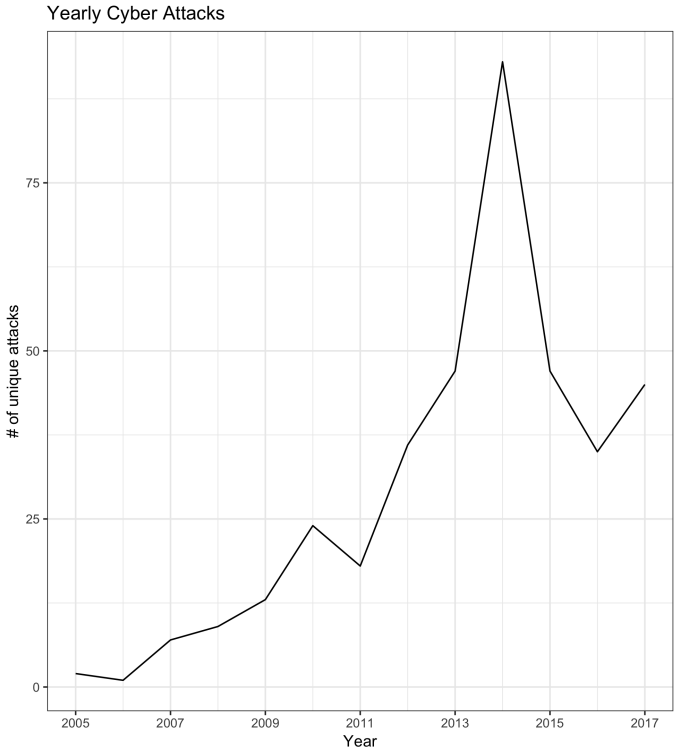


Figure 1: Illustrates the number of distinct cyber attacks launched annually from 2005 until 2017. Each of these attacks might affect multiple targets, but is counted only once.

While there is immense technological understanding of cyber security, the strategic underpinnings of state actors launching and developing cyber weapons has received relatively little attention. I contribute to the debate in three ways. First, I summarize the current literature by enumerating the theories and carefully defining the basic concepts (e.g. cyber attack). Second, to my knowledge I am the first person to conduct a quantitative analysis of state-sponsored cyber attacks, using data acquired from the Council on Foreign Relations (*Connect the Dots on State-Sponsored Cyber Incidents*, n.d.). I use this data to characterize basic properties of cyber weapons. These are the cost, influence of contiguity, the attri-

bution problem, and non-lethality of this technology. This foundational knowledge is not agreed upon in the current literature. I show empirically that cyber weapons are in fact expensive, launched primarily by military superpowers and contrary to widespread agreement, geographic contiguity has a positive effect on state-sponsored cyber attack onset.

Third, I apply this new knowledge to determine the role of cyber weapons in asymmetric warfare as well as to characterize the interaction between traditional and cyber military means. That is, are cyber weapons used alongside conventional weapons or do they substitute for them? I find support for the hypothesis that states with currently powerful militaries are more likely to launch state-sponsored cyber attacks. The results also suggest that cyber weapons are not well suited for asymmetric conflict. I find no support, however, for a complementary nor substitute view on the interaction between traditional and cyber weapons. Hence, it is not necessarily true that cyber weapons always have to be used alongside traditional weapons as Gartzke (2013) suggests. Neither is it true that cyber weapons necessarily replace traditional ones. The compromise seems to be that leaders choose cyber weapons where they are better than traditional ones, such as espionage, and otherwise use cyber weapons in tandem with conventional ones.

2 Theory

First, I highlight some core definitions used in the literature to be precise about what this paper studies. Second, I characterize basic properties of cyber weapons that impact how cyber weapons can be used by nation states. These properties are the high development and deployment costs, the positive impact of contiguity on cyber attacks, the attribution problem and finally the lack of observed fatalities. Third, I focus on what impacts state-sponsored cyber attack onset more generally and the possibility of asymmetric war in cyberspace. I argue that due to high development costs and the direct transfer of strong offensive capabilities into strong defensive capabilities (Lindsay, 2013), the odds are not in favor of militarily

weaker states attacking stronger ones. In other words, I claim asymmetric cyber warfare is not feasible.

2.1 Some Definitions

Due to the novelty of this topic and its complexity it is beneficial to define the concepts thrown around. This has been addressed by Rid (2012), Junio (2013) and Kello (2013) early on in the literature, both by suggesting definitions as well as studying the possible impact of them on possible policies through a constructivist lens. While this paper does not focus on the constructivist questions addressed by Kello (2013), it does use the suggested definitions. The reason being that there is a lack of legal and colloquial definitions for many of the concepts. Making it imperative to establish the precise meaning for *cyberspace*, *cyber weapons*, *cyber attacks* and *state-sponsored cyber attacks*, before delving into the theories.

> *Cyberspace.* Refers to "all computer systems and networks in existence, including air-gapped systems." (Kello, 2013)
>
> *Cyber weapon.* Refers to the code, i.e. a program used in a cyber attack as defined below.
>
> *Cyber attack.* "Refers to the use of code to interfere with the [confidentiality[1], integrity[2], or availability[3]] of a computer system for a political or strategic purpose."[4] (Kello, 2013)
>
> *State-sponsored cyber attack.* Refers to all cyber attacks which have been executed by a government directly or through a third-party hired by a government.

Note that both espionage cyber attacks and non-espionage, which will be discussed in

[1] "Confidentiality is the concealment of information or resources." (Bishop, 2012, Ch. 1.1.1)

[2] "Integrity refers to the trustworthiness of data or resources, and it is usually phrased in terms of preventing improper or unauthorized change." (Bishop, 2012, Ch. 1.1.2)

[3] "Availability refers to the ability to use information or resources." (Bishop, 2012, Ch. 1.1.3)

[4] I adapted Kello's (2013) definition based on the computer security literature, to be more precise about "functionality of a computer system."

more detail shortly, are considered to be cyber weapons since cyber espionage violates confidentiality of a computer system and non-espionage attacks violate integrity and/or availability. Furthermore, given recent coverage of alleged Russia-backed cyber misinformation campaigns, I want to point out that the definition of a cyber weapon does *not* include misinformation campaigns since they are not a piece of code. Thus, this paper does not address misinformation campaigns.

2.2 Properties of Cyber Weapons

In the conflict and security studies literatures, some basic properties of conventional, nuclear, biological, and chemical weapons are accepted. Examples are the high monetary and political cost of developing nuclear weapons or the logistic difficulties of deploying troops across the globe. For cyber weapons there is no consensus about such properties. For instance, Gartzke (2013) argues that weapons in cyberspace carry a high cost, Kello (2013) however, states that cyberspace has a "low barrier of entry." Thus, it is crucial to not only characterize basic properties for cyber weapons, but also to examine the empirical support for them.

2.2.1 Types of Cyber Weapons

Although the preceding definition of a cyber weapon helps narrow the concept, there are still quite a few activities that fit within it. In particular, breaking into an online system to steal trade secrets violates espionage and confidentiality. A distributed disruption of service (DDOS) attack, blocks access to websites and thus violates availability. Both of these satisfy the definition and classify as cyber weapons, but not only are their outcomes different, so are the intentions for launching them in the first place. The same concept is highlighted in the difference between conventional bombs and nuclear bombs. They both satisfy the definition of a bomb, but it would be ludicrous to assume that the strategic considerations of using these two are equivalent.

Table 1 displays the definitions for the different types of cyber attacks used by the Council

of Foreign Relations (*Connect the Dots on State-Sponsored Cyber Incidents*, n.d.), the source of the data used in this paper. The examples shown are brief summaries[5] of actual cyber incidents included in the dataset.

In this paper, I consider 2 groups of attacks, *espionage* and *non-espionage* (defacement, doxing, DDOS, data destruction, sabotage grouped together). The motivation for which is that non-espionage attacks break integrity and availability of computer systems, making them closer to the intuitive idea of what a weapon is supposed to do. On the other hand, espionage attacks break confidentiality, making them primarily an intelligence gathering tool. The majority of the literature does not explicitly classify cyber attacks (Gartzke, 2013; Junio, 2013; D. Betz, 2012; D. J. Betz & Stevens, 2013), but implicitly ignores cyber espionage, effectively addressing only non-espionage attacks.

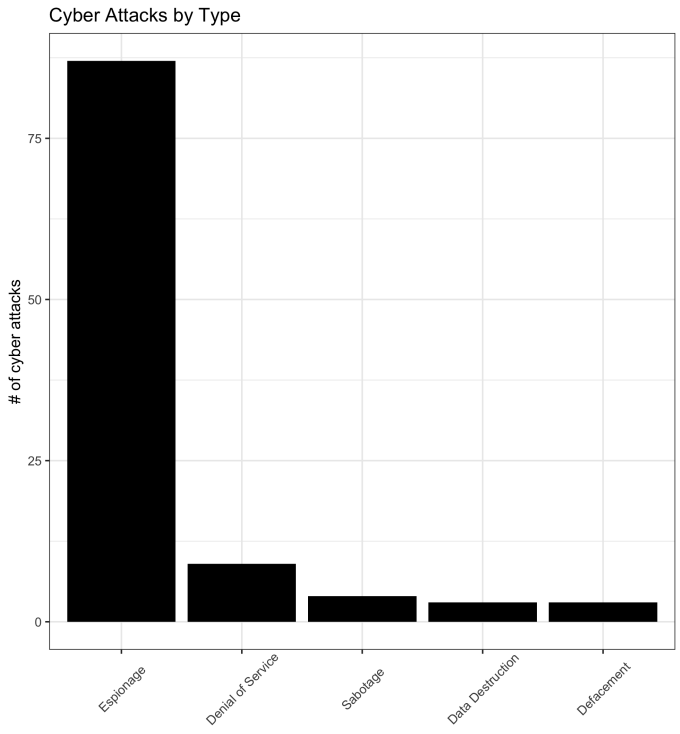


Figure 2: The distribution of unique attacks launched based on their type as defined in table 1. This is taken directly from the data used in the statistical analysis.

[5]Each observation in the dataset includes a description about the incident. I paraphrased this description for brevity.

Definition		**Example**
Defacement	The unauthorized act of changing the appearance of a website or social media account.	Posting of false remarks in support of Saudi Arabia on the Qatar News Agency.
Doxing	The act of searching and publishing private or identifying information about an individual or group on the internet, typically with malicious intent.	Release of employee emails during the Sony Entertainment Pictures hack
Distributed Denial of Service (DDOS)	The intentional paralyzing of a computer network by flooding it with data sent simultaneously from many individual computers.	Multiple large financial institutions' websites became unresponsive and made it unable to do online banking.
Espionage	The act of obtaining confidential information without the information holder's consent.	Compromise of Mitsubishi, a Japanese defense contractors' research and development facilities.
Data Destruction	The use of malicious software to destroy data on a computer or to render a computer inoperable.	NotPetya: Malware encrypted the contents of computers without back up and ability to retrieve, rendering the computers inoperable and costing the world economy billions of dollars.
Sabotage	The use of malware that causes a disruption to a physical process, such as the provision of electricity or normal function of nuclear centrifuges.	Stuxnet: Compromise of industrial control systems that made nuclear centrifuges in Iran fail.

Table 1: Presents the different types of cyber attacks as characterized by the Council of Foreign Relations. In addition to their description, an example from the actual data is portrayed as well. The definitions and examples are taken from the Cyber Operations Tracker project (*Connect the Dots on State-Sponsored Cyber Incidents*, n.d.). The definitions are direct quotes, whereas the examples are paraphrased by me.

To further motivate differentiating types of cyber attacks, it is crucial to look at the degree and type of destruction they can cause. A DDOS attack on Netflix, for instance, would just lightly harm the country. On the other hand, NotPetya, an attack that renders

any computer unrepairable, launched against a hospital, clearly, causes much greater harm. The same comparison can be applied to "conventional bombs" and nuclear bombs where one causes much greater levels of destruction than the other.

Since all cyber weapons are just lines of code, it is easy to treat them all equally. Baliga, de Mesquita, and Wolitzky (Working Paper) formally model the attribution problem in cyberspace in their working paper. There, they use the number of cyber attacks, disregarding the type of attack, as a measure of a state's aggressiveness. Using the previous example, two states are not equally aggressive if they launch the same number of attacks, but one launches DDOS attacks against Netflix and the other NotPetya-type attacks. One causes more destruction than the other. Thus, while it is more intuitive to differentiate types of bombs, it is crucial to be just as vigilant, if not more, in cyberspace. Therefore, in this paper, I differentiate between espionage and non-espionage attacks.

2.2.2 The Cost of Cyber Weapons

The literature is split into two camps: cyber weapons bear low costs (Kello, 2013) or they bear high costs (Gartzke, 2013; Lindsay, 2013). I argue that they bear high costs for multiple reasons. As a summary, cyber weapons need to be backed by conventional military (Gartzke, 2013), can be used a limited number of times (Gartzke, 2013), have an expiration date, and strong offensive capabilities imply strong defensive capabilities (Lindsay, 2013). Thus, I expect mostly militarily strong states to launch cyber attacks as they have the resources to develop and maintain them.

Gartzke (2013) argues that cyber weapons are expensive because they are non-lethal and thus have to be backed up by conventional military means. In other words, only militarily strong actors can afford to launch cyber attacks, because they can resort back to lethal, conventional, means. He also points out that not only are cyber weapons non-lethal, they can also be used only a small number of times since once a vulnerability is exploited, it is likely fixed quickly upon discovery, rendering the weapon useless (Gartzke, 2013). The

assumption that cyber weapons are non-lethal is to date empirically correct, but that does not mean they are non-lethal for technical reasons. I go into more detail on this in the non-lethality section. Nevertheless, the idea that the state has to be able to back up cyber force with traditional force holds even if cyber attacks are lethal.

The opposing view that cyber weapons are cheap, or as Kello (2013) puts it, cyberspace has a "low barrier of entry", is an important alternative to consider. The argument here is that a weaker actor, be it a nation-state, terrorist organization, or lone wolf, can develop cyber capabilities on par with more powerful states, thereby being able to successfully attack them in cyberspace.[6] The classic example used is "cyber pearl harbor", which describes a devastating attack by a terrorist organization on U.S. critical infrastructure costing countless lives (Arquilla, 2012). From a technical point of view, this is a very real possibility and countries like the United States are preparing for it, for instance by creating agencies such as the Cybersecurity and Infrastructure Security Agency (CISA) (*Infrastructure Security*, 2019). However, as I explain in a later section, thus far, no fatalities whatsoever have resulted from cyber attacks. As figure 3 shows, very few attacks are launched by militarily weak state actors. Throwing doubt on the assumption that cyberspace has a "low barrier of entry" (Kello, 2013).

An additional factor driving the costliness is the observation that some cyber weapons have expiration dates.[7] That is, Stuxnet for instance utilized four so-called "zero-days", which are vulnerabilities that no one is aware of. So, once people became aware of those

[6]While this paper only focuses on state-sponsored cyber attacks, the idea that there is a lone wolf capable of executing a disastrous attack is important to address. The same logic applied to weaker states also applies to the lone wolf, in some aspects it is even more extreme. For instance, it is even more unlikely that a lone wolf has the resources to craft a cyber weapon that can be used against a powerful state, since states have very powerful defenses. If the lone wolf has these capabilities and resources, they have an immense skillset, which is very valuable in both the private and public sector, meaning that there is no monetary incentive for a lone wolf to execute such devastating attacks. This leaves other incentives, but these have to be so large that they make up for the very high costs of developing a weapon that has not yet been protected against by the state. This is not impossible, but the likelihood of an individual or a small group having the resources, skills and incentives to do this is very small.

[7]DDOS attacks for instance do not have an expiration date, as a DDOS attack just requires a large number of computers. So, if an actor has constant access to a large set of computers, they can launch a DDOS attack at any point.

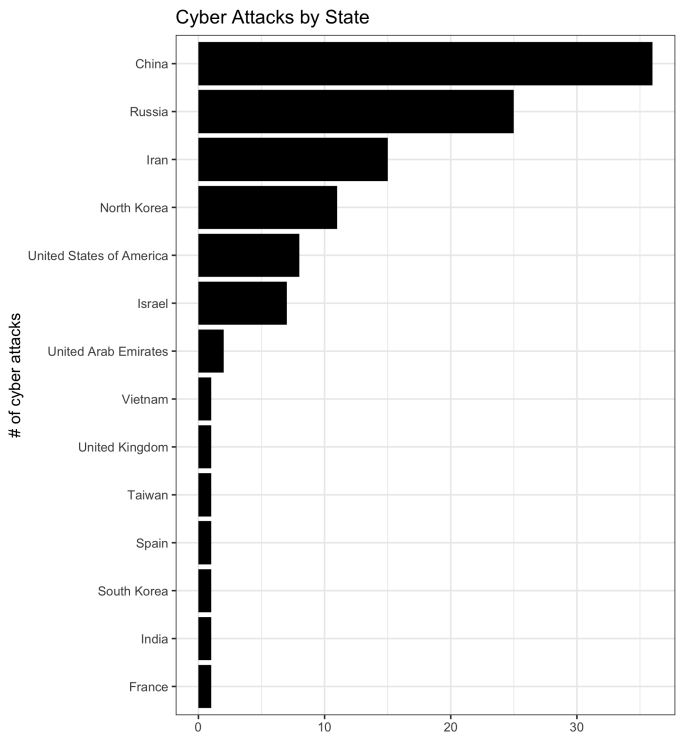


Figure 3: The distribution of unique attacks launched for each attacking country. Just like for figure 1, these attacks can affect multiple targets at once, but are counted only once.

vulnerabilities, they got fixed and Stuxnet was no longer a viable weapon. The reason this drives up the cost is that even if a state had the capability to develop a weapon, it cannot simply stash it, like it could with a nuclear bomb, but has to either decide to use it or accept the possibility that it will lose its functionality.

The final reason mentioned is the relationship between offensive and defensive capabilities. As Lindsay (2013) explains, if a state invests heavily into offensive capabilities, it also secures its systems from those weapons, making it more secure. That is, having better offensive cyber capabilities has a direct positive effect on having stronger defensive cyber capabilities. Considering that the top-5 attackers (illustrated in figure 3) China, Russia, United States, Iran, and Israel have good offensive cyber capabilities also means they have good defensive cyber capabilities, making it harder for weaker states to develop feasible cyber weapons against them.[8]

[8]While this dataset is not exhaustive and might be missing observations and attributions, it is hard to

So far, I have argued that cyber weapons are expensive to develop and maintain. Therefore, I expect that cyber weapons are used by militarily strong states, since they have the resources for them. Figure 3 gives some initial support for this claim, as the top attackers have some of the most powerful militaries in the world, but I also test the hypothesis statistically.

> **Hypothesis (1.1)**: *Due to the high cost needed to develop, deploy and maintain cyber weapons, state-sponsored cyber attacks are mostly launched by militarily strong states.*

2.2.3 Geographic Contiguity and Cyberspace

The result that geographical contiguity has a positive effect on the likelihood of international conflict onset is an important one in the conflict literature (Bremer, 1992; Diehl, 1985).[9] In cyberspace however, intuitively, contiguity ought not matter, as there are no borders in the digital world. This intuition is so strong that only Kello (2013) explicitly mentions it, but the remainder of the literature never addresses it. I argue that geographical contiguity is not associated with state-sponsored cyber attack onset because states have the ability to attack countries across the globe without having to deploy troops or other conventional means that incur a heavy political and/or monetary cost.

If a country attacks another one through cyberspace, not only is there a chance that they will *not* be caught (due to the attribution problem described in more detail in the next section), even if they are caught, the consequences are meager compared to what they would be if for instance troops would be deployed across the globe close to the state.

Figure 4 shows that if a cyber attack occurred and it was attributed, most often there is

---

believe that militarily inferior states would have the capabilities to build "better" cyber weapons that have not been detected by other states or successfully attributed to them. Thereby, supporting the position that cyber weapons require a large amount of military resources to develop.

[9]For this paper I define two states as contiguous if they are at most 400 miles apart by sea. That is, since I am using the Correlates of War Project's Direct Contiguity Data version 3.2 (Stinnett, Tir, Diehl, Schafer, & Gochman, 2002) all possible classifications of contiguity.

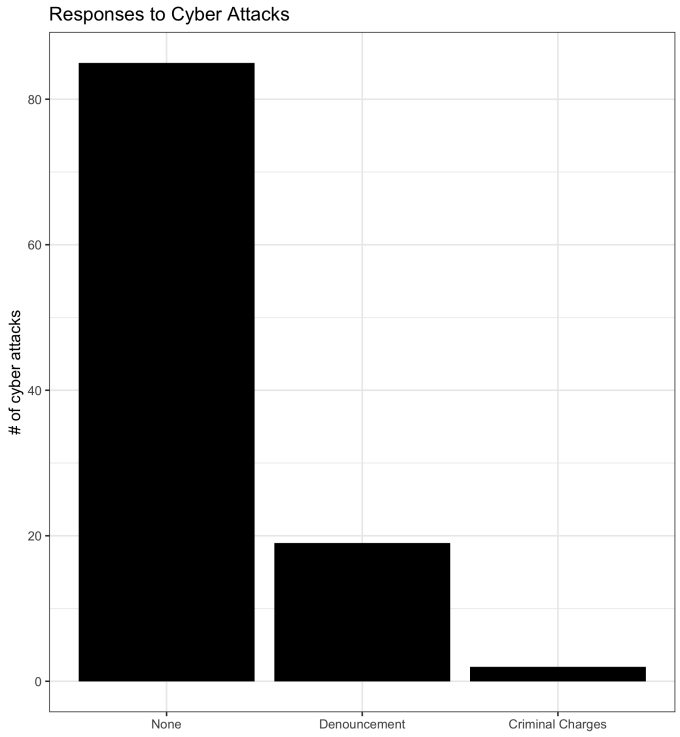


Figure 4: The number of public responses by the targets of cyber attacks against the aggressors. The data is the one used in the statistical analyses, which omits an observed case of sanctions issued. Nonetheless, the trend of no public response is strong.

no public response by the victim. Though secret counter-attacks are a possibility that is not captured in the data, the difficulty of attributing who the attacker is, might make them a less feasible response for nation states. This problem does not exist with conventional means, as there it is often clear who the aggressor is. Thus, making a cyber attack a politically cheaper act of aggression across the world than conventional means allow for.

> **Hypothesis (1.2)**: *Geographical contiguity is not associated with state-sponsored cyber attack onset.*

2.2.4 The Attribution Problem

The attribution problem is the name for the observation that it is hard to identify the perpetrator of a cyber attack. While cyber attribution is possible, it is a difficult task to accomplish (Rid & Buchanan, 2015). The reason it is so important in cyberspace is that

defensive measures are not able to block all cyber attacks, and a show of force is helpful to deter attackers (Wheeler & Larsen, 2003). The problem with a show of force is that it requires a target, which is where the attribution problem comes into play.

Wheeler et al. define attribution in the cyberspace, in their highly technical paper, as "determining the identity or location of an attacker or an attacker's intermediary." (Wheeler & Larsen, 2003) Their paper lists 17 technical techniques of attributing an attack[10], some of which require ongoing surveillance of likely attackers and others constant monitoring of in- and outgoing network traffic to identify anomalies (Wheeler & Larsen, 2003). Both of these require immense resources to execute successfully, leading to Rid and Buchanan's (2015) conclusion that since attribution is so complex, only already militarily strong states have the capability to do it (Wheeler & Larsen, 2003; Rid & Buchanan, 2015). Even if state's have the capabilities, attribution is not an easy task and not guaranteed (Wheeler & Larsen, 2003).

2.2.5 Cyber Non-Lethality

Throughout the short history of cyber conflicts, a striking observation has been made, which holds to this day; cyber attacks have never resulted in any fatalities (Gartzke, 2013; Junio, 2013; Stone, 2013; Kello, 2013; Rid, 2012). This surprising fact sparked debates early on whether or not cyber war can even exist. Rid (2012) and Gartzke (2013) suggest it cannot exist since cyber attacks are non-lethal and even if they are lethal only short-term damage can be induced, hence no ongoing war is feasible. On the other hand, Kello (2013) and Stone (2013) argue that the definitions of war and peace ought to be adapted to account for the non-lethality present in the cyberspace. While these discussions are important, as the word "war" has a strong connotation and just the terminology might affect policies, this paper does not contribute to this debate. Instead, it focuses on the reason why these attacks are non-lethal in the first place.

[10]I will not go into the technical details, but for the interested reader, the article goes into a lot of detail. It requires a good background knowledge in computer networks and the associated protocols used, such as TCP/IP.

There are no technological reasons why cyber weapons are non-lethal, in fact they can be very powerful, as it is possible that a cyber weapon can break into the nuclear attack system of another country and launch its nuclear missiles. Thus, the non-lethality is a result of strategic choices on the side of the attackers. Gartzke (2013) argues that states have no reason to use lethal force without backing it up with traditional force, because the damage would only be short-term.

Another possibility is that it might be cheaper to drop bombs than launch a sophisticated cyber weapon to cause lethal harm. Since cyber weapons exploit physical elements such as blowing up a power plant or cutting off electricity from a hospital to cause harm, whereas conventional means such as airstrikes are a much more direct, and thus easier to deploy, way of harming.

2.3 Asymmetric Cyber Warfare

One of the main questions about cyber war is whether or not it can be asymmetric. The literature is split. On one side, Kello (2013) argues that cyberspace has a low barrier of entry and the cyber pearl harbor scenario (Arquilla, 2012) shows a general fear of a weaker actor causing massive damage. On the other side, Gartzke (2013), Lindsay (2013), and Rid and Buchanan (2015) argue the opposite, claiming that the complexity of a cyber weapon and the amount of resources spent on defense by the powerful state-actors diminishes the possibility of asymmetric war. I argue that asymmetric war in cyberspace is unlikely.

In order to discuss this, I need to do so by considering non-espionage and espionage cyber attacks separately. The current theories on asymmetric cyber warfare, do not address this distinction. In fact, almost all papers only consider non-espionage attacks. I consider cyber espionage separately for two reasons. First, as figure 2 shows, the majority of cyber attacks are espionage. Second, one of the core forms of asymmetry is information asymmetry.

2.3.1 Non-Espionage Cyber Attacks

The cyber pearl harbor scenario includes the important detail that a terror organization, more generally a weaker group or state, would launch such an attack (Arquilla, 2012). That is, since cyber attacks are cheap and can be destructive they allow for asymmetric warfare (Kello, 2013). But as I argued in a previous section, they are not cheap to develop. More importantly, even if they were, the fact that strong offensive capabilities imply strong defensive capabilities, means that major powers not only have strong offensive, but also defensive capabilities (Lindsay, 2013). Thus, either way, it is a difficult task for a weaker state to successfully attack a major power in cyberspace.

On the other hand, a possible argument that makes cyber asymmetric warfare an attractive theory is the attribution problem. As Rid and Buchanan (2015) point out, attribution is possible, but very difficult to do. Hence, a weaker state might look at the attribution problem as lowering the cost of detection and therefore punishment. In the same paper, Rid and Buchanan also point out that due to the complexity of attributing, mostly militarily strong states have these capabilities (Rid & Buchanan, 2015). So, the attribution problem cuts both ways. In one direction, it makes an attack more likely since the weaker attacker can get away with it. On the other hand, it is difficult for a weaker state to estimate the capabilities of the more powerful state to attribute an attack.

So, weaker states face strong defensive cyber capabilities, high costs of developing cyber weapons, and the fear of attribution and hence retribution, by stronger states. Therefore, I expect states with weaker military capabilities tend not to launch non-espionage cyber attacks against states with greater military capabilities.

> **Hypothesis (2.1)**: *Militarily weaker states tend not to launch state-sponsored non-espionage cyber attacks against militarily superior states.*

2.3.2 Espionage Cyber Attacks

For espionage attacks, militarily weaker states have the incentive to attack stronger states to balance out the military capability (information) asymmetry. While Gartzke (2013) believes that cyber attacks might destroy *all* information asymmetries among countries, I argue that the relative military capabilities of the countries involved, affects the likelihood of state-sponsored cyber-attack onset.

In particular, if the two states are relatively equal in power, cyber espionage is more likely because the risk of attribution is lower and the chance of success greater. Whereas if the victim is much more powerful than the attacker, cyber espionage is less likely to occur, because the fear of attribution and thus retribution of the weaker state is greater and the chance of success lower.

> **Hypothesis (2.2)**: *When two states are relatively equal in military capabilities, cyber espionage is more likely to occur.*
>
> **Hypothesis (2.3)**: *When the victim has much greater military capabilities than the attacker, cyber espionage is less likely to occur.*

2.3.3 Cyber and Conventional Means

Gartzke's (2013) argument that cyber attacks need to be backed up by traditional forces is interesting as it not only gives a possible explanation for the high cost of cyber weapons, but also shows a strategic decision made by the attackers. That is, using cyber weapons alongside traditional weapons. All the papers concerning cyber war (Rid, 2012; Stone, 2013; Kello, 2013) consider "pure cyber war", i.e. a war that only happens in cyberspace. But in today's world no one considers "pure naval war", but rather war in general, so why would cyberspace be any different? In other words, the hypothesis boils down to whether or not cyber weapons are complements or substitutes to conventional military means.

Throughout history new technology did not necessarily entirely replace a domain of warfare, i.e. air did not replace ground or naval, it is rather unlikely that cyberspace will replace

any either. The observed lack of fatalities further reduces the possibility of a pure cyber war.

A good example to consider are the state-sponsored cyber attacks by Russia against Georgia. "The attacks, which were carried out by nonstate agents including Russian criminal syndicates, occurred against the backdrop of Russia's ground incursion into Georgia in the summer of 2008." (Kello, 2013) The cyber attacks interfered with the communication infrastructures of the Georgian government as part of a larger military strategy. While this would have been possible with traditional military means, cyber attacks made it easier (Kello, 2013).

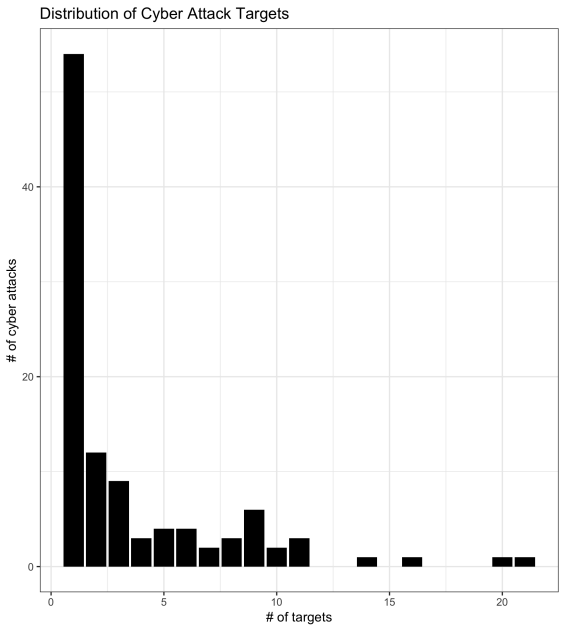
> **Hypothesis (2.4)**: *Cyber weapons are complements not substitutes for conventional military means.*

3 Data and Methods

One of the main contributions of this paper is the inclusion of data and statistical analyses in the study of international state-sponsored cyber conflict, which is the first time this has been done (to my knowledge). Secrecy is the status quo in cyber space so getting reliable data is naturally very difficult. I rely on the "Cyber Operations Tracker" dataset created by the *Council of Foreign Relations* (*Connect the Dots on State-Sponsored Cyber Incidents*, n.d.).[11] The dependent variables of interest are the onset of state-sponsored cyber attacks and the onset of militarized interstate disputes (MIDs). The data includes observations from 2005-2017 and the unit of analysis is a directed-dyad-year.

As figure 5a shows, it is common for cyber attacks to affect multiple victims at once. This is not the case in traditional international conflict. To account for that, I split up each of such incidences into separate observations, illustrated in figure 5b. In other words, while there are only 122 distinct cyber attacks launched in the 12 years, there are 484 observed

[11] In addition to the state-sponsored cyber conflict data, the standard dataset for controls in the literature, such as Correlates of War (Gibler, 2009; Maoz, Johnson, Kaplan, Ogunkoya, & Shreve, 2019; Stinnett et al., 2002) and Polity IV (Marshall & Gurr, 2014) as well as economic and military spending data from the World Bank (*GDP per capita (current US$)*, n.d.; *Military expenditure (% of GDP)*, n.d.) are used as well.



(a) Distribution of # of victims

Attack ID	Sponsor	Victims	...
1	China	United States, United Kingdom	...

Attack ID	Sponsor	Victims	...
1	China	United States	...
1	China	United Kingdom	...

(b) Data transformation

Figure 5: Figure 5a presents the distribution of the number of distinct victims per cyber attack and 5b illustrates how I transformed the data to account for this in the dyadic dataset.

cyber attacks included in the data.[12]

Since only 14 countries are listed as having launched a cyber attack, but 93 have been a victim of one, using a full directed dyad dataset is not feasible, since there would be far more directed dyads where a cyber attack never occurred. At the same time, only using the directed dyads where at least one cyber attack occurred induces selection problems. To counteract that, I randomly sampled 15 countries that have never launched a cyber attack and included them in the directed dyads as attackers, but not as victims.

In addition, I also include 15 hand-picked states[13] that have not launched a cyber attack and are not included in the random sample. These states are considered to be militarily strong and could potentially have cyber capabilities. I do not include these in the main models, but only for robustness checks.

[12]More cyber attacks are included in the original dataset, but due to time constraints and conservative data cleaning methods only a fraction is included in this iteration of the paper.

[13]The states are: Japan, Turkey, Germany, Italy, Egypt, Brazil, Indonesia, Australia, Poland, Algeria, Canada, Saudi Arabia, Thailand, Greece, Ukraine

3.1 Dependent Variables

Cyber attack onset. This is the main dummy variable, which is an indicator whether or not a cyber-attack within the directed dyad at a given year occurred. While this paper differentiates between types of cyber attack, the dependent variable does not account for this, i.e. it is 1 if any type of cyber attack occurred.

Traditional attack onset. This is the other dependent variable, which is also a dummy and is simply an indicator for whether or not a militarized interstate dispute (MID) occurred within the directed dyad in a given year and is taken from the Correlates of War project.[14] Just like for the cyber attack, the type of MID is not accounted for, i.e. it is 1 if any sort of MID occurred.

3.2 Explanatory Variables

I follow the international conflict literature standard by including controls for economic properties, regime type, contiguity, military capabilities and time. To account for time, cubic year splines are included in all models (Carter & Signorino, 2010; Beck, Katz, & Tucker, 1998) and all of the following control variables are lagged by one year to counteract simultaneity concerns.

I use logged GDP per capita data from the World Bank (*GDP per capita (current US$)*, n.d.). Dummy variables accounting for any alliance[15] and contiguity[16] are taken from the Correlates of War project and imputed forward[17]. To control for regime type, this paper uses the Polity IV data set and includes two control variables. First, Abs. Diff. Polity, which is the absolute difference between the polity scores of the states in the dyad. Second, Joint

[14]Uses version 3.1 of the Dyadic Militarized Interstate Disputes dataset (Maoz et al., 2019).

[15]Uses version 4.1 of the Formal Alliances dataset (Gibler, 2009).

[16]Correlates of War Project. Direct Contiguity Data, 1816-2016. Version 3.2 (Stinnett et al., 2002).

[17]E.g. the alliance data ends in 2012, so this paper imputed the years 2013-2017 using the latest possible prior alliance in place, effectively assuming no changes occurred in alliance and contiguity.

Dem., which is a dummy that is 1 if both states are democracies[18] and 0 otherwise (Marshall & Gurr, 2014).

Finally, this paper decided against using the CINC measure from the Correlates of War project as a proxy for military capabilities, because the CINC measure is in large part calculated from for example the number of military personnel and steel production (Singer, Bremer, & Stuckey, 1972). This is of little importance when it comes to the production cyber weapons. The true resources required are highly skilled labor, computational resources as well as high-value intelligence. These, I argue, can be inferred from the combination of the GDP per capita and the proportion of the GDP spent on the military, data for which is taken from the World Bank (*Military expenditure (% of GDP)*, n.d.).

The final explanatory variable of interest is military capability asymmetry. For this measure, I implement the approach used by Gartzke and Westerwinter (2016) on the logged proportion of GDP spent on military instead of CINC. That is, if *milcap* is that measure, the power asymmetry is given by

$$asymmetry_{i,j,t} = \frac{\max\{milcap_i, milcap_j\}}{milcap_i + milcap_j},$$

for all attackers i, victims j, and years t. Thus, as this measure tends towards one, the asymmetry is greater and as it tends towards 0.5 (note it cannot be less than 0.5) the balance is greater (Gartzke & Westerwinter, 2016). The reason for the max statement in the numerator is because I also include a dummy variable in the analyses which identifies whether or not the victim has greater military capabilities than the attacker. Hence, enabling me to make general statements about military capability asymmetry as well as the direction of the asymmetry.

[18] A democracy is defined as having a polity score greater or equal to 7.

3.3 Methodology

The dependent variables are binary and thus I use a logistic regression. I include all the possible dyads between 29 attackers and 93 victims over the 12-year period 2005-2017 and my unit of analysis is directed dyad-years. I have 39,468 observations total, out of which only 377 include a cyber attack and 122 traditional attacks. That is, only 0.96% of all observations include cyber attacks and 0.31% traditional attacks. Since I have about 40,000 observations total but less than 1% of 1's in the dependent variable, I use the rare-events logit with prior correction introduced by King and Zeng (2001) to make sure that my results are not biased. I also include the same models with the standard logistic regression in the appendix.

Since this is a dyadic data set, I use cluster-robust standard errors by undirected dyads in all models. I also include cubic year splines (Carter & Signorino, 2010; Beck et al., 1998), which are the number of years since 2005 in all models.

3.4 Note on Data Accuracy

Two big concerns for using publicly available data on cyber conflicts are the possible inaccuracy due to attribution issues and the selection problem of weaker states. First, while it is impossible to say with certainty that the attributions are correct, the Council of Foreign Relations gathered multiple non-governmental sources to support such an assessment.[19] As the section on the attribution problem explains, attribution is difficult in cyberspace. Countries are trying to exploit this and fake the origins, and can thereby induce measurement error into the data set.

Second, the selection problem that only some attacks, in particular the less sophisticated ones, are recorded in the data set is also a reasonable concern. To alleviate this, consider figure 3. There are for sure many attacks missing, but there is little that can be done about missing data, but the data that is available includes the most powerful states in the world, in fact, the majority of all attacks are launched by them. Thus, the selection seems not to

[19] See (*Connect the Dots on State-Sponsored Cyber Incidents*, n.d.) for more details on their methodology.

be as bad of a problem.

All in all, while the concerns regarding the measurement are valid, but the arguments hopefully helped to counteract them, the value of using empirical evidence in studying international cyber conflict is invaluable as helps to separate the better theories from the worse and allows to study questions more rigorously.

4 Results

Table 2 is the first of two regression tables presented in this paper. The dependent variable is the occurrence of a cyber attack of state A against state B. The unit of analysis is directed dyad-year. As mentioned in the data section, the analysis does not include all countries in the world, since only a fraction actually launched an attack. Instead, all the countries that launched an attack are included as well as a random sample of size 15 from all countries included in the Correlates of War data set that have not launched a cyber-attack. The control variables included are the standard in the literature and control for economic, political, strategic (alliances) factors, and military strength.

4.1 Cost

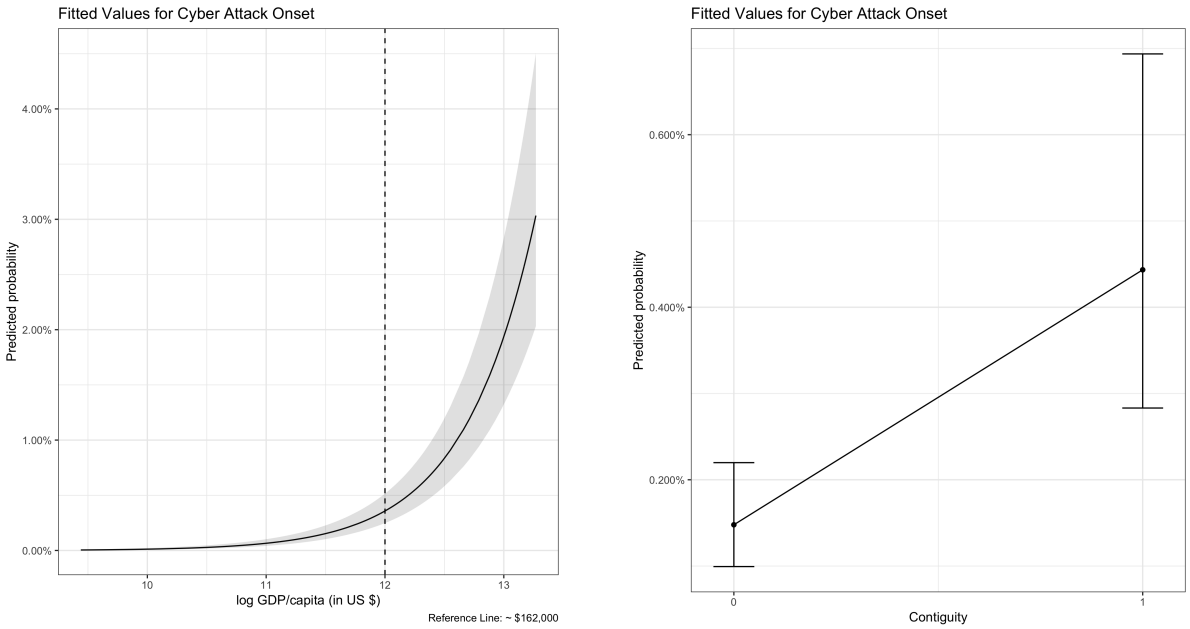
The first observations are the statistically positive effects[20] of the logged GDP and logged military capabilities of the attacker on the likelihood of a cyber attack throughout the different model specifications in table 2. In other words, as the economic strength of a country and the proportion of which is spent on its military increases, so does the probability of them launching a cyber attack. This supports hypothesis 1.1, stating that cyber weapons are used by mostly militarily already strong actors, since they are expensive.

The marginal effects shown in figure 6a are also very telling of the cost behind developing, deploying and maintaining cyber weapons. The exponential increase as the logged GDP of

[20]I am using the word effect in its colloquial sense and am not referring to causal effects throughout this paper.

	Dependent variable:				
	Cyber Attack Onset				
	(1)	(2)	(3)	(4)	(5)
Alliance	1.093***	0.072	−0.022	0.569**	0.573**
	(0.275)	(0.213)	(0.222)	(0.224)	(0.225)
Contiguity	1.643***	1.405***	1.413***	1.105***	1.101***
	(0.343)	(0.239)	(0.241)	(0.231)	(0.229)
log GDP/capita A		1.628***	1.529***	1.705***	1.705***
		(0.130)	(0.147)	(0.139)	(0.140)
log GDP/capita B		1.053***	1.086***	1.231***	1.231***
		(0.151)	(0.154)	(0.116)	(0.116)
log Mil. Capabilities A			2.922***	1.853***	1.792***
			(0.460)	(0.477)	(0.579)
B Mil. Stronger			0.544**	0.200	0.522
			(0.224)	(0.237)	(1.249)
Mil. Cap. Asymmetry			−2.441***	−1.282	−1.109
			(0.925)	(0.814)	(1.084)
Abs. Polity Diff.				0.007	0.007
				(0.017)	(0.018)
Joint Dem.				−2.256***	−2.269***
				(0.268)	(0.282)
B Mil. Stronger × Mil. Cap. Asymmetry					−0.526
					(2.074)
Time Controls	YES	YES	YES	YES	YES
Date Range	2005-2017	2005-2017	2005-2017	2005-2017	2005-2017
Observations	39,468	32,163	24,995	21,847	21,847
Log Likelihood	−1,844.679	−1,319.199	−1,164.103	−1,033.060	−1,032.999
Akaike Inf. Crit.	3,701.358	2,654.398	2,350.205	2,092.121	2,093.998
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 2: The models are rare-events logistic models with prior correction (King & Zeng, 2001; Imai et al., 2008; Choirat et al., 2018) with cluster-robust standard errors at the dyad level. The time controls are cubic polynomials (Carter & Signorino, 2010; Beck et al., 1998) of the years since 2005 are included in all models and the date ranges from 2005-2017.



(a) log GDP/capita A (b) Contiguity

Figure 6: The predicted probabilities for state-sponsored cyber attack onset as logged GDP/capita for the attacking state (6a) and geographic contiguity (6b) change, where the remaining controls are held at their means. The marginal effects are based on model 5 from table 2.

the attacker increases totaling over 3% and really starting to increase at a GDP per capita of about $162,000 further indicates that only the most powerful states really use the cyber weapons.

4.2 Contiguity

The regression results suggest a consistently significant and positive effect of contiguity. In other words, even though cyber weapons are global by definition, they are still more likely used in local conflicts. This is a counter-intuitive result at first and goes against hypothesis 1.2, but not against the traditional conflict literature (Bremer, 1992; Diehl, 1985). There, the positive effect of contiguity on conflict onset is well recorded. Hence, suggesting that leaders do not necessarily launch cyber attacks across the globe because they can, but instead use them strategically in conflicts they are already involved in or have interests in starting, which are more likely to be geographically close (Bremer, 1992; Diehl, 1985).

Figure 6b sheds some more light on the situation as well, showing that while the effect is in fact positive, it is rather small (less than 1%). Furthermore, the robustness check also reports significant positive effects of contiguity, making the result stronger.

4.3 Asymmetric Warfare

Even though table 2 does not separate between espionage and non-espionage attacks and hence does not directly address hypothesis 2.1-2.3, which will be addressed in table 3, it hints at an interesting null effect of military capability asymmetries on state-sponsored cyber attack onset. Model 4 shows a non-significant positive effect of the victim being militarily stronger as well as a non-significant negative effect as the military capability asymmetry grows. To identify whether B Mil. Stronger and Mil. Cap. Asymmetry are correlated and their statistical insignificance is a result of multicollinearity in the model, I run a likelihood-ratio test with the null hypothesis that the coefficients for B Mil. Stronger and Mil. Cap. Asymmetry are jointly zero. I can reject the null hypothesis at 95% confidence (with p-value $p = 0.04438$) that the coefficients are jointly zero, meaning the statistical insignificance is not likely due to multicollinearity in the model.

An increasing value of the military capability ratio variable implies that the military capability ratios among the two states are further apart, i.e. one state is more powerful than the other. The non-significance is unfortunate, but the effect directions fit the hypothesis indicating that a weaker state is more likely to attack a stronger state as long as their military capabilities are relatively close.

The robustness check suggests the results regarding the victim being militarily stronger are weak, because not only do they not gain significance, they also flip signs in table 6's model 4. The asymmetry measure on the other hand, keeps the sign and even gains significance in the robustness checks (see models 4 and 5 in table 6).

	Dependent variable:				
	Cyber Attack Onset				Trad. Attack Onset
	Non-Espionage	Non-Espionage	Espionage		
	(1)	(2)	(3)	(4)	(5)
Alliance	0.188	0.154	0.617**	0.201	0.929
	(0.495)	(0.489)	(0.250)	(0.618)	(0.935)
Contiguity	0.537	0.526	1.238***	0.970*	2.788***
	(0.588)	(0.584)	(0.240)	(0.563)	(0.979)
log GDP/capita A	1.735***	1.745***	1.764***	1.819***	0.868**
	(0.324)	(0.332)	(0.151)	(0.261)	(0.388)
log GDP/capita B	1.193***	1.221***	1.213***	1.538***	0.858**
	(0.214)	(0.214)	(0.127)	(0.365)	(0.430)
log Mil. Capabilities A	14.757***	15.040***	0.585	3.318**	6.600***
	(3.306)	(3.230)	(0.556)	(1.352)	(1.916)
B Mil. Stronger	−0.909	−16.929***	1.728	0.231	1.112**
	(1.104)	(2.168)	(1.315)	(0.487)	(0.437)
Mil. Cap. Asymmetry	−4.989***	−5.562***	−0.668	−0.354	−3.942
	(1.694)	(1.594)	(1.256)	(1.936)	(3.026)
Abs. Polity Diff.	−0.057	−0.060	0.015	−0.021	0.029
	(0.039)	(0.040)	(0.018)	(0.035)	(0.044)
Joint Dem.	−1.356*	−1.332*	−2.379***	−2.876***	301,668.800***
	(0.732)	(0.730)	(0.319)	(0.848)	(0.983)
B Mil. Stronger × Mil. Cap. Asymmetry		26.535***	−2.428		
		(3.405)	(2.195)		
Trad. Attack				0.745	
				(0.573)	
Cyber Attack					0.682
					(0.611)
Time Controls	YES	YES	YES	YES	YES
Date Range	2005-2017	2005-2017	2005-2017	2005-2010	2005-2010
Observations	21,581	21,581	21,809	9,866	9,866
Log Likelihood	−161.588	−158.650	−892.785	−193.803	−264.235
Akaike Inf. Crit.	349.177	345.300	1,813.569	415.606	556.469
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 3: The models are rare-events logistic models with prior correction (King & Zeng, 2001; Imai et al., 2008; Choirat et al., 2018) with cluster-robust standard errors at the dyad level. The time controls are cubic polynomials (Carter & Signorino, 2010; Beck et al., 1998) of the years since 2005 are included in all models. Models 1 and 2 have a perfect prediction problem and model 5's joint democracy coefficient is too large. Table 10 in the appendix uses Firth's (1993) bias reduction approach as implemented in R (Kosmidis, 2019) and recovers the same coefficients across models with the exception of alliance flipping signs in models 1 and 2, joint democracy flipping sign in model 5 and B Mil. Stronger flipping signs in model 4.

4.3.1 Non-Espionage Cyber Attacks

Table 3 includes the full model split by non-espionage and espionage cyber attacks as well as the interaction between traditional military provocations and cyber attacks. I am expecting a negative effect of the victim being stronger, meaning that weaker states tend not to launch state-sponsored non-espionage cyber attacks against militarily stronger states. The results in model 1 and the robustness checks include this effect, but are not statistically significant.

The significant negative effect of military capability asymmetries though, indicates that as one of the states is much more powerful than the other, the likelihood of a state-sponsored cyber attack decreases. This coefficient retains significance in the robustness checks (both with the larger sample and the standard sample but using Firth's (1993) bias reduction). Leading me to believe that the effect is robust. The interpretation of this negative effect is that as one of the states becomes military much stronger than the other, cyber attacks are less likely to occur. Since the coefficient for B Mil. Stronger is not significant, hypothesis 2.1 cannot be supported, but the result is suggestive nonetheless.

In model 2, the negative effect of the victim being stronger is now significant, as is the positive interaction term, which can be interpreted as how much weaker the attacker is, gives a very counter-intuitive result. As the attacker becomes weaker, he is more likely to launch a non-espionage attack. Examining figure 7 yields that the positive effect, while significant is actually negligible. This is not enough to support hypothesis 2.1, but it further adds some weight to it.

4.3.2 Espionage Cyber Attacks

Model 3 only looks at espionage cyber attacks. Here, I am expecting a negative effect for the asymmetry measure for both hypothesis 2.2 and 2.3 and a negative effect on the interaction term for hypothesis 2.3. The regression results match my expectation, but lack the statistical significance. To check whether the statistical insignificance comes from multicollinearity of the two variables I conduct a likelihood-ratio test with the null hypothesis that B Mil.

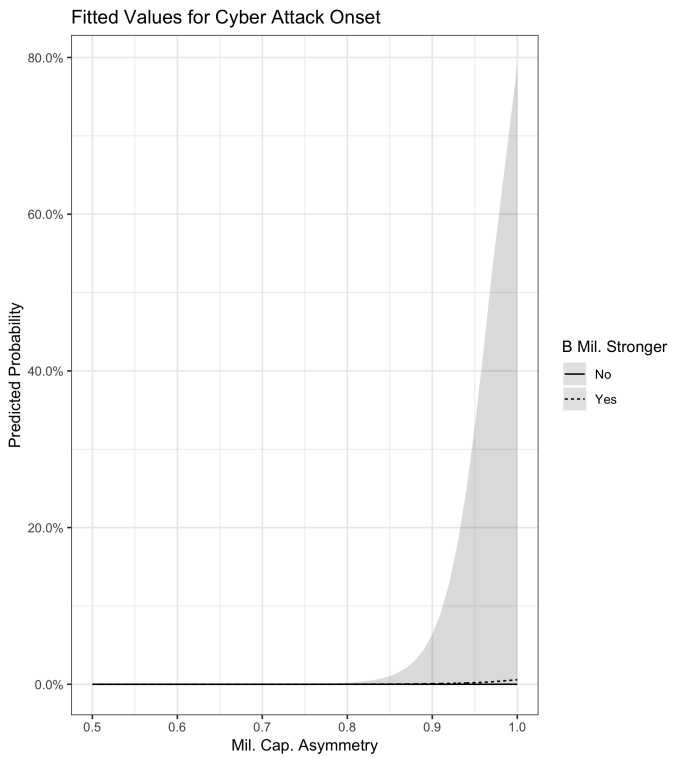


Figure 7: The predicted probability on state-sponsored cyber attack onset as the military capability asymmetry changes for the cases where the victim is weaker and stronger. The marginal effects are taken from model 2 in table 3.

Stronger and Mil. Cap. Asymmetries are jointly zero. I can reject the null hypothesis with 95% confidence (with p-value $p = 0.01836$), meaning the statistical insignificance is not likely due to multicollinearity in the model.

The robustness checks where I include the selected countries into the set of attackers and where I use Firth's (1993) approach both yield the same results (see tables 8 and 10), but the asymmetry measure is now statistically significant. Nonetheless, similar to non-espionage attacks, while the effect is suggestive, it is not statistically significant and robust enough to support hypothesis 2.2 and 2.3.

4.4 Cyber vs. Traditional

Considering the interaction between traditional and cyber military means, it is important to note that the time period is very short, especially for international conflict, since it only

covers 2005-2010. This might also partially explain the exorbitantly large effect for the joint democracies variable in the last model of table 3. To show support for hypothesis 2.4, the coefficients for both cyber (in model 4) and traditional attack (in model 5) ought to be positive and significant. Although they are both positive, neither is statistically significant. Both robustness checks yield the same results, with the exception that the cyber attack variable in model 5 is now significant. Thus, hypothesis 2.4 cannot be supported.

It is possible that the lack of significance comes from the very short observation period, as there have been few MIDs and cyber attacks. If this is not the case, then this result does not imply that cyber weapons are substitutes for traditional military means, but rather are simply used independently for different tasks. Considering that the majority of cyber attacks are espionage, this result would make sense. It would be rare to observe an attributed cyber espionage attack alongside a traditional military operation.

5 Conclusion

The goal of this paper is not only to make progress on understanding how cyber weapons are used in international relations, but also to provide the data and baseline results to further study this phenomenon. This being the first paper (to my knowledge) to have empirical evidence, I focus on establishing some basic results that settle ongoing debates in the literature and introduce some trends that have not been addressed yet.

I find that cyber weapons are expensive and suggest that the resource-intensive process, the perishability (Gartzke, 2013) and the non-lethality is what drives the cost up. It is resource-intensive because powerful states keep improving their offensive capabilities and thereby also improving their defensive ones (Lindsay, 2013). This creates a high barrier of entry into cyberspace, meaning that asymmetric cyber warfare is less likely.

The non-lethality of cyber weapons being a strategic choice rather than a limitation of technology is an important aspect to be studied further. It is also important to understand

whether cyber attacks lead into a downward spiral, i.e. getting more and more violent or if they are stagnant. That is, do states simply spy on each other in cyberspace, but rarely go further than that, and if they do, is it ever bad enough to cause full-out war?

Another important aspect to consider when studying international cyber conflict is the vast diversity of types of attacks possible. I consider the relationship between traditional military means and cyber weapons from a complement or substitutes point of view, but another way of looking at it is that cyber weapons can extend the range of aggressiveness open for a leader to show. That is, while traditional military means only allow for a very strong show of force, for instance by positioning troops on a border. The possibility of defacing websites on one end and the ability to delay the nuclear program on the other, allow the leader of a state a more fine-grained dialing in of the aggressiveness for a show of force. To study this, one might classify the observations in the data based on "perceived aggressiveness" and test the relationship that way.

To sum up, this paper is hopefully just the beginning of studying the role of cyber weapons in international conflict. It tries to dissuade the reader of being afraid of a cyber pearl harbor event, but instead think of cyber weapons as a more flexible medium for a leader to show strength and gain an information advantage if the state already has the military and economic prowess to develop new weapons and keep up with the other powerful states.

References

Arquilla, J. (2012, Nov). Panetta's wrong about a cyber 'pearl harbor'. *Foreign Policy*. Retrieved 04/15/2019, from https://foreignpolicy.com/2012/11/20/panettas-wrong-about-a-cyber-pearl-harbor/

Baliga, S., de Mesquita, E. B., & Wolitzky, A. (Working Paper). Deterrence with imperfect attribution. *Unpublished*. Retrieved from http://economics.mit.edu/files/14938

Beck, N., Katz, J. N., & Tucker, R. (1998). Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science*, *42*(4), 1260-1288.

Betz, D. (2012). Cyberpower in strategic affairs: Neither unthinkable nor blessed. *Journal of Strategic Studies*, *35*(5), 689-711. doi: 10.1080/01402390.2012.706970

Betz, D. J., & Stevens, T. (2013). Analogical reasoning and cyber security. *Security Dialogue*, *44*(2), 147-164. doi: 10.1177/0967010613478323

Bishop, M. (2012). *Computer security: art and science*. Addison-Wesley.

Bremer, S. A. (1992). Dangerous dyads: Conditions affecting the likelihood of interstate war, 1816-1965. *Journal of Conflict Resolution*, *36*(2), 309-341. doi: 10.1177/0022002792036002005

Carter, D. B., & Signorino, C. S. (2010). Back to the future: Modeling time dependence in binary data. *Political Analysis*, *18*(3), 271–292. doi: 10.1093/pan/mpq013

Choirat, C., Honaker, J., Imai, K., King, G., & Lau, O. (2018). Zelig: Everyone's statistical software [Computer software manual]. Retrieved from http://zeligproject.org/ (Version 5.1.6)

Connect the dots on state-sponsored cyber incidents. (n.d.). Council on Foreign Relations. Retrieved 04/15/2019, from https://www.cfr.org/interactive/cyber-operations

Diehl, P. F. (1985). Contiguity and military escalation in major power rivalries, 1816-1980. *The Journal of Politics*, *47*(4), 1203-1211. doi: 10.2307/2130814

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, *80*, 27-38. doi: 10.1093/biomet/80.1.27

Gartzke, E. (2013). The myth of cyberwar: Bringing war in cyberspace back down to earth. *International Security*, *38*(2), 41-73. doi: 10.1162/ISEC_a_00136

Gartzke, E., & Westerwinter, O. (2016). The complex structure of commercial peace contrasting trade interdependence, asymmetry, and multipolarity. *Journal of Peace Research*, *53*(3), 325-343. doi: 10.1177/0022343316637895

Gdp per capita (current us$). (n.d.). Retrieved 04/15/2019, from `https://data.worldbank.org/indicator/NY.GDP.PCAP.CD`

Gibler, D. M. (2009). *International military alliances, 1648-2008.* CQ Press.

Hlavac, M. (2018). stargazer: Well-formatted regression and summary statistics tables [Computer software manual]. Bratislava, Slovakia. Retrieved from `https://CRAN.R-project.org/package=stargazer` (R package version 5.2.2)

Imai, K., King, G., & Lau, O. (2008). Toward a common framework for statistical analysis and development. *Journal of Computational Graphics and Statistics*, *17*, 1–22.

Infrastructure security. (2019, Mar). Retrieved 04/15/2019, from `https://www.dhs.gov/topic/critical-infrastructure-security`

Junio, T. J. (2013). How probable is cyber war? bringing ir theory back in to the cyber conflict debate. *Journal of Strategic Studies*, *36*(1), 125-133. doi: 10.1080/01402390.2012.739561

Karnouskos, S. (2011). Stuxnet worm impact on industrial cyber-physical system security. *IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society*, 4490-4494.

Kello, L. (2013). The meaning of the cyber revolution: Perils to theory and statecraft. *International Security*, *38*(2), 7-40. doi: 10.1162/ISEC_a_00138

King, G., & Zeng, L. (2001, Spring). Logistic regression in rare events data. *Political Analysis*, *9*, 137–163.

Kosmidis, I. (2019). brglm: Bias reduction in binary-response generalized linear models [Computer software manual]. Retrieved from https://cran.r-project.org/package=brglm (R package version 0.6.2)

Lindsay, J. R. (2013). Stuxnet and the limits of cyber warfare. *Security Studies*, *22*, 365-404. doi: 10.1080/09636412.2013.816122

Maoz, Z., Johnson, P. L., Kaplan, J., Ogunkoya, F., & Shreve, A. P. (2019). The dyadic militarized interstate disputes (mids) dataset version 3.0: Logic, characteristics, and comparisons to alternative datasets. *Journal of Conflict Resolution*, *63*(3), 811-835. doi: 10.1177/0022002718784158

Marshall, M. G., & Gurr, T. R. (2014). *Polity iv project: Political regime characteristics and transitions, 1800- 2013.* Retrieved 04/15/2019, from http://www.systemicpeace.org/polity/polity4.htm

Military expenditure (% of gdp). (n.d.). Retrieved 04/15/2019, from https://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS

Rid, T. (2012). Cyber war will not take place. *Journal of Strategic Studies*, *35*(1), 5-32. doi: 10.1080/01402390.2011.608939

Rid, T., & Buchanan, B. (2015). Attributing cyber attacks. *Journal of Strategic Studies*, *38*(1-2), 4-37. doi: 10.1080/01402390.2014.977382

Singer, D. J., Bremer, S., & Stuckey, J. (1972). Capability distribution, uncertainty, and major power war, 1820-1965. *Peace, war, and numbers*, *19*, 48.

Stinnett, D. M., Tir, J., Diehl, P. F., Schafer, P., & Gochman, C. (2002). The correlates of war (cow) project direct contiguity data, version 3.0. *Conflict Management and Peace Science*, *19*(2), 59-67. doi: 10.1177/073889420201900203

Stone, J. (2013). Cyber war will take place! *Journal of Strategic Studies*, *36*(1), 101-108. doi: 10.1080/01402390.2012.730485

Wheeler, D. A., & Larsen, G. (2003). Techniques for cyber attack attribution. *Defense Technical Information Center*, 82. Retrieved from

https://apps.dtic.mil/docs/citations/ADA468859

A Appendix

	Dependent variable:				
	Cyber Attack Onset				
	(1)	(2)	(3)	(4)	(5)
Alliance	1.093***	0.072	−0.022	0.569**	0.573**
	(0.275)	(0.213)	(0.222)	(0.224)	(0.225)
Contiguity	1.643***	1.405***	1.413***	1.105***	1.101***
	(0.343)	(0.239)	(0.241)	(0.231)	(0.229)
log GDP/capita A		1.628***	1.529***	1.705***	1.705***
		(0.130)	(0.147)	(0.139)	(0.140)
log GDP/capita B		1.053***	1.086***	1.231***	1.231***
		(0.151)	(0.154)	(0.116)	(0.116)
log Mil. Capabilities A			2.922***	1.853***	1.792***
			(0.460)	(0.477)	(0.579)
B Mil. Stronger			0.544**	0.200	0.522
			(0.224)	(0.237)	(1.249)
Mil. Cap. Asymmetry			−2.441***	−1.282	−1.109
			(0.925)	(0.814)	(1.084)
Abs. Polity Diff.				0.007	0.007
				(0.017)	(0.018)
Joint Dem.				−2.256***	−2.269***
				(0.268)	(0.282)
B Mil. Stronger × Mil. Cap. Asymmetry					−0.526
					(2.074)
Year	0.455*	0.148	0.016	0.105	0.105
	(0.272)	(0.289)	(0.266)	(0.288)	(0.288)
Year squared	0.017	0.062	0.085**	0.057	0.057
	(0.039)	(0.042)	(0.039)	(0.043)	(0.043)
Year cubed	−0.003	−0.005***	−0.006***	−0.004**	−0.004**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	−7.782***	−38.136***	−31.325***	−37.061***	−37.265***
	(0.677)	(2.406)	(3.215)	(2.709)	(3.100)
Time Controls	YES	YES	YES	YES	YES
Date Range	2005-2017	2005-2017	2005-2017	2005-2017	2005-2017
Observations	39,468	32,163	24,995	21,847	21,847
Log Likelihood	−1,844.679	−1,319.199	−1,164.103	−1,033.060	−1,032.999
Akaike Inf. Crit.	3,701.358	2,654.398	2,350.205	2,092.121	2,093.998
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 4: Extended regression table 2 which includes all of the coefficients not shown in table 2.

	Dependent variable:				
	Cyber Attack Onset				Trad. Attack Onset
	Non-Espionage	Non-Espionage	Espionage		
	(1)	(2)	(3)	(4)	(5)
Alliance	0.188	0.154	0.617**	0.201	0.929
	(0.495)	(0.489)	(0.250)	(0.618)	(0.935)
Contiguity	0.537	0.526	1.238***	0.970*	2.788***
	(0.588)	(0.584)	(0.240)	(0.563)	(0.979)
log GDP/capita A	1.735***	1.745***	1.764***	1.819***	0.868**
	(0.324)	(0.332)	(0.151)	(0.261)	(0.388)
log GDP/capita B	1.193***	1.221***	1.213***	1.538***	0.858**
	(0.214)	(0.214)	(0.127)	(0.365)	(0.430)
log Mil. Capabilities A	14.757***	15.040***	0.585	3.318**	6.600***
	(3.306)	(3.230)	(0.556)	(1.352)	(1.916)
B Mil. Stronger	−0.909	−16.929***	1.728	0.231	1.112**
	(1.104)	(2.168)	(1.315)	(0.487)	(0.437)
Mil. Cap. Asymmetry	−4.989***	−5.562***	−0.668	−0.354	−3.942
	(1.694)	(1.594)	(1.256)	(1.936)	(3.026)
Abs. Polity Diff.	−0.057	−0.060	0.015	−0.021	0.029
	(0.039)	(0.040)	(0.018)	(0.035)	(0.044)
Joint Dem.	−1.356*	−1.332*	−2.379***	−2.876***	301,668.800***
	(0.732)	(0.730)	(0.319)	(0.848)	(0.983)
B Mil. Stronger × Mil. Cap. Asymmetry		26.535***	−2.428		
		(3.405)	(2.195)		
Trad. Attack				0.745	
				(0.573)	
Cyber Attack					0.682
					(0.611)
Year	3.020**	3.130**	−0.391	0.817	−1.153*
	(1.296)	(1.324)	(0.260)	(1.077)	(0.647)
Year squared	−0.466**	−0.484**	0.150***	−0.269	0.691**
	(0.190)	(0.194)	(0.043)	(0.428)	(0.323)
Year cubed	0.021**	0.021**	−0.009***	0.036	−0.097**
	(0.009)	(0.009)	(0.002)	(0.050)	(0.042)
Constant	−20.957***	−20.816***	−39.957***	−40.289***	−14.099*
	(6.479)	(6.518)	(3.390)	(8.236)	(7.746)
Time Controls	YES	YES	YES	YES	YES
Date Range	2005-2017	2005-2017	2005-2017	2005-2010	2005-2010
Observations	21,581	21,581	21,809	9,866	9,866
Log Likelihood	−161.588	−158.650	−892.785	−193.803	−264.235
Akaike Inf. Crit.	349.177	345.300	1,813.569	415.606	556.469
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 5: Extended regression table 3 which includes all of the coefficients not shown in table 3.

	Dependent variable:				
	Cyber Attack Onset				
	(1)	(2)	(3)	(4)	(5)
Alliance	0.856***	0.060	−0.167	0.517***	0.521***
	(0.223)	(0.182)	(0.184)	(0.191)	(0.193)
Contiguity	1.639***	1.476***	1.522***	1.224***	1.223***
	(0.271)	(0.213)	(0.205)	(0.197)	(0.197)
log GDP/capita A		1.825***	1.864***	2.093***	2.095***
		(0.150)	(0.141)	(0.136)	(0.138)
log GDP/capita B		0.958***	1.083***	1.264***	1.262***
		(0.135)	(0.138)	(0.109)	(0.109)
log Mil. Capabilities A			3.375***	1.926***	1.859***
			(0.287)	(0.318)	(0.407)
B Mil. Stronger			0.232	−0.203	0.192
			(0.214)	(0.222)	(1.282)
Mil. Cap. Asymmetry			−3.883***	−2.855***	−2.624**
			(0.818)	(0.742)	(1.058)
Abs. Polity Diff.				−0.006	−0.006
				(0.015)	(0.015)
Joint Dem.				−2.588***	−2.604***
				(0.246)	(0.252)
B Mil. Stronger × Mil. Cap. Asymmetry					−0.644
					(2.153)
Year	0.437*	0.157	0.012	0.096	0.096
	(0.224)	(0.232)	(0.218)	(0.238)	(0.238)
Year squared	0.022	0.059*	0.085***	0.059*	0.059*
	(0.033)	(0.034)	(0.033)	(0.036)	(0.036)
Year cubed	−0.003**	−0.005***	−0.006***	−0.005***	−0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	−8.098***	−40.065***	−34.046***	−41.073***	−41.320***
	(0.560)	(2.534)	(2.680)	(2.397)	(2.705)
Time Controls	YES	YES	YES	YES	YES
Date Range	2005-2017	2005-2017	2005-2017	2005-2017	2005-2017
Observations	66,768	57,408	47,525	43,252	43,252
Log Likelihood	−2,524.353	−1,901.639	−1,677.204	−1,478.982	−1,478.862
Akaike Inf. Crit.	5,060.705	3,819.278	3,376.407	2,983.963	2,985.724
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 6: Robustness check for table 2 that includes additional 15 hand-picked countries as attackers in the data set in addition to the countries which launched cyber attacks and the 15 randomly sampled ones.

	Dependent variable:				
	Cyber Attack Onset				
	(1)	(2)	(3)	(4)	(5)
Alliance	1.091***	0.070	−0.024	0.570**	0.574**
	(0.275)	(0.213)	(0.222)	(0.224)	(0.225)
Contiguity	1.643***	1.406***	1.416***	1.108***	1.104***
	(0.343)	(0.239)	(0.241)	(0.231)	(0.229)
log GDP/capita A		1.632***	1.535***	1.712***	1.712***
		(0.130)	(0.147)	(0.139)	(0.140)
log GDP/capita B		1.056***	1.089***	1.236***	1.237***
		(0.151)	(0.154)	(0.116)	(0.116)
log Mil. Capabilities A			2.930***	1.862***	1.798***
			(0.460)	(0.477)	(0.579)
B Mil. Stronger			0.544**	0.199	0.545
			(0.224)	(0.237)	(1.249)
Mil. Cap. Asymmetry			−2.456***	−1.294	−1.118
			(0.925)	(0.814)	(1.084)
Abs. Polity Diff.				0.007	0.007
				(0.017)	(0.018)
Joint Dem.				−2.272***	−2.286***
				(0.268)	(0.282)
B Mil. Stronger × Mil. Cap. Asymmetry					−0.567
					(2.074)
Year	0.486*	0.178	0.040	0.128	0.128
	(0.272)	(0.289)	(0.266)	(0.288)	(0.288)
Year squared	0.013	0.058	0.082**	0.055	0.055
	(0.039)	(0.042)	(0.039)	(0.043)	(0.043)
Year cubed	−0.003	−0.005**	−0.006***	−0.004**	−0.004**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	−7.869***	−38.205***	−31.204***	−36.883***	−37.106***
	(0.677)	(2.406)	(3.215)	(2.709)	(3.100)
Time Controls	YES	YES	YES	YES	YES
Date Range	2005-2017	2005-2017	2005-2017	2005-2017	2005-2017
Observations	39,468	32,163	24,995	21,847	21,847
Log Likelihood	−1,844.679	−1,319.199	−1,164.103	−1,033.060	−1,032.999
Akaike Inf. Crit.	3,701.358	2,654.398	2,350.205	2,092.121	2,093.998
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 7: Robustness check for table 2 that includes the same sample, but instead of a rare-events logit uses the standard logistic regression.

	Dependent variable:				
	Cyber Attack Onset				Trad. Attack Onset
	Non-Espionage	Non-Espionage	Espionage		
	(1)	(2)	(3)	(4)	(5)
Alliance	−0.212	−0.280	0.586***	0.283	0.308
	(0.453)	(0.462)	(0.206)	(0.503)	(0.626)
Contiguity	0.364	0.345	1.325***	0.715	3.315***
	(0.459)	(0.456)	(0.205)	(0.496)	(0.721)
log GDP/capita A	1.854***	1.839***	2.133***	2.173***	0.778**
	(0.231)	(0.234)	(0.146)	(0.241)	(0.364)
log GDP/capita B	1.162***	1.188***	1.258***	1.582***	0.698**
	(0.204)	(0.203)	(0.114)	(0.341)	(0.323)
log Mil. Capabilities A	8.099***	8.746***	1.092**	2.659***	1.959*
	(0.988)	(0.939)	(0.425)	(0.789)	(1.106)
B Mil. Stronger	−0.531	−10.225***	0.980	−0.321	0.602**
	(0.621)	(3.629)	(1.302)	(0.504)	(0.291)
Mil. Cap. Asymmetry	−5.325***	−6.617***	−2.138*	−2.461*	−2.851
	(1.868)	(1.752)	(1.139)	(1.477)	(2.460)
Abs. Polity Diff.	−0.127***	−0.128***	0.007	−0.045	0.012
	(0.034)	(0.034)	(0.016)	(0.030)	(0.040)
Joint Dem.	−2.024***	−1.873***	−2.627***	−3.553***	−1.874*
	(0.617)	(0.624)	(0.271)	(0.722)	(0.999)
B Mil. Stronger × Mil. Cap. Asymmetry		16.097***	−1.927		
		(5.510)	(2.202)		
Trad. Attack				0.817	
				(0.585)	
Cyber Attack					1.035*
					(0.620)
Year	3.171**	3.163**	−0.332	1.185	−0.953**
	(1.417)	(1.419)	(0.213)	(1.187)	(0.428)
Year squared	−0.473**	−0.472**	0.138***	−0.357	0.395*
	(0.200)	(0.200)	(0.036)	(0.428)	(0.231)
Year cubed	0.020**	0.020**	−0.008***	0.040	−0.048
	(0.009)	(0.009)	(0.002)	(0.048)	(0.032)
Constant	−30.893***	−29.282***	−43.282***	−44.809***	−18.748***
	(5.888)	(5.962)	(2.867)	(6.600)	(6.808)
Time Controls	YES	YES	YES	YES	YES
Date Range	2005-2017	2005-2017	2005-2017	2005-2010	2005-2010
Observations	42,894	42,894	43,205	19,721	19,721
Log Likelihood	−239.418	−236.513	−1,301.722	−263.230	−496.826
Akaike Inf. Crit.	504.836	501.025	2,631.444	554.460	1,021.652
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 8: Robustness check for table 3 that includes additional 15 hand-picked countries as attackers in the data set in addition to the countries which launched cyber attacks and the 15 randomly sampled ones.

	Dependent variable:				
	Cyber Attack Onset				Trad. Attack Onset
	Non-Espionage	Non-Espionage	Espionage		
	(1)	(2)	(3)	(4)	(5)
Alliance	0.155	0.132	0.619**	0.172	0.931
	(0.495)	(0.489)	(0.250)	(0.618)	(0.935)
Contiguity	0.474	0.472	1.242***	0.983*	2.827***
	(0.588)	(0.584)	(0.240)	(0.563)	(0.979)
log GDP/capita A	1.800***	1.816***	1.773***	1.857***	0.879**
	(0.324)	(0.332)	(0.151)	(0.261)	(0.388)
log GDP/capita B	1.231***	1.257***	1.219***	1.567***	0.870**
	(0.214)	(0.214)	(0.127)	(0.365)	(0.430)
log Mil. Capabilities A	15.354***	15.816***	0.592	3.379**	6.684***
	(3.306)	(3.230)	(0.556)	(1.352)	(1.916)
B Mil. Stronger	−1.357	−19.881***	1.752	0.223	1.125**
	(1.104)	(2.168)	(1.315)	(0.487)	(0.437)
Mil. Cap. Asymmetry	−5.055***	−5.674***	−0.683	−0.371	−4.007
	(1.694)	(1.594)	(1.256)	(1.936)	(3.026)
Abs. Polity Diff.	−0.061	−0.064	0.015	−0.021	0.029
	(0.039)	(0.040)	(0.018)	(0.035)	(0.044)
Joint Dem.	−1.430*	−1.397*	−2.397***	−2.997***	−17.753***
	(0.732)	(0.730)	(0.319)	(0.848)	(0.983)
B Mil. Stronger × Mil. Cap. Asymmetry		29.486***	−2.467		
		(3.405)	(2.195)		
Trad. Attack				0.724	
				(0.573)	
Cyber Attack					0.686
					(0.611)
Year	3.310**	3.409**	−0.369	0.993	−1.165*
	(1.296)	(1.324)	(0.260)	(1.077)	(0.647)
Year squared	−0.503***	−0.519***	0.148***	−0.322	0.701**
	(0.190)	(0.194)	(0.043)	(0.428)	(0.323)
Year cubed	0.022**	0.023**	−0.009***	0.041	−0.099**
	(0.009)	(0.009)	(0.002)	(0.050)	(0.042)
Constant	−22.035***	−21.671***	−39.772***	−40.814***	−13.332*
	(6.479)	(6.518)	(3.390)	(8.236)	(7.746)
Time Controls	YES	YES	YES	YES	YES
Date Range	2005-2017	2005-2017	2005-2017	2005-2010	2005-2010
Observations	21,581	21,581	21,809	9,866	9,866
Log Likelihood	−161.588	−158.650	−892.785	−193.803	−264.235
Akaike Inf. Crit.	349.177	345.300	1,813.569	415.606	556.469
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 9: Robustness check for table 3 that includes the same sample, but instead of a rare-events logit uses the standard logistic regression.

	Dependent variable:				
	Cyber Attack Onset				Trad. Attack Onset
	Non-Espionage	Non-Espionage	Espionage		
	(1)	(2)	(3)	(4)	(5)
Alliance	−0.214	−0.281	0.585***	0.286	0.308
	(0.393)	(0.399)	(0.203)	(0.447)	(0.596)
Contiguity	0.361	0.342	1.325***	0.719	3.303***
	(0.404)	(0.398)	(0.201)	(0.442)	(0.678)
log GDP/capita A	1.855***	1.840***	2.132***	2.166***	0.775**
	(0.204)	(0.205)	(0.143)	(0.215)	(0.344)
log GDP/capita B	1.162***	1.188***	1.257***	1.577***	0.695**
	(0.178)	(0.176)	(0.112)	(0.304)	(0.307)
log Mil. Capabilities A	8.101***	8.751***	1.093***	2.659***	1.956*
	(0.841)	(0.788)	(0.416)	(0.702)	(1.050)
B Mil. Stronger	−0.546	−10.199***	0.975	−0.314	0.601**
	(0.509)	(2.694)	(1.271)	(0.446)	(0.278)
Mil. Cap. Asymmetry	−5.324***	−6.620***	−2.136*	−2.445*	−2.817
	(1.646)	(1.553)	(1.115)	(1.301)	(2.324)
Abs. Polity Diff.	−0.127***	−0.128***	0.007	−0.045*	0.012
	(0.030)	(0.029)	(0.016)	(0.027)	(0.038)
Joint Dem.	−2.029***	−1.877***	−2.623***	−3.527***	−1.854**
	(0.536)	(0.539)	(0.265)	(0.620)	(0.938)
B Mil. Stronger × Mil. Cap. Asymmetry		15.995***	−1.917		
		(4.061)	(2.148)		
Trad. Attack				0.828	
				(0.530)	
Cyber Attack					1.079*
					(0.587)
Year	3.175***	3.167***	−0.339*	1.136	−0.949**
	(1.057)	(1.047)	(0.200)	(0.873)	(0.405)
Year squared	−0.473***	−0.472***	0.139***	−0.343	0.391*
	(0.150)	(0.148)	(0.034)	(0.335)	(0.219)
Year cubed	0.020***	0.020***	−0.008***	0.039	−0.047
	(0.007)	(0.007)	(0.002)	(0.039)	(0.030)
Constant	−30.993***	−29.378***	−42.923***	−44.308***	−17.621***
	(4.859)	(4.872)	(2.809)	(5.939)	(6.464)
Time Controls	YES	YES	YES	YES	YES
Date Range	2005-2017	2005-2017	2005-2017	2005-2010	2005-2010
Observations	42,894	42,894	43,205	19,721	19,721
Log Likelihood	−239.903	−237.124	−1,301.800	−263.648	−496.987
Akaike Inf. Crit.	505.807	502.249	2,631.601	555.296	1,021.974
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 10: Robustness check for table 3 that includes the same sample, but instead of a rare-events logit uses Firth's (1993) bias-reduction approach as implemented in the R (Kosmidis, 2019).