

Dyad Geopolitical Scores - Variable Codebook

Dataset Overview

File Name: dyad_geopolitical_scores.csv
Unit of Observation: Dyad-year
Time Coverage: 1950-2024
Source Data: GGE_1950_2024.csv (Global Geopolitical Events)

Variable Definitions

Identifiers

Variable	Type	Description
dyad	String	Country pair identifier in alphabetical order (ISO3-ISO3 format, e.g., "CHN-USA")
country1_code	String	ISO 3-letter country code for first country in dyad (alphabetically)
country2_code	String	ISO 3-letter country code for second country in dyad (alphabetically)
year	Integer	Calendar year of observation

Primary Geopolitical Measures

Variable	Type	Range	Description
geo_score	Float	[-1, 1]	Static geopolitical score. Annual average of Goldstein Scale values divided by 10 for all events between dyad countries in given year. Missing if no events observed.
geo_score_dyn	Float	[-1, 1]	Dynamic geopolitical score. Exponentially weighted score incorporating historical information with decay parameter $\lambda = 0.3$. Accounts for temporal persistence in bilateral relations.
geo_score_ma	Float	[-1, 1]	Moving average geopolitical score. Event-weighted average of geo_score over 3-year window (t-2 to t). Missing if no valid observations in window.

Event and Weight Variables

Variable	Type	Range	Description
number_of_events	Integer	$[0, \infty)$	Count of geopolitical events between dyad countries in given year
cum_events	Float	$[0, \infty)$	Cumulative event count with exponential decay. Calculated as: $\text{cum_events}(t) = (1-\lambda) \times \text{cum_events}(t-1) + \text{events}(t)$ where $\lambda = 0.3$
weight	Float	$[0, 1]$	Weight assigned to current year's geo_score in dynamic calculation. Equals $\min(1, \text{events}(t) / \text{cum_events}(t))$ when events > 0, else 0

Score Interpretation

Goldstein Scale Mapping

The underlying Goldstein Scale ranges from -10 (maximum conflict) to +10 (maximum cooperation). All geo_score variants are scaled by dividing by 10:

- **-1.0 to -0.6:** Severe conflict (military attacks, war, use of force)
- **-0.6 to -0.3:** Significant hostility (military threats, sanctions, breaks in relations)
- **-0.3 to 0.0:** Minor conflict (diplomatic protests, recalls, verbal hostility)
- **0.0 to 0.3:** Neutral to mild cooperation (routine diplomatic exchanges, visits)
- **0.3 to 0.6:** Moderate cooperation (agreements, treaties, joint statements)

- **0.6 to 1.0:** Strong cooperation (military alliance, economic integration, extensive aid)

Missing Values

- `geo_score`: Missing (NaN) for dyad-years with no observed events
- `geo_score_dyn`: Never missing; equals 0 for dyads with no prior events, decays toward 0 during periods without events
- `geo_score_ma`: Missing (NaN) when no valid observations exist in 3-year window

Methodological Notes

Dynamic Score Calculation (`geo_score_dyn`)

The dynamic score incorporates temporal persistence using an exponential decay model:

1. **Initialization ($t=1$):** `geo_score_dyn` = `geo_score` if events > 0, else 0
2. **No events ($\text{events}(t) = 0$):** `geo_score_dyn`(t) = `geo_score_dyn`($t-1$) \times (1- λ)
3. **With events ($\text{events}(t) > 0$):** `geo_score_dyn`(t) = (1- w) \times `geo_score_dyn`($t-1$) + $w \times \text{geo_score}(t)$
 - where $w = \min(1, \text{events}(t)/\text{cum_events}(t))$

Moving Average Calculation (`geo_score_ma`)

- **Window:** 3 years (current year plus 2 prior years)
- **Weighting:** Event-weighted average using `number_of_events` as weights
- **Treatment of gaps:** Only non-missing `geo_score` values included in calculation

Key Parameters

Parameter	Value	Description
DECAY (λ)	0.3	Exponential decay rate for dynamic scores and cumulative events
MA_YEARS	3	Moving average window length (years)

Usage Notes

1. **Balanced Panel:** Dataset includes all possible dyad-year combinations within the time range, including years with no events (`number_of_events = 0`)
2. **Dyad Directionality:** Dyads are undirected; "CHN-USA" and "USA-CHN" are treated as the same dyad with countries ordered alphabetically
3. **Temporal Coverage:** While the panel is complete, actual event coverage may vary by dyad depending on data availability in the source GGE dataset
4. **Recommended Usage:**
 - Use `geo_score` for cross-sectional analysis or when focusing on contemporaneous relations
 - Use `geo_score_dyn` for time-series analysis requiring smooth transitions and accounting for relationship persistence
 - Use `geo_score_ma` for robustness checks or to reduce measurement error from year-to-year volatility
5. **Interpretation of Sign:**
 - **Negative values:** Indicate conflict, hostility, or deteriorating relations
 - **Positive values:** Indicate cooperation, partnership, or improving relations
 - **Values near zero:** Indicate neutral or minimal interaction