

# External Threats, Linguistic Homogeneity, and Political Unification

Chen Zeng

Department of Political Science  
Renmin University of China

March 14, 2022

# Introduction



Niccolò Machiavelli, a Florentine in the 16th century, argued for the unification of Italy, which didn't happen until the *Risorgimento* in 19th century, citing the following reasons:

- Florentine lifestyle
- The Italian language, i.e. Tuscan dialect
- Women, god, and heroes
- External interference

Today we look at two key elements in the political unification of states — external security threats and linguistic homogeneity.

## 1 Conceptualization and Operationalization

- Political Unification
- External Security Threats
- Linguistic Homogeneity

## 2 Exploratory Analysis and Causal Inference

## 3 Reflections

- The Fundamental Problem of Causal Inference
- Selection Bias
- Confounding Factors – Geographical Proximity
- Alternative Research Design – fsQCA & Logistic Regression

## Definition

**Political unification** occurs when two or more sovereign states merge into one.

- Political unification has been one of the most debated areas in the study of political science.
- Notable instances included the US (American Revolutionary War), Germany (*Sonderweg*), and Italy (*Risorgimento*).
- Does the unification happen incrementally, or is there a clear moment for such unification?
- Recent development of the EU, free trade areas, and customs unions has led political scientists to examine if there are limits to such political and economic integration. To which point do they stop being just “integration” and become unification?

# Political Unification

- An *operational* definition of political unification used in this article is that two sovereign states voluntarily merge into a federation or unitary state. Conquests/accessions are excluded.

**Unit of analysis** Political unification (of a state dyad)

**Unit of observation** Inter-state actions (dichotomous variable: unification, linguistic homogeneity); and individual states (continuous variable: threat levels)

- This research leverages datasets from the Correlates of War project. Qualifying states have to meet at least one of the following criteria.

**Criterion 1** Before 1920, a population of at least 500,000 and establishment of diplomatic missions at or above the rank of *chargé d'affaires* by Britain and France.

**Criterion 2** After 1920, membership in the League of Nations or United Nations *or* a population of at least 500,000 and establishment of diplomatic missions from any two major powers.

# Research Hypotheses

**Hypothesis 1** External threats are *necessary* for political unification.

**Hypothesis 2** Linguistic homogeneity is *necessary* for political unification.

- The article takes on the form of a *falsification* probe and focuses on only *necessary* conditions for political unification.
- That is to say that security threats or linguistic homogeneity may not be *sufficient* for political unification, but the lack of either condition cannot result in the expected outcome.
- This method is prone to many methodological and substantive vulnerabilities. We will cover them in the final section, but now, let's take a look into why the author sets those two key explanatory variables.

# External Security Threats

- Riker (1975) argues that states unify for security reasons; they desire protection from external threats

hi



In the following subsections, we set let  $U$  denote political unification,  $S$  external security threats, and  $L$  linguistic homogeneity.

- From previous analysis, the author detected 12 cases of political unification from the MID dataset, where linguistic homogeneity is present in all cases. Incorporating the Switzerland case as an outlier, we have 13 cases in total. Mathematically, we have

$$P(L|U) = \frac{12}{13} \approx 0.92 \quad (1)$$

- According to Bayes theorem, we have

$$P(U|L) = \frac{P(U)P(L|U)}{P(L)} \quad (2)$$

- Since  $P(U)$  and  $P(L)$  have the same denominator, we can simplify the equation as

$$P(U|L) = \frac{N(U)P(L|U)}{N(L)} \quad (3)$$

whereas  $N(U) = 12$ , as per previous discussion.

- Only considering all 6 UN official languages, we have

$$N(L) < N(EN)^2 + N(ES)^2 + N(FR)^2 + N(ZH)^2 + N(RU)^2 + N(AR)^2 = 2115 \quad (4)$$

- Furthermore, we calculate the probability of political unification given the presence of linguistic homogeneity.

$$P(U|L) = \frac{N(U)P(L|U)}{N(L)} < \frac{12 \times 0.92}{2115} \approx 0.52\% \quad (5)$$

# Selection bias

hi

# Confounding factors – geographical proximity

$$x \sim \mathcal{N}(\mu, \sigma^2)$$

$$x \sim \mathcal{N}(\mu, \sigma^2)$$

# Alternative research design – Logistic Regression

$$\textit{logit}(P(U)) = \beta_0 + \beta_1 \times S + \beta_2 \times L \quad (6)$$

$$\textit{logit}(P(U)) = \beta_0 + \beta_1 \times S + \beta_2 \times L + \beta_3 \times S \times L \quad (7)$$