Factionalism and the Red Guards under the Cultural Revolution: Ideal Point Estimation Using Text-as-Data Scaling Method

Yi-Nung Tsai (NCCU), Dechun Zhang (Leiden) and David, Yen-Chieh Liao (Essex)

Abstract

This article estimates the spatial positioning of the political elites and participants in the Cultural Revolution through analyzing expressed political views in propaganda publication in China. Prior theoretical research indicates that the political elites and the Red Guards lost control of the social movement, and it evolved into sheer verbal violence and physical skirmishes across all provinces in the end (MacFarquhar and Schoenhals 2006). We show that in this political chaos, the Red Guards and the elite were dynamically fractioned, as is reflected in self-printed propaganda, such as big-character posters (Dazibao) and tabloids (Xiaobao). This paper develops a novel approach that augments TextRank algorithm to extract keywords and Chinese phrases, on top of the Poisson scaling model (Wordfish) to estimate the differences of spatial positions using the extracted textual key terms. Results are shown to be consistent with the literature.

Introduction

This article proposes an estimation strategy to measure politicians' ideological positions from textual data, together with an application to study the factionalism of the Red Guards during the Cultural Revolution. One of the most intriguing features of the mass movements that sparked the Cultural Revolution is the intensity of factionalism that has drawn a series of studies' attention to exploring the nature and the origin of factionalism. The main question of our interest is the overall ideological dissimilarity between the political elite and the Red Guards: if overall, the political elite placed themselves differently vis-à-vis the Red Guards along the spectrum of ideological positions. Afterwards, we study, at the time interval of each major incident, the degree of disagreement among different factions in the Red Guards and investigate its spatial positioning in correspond to the political elite.

Estimation Strategy

We seek to find an algorithm that serves two main functions: i) extract key Chinese words and phrases (slogans); ii) estimate the differences of spatial positions using the extracted textual data. To this end, we develop a new strain of Wordfish algorithm that takes advantage of the facilitation of TextRank algorithm to extract Chinese characters and phrases. It not only extracts sole key characters (as in Wordfish) but also considers semantic structures of Chinese characters to extract key phrases. The consideration of semantic structures of our proposed algorithm mitigates the biasedness and errors that are often found in the original Wordfish caused by counting sole words and neglecting the linguistic contexts.

Automated Keyword Extraction

Particularly, in the first step, the latest Universal Dependencies 2.5 model tokenizes Simplified Chinese documents, based on linguistic usage patterns of Chinese language in natural language processing task, and TextRank, developed by Mihalcea and Tarau (2004), effectively extracts a set of representative terms from the input text, based on values of importance they carry, to compose domain-specific dictionaries. Mathematically, we can calculate the value of importance of a vertex i by

$$V(i) = (1-d) + d \sum_{V(j) \in Link(i)} \frac{p_{ji}}{\Sigma V(l) \in Link(i) \, p_{jl}} V(j), \quad (1)$$

where $V(i)$ is the value of importance of vertex n-specific dictionaries.

Figure 2: Smoothed Density Distributions of Estimated Positions for the Political Elites and the Red Guards
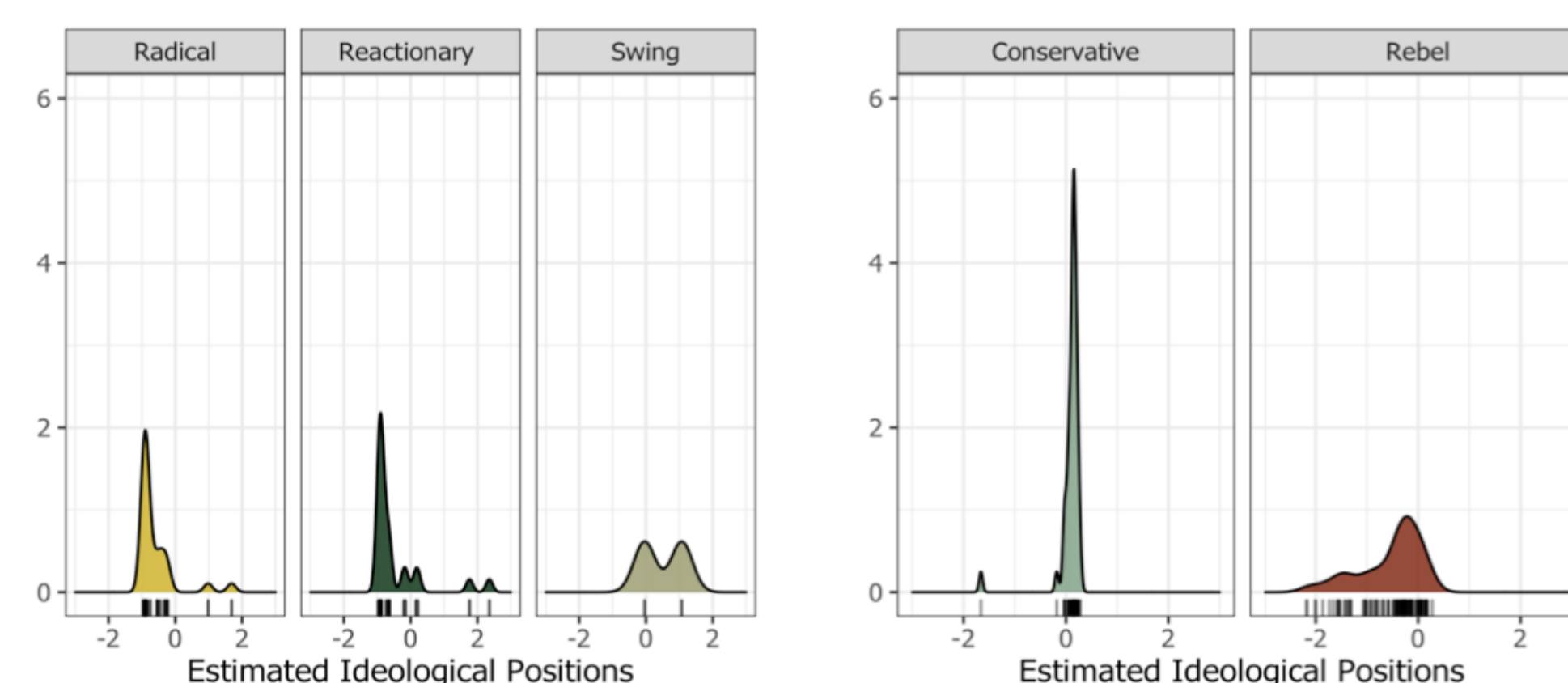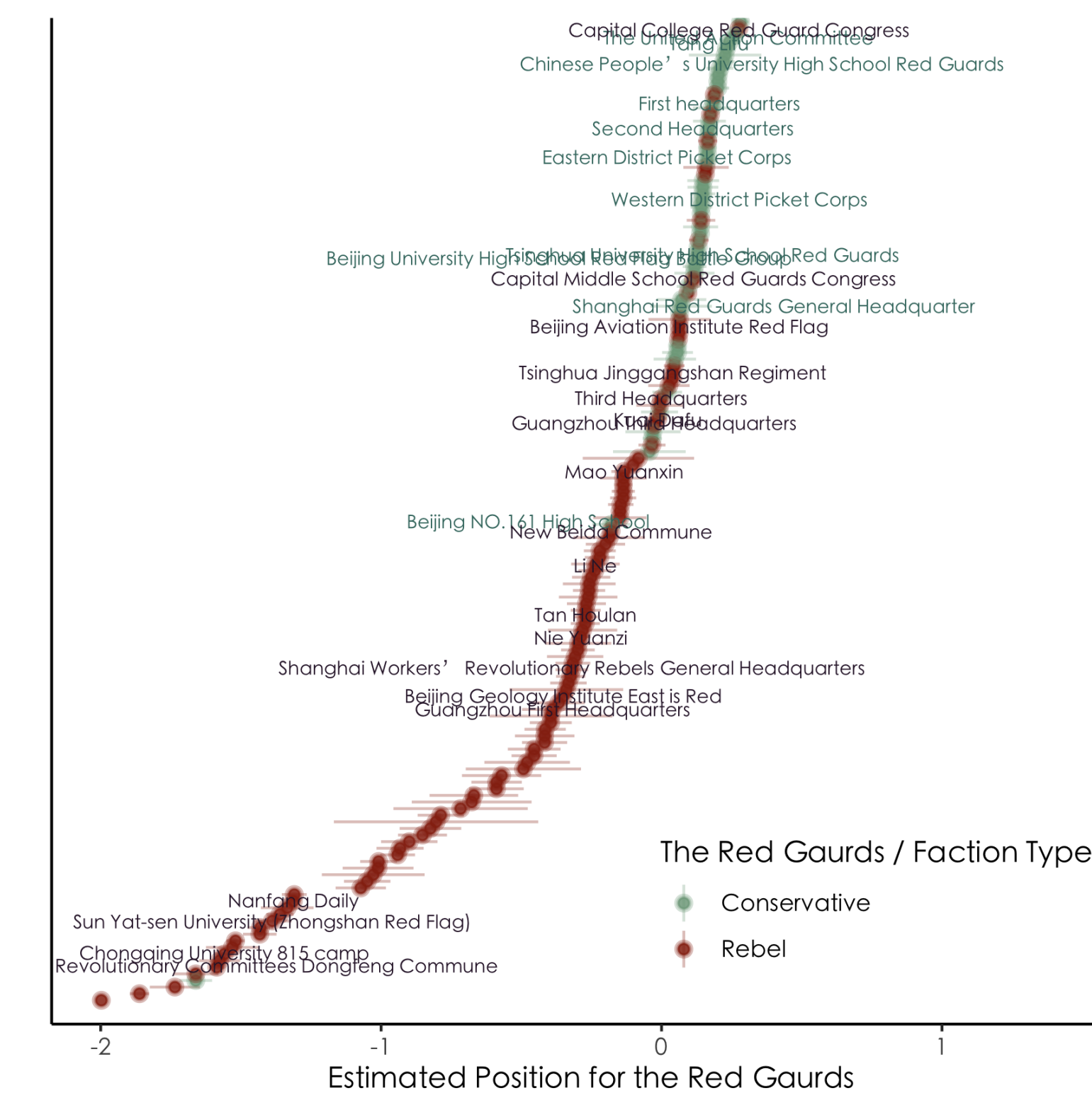

Figure 3: Estimated Positions for Individual Members and Organizations in the Red Guards

Wordfish Poisson Scaling Model

In the second stage, the model proposed by Slapin and Proksch (2008) produces the estimated positions by analyzing frequencies of extracted keywords and phrases from dictionaries constructed in the first stage. For simplicity, the distribution of the frequencies is assumed to be a Poisson process, as the Poisson process only has a single parameter, θ, for the model to estimate. θ represents both the mean and the variance of the distribution. The model takes the following functional form:

$$y_{jm} \sim Poisson(\theta_{jm}) \quad (2)$$

$$\theta_{jm} = exp(a_m + b_j + \beta_j * w_i^m), \quad (3)$$

To complete the model, Bayesian formulation is introduced by the specification of prior multi-normal distribution for model parameters and w_i^m, for any document m:

$$\begin{pmatrix} p(a_m) \\ p(b_j) \\ p(\beta_j) \\ p(w_i^m) \end{pmatrix} \sim IIN \left(\begin{pmatrix} \mu_a \\ \mu_b \\ \mu_\beta \\ \mu_w \end{pmatrix}, \begin{pmatrix} \sigma_a^2 & 0 & 0 & 0 \\ 0 & \sigma_b^2 & 0 & 0 \\ 0 & 0 & \sigma_\beta^2 & 0 \\ 0 & 0 & 0 & \sigma_w^2 \end{pmatrix} \right) \quad (4)$$

Conclusion

Accurate estimation of ideological positions is the primary and fundamental step towards the understanding of factionalism and cleavage in the Red Guards and the political elite. Yet, appropriate numerical data were seldom recorded and are rather scare at that point of time in China. Estimation outcomes of ideological positions demonstrate that the majority of the Red Guards factions and the elite factions are contrasting groups for most of time. This seems sensible and matches many evidences from previous literature (e.g., Walder and Su 2003; Wang 2019). This generally in alignment with Walder (2004), Walder (2006), and Walder (2009). Moreover, when the Conservative Red Guards were deemed as counterrevolutionaries and under arrest, we indeed record less factions from the Conservative.

Figure 7: The Estimated Positions of Each Major Historical Incident for Individuas of the Red Guards and the Political Elite
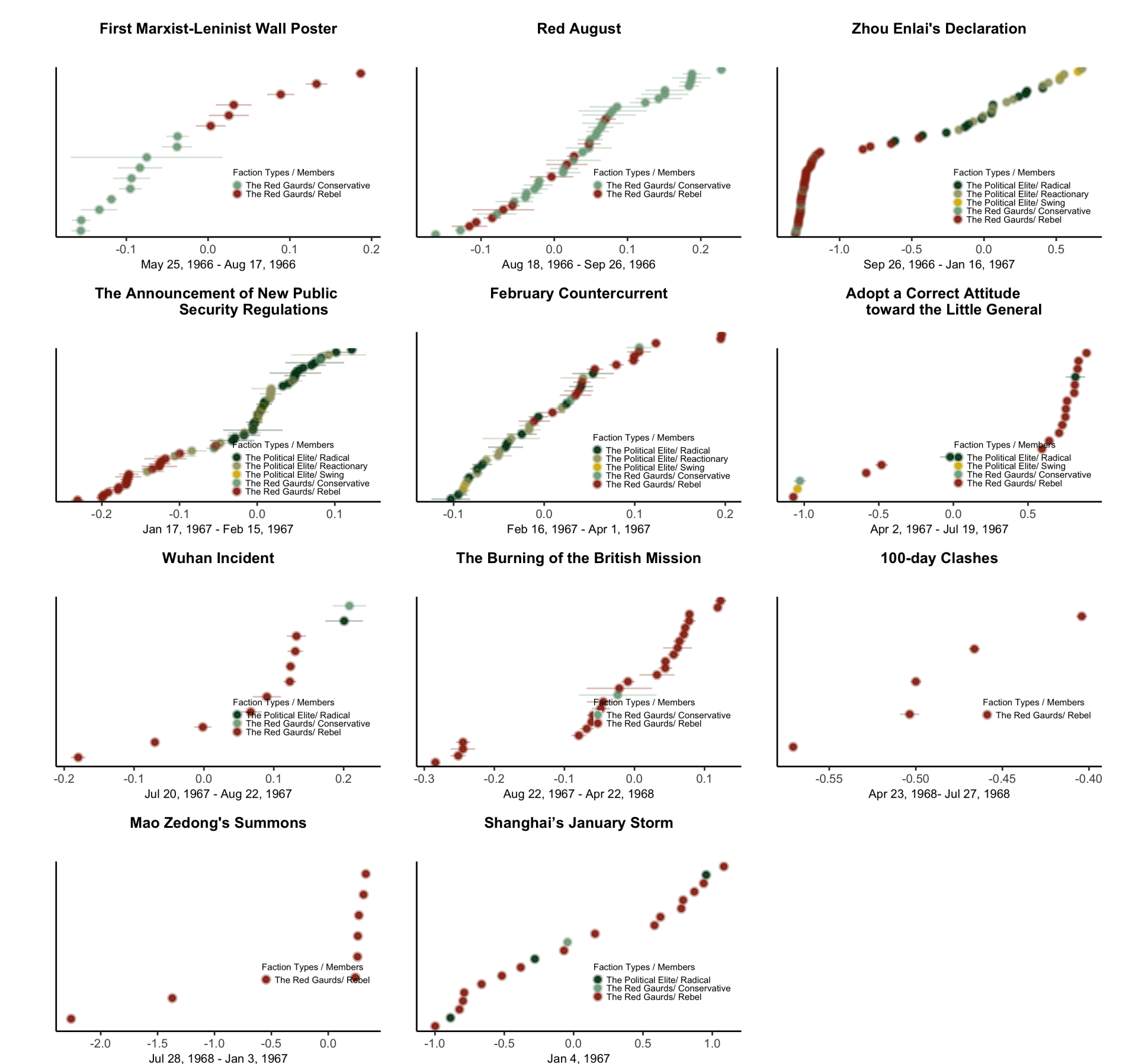
Contact Information

- Corresponding author: David, Yen-Chieh Liao
- Github:davidycliao/redgaurds
- Email: yl17124@essex.ac.uk