Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 7-8, 2015

Uncertainties in Nuclear Proliferation Modeling

Chul Min Kim, Man-Sung Yim*, Hyeon Seok Park
Korea Advanced Institute of Science and Technology
**Corresponding Author: msyim@kaist.ac.kr*

1. Introduction

There have been various efforts in the research community to understand the determinants of nuclear proliferation and develop quantitative tools to predict nuclear proliferation events. Such systematic approaches have shown the possibility to provide warning for the international community to prevent nuclear proliferation activities. However, there are still large debates for the robustness of the actual effect of determinants and projection results. Some studies have shown that several factors can cause uncertainties in previous quantitative nuclear proliferation modeling works [5,6]. This paper analyzes the uncertainties in the past approaches and suggests future works in the view of proliferation history, analysis methods, and variable selection.

2. Problems of Nuclear Proliferation Modeling

Previous models considered various determinants to project historical nuclear proliferation activities by using statistical analysis methods [1,2,3,4]. Singh and Way [1] defined four levels of proliferation history.

Table I: Four Levels of Nuclear Proliferation

Level	Name	Description
0	No interest	No proliferation attempts
1	Exploration	Country considered nuclear weapons and conducted some exploratory work
2	Pursuit	Country started a nuclear weapons development program
3	Acquisition	First explosion/assembly of nuclear weapon

They analyzed the proliferation history data of all countries between 1945 and 2000. To calculate the proliferation risk of each country-year, various determinants are analyzed as independent variables while the levelized historical nuclear weapons development data is used as dependent variables. The coefficient and variance of each determinant is calculated based on the regression and survival analysis. After this, efforts have been made to determine the cause of nuclear weapons proliferation for 10 years. Various determinants have been analyzed in this process.

3. Sources of Uncertainty in Previous Modeling

After 2009, some arguments have been made for several problems in current modeling. Montgomery and Sagan pointed out several problems of current modeling methods: various history coding, overlooked bureaucratic power, poor nonproliferation regime, variables are chosen for convenience, etc. [5] Sagan pointed out additional problems in his paper in 2011: inaccurate capability data, very small number of successful proliferation cases, etc. [6]

Although recent study of nuclear proliferation model tries to accept Sagan's argument and increase their robustness, some problems still remain. First, the proliferation history coding rule cannot distinguish among different proliferation levels. Second, projection results of previous modeling largely depend on the form of independent variables and history coding rule. Third, the effectiveness of the determinants varies when different variable set is used for analysis.

This paper tries to categorize current problems into three topics: proliferation history data, analysis methods and variable selection. In each category, uncertainty of the projection result of nuclear proliferation risk is analyzed.

3.1 Uncertainties of Proliferation History Data

Each levels of nuclear program history variable are coded as 1 if the country achieved the level and 0 if not. However, there has been some difference between researchers. From the first proliferation dataset from Singh and Way [1], Jo and Gartzke [2], and Bleek [3] made their own proliferation datasets. However recent studies perform analysis with various history data and argue that their target determinants are robust [5,6]. We analyzed the projected risk using Singh and Way's dataset, and Bleek's dataset to compare the coefficients of the determinants. Different history coding changes the effect of proliferation determinants, but does not change the proliferation risk projection result of countries. The coefficients changed less than 30% while the effect of all determinants did not change.

3.2 Uncertainties of Analysis Methods

Regression analysis and event history analysis (also known as survival analysis) have been used for the analysis method for nuclear proliferation modeling. In regression analysis, logistic and Rare Events Logistic

(RELogit) regression have been used for proliferation modeling. They have an advantage in reducing the number of blank data points and projecting the years after the achievement of each level. However, the result is largely dependent on the form of proliferation history variables.

In event history analysis, Weibull and Cox method have been used. The definition of survival analysis fits well with analyzing the causes of nuclear proliferation since the data "exits" the analysis when it achieves each level and the data is censored if it does not have nuclear proliferation attempt. However, this characteristic decreases the number of observations and increases statistical error. Time independent nature of survival analysis also limits the flexibility of the model. It only considers the time used to achieve the level rather than considering the changing nature of the determinants.

In addition, two common things have been recognized as the major problem of current analysis methods. First, current analysis methods assume that the influence of the determinants is constant over time. This causes significant uncertainty of the coefficient of the determinants. Second, there are only small numbers of successful proliferation cases. Only 23~30 states explored, 15~16 states pursued, and 9 states acquired nuclear weapons in various proliferation history data.

To analyze the uncertainty caused by the problems above, we compare the projection results of various time range and target countries based on different forms of proliferation history coding rules. NPT (Treaty on the Non-Proliferation of Nuclear Weapons) has been opened for signature in 1968, and it was the most important event for nuclear nonproliferation after 1945. Therefore, we analyzed the proliferation risk for three time ranges: 1945-2000, 1945-1967, 1968-2000.

When the target year changed, the influence of variables related to domestic politics changed while the influence of international politics variables and capacity variables were changed less than 30%. This result suggests that further time dependent study should focus on the influence of domestic politics determinants including leader's characteristics and political system of a state.

In terms of proliferation history coding rules, only regression analysis methods are influenced since event history analysis automatically "exits" after the achievement. We check three possible variations: give "1" after achieving the level, give "1" for current level while give "blank" for lower levels, and give "blank" after achieving the level. First rule tries to explain current nuclear program development state of a country. Second rule is basically same with the first rule, but it only allows a state be a certain level at a moment. Third rule inhibits "annually repeated proliferation" because all "1"s is recognized as proliferation event in regression analysis.

There is a significant difference in the projection result when different history coding rule is used. The influence of capacity variables including GDP, electricity is changed since the number of explored, pursued, or acquired state has huge difference. The result from different coding rule leads at least one determinant to have opposite influence compared with survival analysis cases.

3.3 Uncertainties of Variable Selection

Various variable set has been developed and expanded to add robustness of nuclear proliferation models [7]. They can be categorized into three types: domestic politics, international politics, and capability. In current model, the democracy score of the country represents the decision for nuclear programs. However, it is not an appropriate tool since the leader's characteristics should be considered with political structure of a country. The capability category includes only economic variables and does not consider the nuclear latency, while the nuclear decision is made based on both economic and nuclear capabilities. Also, correlations between variables of different categories or among same categories increase the uncertainties for the influence of proliferation determinants.

There could be some combinations of variables for the determinants. As Sagan argued, no democracy state with NPT ratification initiated or developed nuclear program. Current variable set does not consider the possibility of combined variables and often make counter-intuitive conclusions. For example, democracy score and NPT ratification variable are often regarded as ineffective, or even have positive effect on nuclear programs.

4. Suggestion for Future Works

The research community still lacks the knowledge for the source of uncertainty in current models. Fundamental problems in modeling will remain even other advanced modeling method is developed. Before starting to develop fancy model based on the "time-dependent proliferation determinants" hypothesis, using graph theory, etc., it is important to analyze the uncertainty of current model to solve the fundamental problems of nuclear proliferation modeling.

The uncertainty from different proliferation history coding is small. Serious problems are from limited analysis methods and correlation among the variables. Problems in regression analysis and survival analysis cause huge uncertainties when using the same dataset, which decreases the robustness of the result. Inaccurate variables for nuclear proliferation also increase the uncertainty. To overcome these problems, further quantitative research should focus on analyzing the knowledge suggested on the qualitative nuclear proliferation studies.

Previous models shared fundamental assumption: future nuclear proliferation events could be predicted if the determinants are well analyzed based on historical events. However, this assumption has fundamental limitations since they do not consider the importance of determinants can change over time. Therefore, "predicting nuclear proliferation" is hard to be realized based on current methods even if we have accurate proliferation history data or proliferation determinants. Further efforts should be made based on the consideration of limitations and uncertainties of current modeling methods.

REFERENCES

[1] S. Singh, C. R. Way, The Correlates of Nuclear Proliferation: a quantitative test, Journal of Conflict Resolution, Vol.48, No.6, p. 859-885, 2004.

[2] D-J Jo and E. Gartzke, Determinants of Nuclear Weapons Proliferation, Journal of Conflict Resolution, Vol.51, No.1, p. 167-194, 2007.

[3] Jun Li, Man-Sung Yim, David N. McNelis, Model-based calculations of the probability of a country's nuclear proliferation decisions, Progress in Nuclear Energy, Vol. 52, No. 8, p. 789-808, 2010.

[4] Philipp C. Bleek, Why Do States Proliferate? Quantitative Analysis of the Exploration, Pursuit, and Acquisition of Nuclear Weapons, Chap. 8 in Forecasting Nuclear Proliferation in the 21st Century: The Role of Theory, Volume 1. Stanford, Calif.: Stanford University Press, 2010.

[5] C. R. Way and J. Weeks, Making it Personal: Regime Type and Nuclear Proliferation, American Journal of Political Science, Vol. 58, No. 3, p. 705–719, 2014.

[6] P. C. Bleek and E. B. Lorber, Security Guarantees and Allied Nuclear Proliferation, Journal of Conflict Resolution, Vol. 58, No. 3, p. 429–454, 2014.