

# Defense Spending Spillovers

Isaac Liu

May 6, 2021

## Idea

We examine spillovers/relationships between military spending among countries using the PCA/Factor methods and potentially the LASSO. Using PCA we can summarize the relationships in spending among countries or over time and possibly attach interpretations or labels to key factors. Using the LASSO, regressing an individual country's spending on that of other countries, we can see which countries impact or are impacted by others.

A hypothesis is that spending in neighboring countries increases local spending. Yet another hypothesis is that ideological differences/alignment are somehow related. We can also check to see if the US, Russia, or other countries have disproportionate influence. It's also possible that key factors change from period to period- post Cold War to Post 9/11, for example.

## Literature

In the literature, the main approach to examining military spending appears to be spatial autoregression: <https://www.sciencedirect.com/science/article/pii/S0176268017301581>. I haven't seen anything using machine learning methods, even though I flipped through 10 Google Scholar pages. Though, here's a triple LASSO involving military spending but the question is different (it's looking at the fiscal multiplier): [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3819192](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3819192).

Our approach will allow us to look at potentially both spatial and ideological factors. Machine learning methods will also allow us to take a unique approach of imposing sparsity and dimension reduction (a country's spending is only affected by a few others, or is explained by only a couple key factors).

This question presents some unique challenges as to dealing with pretty limited panel data: a small sample size with a lot of missing values, which provide space for unique methods.

## Military Expenditure Data

Our data on military expenditure as a percentage of GDP comes from SIPRI, the Stockholm International Peace Research Institute. There are several other series available, such as spending in absolute terms or as a share of the government budget. We have also download the Varieties of Democracy dataset, which contains information on hundreds of political variables and indicators for ethnic conflict and other measures, which could lead to heterogenous impacts on spending.

Figure 1: Military Expenditure as Percentage of GDP Time Series for Countries

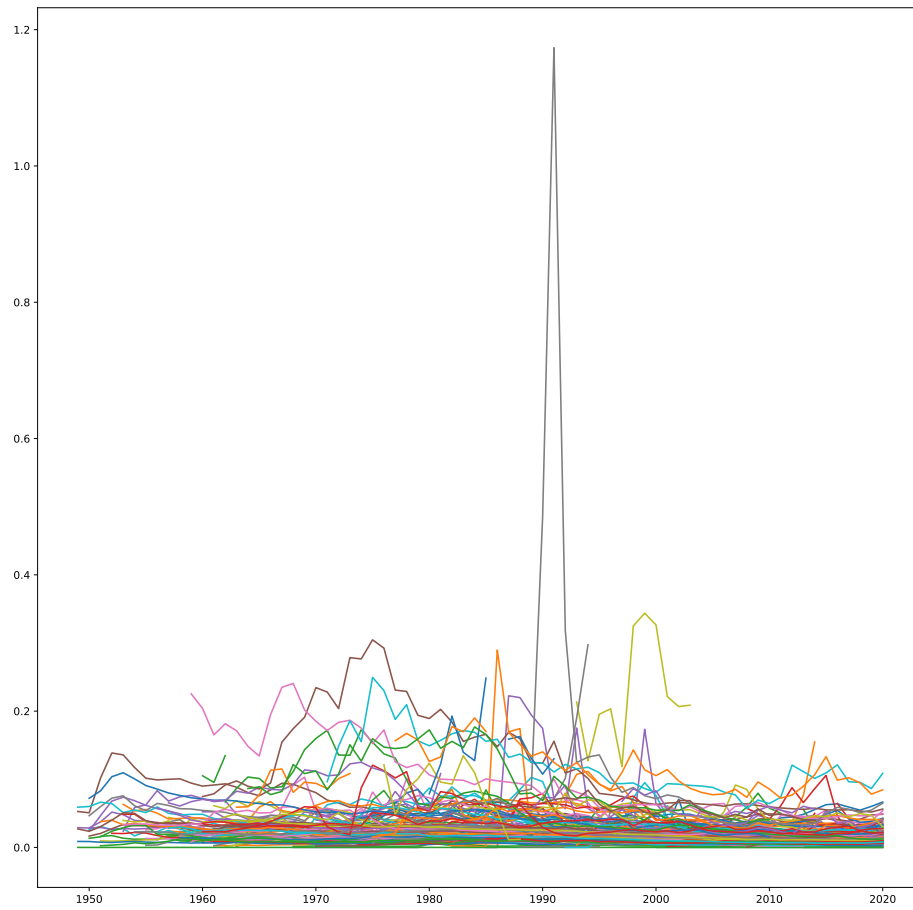
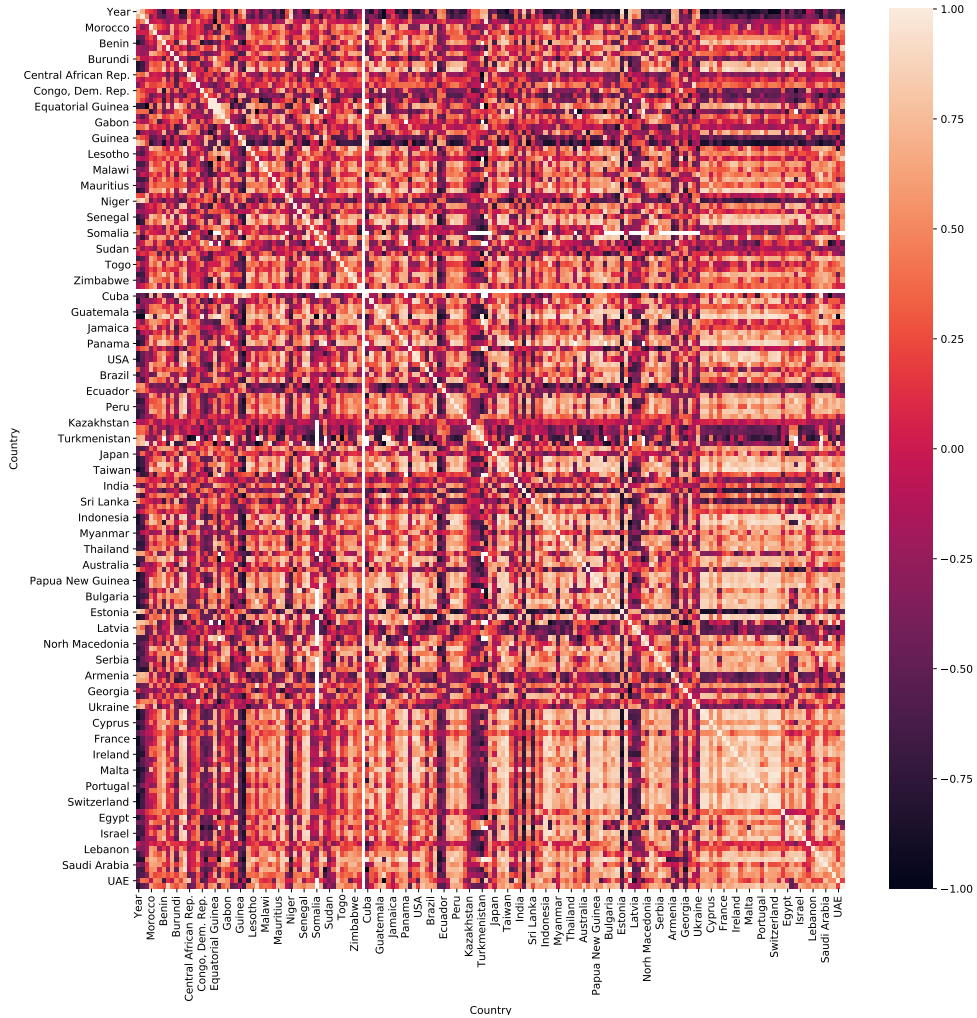


Figure 2: Correlations Between Country Military Spending



## The Missing Values Problem

In order to use PCA and the LASSO, we need to figure out a way to handle missing values in the panel. Here I interpolate values between data points where we have spending information for a country, and also restrict the sample to observations after 1995. Then I delete any countries which still have a missing value for any year. A lot more could probably be done to optimize the size of the

balanced panel; we could try to maximize the rectangular filled area in the charts below.

Figure 3: Missing Values Before Interpolation

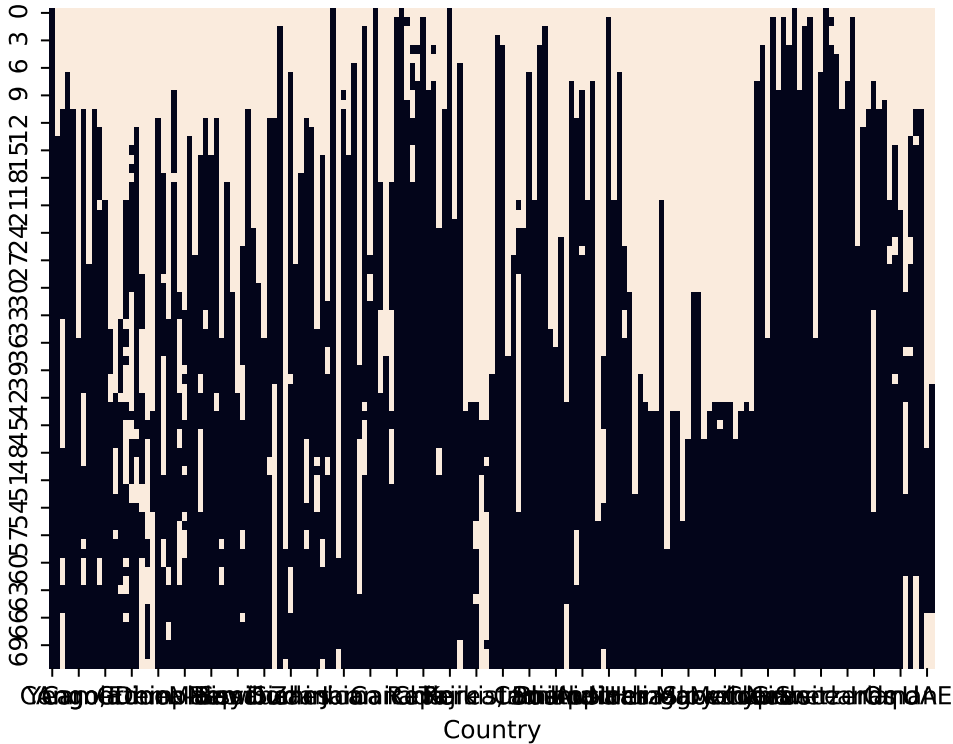
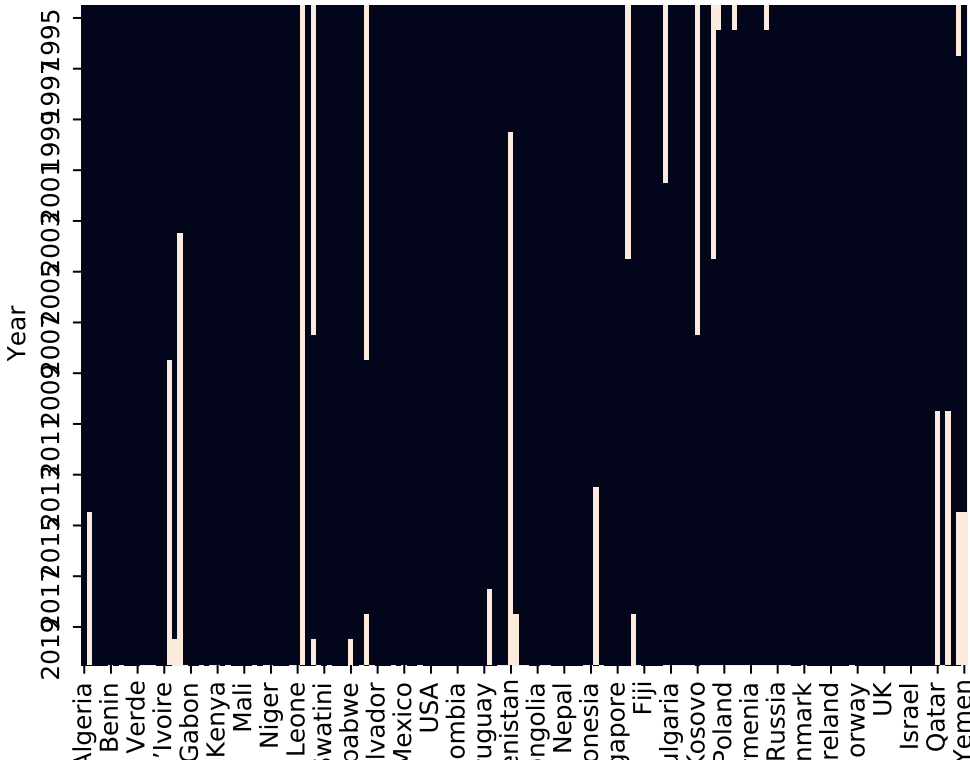


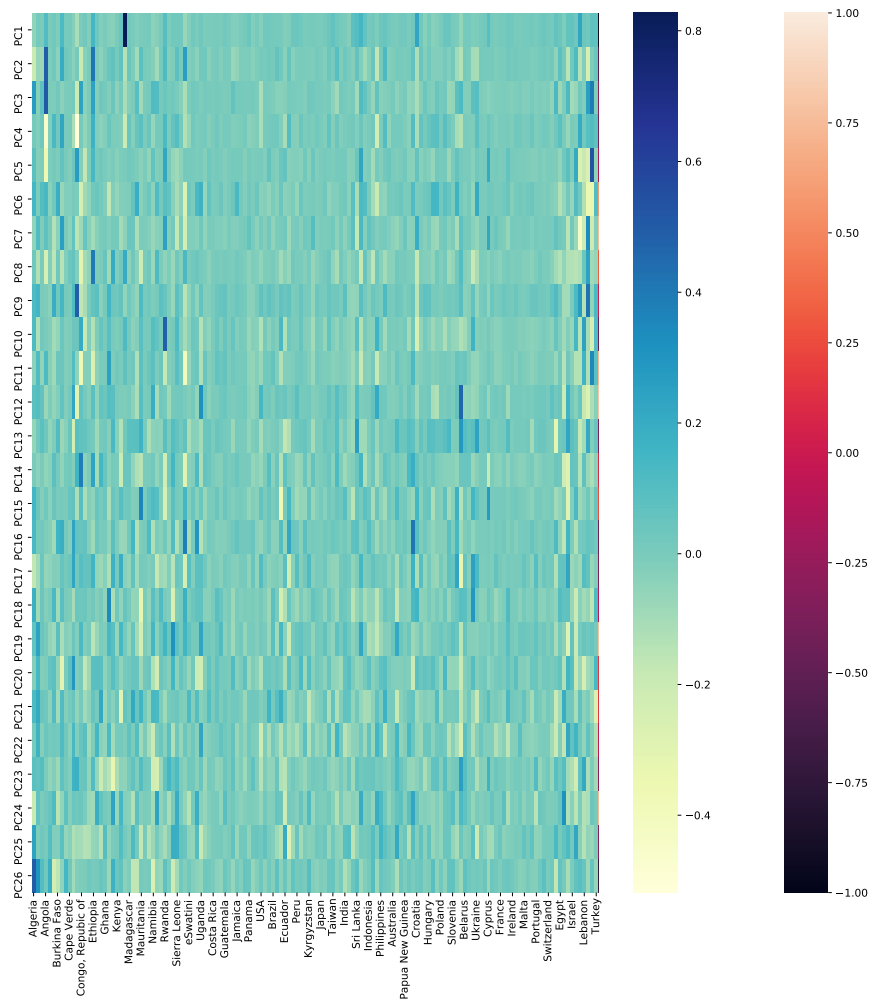
Figure 4: Missing Values After Interpolation and Sample Restriction



## PCA Results

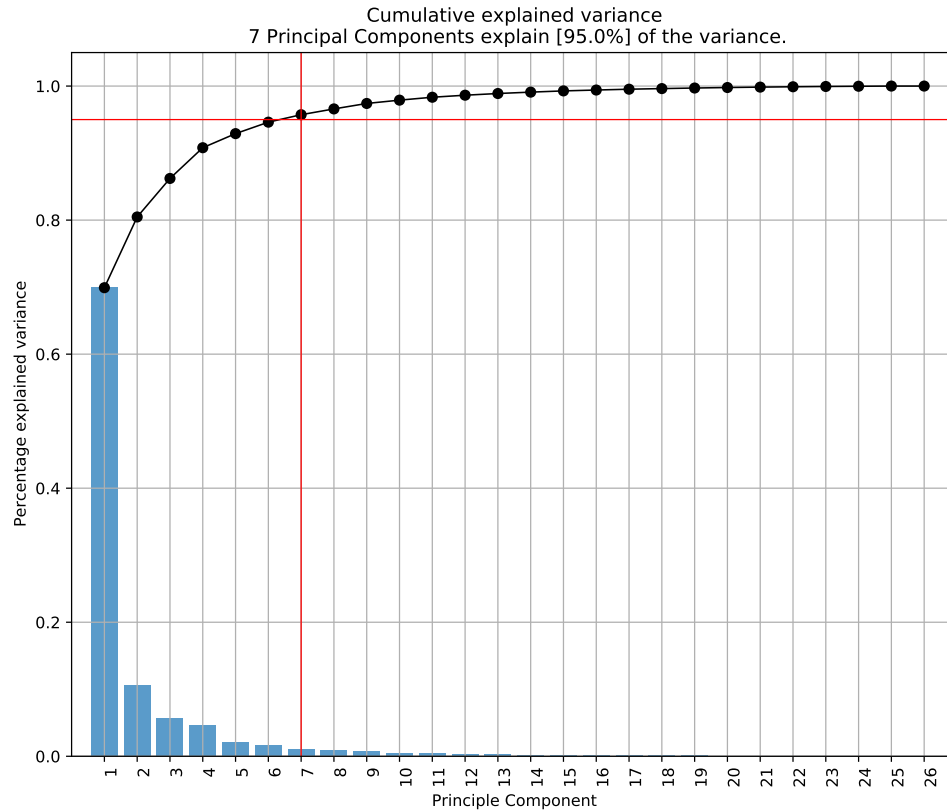
I performed a PCA to see if there were components which explained a large amount of the variation in spending across countries and over time.

Figure 5: Military Spending PCA Loadings



The data is very well summarized with very few principal components.

Figure 6: Military Spending PCA Share of Variance Explained



## LASSO Results

Doing the LASSO feels a bit strange, given that we don't really have a "different" dependent variable; we are merely regressing a country's spending on that of other countries. This means that we must impose a bunch of potentially dicey sample restrictions.

At the moment, I am running a LASSO regression for each individual country across time (so there is a single observation for each time period). The dependent variable is that country's spending per GDP and the independent variables (which may have penalized coefficients) are spending for all the other countries.

Here are the average coefficients across countries.

I am not terribly confident that these results are not just pure noise, because I do not see any clear patterns. The sample size restriction to about 25 observations seems too harsh, but necessary in



order to include a large amount of other country independent variables in every regression.

## **Extension: Textual Analysis**

If time allows, we could also look into a quick textual analysis of UN resolutions or NATO statements to validate any findings about the importance of ideological or geographic factors.