Deconstructing the Gun: Tracing Technologies through the Linkage Coefficient

Ethan Chu   Maryam Gooyabadi

Wesleyan University, Center for Study of Guns and Society, Quantitative Analysis Center

Introduction

One of the understudied aspects of firearm innovation is the role that patent classifications play in revealing areas of firearm technology that have experienced growth. Despite their potential, there has been no established metric or coefficient to quantify how innovation fluctuates across firearm patent technologies over time. Our analysis will provide a framework for understanding how patent technologies are linked together through the Linkage Coefficient. We then apply a time series–based approach to explain changes in the Linkage Coefficient over time in certain classifications, offering a new framework for understanding technological evolution within the firearm domain.

Classifications Codes and Firearm Technologies:

The master dataset contains 52,421 firearm patents that have classification codes revealing information about the technology behind the patent. Classification codes are of the form:

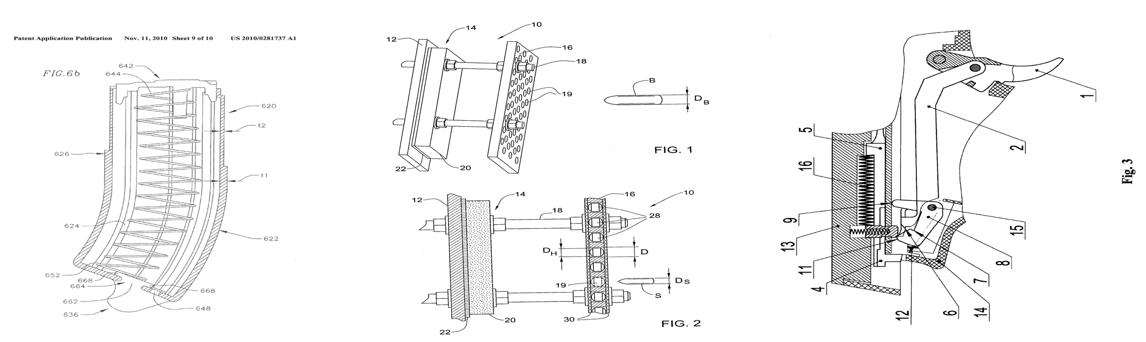
$$(\mathcal{L}_1)\text{-}(n_1)\text{-}(\mathcal{L}_2)\text{-}(n_2)/(n_3)$$

where $\mathcal{L}$ represents a letter $[A-H,Y]$ and n represents an integer.

Classification codes reveal the underlying technology behind patents, and firearm patents are all coded under **F41-$\mathcal{L}$-00/00**, indicating that they are weapons. Our main interest are patents which contain sub-classifications–describing specific technological features behind the patent. The **sub-classification reveals the underlying technology behind the firearm** that guides us in the analysis behind this study. Below are a few examples of patents with sub-classification codes.

Table 1. Classification Codes and Sample Data

Classification	Description	Example Patent
F41A9/62	Feeding & Loading Ammunition Devices	(a) US8839543B2
F41H5/013	Armor plates and configuration of armor	(b) US20060213360A1
F41A19/10, F41A17/72, F41A19/14	Firing or trigger mechanisms; cocking mechanisms	(c) US20160209142A1



(a) Gun Magazine (b) Perforated Armor (c) Handgun Triggers

Figure 1. Images of Patents that Contain Prevalent Classification Codes

For our analysis, we will analyze patents containing multiple sub-classifications [n = 20,644] to understand the movement of technologies across time together.

Zooming Out: Micro to Macro Trends

Our main interest is finding how certain technologies relate to others through classification codes, but since these classifications describe very specific technologies, it would be a daunting task to analyze these micro level trends. Thus, we take another approach by looking at classification codes which do not contain the full path, level 3 and level 4 codes, that show how general technological advances fluctuate overtime.

$$\text{[Level 3]}\ (\mathcal{L}_1)-(n_1)-(\mathcal{L}_2)-00/00 \quad \text{[Level 4]}\ (\mathcal{L}_1)-(n_1)-(\mathcal{L}_2)-(n_2)/00$$

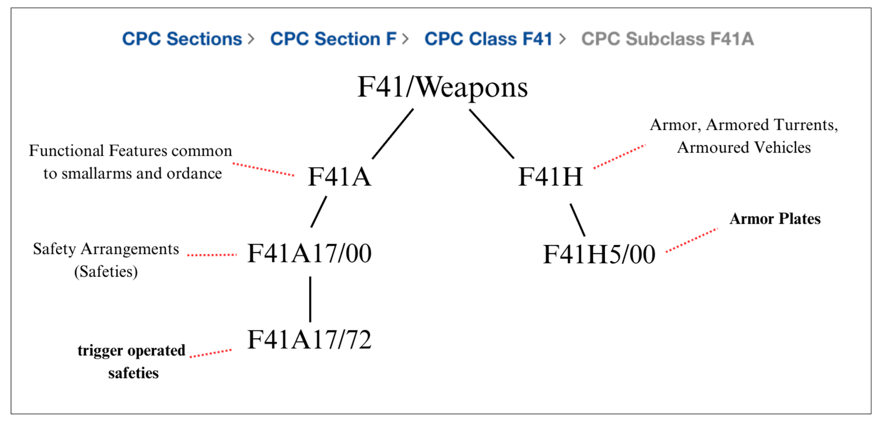
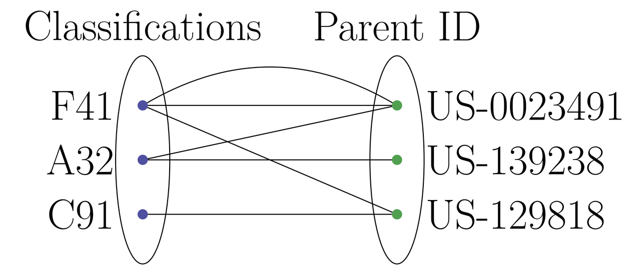


Figure 2. Sample CPC Scheme Layout

The Methodology Behind the Linkage Coefficient

The Parent-Patent Child-Classification Framework:

The notion behind our coefficient can be seen through the structure of the following diagram.



We essentially find classification codes which are 'siblings' through a shared 'parent patent' and observe the number of edges between parents. This process is defined more rigorously in the next few sections through a combinatorial framework

Co-occurrence Weights

Given j, k are classifications and N_j and N_k represent the neighbors (parent patents), we use the following process to measure the level of co-occurrence between classifications.

For each given $v \in N_j \cap N_k$, let $J_v = \{\omega_v^1, \omega_v^2, ..., \omega_v^n\}$ be the set of edges between node j and v. Likewise for $K_v = \{\kappa_v^1, \kappa_v^2, ..., \kappa_v^n\}$. To measure the co-occurrence of j, k, we calculate the following sum and denote the co-occurrence to be $\alpha_{i,j}$.

$$\alpha_{i,j} = \sum_{v \in N_j \cap N_k} |\{\{x, y\} : x \in J_v, y \in K_v\}|$$

The Linkage Coefficient

Let i, j be classifications. If K is the set of all classifications which occur with i, we find the frequency of i, j appearing together relative to i, denoted by $s(i, j)$, through the formula.

$$s(i,j) = \frac{\alpha_{i,j}}{\sum_{x \in K} \alpha_{i,x}}$$

and take an average of $s(i, j)$ and $s(j, i)$ to create our Linkage Coefficient, LC:

$$LC(i,j) = LC(j,i) = \frac{s(i,j) + s(j,i)}{2}$$

Visualizing Networks Across Time

Using the Linkage Coefficient, we can see how certain level 3 classifications fluctuate overtime visually to observe technological changes. We place focus on $\text{F41}(\mathcal{L}) - 00/00$, which contains the most frequent sub-classifications in the data. Our main findings from our network are:

- F41A innovation is almost strictly within F41A with some movement towards F41F.
- 1986-1989 had much more dispersed technologies developing.

Strong Linkage Moderate Linkage Low Linkage

F41A F41B F41C F41F F41G F41H F41J

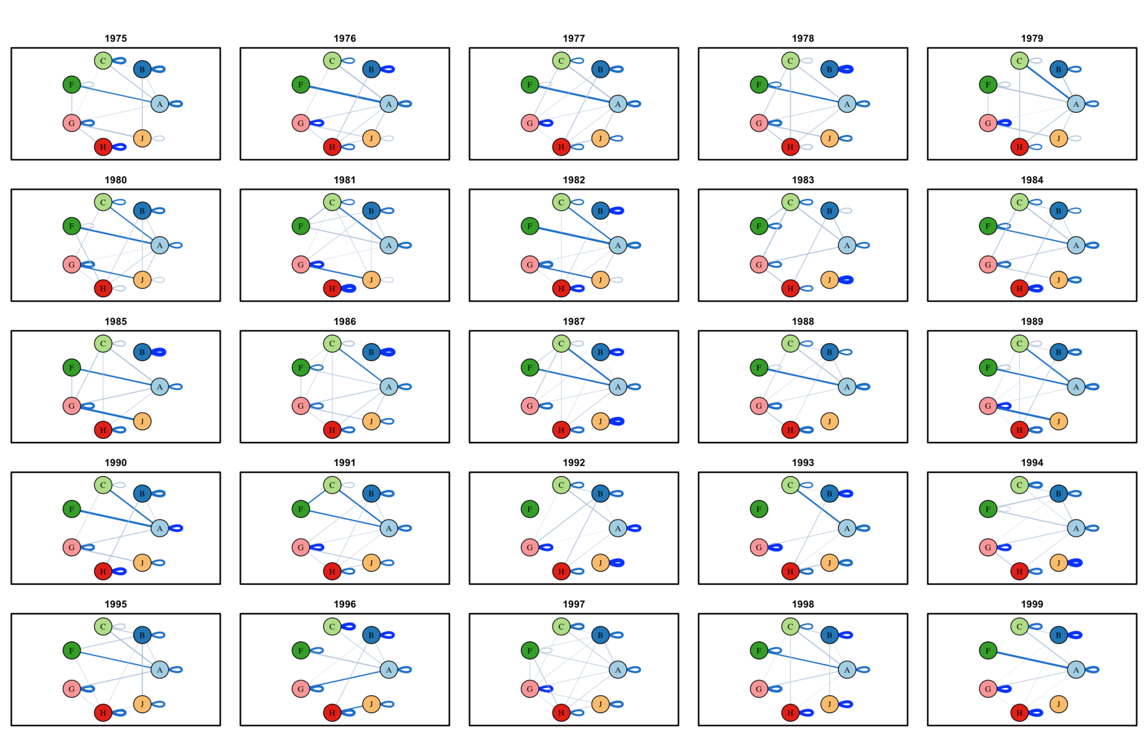
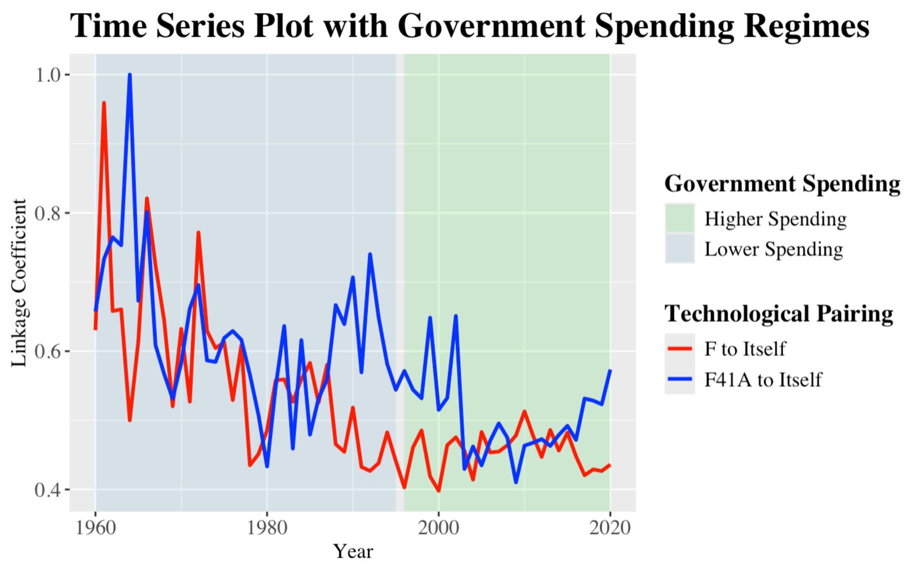


Figure 3. Networks Graphs (1975 to 1999) for $\text{F41}\mathcal{L} - 00/00$ Classes

Time Series Based Approaches

The goal of time series modeling is to predict and explain relationships in a dependent variable fluctuating overtime. Our analysis delves into how $LC(F, F)$ and $LC(F41A, F41A)$ fluctuate overtime, and we consider government spending aggregates from 1960 to 2020 as an exogenous variable to analyze. To prepare the data, we took a first difference for government spending variable and for $LC(F41A, F41A)$ to make the data stationary for modeling.



Modeling and Results: We run naive, mean, exponential smoothing, and drift models alongside ARX, MAX, ARMAX, and ARIMAX models to model our data. Running these models, we used RMSE as a metric to determine the best model fit. We find that $LC(F, F)$ is best modeled by the ARX model given by the following equation

$$Y_t = 0.54 + 0.25Y_{t-1} + 0.22Y_{t-2} - 0.25Y_{t-3} + 0.13Y_{t-4} + 0.49Y_{t-5} - 0.05X_t + \varepsilon_t$$

where X_t is the difference in standard deviation of government spending. Through this model, differenced standardized government spending is significant at the 10% significance level, and the first lag ($\phi_1 = 0.25$) and fifth lag ($\phi_5 = 0.49$) were statistically significant as well (p values < 0.01). On the other hand, for $LC(F41A, F41A)$ the ARIMAX was the best model chosen:

$$\Delta Y_t = 0.31\Delta Y_{t-1} + 0.37\Delta Y_{t-2} - \varepsilon_{t-1} - 0.04X_t + \varepsilon_t$$

In our ARIMAX, we have that the change of previous observations (ΔY_{t-n}) and random shocks from previous observations (ε_{t-1}) are significantly associated with the change in $LC(F41, F41)$ today (p values < 0.05). However, the difference in standard deviation of government spending is not significant.

Discussion and Limitations of Study

The association between $LC(F, F)$ and differenced standardized government spending show changes to government spending within the year might play a role in the levels to which firearms develop technologies contained **only in** mechanical engineering, weaponry, and blasting technologies. However, insignificant results for $LC(F41A, F41A)$ show government spending might not have an impact on firearm patents that focus on small firearm development technologies.

Despite these findings, there are still a few limitations to consider behind the study:

- **Lack of an error measure:** $LC(i, j)$ has not yet been tested against traditional cost/loss functions such as RMSE or rank ordering.
- **Other Viable Coefficients:** Could be the case different ways to weight the coefficient are more appropriate.

More research is ultimately required to fully understand these relationships between firearm technologies and bring us closer to a more comprehensive understanding of how technological domains evolve in response to historical events such as war and scientific developments.

Acknowledgments

I'd like to thank the Ronald E. McNair Fellowship program and Baker 64' Collabria Fellowship in Data Analysis for the funding behind this project. I'd like to also acknowledge Maryam Gooyabadi, my faculty advisor, for pushing and encouraging me throughout this whole process. Much of the work done could not have been achieved without collaboration with her. Special thanks also to my McNair advisor, Ronnie Hendrix, for his support throughout this whole research process.