Greek – Turkish Arms Race: An Approach Using Neural Networks

Spyridon Ntellis

Master's Thesis Defence
Master's Program in Accounting & Finance
Supervisor: Prof. Achilleas Zapranis

Rhodes, 27 January 2022

Outline

1. Introduction
2. Literature Review and Background
3. Methodology
4. Results
5. Discussion
6. Conclusions

1. Introduction

- ✓ Researchers commonly portray Greece and Turkey as a prime example of an arms race relationship due to their tense relations and their level of military expenditure.
- ✓ Research on arms races has ultimately concerned itself with whether arms races could lead to war.
- ✓ There has been an international expansion of defence spending in recent years: Global spending in 2019 was estimated at about **2 trillion USD** (const. 2018) or **5.27 bn USD per day**.
- ✓ The issue of defence spending in Greece was brought under the spotlight after the country's recent economic crisis.

1. Introduction

Subject of the thesis

- ✓ This thesis revisits arms race considerations in the context of Greek – Turkish relations for years 1963–2018.
- ✓ Greek – Turkish defence expenditure dynamics are analysed through the use of artificial neural networks and state-of-the-art techniques for model estimation and input significance evaluation.

Aim

To provide researchers with a solid framework for the study of arms races in the specific case of Greece and Turkey, shedding light on oft-overlooked —yet critical— methodological issues.

1. Introduction

Objectives

The objective of the thesis was three-fold:

- **To establish whether there is an arms race** between Greece and Turkey and provide Greek financial administrators and decision-makers with valuable insights into the interaction of Greek–Turkish armaments
- **To regenerate interest** in the use of machine learning to examine defence spending dynamics (where conventional statistics prevail)
- **To highlight substantial issues** involved in designing, executing and evaluating arms race research using neural networks

2. Literature Review and Background

Greek – Turkish relations

- ✓ Greek – Turkish relations have always been tense, despite temporary periods of cordiality.
- ✓ Relations deteriorated significantly in the aftermath of the Turkish invasion of Cyprus in 1974. A war between the two countries was only narrowly averted in 1987 and 1996.
- ✓ Relations are entangled by a series of **unresolved disputes**:
  - The Cyprus Issue
  - The breadth of territorial waters in the Aegean Sea
  - The extent of EEZ, illegal immigration, etc.

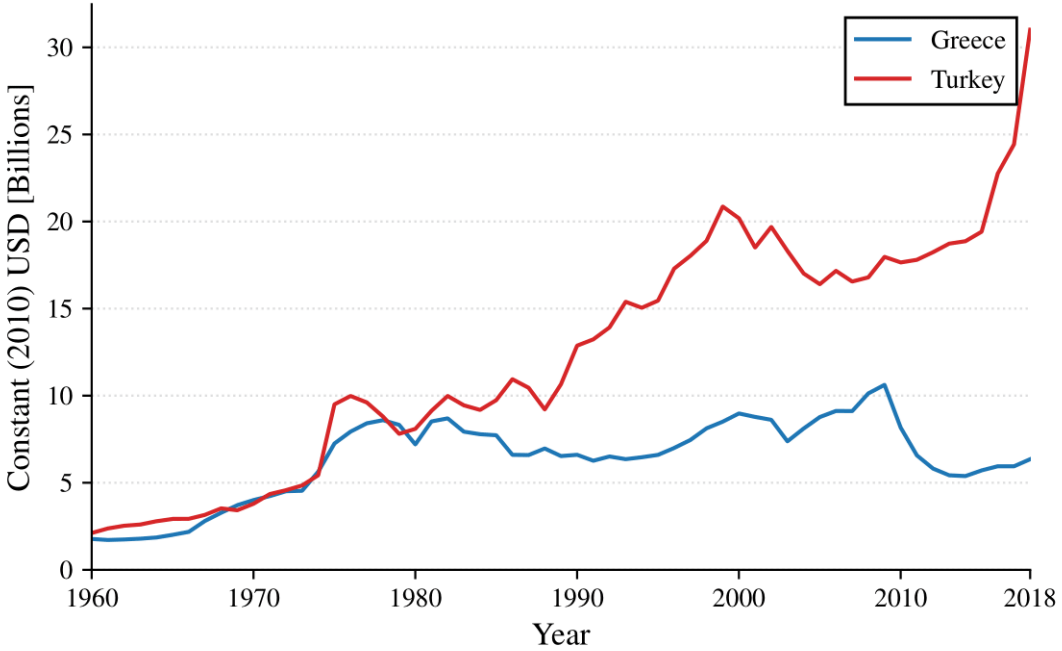
2. Literature Review and Background

Greek – Turkish relations

- ✓ The Greek–Turkish conflict is argued to be an **identity-based conflict**, stemming from both countries' collective identities and national historical narratives, which are built upon a vilification of one another (see e.g. Heraclides, 2012).
- ✓ Incapable and populist leaders as well as the reproduction of conflict through education serve to exacerbate disputes.

2. Literature Review and Background

Greek – Turkish annual defence expenditure (level data)

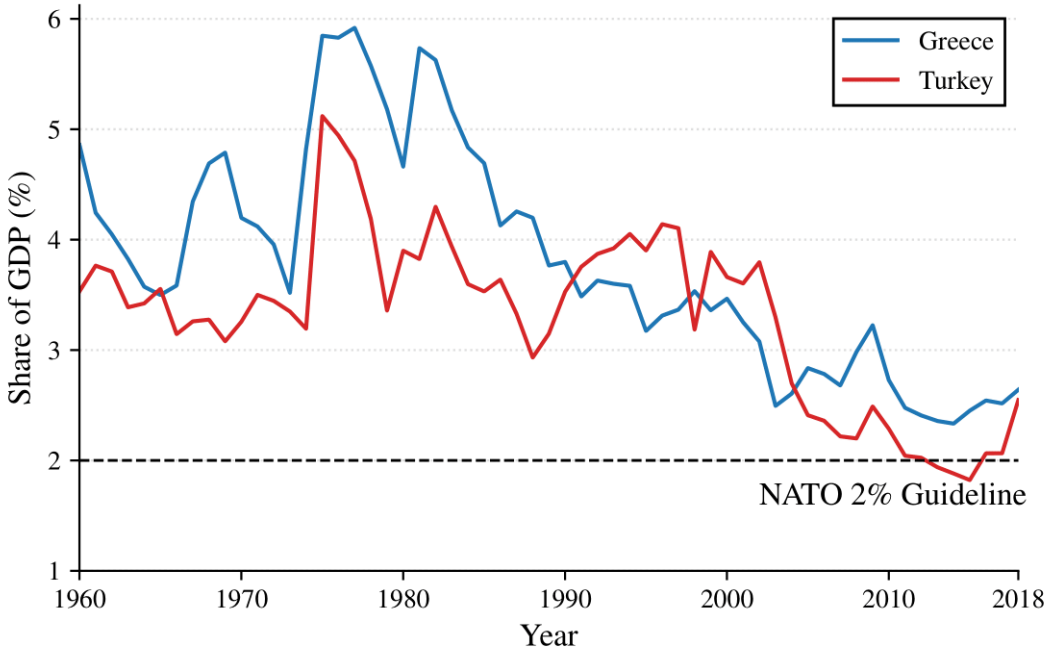


Avg. annual change of spending for period 1963 – 2018:
Greece: 2.7%
Turkey: 5.3%

Source: Author's calculations based on Stockholm International Peace Research Institute [SIPRI] (2020)

2. Literature Review and Background

Greek – Turkish annual defence expenditure (share data)



Source: SIPRI (2020). The 2% guideline reflects security considerations made in 2006 and *is not* indicative of past conditions.

2. Literature Review and Background

The notion of arms races

- ✓ Military arms races bear a mostly negative connotation, since they are commonly associated with conflict and war.
- ✓ Huntington (1958) argues that they have become a distinctive phenomenon only since the industrial revolution.
- ✓ They are a key issue lying at the root of **strategic studies** and, as such, they are closely linked to the notion of national security.
- ✓ Arms races may be either **quantitative** (absolute armament levels) or **qualitative** (armament characteristics – technological advances).

2. Literature Review and Background

Defining an ‘arms race’

- ‘a progressive, competitive peacetime increase in armaments by two states or coalitions of states resulting from conflicting purpose or mutual fears’ (Huntington, 1958)
- ‘two or more parties perceiving themselves to be in an adversary relationship, who are increasing or improving their armaments at a **rapid** rate and structuring their respective military postures with a **general** attention to the past, current, and anticipated military and political behaviour of the other parties’ (Gray, 1971)
- ✓ Researchers have criticized the **ambiguity** of proposed definitions

2. Literature Review and Background

Quantitative models of arms races

- ✓ Two types of models are prominent in early literature: **action–reaction** and **domestic structure** models.
- ✓ Complex models that incorporate factors from both types of models prevail in recent research.

Richardson's model

- ✓ Richardson's model, which was published posthumously (1960), is one of the oldest and the most influential model of arms races.
- ✓ Arms races are modelled using a system of **linear differential equations** in an action–reaction context.

2. Literature Review and Background

Beyond Richardson: alternative models and results

- ✓ Empirical applications of Richardson's model have not been satisfactory.
- ✓ Researchers adopted several interpretations of Richardson's equations and proposed **new models** in an attempt to overcome shortcomings.
- ✓ Research has predominantly been based on traditional **econometrics**.
- ✓ Given the poor performance of action–reaction models, recent research focusses on answering the arms race question within the wider context of the **determinants of military expenditure**.
- ✓ Results in most cases have been **inconclusive**.

2. Literature Review and Background

Statistical considerations

- ✓ Questionable validity of models due to serial correlation and the large number of parameters.
- ✓ Concerns about excessive data manipulation
- ✓ Overlooked assumptions of methods used
- ✓ ‘Good measures of fit do not necessarily imply the right model’
- ✓ Drawing specific boundaries between the effects of explanatory variables is not always possible, as these effects may be indirectly influenced by several complex interactions.

2. Literature Review and Background

Choosing representative and reliable data

- ✓ Changing data sources may produce a substantial discrepancy in the results obtained — even between revisions of the same source.
- ✓ Institutions reporting military expenditure depend on government data beyond their control.
- ✓ Corrections for inflation, exchange rates, etc.
- ✓ Data reported by countries may intentionally be inaccurate in order to mislead rivals or for domestic reasons

2. Literature Review and Background

Choosing a suitable proxy of armaments

- ✓ There is a wide consensus among researchers that **defence expenditure** is the best available proxy of armaments.
- ✓ However, it is unclear whether levels, logarithms, the change or share of military expenditure out of GDP are more suitable measures.
- ✓ The use of aggregate expenditure data is problematic in cases when non-correlated races occur simultaneously.

Wider considerations

Researchers who have attempted to summarize arms race literature seem to be dissatisfied by the approaches adopted and the knowledge gained.

2. Literature Review and Background

Using neural networks for arms race research

- ✓ Neural networks have only **scarcely** been adopted in literature as an approach to arms race research.
- ✓ All relevant papers are related to the Greek – Turkish conflict.
- ✓ In all papers, factors related to Turkey emerge as **prime determinants** of Greek spending.

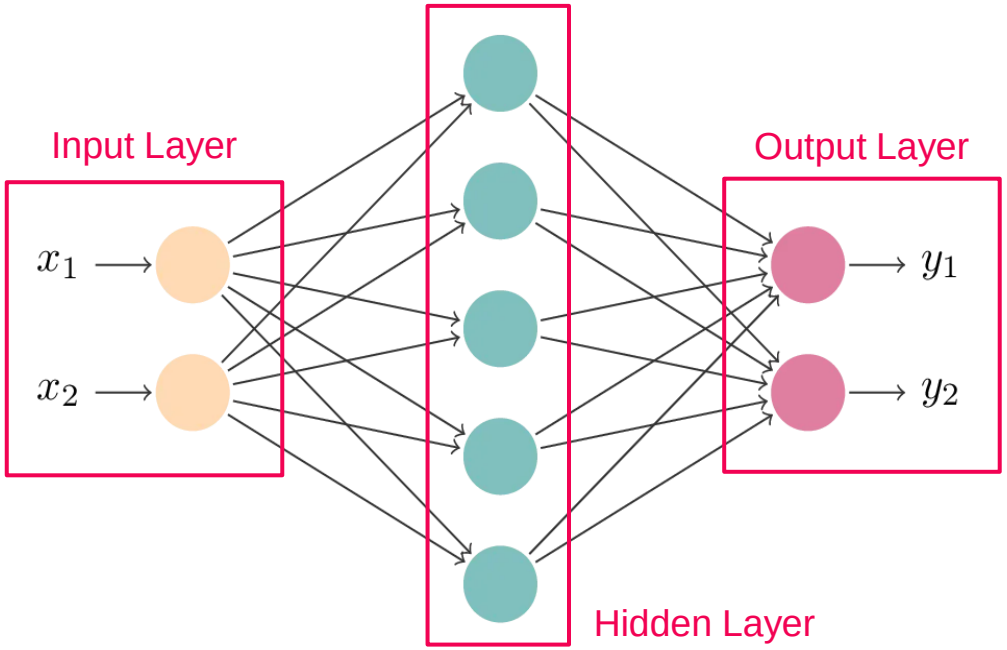
2. Literature Review and Background

Neural networks

- ✓ Artificial neural networks are a **machine learning method** inspired by the functioning of biological neurons, originally formalized by McCulloch and Pitts (1943).
- ✓ They have enjoyed great popularity in recent years due to a growing interest in data analysis.
- ✓ Their applications extend over a wide range of disciplines.
- ✓ They are proved to be **universal approximators** of functions and their derivatives, requiring almost no a priori assumptions about the underlying process being modelled.

2. Literature Review and Background

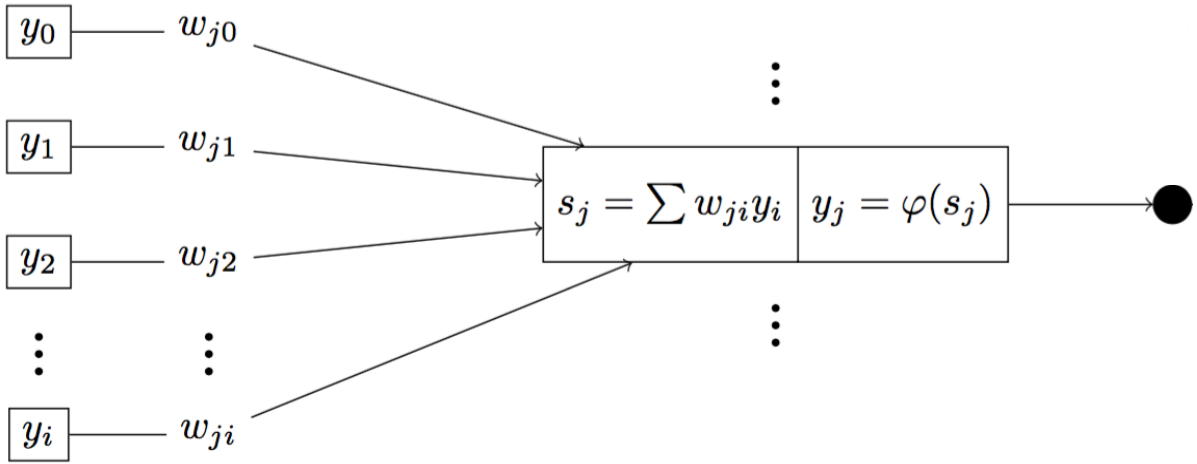
Example of a 2–5–2 neural network



Source: Adapted from latexdraw.com

2. Literature Review and Background

Signal flow inside a neuron



Source: Adapted from newbiettn.github.io

2. Literature Review and Background

Types of neural networks

- ✓ Modern neural networks appear in a wide **variety of topologies**: Convolutional NNs, Recurrent NNs, Wavelet NNs, ensembles, etc.

Learning

- ✓ Learning in neural networks occurs through an iterative process of **readjusting the weights** until the defined cost function is minimized.
- ✓ **Categories**: Supervised, unsupervised, reinforcement, semi-supervised
- ✓ The most popular learning algorithm for supervised learning tasks is **backpropagation of errors**.
- ✓ A significant concern when learning is **overfitting**.

2. Literature Review and Background

SHAP values

- ✓ SHAP (short for **Shapley Additive ExPlanations**) is a novel model-agnostic, perturbation-based method for input significance estimation, with a theoretical foundation in game theory (Shapley values).
- ✓ SHAP was proposed by Lundberg and Lee (2017).
- ✓ The main idea is to **calculate the contribution** of each input value to the prediction made by the underlying model in a similar way to Shapley values, which determine the contribution of each agent in a game.
- ✓ SHAP values demonstrate how to get to the actual output starting from a base value, which would be predicted if no features were known.

2. Literature Review and Background

SHAP values

- ✓ The exact calculation of SHAP values is **time-demanding** due to averaging over all possible orders of input features.
- ✓ Lundberg and Lee (2017) have proposed approximation methods, either model-agnostic or model-specific, in order to accelerate calculations. The most prominent method is **Kernel SHAP**.
- ✓ **Advantages** of SHAP: Theoretical foundation in game theory, measures of global and local interpretability, implementation in Python
- ✓ **Disadvantages** of SHAP: Feature independence, considerable computation time, indiscriminate use

3. Methodology

Process summary

A model to **predict armaments** is designed and trained on available data, until an acceptable performance is obtained. **Input significance analysis** is then performed on the trained model to establish whether an arms race exists.

Proposed definition of arms races

An arms race is defined as a specific pattern of arms acquisition, where the armaments level of each country **is primarily explained** by factors that refer to the military capabilities of its rival.

Input Variables

MODEL	VAR	DESCRIPTION	NOTES
Model A	A1	Turkish military expenditure	Based on Richardson (1960)
	A2	Greek military expenditure	
Model B	B1	Ratio of Greek to Turkish armed forces personnel	Based on original neural networks approach (Refenes et al., 1995)
	B2	Greek military expenditure per person in the armed forces	
	B3	Turkish military expenditure per person in the armed forces	
	B4	Greek military expenditure	
	B5	Turkish military expenditure	
Model C	C1	Turkish military expenditure	Self-developed model — Variables are chosen on an ad hoc basis
	C2	Lagged change of Turkish military expenditure	
	C3	Lagged change of Greek military expenditure	
	C4	Aggr. change of Greek public debt over the previous 2 yrs	
	C5	Aggr. change of Greek gross capital formation over the previous 2 yrs	
	C6	Lagged change of Greek imports	

Output variable: Change in defence expenditure

3. Methodology

Data reliability

Data were gathered from readily available and **reputable sources**, such as the World Bank and the IMF. Defence expenditure data were drawn from the **SIPRI database**, which is reportedly considered a gold standard.

Model selection and data preparation

- The model selection process involved starting with the simplest class of models (1 hidden layer and 1 hidden unit) and incrementing units; the model with the **least associated validation loss** was chosen.
- Data were **split** into train – validation – test sets.
- Data were **scaled** by subtracting the median and scaling based on IQR.

3. Methodology

Basic model parameters

- ✓ Fully interconnected multilayer feedforward networks were used.
- ✓ Error measure: **Mean Squared Error (MSE)**.
- ✓ Model topologies and parameters:

Model	Units	Activation	Optimiser	Epochs
Model A	2–8–1	Sigmoid–Tanh	Adam	400
Model B	5–8–1	Sigmoid–Tanh	Adam	300
Model C	6–7–1	Sigmoid–Tanh	Adam	700

3. Methodology

Input significance analysis

- ✓ Input significance analysis is performed using **Kernel SHAP**.
- ✓ Since Kernel SHAP assumes feature independence, **correlation analysis** (using **Pearson's** and **distance correlation** coefficient) was performed before evaluating input significance.

Evaluation of predictive power

- ✓ Evaluation was performed using Root Mean Squared Error (**RMSE**), Mean Absolute Error (**MAE**) and Coefficient of Determination ($\mathbf{R^2}$).
- ✓ Three **baseline models** were used as benchmarks.

4. Results

Test-set results of all models

Model	Performance metric		
	RMSE	MAE	R^2
Best baseline ($\Delta Y_{t+1} = \Delta Y_t$)	0.1008	0.0766	0.0442
Model A	0.1025	0.0864	0.0106
Model B	0.0957	0.0707	0.1373
Model C	**0.0798**	**0.0629**	**0.4006**

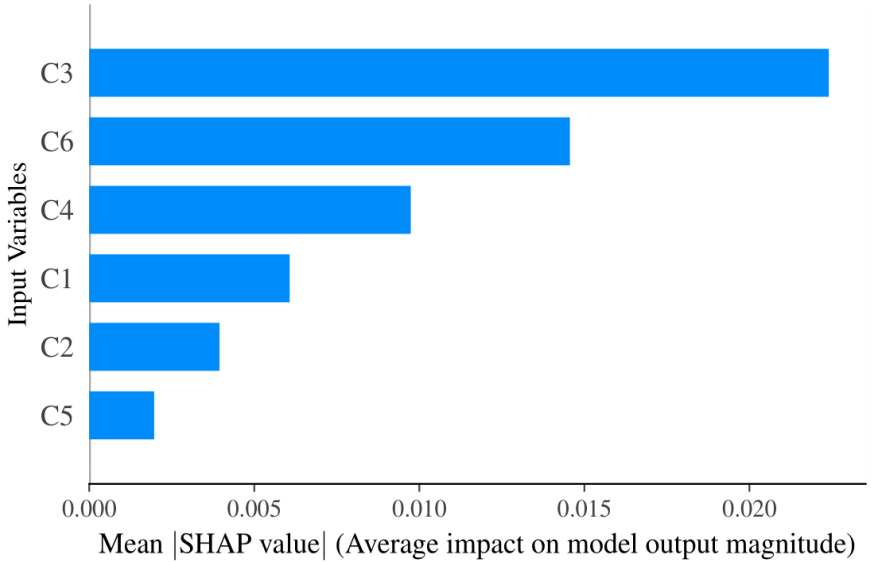
4. Results

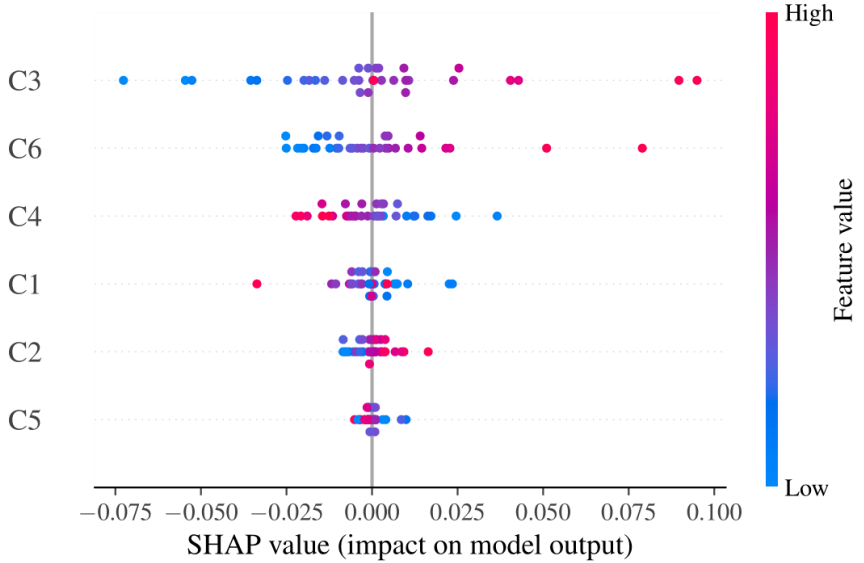
Correlations

- ✓ **Model A**: Both Pearson's and distance correlation coefficients are statistically significant at any reasonable level of significance.
- ✓ **Model B**: 8 of 10 Pearson's coefficients are statistically significant at an 8% level. Distance correlation coefficients are all statistically significant at an 8% level.
- ✓ **Model C**: 7 out of 15 Pearson's coefficients and 5 out of 15 distance correlation coefficients are statistically significant at an 8% level.
- ✓ Since Model C exhibits notably reduced correlation issues, **input significance analysis was performed only on Model C**.

4. Results

Input significance analysis (SHAP values)





5. Discussion

Interpretation of results — Comparison to previous research

- ✓ Among Models A, B and C, it is **only Model C** that exhibits a performance that is substantially superior to that of the baseline models.
- ✓ Given the SHAP values of Model C, it appears that the change in Greek defence expenditure is best predicted by internal factors, which suggests that Greece **does not engage in an arms race** with Turkey.
- ✓ This conclusion is **in contrast** to previous research.
- ✓ The R^2 value of Model C (40%) seems notably inferior to that reported in Refenes et al. (1995) (85%). However, the 85% value was calculated on their combined training and test set.

5. Discussion

Considerations

- ✓ The **variance of the output** variable places a constraint on the optimal level of error that can be achieved by any model and should raise suspicion against seemingly perfect predictions.
- ✓ When evaluating the predictive power of Model C, the RMSE of 7.47% (or 475 mln EUR for Greek spending in 2018) **may not be acceptable** depending on the task at hand.
- ✓ **Correlation issues** (particularly evident in Models A and B): Input significance measures are affected by correlation; previous research has not concerned itself with this effect.

5. Discussion

Considerations

- ✓ The issue of choosing appropriate **levels of significance**
- ✓ The issue of choosing **input variables** in the absence of a theoretical framework and given that ‘post hoc rationalization of one’s findings is very easy’ (Brauer, 2002).
- ✓ **Availability and reliability of data** — No reliable information about the estimation process used for historical data
- ✓ The **scarcity of data** due to their annual frequency poses statistical limitations; previous research using neural networks does not make use of a separate validation set.

5. Discussion

Considerations

- ✓ **Absence** of standardization and specification bias
- ✓ **Incorrect understanding** of SHAP: SHAP produces feature attributions that are true to the model, but not necessarily true to the data

Final remarks

- ✓ The statistical considerations outlined may be the **underlying reason** for the **contradictory results** reported by researchers.
- ✓ The **diminished interest** in arms race research since the end of Cold War is not conducive to the emergence of new approaches and to a critical approach to already existing ones.

6. Conclusions

- ✓ Arms races are a highly **complex issue**, especially given the fact that human actors and their decisions are involved.
- ✓ Greece apparently (given limitations) **does not engage** in an action–reaction relationship with Turkey, since the annual change in Greek defence spending is principally determined by internal factors.
- ✓ **Methodological and statistical issues** exert considerable influence on the results, but they have not been given proper attention by researchers. This explains why research has been inconclusive.
- ✓ Research is needed to establish algorithms that are able to **leverage small datasets**. Traditional econometrics should not be disregarded.

6. Conclusions

Recommendations for future research

- ✓ Incorporating social and psychological factors into arms race models
- ✓ The digitization of public administration enables researchers to obtain more detailed data on public expenditure
- ✓ Critical analysis of theoretical aspects, in the light of modern political science theories
- ✓ Development and use of methods for input significance analysis which are not affected by correlations between variables
- ✓ New methods that are capable of leveraging small datasets

Thank you for your attention!