STRATEGY 21
통권34호 Vol.17 No.2, 2014

Quantifying Naval Power and Its Implications

Bae Hack-Young *

Ⅰ. Purpose and Structure

The purpose of this article is to introduce a quantitative approach to study the effects of naval power on various conflicts between countries

* Ph.D. at Florida State University and Lieutenant Commander in ROK Navy.

along with their implications. How can we understand the effect of naval power on conflicts between or among countries? What is the significance of naval power development across countries and over time? There have been many attempts to understand the variations of naval power and its effect on conflicts. Many qualitative and historical approaches have been used to analyze it. Even though it allows us to control many different factors affecting conflicts between countries, it has a limitation to reveal a general causality between a conflict and naval power. For example, if I choose conflicts between two Koreas as a case, it allow me to control the origin of rivalry (Korean War), post-1945 minor power rivalry, territorial dispute (especially with the conflicts over the Northern Limit Line (NLL) in the West Sea), and a mixed regime dyad (South Korea made a transition to a full-grown democracy during the late 1980s and early 1990s, whereas the North has remained an authoritarian non-democratic regime during the whole period of rivalry). However, it does not allow me to generalize a theory or causality from the study to other conflicts or conflictual dyadic relationships such as Israel-Iran and China-US.[1] Thus, a quantitative study supplements in-depth case study. However, few quantitative attempts exist in a naval power study. The article is structured as follows to introduce a quantitative approach for a naval power study. It starts in the second section with the introduction of quantitative approach to conflict study. What is the benefit and what inferences can we draw from it? The third to the fifth sections introduce a way to quantify naval power based on Crisher and Souva's recent article about naval power data from 1865 to 2011; this part suggests implications and potential future research. The article concludes with a broader discussion relating to a quantitative approach to studying conflict and naval power.

---

1) George, A.L., 2005. Comparative methods: Controlled comparison and within-case analysis. In A. L. George & A. Bennett, Eds., *Case studies and theory development in the social sciences*. MIT Press.

Ⅱ. What Is the Quantitative Approach?[2)]

Research methods in conflict study are often divided into two main types—quantitative and qualitative methods. This article will study conflict and naval power using the quantitative method. When it comes to quantitative methods, you will probably be thinking of statistics or numbers. They are part of the quantitative method and capture some of the essence of quantitative methods. The following definition, taken from Aliaga and Gunderson (2000), describes what we mean by quantitative research methods: quantitative research is "explaining phenomena by collecting numerical data that are analyzed using mathematically based methods (in particular statistics)"[3)]. The first part of the definition is about 'explaining phenomena'. This is a key element of all research and all scholars do. The goal is to explain some variation of political phenomena (dependent variable) in the world with factors (independent variables) that drive the variation. For example, in a security study, this could be questions like "Why do some countries go to war more than others do?", and "How much does naval power affect war and conflict between countries?", "Does an increase of naval power deter the adversary?", and so on.[4)] In quantitative research, we collect numerical

2) A detailed overview of the extent and the basis of quantitative approach see Aliaga, M. & Gunderson, B., Interactive statistics (Prentice Hall: 2000) and Kennedy, Peter. A Guide to Econometrics (Blackwell Publishing: 2008).

3) Ibid., pp. 4–26.

4) There are two major arguments of naval power on variation of conflicts: positive and negative. On the one hand, naval power enables countries to project their power to other countries. Thus, it encourages countries to project its military power on other country when disagreement on controversy issue between two countries (see Crisher, Brian Benjamin, and Mark Souva. "Power Sea: A Naval Power Dataset, 1865-2011". International Interactions, no. just-accepted (2014)). On the other hand, naval power deters adversaries' military action. The navy-oriented second-strike capability changes the calculation of first strike payoff. Thus, it makes conflicts less likely (see Huth, Paul, and Bruce Russett. "Deterrence failure crisis escalation". International Studies Quarterly (1988): 29-45). It has a limitation to apply these theories to countries in Northeast Asia because they focus on super power rivalries. However, I introduce these theories to show that

data. This is closely related to the final part of the definition, which is "analysis using mathematically based methods". In order to be able to use mathematically based methods, the data has to be in numerical form. This is why many people think quantitative research is not an easy way to analyze security studies. Because quantitative research is essentially about collecting numerical data to explain a particular phenomenon (variation of dependent variable), particular questions seem immediately suited to being answered using quantitative methods. For example, how many wars occurred between superpower countries? What percentage of countries with large naval power experience war and conflict? Have countries with large naval power experienced more wars or conflicts than countries with lesser naval power? These are all questions we can look at quantitatively, as the data we need to collect is already available to us in numerical form. There are many questions about naval power (as a dependent variable and also as an independent variable) we might want to look at, but which do not seem to produce any quantitative data. In fact, many scholars are skeptical about quantifying naval power and little research actually occurs in the form of "naturally" quantitative data about the topic. Much data that does not naturally appear in a quantitative form can be collected in a quantitative way, e.g., naval power. Crisher and Souva (2014), for example, attempt to quantify a naval power and suggest several theoretical and practical implications from the data. This is possible by designing research instruments aimed specifically at converting elements of naval power that don't naturally exist in a quantitative form into quantitative data, which I will introduce this article. The last part of the definition refers to the mathematical method, which is what people usually think about when they think of quantitative research, and is often seen as the most important part of

---

there are conflicting theories in effect of naval power on war. As you see these two conflicting theories about the effect of naval power on conflicts, we need to test these theories employing scientific method and data to see which one is more valid in real world.

quantitative studies. However, using the right research design and data collection instruments is actually more crucial than using the right data analysis tools like statistics. However, most researchers do not really have to be particularly expert in the mathematics underlying the methods, as computer software (e.g., STATA or R, etc.) allows us to do the analyses quickly and relatively easily. The quantitative method is pragmatic in descriptive and inferential questions to support a causal relationship, which I will show with naval power data in this article.

Ⅲ. Quantifying Naval Power[5)]

Crisher and Souva (2014) recently attempted to quantify state naval power, which they define as a state's ability to use sea-based weapons to inflict physical damage on other states' people, territory, structures, and weapons systems. This conceptual definition should be operationalized by observable and measurable indicators for a state's naval power. It is too ideal to measure each ship's ability and combination of ship's ability to inflict damage on an adversary's territory or weapons systems. Such an assessment would consider a ship's displacement, weapons systems, total firepower, speed, armor, maneuverability, and command, control, communications, computers, and intelligence (C4I) capability, among other capabilities. Unfortunately, such an assessment is not possible. Instead, Crisher and Souva propose to calculate the total tonnage of a country's primary warships. They recognize the difficulty of operationalization and note that the aggregate tonnage of a state's navy may tell us about

5) For a detailed overview of the extent and the basis of quantifying naval power, see Crisher, Brian Benjamin, and Mark Souva. "Power Sea: A Naval Power Dataset, 1865-2011". *International Interactions, Just-accepted* (2014). The purpose of this section is to introduce and summarize Crisher and Souva's article about quantifying naval power. Most of the chapter heavily relies on their article for implications I will introduce in the next section.

its overall strength, but it will not tell us how well their navy will perform in a combat capacity or whether their navy is qualitatively better than a navy of comparable tonnage. Despite the limitation, it tells countries' naval power that we need to analyze the effect on a various conflicts between countries. For coding, they need to decide what ships to include in the dataset before computing tonnages. In general, their data includes all ships with the capability of using kinetic force to damage targets for purposes beyond self-protection and that can operate outside of their littoral waters. The problem is ships satisfying this criterion will vary over time. Given these problems, they took several steps to satisfy the conceptual definition of naval power.

Three steps have been taken to quantify naval power from 1865 to 2011 by Crisher and Souva.[6] First, they distinguish naval periods that allow them to capture changes to the dominant warship over time.[7] The example they gave in the article is that a pre-Dreadnought battleship was not the most capable ship type in 1910 (the super-Dreadnought class battleships were), but compared to the premier warships twenty years earlier, it is at least as capable. Further, as we previously noted, no single dimension allows for a perfect distinction between warships. Because of changes in naval technology and the multiple dimensions that comprise warship capability, Crisher and Souva distinguish five different naval periods. Their first period extends from 1860 to 1879. This is a transitional period as ship designers began coming to grips with the technological leaps in terms of hulls, guns, and munitions. Hulls were made thicker, sometimes out of iron and sometimes out of wood. The second period was from 1880 to 1905. The pre-Dreadnought emerges as the dominant warship. The British HMS *Royal Sovereign* launched in 1891 is an example of a pre-Dreadnought from the period. Whereas the *Agamemnon* displaced 8,510 tons, the *Royal Sovereign* displaced 15,580 tons. Additionally, the primary guns of the *Royal Sovereign* were four

6) Ibid., pp. 9-19.
7) Ibid., pp. 12-16.

13.5 inch breech-loading guns capable of firing a 1,250 pound shell 12,000 yards, while the guns of the *Agamemnon* could only reach 6,500 yards. Lastly, despite being vastly heavier than the *Agamemnon*, the *Royal Sovereign* had a maximum speed of 15.7 knots, which was 2.7 knots faster than the *Agamemnon*. In sum, the pre-Dreadnoughts were faster, heavier, and more powerful than the battleships of the preceding period. The third period covers the years 1906 to 1945. The launch of the HMS *Dreadnought* in 1906 ushered in the era of the battleship. The *Dreadnought* at its launching was the fastest battleship in the world and could reach a speed of 21 knots (roughly 39 km/h). Additionally, she displaced over 20,000 tons when fully loaded and was armed with ten 12-inch guns. Another notable battleship of this period was the German battleship *Bismarck*. At the time of its launch in 1939, the *Bismarck* displaced over 50,000 tons and carried eight 15-inch guns. These 15-inch guns were capable of firing 1,800-pound shells. Clearly, during this time battleships became bigger and more powerful. This period also saw the development of the aircraft carrier, which began to displace the battleship as the capital warship during World War II. The worth of the aircraft carrier was shown during the sinking of the *Bismarck*. In a battle with the HMS *Hood*, one of Great Britain's major battle cruisers, the *Bismarck,* sank the *Hood* and proceeded to head back to port for repairs. However, torpedo-bombers launched from the HMS *Royal Ark* intercepted the *Bismarck* and badly damaged her rudders, making her virtually unmaneuverable. This allowed other British battleships to catch up, and eventually sink the *Bismarck*. The fourth period is the first post-World War II period and extends from 1946 to 1958. As the primary naval power in this period, the U.S. Navy focused on projecting power inland. This led to an era where technological advances in armaments outpaced advances in ship design - notably the improvement in missile technology. Lastly, the fifth period deals with warships between 1959 and 2011. Two major technological innovations mark the beginning of this final period. Both of these innovations

highlight the U.S. Navy's focus on using the navy to project power inland in the post-WWII world. The first occurs in 1959 with the launching of the George Washington class nuclear submarines. These were the first submarines to carry Polaris nuclear missiles. Additionally, the launch of the USS *Enterprise* in 1960 marked the launch of the first nuclear powered aircraft carrier. Ships could now inflict an incredible amount of damage on an enemy state and stay afloat or submerged as long as they had the necessary supplies to sustain their crew. These innovations create a natural cutoff point to mark the late period of naval technology.[8)]

The second step in order to quantify naval power is to record individual ships.[9)] After establishing the naval periods, Crisher and Souva record all ships and their ship types that meet minimum criteria. Periods One (1865-1879) and Two (1880-1905) include the least amount of variation among the types of warships available to all the world's navies. As such, their minimum criterion for recording a ship is straightforward for these two periods. In Period One, they record all ships if they displace at least 1,000 tons.

For Period Two they add a gun-size requirement and record all ships if they displace at least 2,000 tons and have a 5-inch primary gun or greater. Due to the lack of variation in ship types in these periods, they only record a ship's displacement, not their ship type. By Period Three (1906-1947), as they noted previously, the landscape of naval technology had been dramatically altered. Because of this, there was a need to alter the minimum criteria for recording ships as well. In particular, Crisher and Souva have minimum criteria for aircraft carriers, non-carrier warships, and submarines. They record all aircraft carriers that are designated as such. However, when recording the ship

8) See more Polmar, Norman. *The Naval Institute Guide Ships Aircraft U.S. fleet.* Naval Institute Press, 2000, Polmar, Norman. *The Naval Institute Guide Soviet Navy.* Naval Institute Press, 1991, and Prezelin, Bernard. *The Naval Institute Guide Combat Fleets World.* Naval Institute Press, 1995.

9) Ibid., pp. 16-17.

type for these carriers, they make a distinction between major and minor aircraft carriers.[10] Major aircraft carriers have at least 10,000 tons displacement, while minor aircraft carriers have less than 10,000 tons displacement. Next, they record all submarines that are designated as such. In this case, they consider submarines "major" that displace at least 1,000 tons submerged and have four torpedo tubes, while submarines that displace less than 1,000 tons submerged are considered "minor". Lastly, we record all non-carrier warships that have at least 2,000 tons displacement and 5-inch guns, or ships with 1,000 tons of displacement and at least 3 torpedo tubes. Among non-carrier warships, they do distinguish between major and minor battleships. Ships that are designated as battleships and have at least 20,000 tons of displacement and 12-inch guns are considered major battleships, while battleships that do not meet these requirements are considered minor battleships. They record ships in Period Four (1947-1958) similar to Period Three. They have minimum criteria for aircraft carriers, non-carrier warships, while making some additional distinctions among certain ship types. Because there was little development in ship design during this period, the coding system is similar to that of Period Three, but with some increases in the minimum displacements. Ships designated as aircraft carriers are recorded as a major aircraft carrier if they displace at least 20,000 tons and have at least 10 jet fighters. Aircraft carriers with less than 20,000 tons of displacement are considered minor aircraft carriers. Submarines with at least 2,000 tons displacement submerged and four torpedo tubes are considered major submarines, while submarines with less than 2,000 tons of displacement are considered minor. Lastly, they record non-carrier warships that have at least 2,000 tons of displacement and 5-inch guns or six torpedo tubes.[11] In

10) The Conway series makes a similar distinction for other types of ships. For example, armored cruisers are classified either as an armored cruiser or as light armored cruisers. Essentially, they are making the same distinction among ship types as the Conway series but applying it to more ship types, (e.g., battleships, aircraft carriers, and submarines).

Period Five (1959-2011) they record aircraft carriers with at least 30,000 tons of displacement and 10 jet fighters as major aircraft carriers, while minor aircraft carriers are those with less than 30,000 tons of displacement. For non-carrier warships, they record ships that have at least 3,000 tons of displacement and 5-inch guns, at least 6 torpedo tubes, or missile capability. In terms of submarines, they consider those submarines that are capable of launching nuclear ballistic missiles separately from other submarines. However, conventional submarines with at least 3,000 tons of displacement submerged and four torpedo tubes are classified as major submarines, while submarines with between 2,000-3,000 ton displacement submerged and four torpedo tubes are classified as minor submarines.

The third step is to count the total number of ships that satisfy criteria for each state in a given year.[12] If a ship satisfies the criteria for counting, sum the total tonnage of the ships in given year. Based on this quantified naval power data, I can move on to the analysis part.

Ⅳ. Descriptive Analysis

Based on the data above, I can show the trend and change of each country's naval power over time and try to draw inferences from it.

11) They drop the distinction between major and minor battleships in this period as no battleships were launched in this period.

12) Ibid., p. 18.

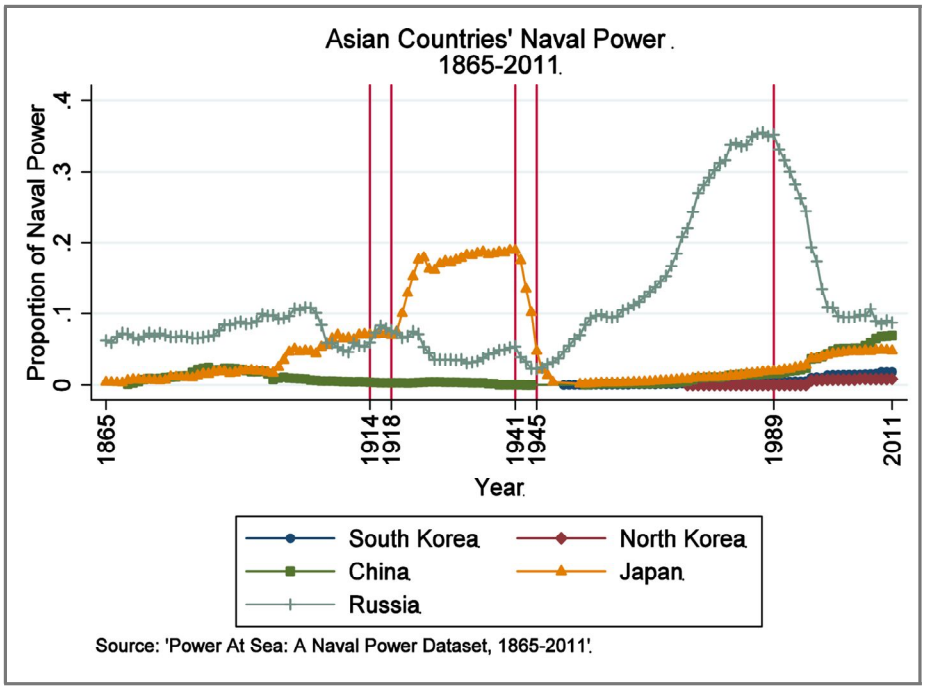

Figure 1. Northeast Asian Countries' Naval Power Change from 1865 to 2011.

The figure above shows the change of Northeast Asia countries' total tonnage from 1865 to 2011. In the figure, the horizontal axis represents the temporal domain of the dataset for naval tonnage from 1865 to 2011. The domain covers WWI and WWII and many other wars and conflict periods in the world. The vertical axis represents the proportion of total naval tonnage of each Northeast Asia countries. I include Russia because it affects Asian security as found in the Korean War, the Russo-Japanese War and so on. Before the two World Wars and during Cold War, there are few drastic changes but steady increase of naval powers among Northeast Asian countries in naval power. In particular, during the Cold War, Russia beefed up its naval power. The descriptive statistics show that when naval power increases, there is no war. However, once naval power is saturated, countries are more likely to crush each other because they have more power-projection capability. For example, before two world wars and the collapse of the

Berlin Wall, countries kept increasing their naval power. It could be interpreted both that naval power keeps peace but also encourages conflicts. It is limited to interpret whether a naval power increase or decrease a probability of war between countries with this descriptive statistics. One more interesting trend of naval power in Northeast Asia is all four countries built more ships over the last two decades. China and Japan each built almost ten percent of the total tonnage produced in the world. China launched its first aircraft carrier (*Liaoning*) in 2012. On top of that, China is developing its own aircraft carrier program, including China's first domestically built aircraft carrier, which will be a larger version of Liaoning. The design is reportedly based on drafts of a Soviet-era, nuclear-powered, 80,000-ton vessel capable of carrying 60 aircraft. Japan, on the other hand, commissioned its first Aegis-equipped destroyer, the DD173 *Kongō*, in 1993. Japan has commissioned three *Ōsumi* class LSTs (型輸送艦), whose displacement will be 8,900 tons standard, and 14,000 tons full load. This trend in Northeast Asia might keep the "balance of terror", which would mean that none of the countries could initiate a war based on its rival's naval power-specifically, how much these countries have increased their respective naval powers and how much this change has affected the naval power proportion over time.

Figure 2 shows two types of information in a graph-the change of total tonnage of each country and the change of proportion of tonnage each year. The horizontal axis represents the years from 1950 to 2011 (the temporal domain limited including post-WWII period). The vertical axis shows two types of information; the left axis shows the level of total tonnage and the right axis shows the level of increase or decrease in proportion of a given year in comparison to the previous year.
Despite the different degree of changes, all four countries put out the resources to build more naval ships. Particularly in the late 1990s and early 2000s, the increase in naval power of all four countries was drastic.

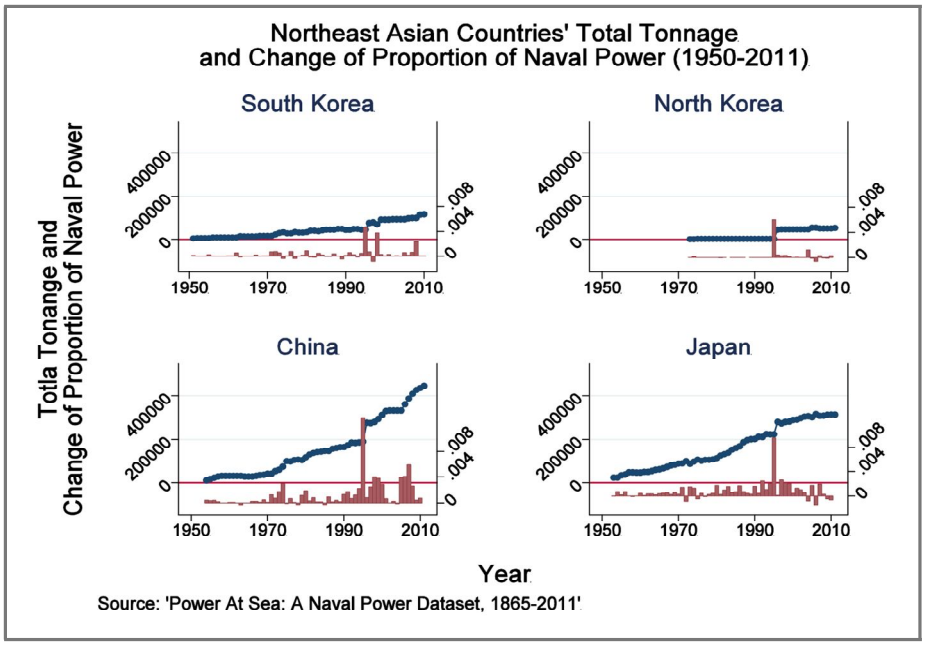

Figure 2. Northeast Asian Countries' Total Tonnage and Change of Proportion of Naval Power (1950–2011).

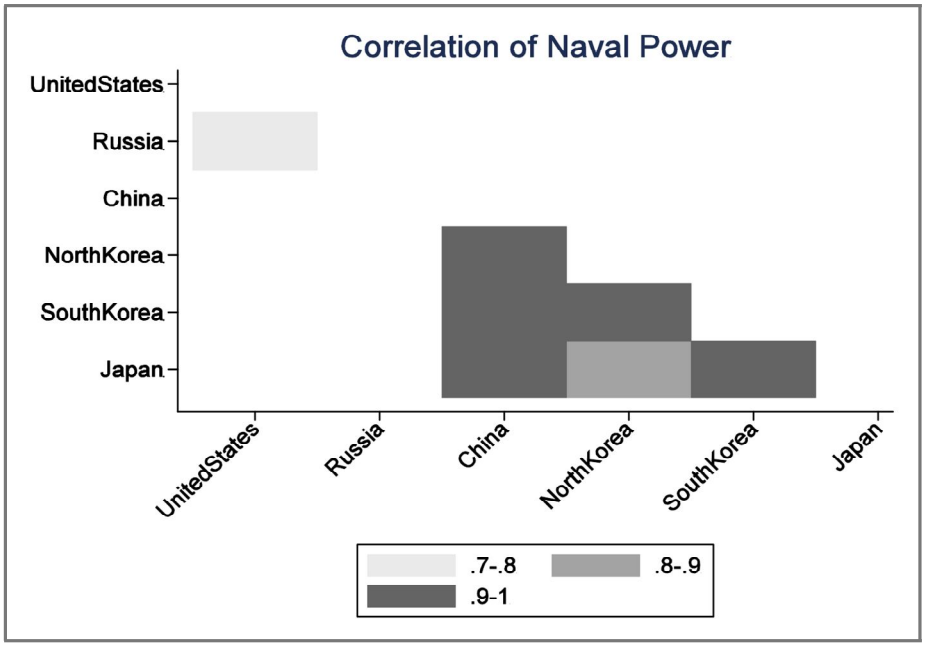

Figure 3. Correlation of Naval Power among Northeast Asian Countries.

One of the reasons for building additional naval ships is to defend oneself from outside threats. However, not all countries pose a threat to another country. For example, can you imagine a threat between Finland and Guatemala? They might have some issues, but they are not serious enough to result in conflict between these two countries. Additionally, it would be hard to find any serious conflictual interaction between the two countries, unlike the relations between North and South Korea or China and Japan. Thus, these international rivals[13] are more likely to experience more disputes, so that they are more likely to develop naval power in response to a rival. Figure 3 shows the correlation of naval power among Northeast Asian countries. Not surprisingly, North and South Korea, Japan and South Korea, Japan and China's naval power appear highly correlated, while Russia and other Asian countries are not. This empirical correlation does not tell whether this naval power arms race between Asian rivals increases or decrease conflicts between them. However, it tells that rivals countries, which have experienced severe disputes, are more likely to develop their naval power when its adversary does. This empirical observation leads to a theoretical question: when two countries are in a naval arms race, are they more likely to experience war?

13) For more deep discussion about international rivalry, see Diehl, Paul Francis, and Gary Goertz. *War and peace International rivalry* (University of Michigan Press: 2000). The definition of international rivalry is two countries that are not satisfied with the current status quo between the countries. Thus, they keep trying to change it by using methods such as diplomatic coercion or military force. This leads to more disputes and wars between international rivals. There are many empirical definitions about international rivalry. In this article, I use Klein, James P, Gary Goertz, and Paul F Diehl's empirical definition: more than six militarized disputes within twenty years. For more details, see Klein, James P, Gary Goertz, and Paul F. Diehl. "The New Rivalry Dataset Procedures Patterns". *Journal of Peace Research* 43, no. 3 (2006): 331-348.

Ⅴ. Inferential Analysis[14)]

A goal always determines a method of empirical analysis. The key theoretical question is whether the change of naval power deters a war between countries and its implication to conflicts in Northeast Asia. It has a limitation of large-N study to apply its result in area or case study. However, this attempt allows us to understand regional conflict and naval power with a theory from a theory drawn from many cases. As I introduced (see note 4), there are conflicting arguments on the effect of change in naval power. In the previous section, descriptive statistics support both arguments. This leads me to one hypothesis:

H1: If the naval power increases, a war between countries is less likely to occur.

If I can't nullify this hypothesis with statistical result, an increase of naval power makes war less likely. The theoretical question and hypothesis in this article lead me to selecting the unit of analysis, variables, and method I will use to explore the question. First, the unit of analysis is rivalry year. While extant empirical evidence appears to show Asian rivals experience high level of a naval arms race, the question is about the effect of the arms race on war. Second, the dependent variable is war, rather than all militarized disputes.

Most studies collapse different levels of dispute together, even though the cause and consequence of war in one dispute might be very different from that in other war. Third, the common probit model will

14) 12. The data in this section is partly derived from my dissertation (War and Rivalry: Political Shock and Bargaining, Florida State University: 2014)). For a broader and deeper discussion on data choice, see my dissertation.

be used. This allows me to make a direct comparison of effect of naval arms race on war between rivals with other factors affecting war.

1. Dependent Variable

The dependent variable is war onset between rivals. The term war is generally used to refer to armed conflict between organized groups. However, this leads me to more questions: what organized groups should be included? How severe should the conflict be? When does it start? The first question is about participants of the armed conflicts. The wars in my data set are between states rather than among states; between /among a state(s) and a non-state entity; and within states. In other words, wars take place between states (members of the interstate system). The second question is about the level of severity within armed conflict. A battle-related death has been used as a criterion-in other words, how many battle-related persons die in a conflict within a certain period. There is a tradeoff on the level of the criterion. On the one hand, a high threshold will exclude many forms of conflict not involving any casualties or minor armed disputes with few fatalities. Additionally, it will miss the latent tension between rivals and lower level conflicts. On the other hand, the low-level threshold allows us to include more cases. However, it is also less likely to be reliable information about the broader set of such events for different times and regions of the world.[15] The third question is about when the war onset should be coded as 1. Maoz argues that we should code when there is an event, such as actual battle between two countries, which will satisfy the war criteria (Maoz 2005). However, once a rival declares a war or decides to fight against a country, it is the war regardless of whether the rivals fight.

There are three war data sets I will compare to determine which

15) Gleditsch, Kristian. "A Revised List of Wars of Independent States, 1816-2002". *International Interactions* 30, no. 3 (2004): 231-262.

data set is more appropriate to determine war onset between rivals. One representative war data set is the Correlates of War (COW) Project[16] Ver. 4.0. The COW Project's International War data (Singer 1972) has served as a standard source for research on armed conflict between states. Its empirical definition of interstate war is a military conflict between international member states, resulting in a minimum of 1,000 battle-related combatant fatalities within a twelve-month period.[17] The other data set is Gleditsch's revised war data set.[18] The purpose of the data set is to re-examine existing COW data on wars between independent nations rather than to advocate a new approach to the study of conflict or the definition of war. Primarily, he focuses on finding missing independent states in the COW project[19]. He includes many political entities that were not colonies and that were commonly regarded as independent states that were excluded from the COW interstate war list. Thirty-five wars were added and reclassified from extra-systemic or civil wars to interstate wars. In terms of the number of wars in the data, the Gleditsch data set has more war cases than COW.

16) The latest publicly released version of the COW war data (Version 4.0) is available from http://www.correlatesofwar.org/. This contains data on the participation of nation states in international wars between 1816 and 2007

17) Sarkees, Meredith Reid, and Frank Wayman. *Resort War: 1816-2007* (CQ Press, 2010).

18) Gleditsch, Kristian. "A Revised List of wars of Independent states, 1816-2002". *International Interactions* 30, no. 3 (2004): 231-262.

19) COW project has different criteria before and after 1920. Before 1920, candidates had a population greater than 500,000 and whether entities were "sufficiently unencumbered by legal, military, economic, or political constraints to exercise a fair degree of sovereignty and independence". After 1920, Singer and Small relied on whether a nation "either (a) was a member of the League of Nations or the United Nations at any time during its existence, or (b) met the half-million population minimum and received diplomatic missions from any two (rather than the specific two) major powers (Singer and Small 1972, p. 21)

Table 1. The Difference among COW's, Maoz's And Gleditsch's War Data Set

	COW (Ver. 4.0)	Maoz (Ver. 2.0)	Gleditsch
Temporal Domain	1816-2007	1816-2001	1816-2002
Empirical Definition	More than 1000 battle-related death in a year between countries	More than 1000 battle-related death in a year between countries	More than 1000 battle-related death in a year between countries
Features	The first war data set	Correction of multilateral wars in COW	Correction of participant in COW
War Years from 1816 to 2000 (War Onset)	808 (244)	718 (271)	979 (263)
War Years B/W Rivals from 1816 to 2000 (War Onset)	308 (76)	316 (113)	348 (81)

Source: COW MID Ver. 4.0, Maoz War data Ver. 2.0, and Gleditsch's revised war data.

The last data set is Maoz's Dyadic Militarized Interstate Disputes Dataset Version 2.0 (DYMID2.0) (Maoz 2005). Maoz also used COW's war data set and corrected several COW errors caused by transforming multilateral disputes to a dyadic record.[20] COW coded a war between

20) First, some of the dyads in a multilateral dispute may not be actually valid. He gives dispute #257, World War I, as an example. In this war, there are quite a few states that are listed on opposing sides of the dispute, but never threatened, displayed, let alone used force against each other. There is war between Japan and Bulgaria only because they were on the opposite side. Second, it measure inaccurate levels of hostility. He gives dispute #258, World War II, as an example. Even though the U.S. and Hungary declare war on each other, they did not fight each other. The dyadic hostility level was plugged in by the level between Axis powers (Germany, Italy, Japan, Hungary, Romania, Bulgaria) versus Allies (U.S., Britain, France, USSR, Australia, Belgium, Brazil, Canada, China, Denmark, Greece, Netherlands, New Zealand, Norway, Poland, South Africa, Yugoslavia). Third, the disputed year is not accurate. If a dispute had more than one state on a given side of the dispute, but started and ended within the same year, it gets one record for every valid dyad in the dispute. If the dispute had m initiators and n targets, it may-but does not necessarily-have as many as m×n records. He attempted to include only dyads for which there is evidence of actual exchange of military hostility acts (Dyadic Militarized Interstate Disputes (DYMID2.0) Dataset-Version 2.0
http://psfaculty.ucdavis.edu/zmaoz/dyadmid.html

countries even when there was no physical fight between them. Maoz has more strict criteria on war, adding the requirement of a physical fight. Table 1 shows how different these datasets are in terms of coding war and war onsets. Among these three datasets, I will use Gleditsch's war dataset because it includes the most comprehensive war onsets. My dependent variable is dichotomous between war onsets and others, which produce two types of zero in the variable: no war and ongoing war. I will drop ongoing war cases for two reasons. First, the comparison for the test should be conducted between war initiation and no war. The focus of my theory is about the beginning of war between rivals, rather than ongoing war. How long a war lasts or whether or not war is ongoing is out of my scope. By excluding ongoing war between rivals, I can compare between war initiation and no war. Second, I can avoid reverse causality (endogeneity). There is concern that some political shocks are caused by war because of political shocks that occur after war initiation and during an ongoing war. By eliminating ongoing war observations, I can prevent the naval arms race caused by war.

2. Key Independent Variable

Naval Power. The variable is created by summing up the total tonnage of rivals. It captures how much more naval power is built by two countries than previous year. For example, North Korea's total tonnage is 53,274 tons and South Korea's is 118,180 tons in 2010. The value in 2010 for this variable will be 171,454 tons.

3. Control Variables

Joint Democracy.[21] It indicates whether each dyad is in a joint democracy. Employing Polity III allows me to identify joint democracies

21) For more details see Conrad, Justin, and Mark Souva. "Regime Similarity Rivalry". *International Interactions* 37, no. 1 (2011): 1-28.

based on the degree of institutionalization. I use democracy-autocracy score from −10 to +10, with −10 being full autocracy and +10 being full democracy. Six will be the cutoff of between democracy and non-democracy. If two rivals are in a democracy, Joint Democracy is coded as 1, 0 otherwise expecting negative coefficient.

Joint Satisfaction. It indicates whether rivals are satisfied with the status quo. It is captured by measuring the similarity of each rival's entire alliance portfolio with that of the international system's dominant power. It assumes that if rivals have a similar alliance pattern they are less likely to become rivals. I measure similarity with Signorino and Ritter's SC Coefficient (Signorino & Ritter 1999) expecting a negative coefficient in parallel fashion to my Joint Democracy variable.

Power Parity. It indicates the balance of power between rivals. I use COW's Composite Capabilities Index dividing the weaker rival's Composite Capabilities Index by the stronger rival's (Singer 1987). It results in ranges between zero and one. The highest possible value is an exact equality in terms of balance of power while the lowest value means a large national capability gap.

System Shock and Domestic Shock. Goertz and Diehl argued that the presence of system or domestic shocks is a necessary condition for the onset of rivalry (Goertz & Diehl 1995). System shock is coded, as Goertz and Diehl argued, as world wars, major change in the distribution of territory, and/or change of power distribution. As they coded, I also code the years and the periods for ten years after the shock ended. The domestic shock variable will be coded similarly to that of Goertz and Diehl. When either rival experienced a COW-recognized civil war, government malfunctioning, and irregular leadership change, it is coded as 1, and 0 otherwise.

Contiguity.[22] This indicates whether the rivals are sharing a border (land or river or less than 12 NM at sea). I use COW's contiguity data

22) For more detail see Gochman, Charles S. "Interstate metrics Conceptualizing, operationalizing, measuring Geographic proximity states Congress Vienna". 17, no. 1 (1991): 93-112.

set by coding 1 (separated by a land or river border) and 2 (separated by 12 miles of water or less) as 1 and 3 (separated by 24 miles of water or less but more than 12 miles), 4 (separated by 150 miles of water or less but more than 24 miles), 5 (separated by 400 miles of water or less but more than 150 miles) are coded as 0 (Gochman 1991).

Joint Autocracy To control the democratic peace theory and joint institutional similarity mechanism, a joint autocracy variable will be added. Joint democracy variable is not used because there are few joint democracies in the rivalry sample. I use the same COW Polity data to generate the variable using the same cut-off (6) for identify democratic and non-democratic countries. If two rivals are jointly nondemocratic, it will be coded as 1, and 0 otherwise.

Power Parity. To control the balance of power hypotheses, a change of balance of power between rivals will be added. This is the same variable I used for the selection model expecting null effect on war onset.

Contiguity[23]. Vasquez's argument will be controlled by a contiguity variable (Vasquez & Henehan 2001). Shared access to a physical area can lead directly to interstate friction, although the rivals involved agree as to where the border lies between them. He operationalizes the territorial disputes between rivals as an indirect measure; contiguity (1 = yes, 0 = no) is coded as 1 if an adjacent land border exists or there is a separation between countries of less than 150 miles of water, and as 0 otherwise. However, the coding rule for contiguity is stricter in my dissertation. I used the same variable for my selection model: either sharing a border by land or river or less than 12 NM at sea will be coded as 1, and as 0 otherwise.

Peace Year To control the temporal dependency, variable Peace Year

23) There is controversy of this variable regarding whether contiguity may be associated with war because territorial issues are more prone to violence or because contiguity provides an opportunity for war due to proximity or frequent interactions. As Vasquez mentioned, the analysis in my paper is not intended to provide such an argument, but to assess a political shock explanation of why rivals go to war.

is added by counting the years since either the start of rivalry or the previous occurrence of a war. I generate this variable by using Beck, Katz and Tucker's BTSCS STATA code.[24] It also produces dummy variables such as spline1, spline2, and spline3.

4. Model

$$Y\ (The\ Likelihood\ of\ war\ between\ rivals\ in\ a\ Given\ Year)_{it} = a + \beta_1\ (Joint\ Democracy)_{it} + \beta_2\ (Joint\ Satisfaction)_{it} + \beta_3\ (System\ Shock)_{it} + \beta_4\ (Domestic\ Shock)_{it} + \beta_5\ (Parity)_{it} + \beta_6\ (Contiguity)_{it} + \beta_7\ (Peace\ Years)_{it} + \beta_8\ (Spline1)_{it} + \beta_9\ (Spline2)_{it} + \beta_{10}\ (Spline3)_{it} + \varepsilon_{it}$$

In Table 2, I also summarize expectations about each variable. The primary dependent variable for each model is war onset between rivals; the independent variable is the summation of naval power of rivals.

Table 2. Expectations About βs

Regressors	Expectation
Joint Democracy	**− (Significant)**
Joint Satisfaction	**− (Significant)**
System Shock	**+ (Significant)**
Domestic Shock	**+ (Significant)**
Parity	**+ (Significant)**
Contiguity	**+ (Significant)**
Peace Year	**− (Significant)**

24) For more detail, see Beck, Nathaniel, Jonathan N Katz, and Richard Tucker. "Taking Time Seriously: Time Series Cross-section analysis binary dependent variable". *American Journal Political Science* 42, no. 4 (1998): 1260-1288.

Table 3. War Onset and Summation of Naval power (KDG's Rivalry Data[25])

	(1) Model 1	(2) Model 2	(3) Model 3
War			
Sum of Naval Power	**−0.000000256***	**−0.000000287***	**−0.000000371***
	(0.000000110)	(0.000000140)	(0.000000155)
Joint Demo		0.740*	0.734*
		(0.313)	(0.330)
Joint Satisfaction		−0.0221	−0.147
		(0.363)	(0.377)
System Shock		0.645*	0.630
		(0.317)	(0.361)
Domestic Shock		0.647**	0.647**
		(0.211)	(0.218)
Parity		−0.912	−1.464*
		(0.576)	(0.680)
Contiguity		0.0356	0.0678
		(0.229)	(0.238)
Peace Years			0.0111
			(0.0724)
spline1			0.000135
			(0.000460)
spline2			−0.0000760
			(0.000259)
spline3			0.0000107
			(0.0000448)
_cons	−2.055***	−2.738***	−2.567***
	(0.103)	(0.398)	(0.583)
N	1638	1623	1621
adj. R^2			

Standard errors in parentheses * ** ***
$p < 0.05$, $p < 0.01$, $p < 0.001$

25) There are two representative rivalry datasets: Thompson (2007) and Klein, Diehl and Goertz (2006). The former dataset was constructed based on operational definition: when two countries perceive each other as a threat, it is a rivalry. The latter follow Goertz and Deihl's definition: if two countries experienced more than 6 militarized disputes, it is a rivalry. In this article, I use KDG's rivalry dataset for analysis because it is less subjective and fit in to purpose of this article to introduce a qualitative approach.

Table 3 displays three major results, including a result based on the statistical estimation. The table shows that the naval power between rivals makes war less likely between rivals. The first column of results in Table 3 reports the estimated impacts of the naval power without other control variables. The second column of results shows the estimates when I include control variables measuring other competing theories: whether both rivals in the dyad are democratic countries, whether or not they have similar alliance, whether or not there is system or domestic shock, whether or not they are sharing a border, and the power parity. The third column shows the results taking into account time dependency. One of the assumptions of this model is that each war occurs independently. By controlling the time dependency among wars, I can predict the effect of naval power on war between rivals more accurately.

With these results in the three models, increase of naval power makes war between rivals less likely, which supports hypothesis 1. The results consistently show that the effect of naval power on war between rivals is negatively related to the likelihood of war onset between rivals. The coefficient of the "naval power between rivals" variable is negative and statistically significant across the models. These results support the hypothesis that rivals are less likely to go to war against a rival who has built up their navy.

Ⅴ. Concluding Discussion

The goal of this article is to introduce the idea of quantifying naval power to study its effect on various conflict phenomena and the development of naval power. Using a quantitative approach to study naval power has both strengths and weaknesses. On the one hand, it has several effective ways to study the effect of naval power. First, we can test and validate already constructed theories about naval power.

Scholars have accumulated theoretical arguments on naval power. These indicate the effect of naval power on different international and domestic conflicts as well as what drives the variation of naval power across countries over time. We can test previous these theories by testing hypotheses that are constructed before the data is collected. Through the hypotheses test, we can generalize research findings because it can be replicated on many different populations and sub-populations. Second, it is useful for making quantitative predictions. The researcher may construct a situation that eliminates the confounding influence of many variables, allowing one to establish cause and effect relationships more credibly. Given a model, we can predict a probabilistic change of war when a naval power changes across countries and over time. Third, it also provides outcomes that are more credible for policy practitioners such as administrators, politicians, and strategists. The outcomes from quantitative analysis are not based on selective cases, but on large number of observations. Its external validation is more credible than a small number of case studies. On the other hand, the analysis has weaknesses as well. First, too much generalization poses limitations to a country-specific context. The researcher's categories, which are used for data construction or statistical analysis, might not reflect local constituencies' understandings. In other words, it suffers from lack of internal validation. Second, it may not provide an outcome appropriate to a specific case. Knowledge produced from the approach might be too abstract and general for direct application to specific local situations, contexts, and individual country.

Given the strengths and weaknesses, I argue that using the quantitative approach to study naval power is valuable by throwing a new light on the new aspect of conflict study that qualitative study cannot show. One of important agendas in South Korea is about how we can deter North Korea.[26] As I showed in the statistics model, the increase of naval power makes war less likely between rivals. However, the temporal

26) Kim, Hong-Cheol. "How to Deter North Korea? Military Provocations". *The Korean Journal International Studies* 10, no. 1 (2012): 63-93.

domain is too broad and the indicator of naval power is too limited as a tonnage. For a more validated inference, the model need to use a narrower temporal domain and more sophisticated indicator for a naval power. By improving these research design data, I can draw more validated inference and provide a more useful implication to policy practitioners in South Koreas as well as ROK Military and Navy.

Reference

Aliaga, Martha, and Brenda Gunderson. *Interactive statistics*. Prentice Hall, 2000.

Bae, Hackyoung. *War and Rivalry: Political Shock and Bargaining*. Ph.D. Diss. Florida State University, 2014.

Crisher, Brian Benjamin, and Mark Souva. "Power Sea: A Naval Power Dataset, 1865-2011". *International Interactions*, Just accepted, 2014.

Correlates of War Project. State System Membership List, v. 2008.1, 2008. Available at http://correlatesofwar.org.

Ganguly, Sumit, and William Thompson. *Asian Rivalries Conflict, Escalation, Limitations Two Level Games*. Stanford University Press, 2011.

Gleditsch, Kristian. "A Revised List of Wars Independent States, 1816-2002". *International Interactions* 30, no. 3 (2004): 231-262.

Gochman, Charles S. "Interstate Metrics Conceptualizing, Operationalizing, Measuring Geographic Proximity States Congress Vienna". *International Interactions* 17, no. 1 (1991): 93-112.

Goertz, Gary, and Paul F Diehl. "The Initiation Termination Enduring Rivalries The Impact Political Shocks". *American Journal Political Science* (1995): 30-52.

Maoz, Zeev. "DyadicMID Dataset (version2.0)", 2005. Available at http://psfaculty.ucdavis.edu/zmaoz/dyadmid.html.

Polmar, Norman. *The Naval Institute Guide Ships Aircraft U.S. fleet*. Naval Institute Press, 2005.

———. *The Naval Institute Guide Soviet Navy*. Naval Institute Press, 1991.

Prezelin, Bernard. *The Naval Institute Guide Combat Fleets World*. Naval

Institute Press, 1995.

Sarkees, Meredith Reid, and Frank Wayman. *Resort War: 1816 – 2007*. CQ Press, 2010.

Signorino, Curtis S, and Jeffrey M Ritter. "Tau-b Not Tau-b: Measuring Similarity Foreign Policy Positions". *International Studies Quarterly* 43, no. 1 (1999): 115-144.

Singer, J. David. "The 'correlates war project': Interim Report Rationale". *World Politics* 24, no. 2 (1972): 243-270.

Singer, J. David. "Reconstructing Correlates War Dataset Material Capabilities States, 1816-1985". *International Interactions* 14 (1987).

Vasquez, John, and Marie T. Henehan. "Territorial Disputes and Probability War, 1816-1992". *Journal Peace Research,* 38, no. 2 (2001): 123-138.

요 약

해군력의 정량화와 함의

배 학 영 *

이 논문의 목적은 해군력 개량화를 소개하고 그 활용에 대하여 제안함에 있다. 어떻게 하면 여러 국가 간의 다양한 분쟁에 대한 해군력의 효과를 효과적으로 이해할 수 있을까? 혹은, 어떻게 하면 다양한 해군력의 나라별, 시간별 변화를 이해를 할 수 있을까? 지금까지 많은 학자들이 해군력의 변화와 그 변화에 따른 해군력이 분쟁에 미치는 영향을 규명하려고 많은 노력을 해왔다. 그 중의 한 방법이 정성적인 방법이나 아직 정량적인 시도는 매우 적다. 이 글은 해군력을 정량화하는 방법과 그 데이터를 이용하여 여러 기존 이론을 검증하고 여러 다른 연구주제를 연구하는데 어떻게 이용이 될 것인지를 소개를 하는 글이다.

본 논문의 주요 쟁점은 다음과 같다. 첫째, 계량화적 접근이란 무엇인가에 대해 논의 해 본다. 계량화란 무엇이며 정성적인 방법과의 차이는 무엇인지를 통해 정량화의 이용 가치에 대해 논의해 본다. 둘째, 해군력의 정량화이다. 해군력의 정량화를 위해 어떠한 기준들을 세우고, 그 기준에 따라 함정들을 코딩하고 톤수를 세는 과정을 설명한다. 셋째, 정량화된 해군력을 바탕으로 동북아시아 국가들의 해군력 변화를 서술적으로 분석한다. 이제 주어진 해군력 데이터(주요 함정의 톤수)를 가지고 각 동북아 국가별 시간별로 어떠한 변화를 거쳐 왔고, 각 분쟁들 (1,2차 세계대전 등)에는 어떠한 상관관계가 있는지를 단순 통계적 방법을 이용하여 알아본다. 넷째, 해군력의 변화가 경쟁국가 간의 전쟁 발발에 있어서 어떤 영향을 미치는 지에 대하여 통계적인 방법을 이용하여 검증해 본다. 묘사적인 방법은 다른 요소들에 대한 통제가 이루어 지지 않아, 정확히 해군력과 경쟁국가 간의 전쟁에 대한 인과적인 관계를 증명하기에는 한

* 해군소령, 해군본부

계가 있다. 따라서, 다른 경쟁적 이론들을 (예를 들어 민주평화론 등) 통제하여 해군력이 숙적국가 간의 전쟁 발발에 미치는 영향을 검증하였다. 상호 해군력의 증가는 경쟁국가 간에는 전쟁을 덜 일으키는 요인으로 작용하였으며, 이는 해군력이 경쟁국가 간에는 억제력이 있다고 추론 할 수 있다.

궁극적으로 해군력의 영향에 대한 정량적인 접근은 기존 연구의 검증, 미래 예측, 국가의 정책결정자들에게 보다 신뢰가 가는 자료를 제공하는 장점들이 있다. 이러한 장점들을 바탕으로 해군력의 영향에 대한 연구는 분쟁분야에 있어서 학술적이나 실용적인 측면에서 많은 이점이 있다.

키워드 : 정량화, 해군력 건설, 해군력 평가, 해군전략, 체계분석

투고일 : 2014년 5월 7일 | 심사일 : 2014년 7월 3일 | 심사완료일 : 2014년 7월 17일