

Measuring Geopolitical Alignment from Historical Events

Tianyu Fan

Yale University

NBER AI and Economic Measurement, Spring 2026

LLM as Coder in a Measurement Framework

Object

- Bilateral *geopolitical alignment*
- Not directly observed: *formed, revealed, and updated* through events
- Six dimensions: economic; diplomatic; security; legal / territorial; multilateral; other

LLM implementation (3 sequential tasks)

- **Task 1:** search authoritative sources for bilateral events
- **Task 2:** classify each event under *CAMEO* (cooperation-conflict)
- **Task 3:** assign *Goldstein* score on $[-10, +10]$

Measurement principle

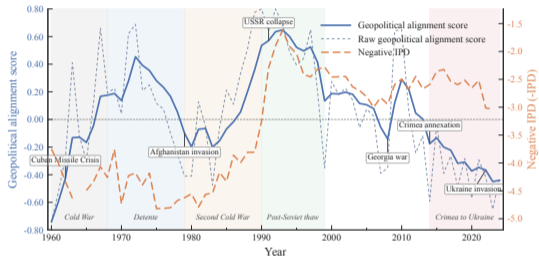
Geopolitical events are well-documented historical records that characterize bilateral geopolitical alignment.

833,485 events · 193 countries · 18,528 bilateral pairs · 1950–2024

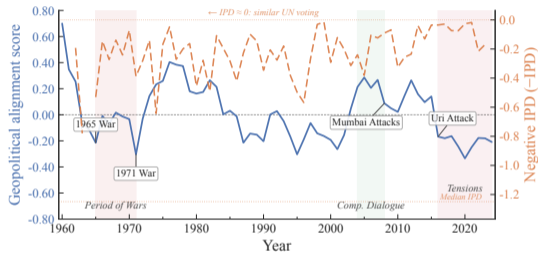
→ The researcher fixes the framework; the LLM applies it.

The Measure Tracks Bilateral History

U.S.–Russia, 1960–2024



India–Pakistan, 1960–2024



- **U.S.–Russia:** Cold War → Détente → post-Soviet thaw → Crimea/Ukraine
- **India–Pakistan:** persistent rivals, yet always vote similarly in UN

→ event-based alignment captures bilateral relations; UNGA captures multilateral affinity.

Replicability: Not a Model-Specific Artifact

Across models (4 frontier configs, 300 random dyad-years)

- Cross-model agreement: $\rho = 0.88$

Within model (same 300 dyad-years, repeated runs)

- Replicability: $\rho = 0.93$

Across pipeline (different LLM + prompt, 12,480 country-years)

- Country-year index: $r = 0.87$
- Downstream growth IRFs: $r = 0.97$

Human audit in progress (stratified random sample)

→ Variation comes from well-documented historical events, not hidden model judgment.

Models tested

- Gemini 2.5 Pro (baseline)
- GPT-5.4
- Claude Opus 4.6
- Claude Sonnet 4.6

Application: The External Political Environment of Development

Countries develop within an *external geopolitical environment* that shapes stability, access to markets, capital, and technologies.

Headline

- One-SD permanent improvement in alignment: +10% GDP per capita over 25 years
- Implied magnitudes range -30% to $+30\%$ across countries and time periods

Robustness

- Stable across partner groups, historical periods, and interaction types; robust to controls and two complementary IV checks

Applicability

- Companion paper: how geopolitics shapes post-Cold War trade globalization

→ LLMs do not replace researcher judgment, but they make new forms of disciplined measurement feasible.

Backups

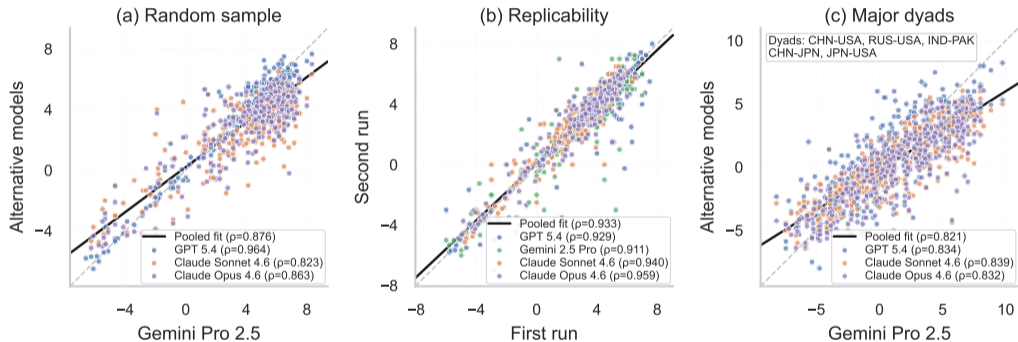
- Sample event coding (U.S.–Russia 2022)
- Cross-model 3-panel scatter
- Full-pipeline IRF overlay
- Growth baseline IRF
- Greece–Turkey validation (second UNGA-inversion case)
- Prompt structure + event categories

Backup: Sample Event Coding (U.S.–Russia, 2022)

Event	CAMEO Class.	Econ. Type	Goldstein
Military Aid to Ukraine	Material Conflict	Not econ.	-8.0
Sweeping Sanctions on Russia	Material Conflict	Sanctions	-7.5
Russia Suspended from UN HRC	Material Conflict	Not econ.	-6.5
Biden Labels Putin “War Criminal”	Verbal Conflict	Not econ.	-5.5
UN General Condemnation	Verbal Conflict	Not econ.	-4.5
Griner–Bout Prisoner Exchange	Material Cooperation	Not econ.	+7.0

Each dyad-year returns structured event records under CAMEO + Goldstein, aggregated to a yearly bilateral score.

Backup: Cross-Model Agreement (Full)

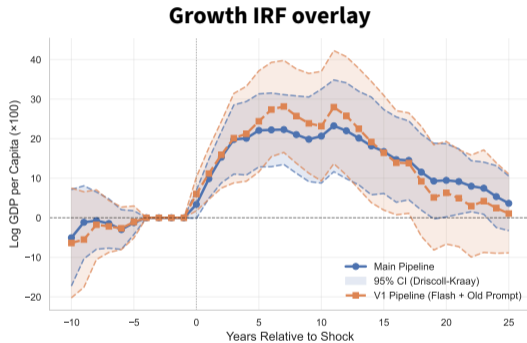
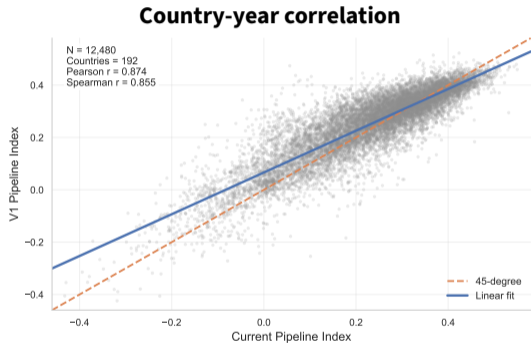


(a) random sample, $\rho = 0.88$

(b) within-model, $\rho = 0.93$

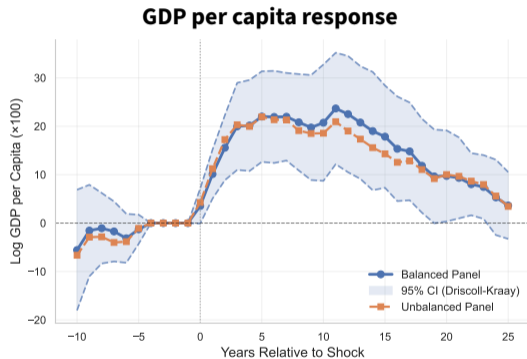
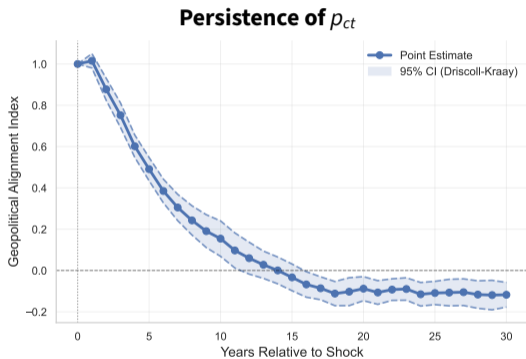
(c) major dyads, $\rho = 0.82$

Backup: Full-Pipeline IRF Overlay



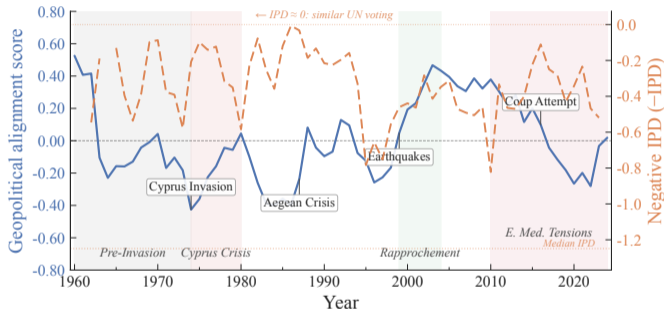
Different LLM (Gemini 2.5 Flash) + different prompt structure: $r = 0.87$ country-year, $r = 0.97$ across IRF horizons.

Backup: Growth Baseline IRF



Local projections, country FE + region-year FE; Driscoll-Kraay SEs.

Backup: Greece–Turkey (Where UN Voting Fails Again)



NATO allies in chronic dispute (Cyprus, Aegean, Eastern Mediterranean). UN voting again shows them as aligned despite real bilateral hostility.

Backup: Prompt Structure + Event Categories

Six event dimensions (Appendix A.2 Table A1):

- **A. Economic relations:** trade, sanctions, investment, energy, technology, infrastructure
- **B. Diplomatic / political:** summits, embassy, ministerial meetings, public diplomacy
- **C. Security / defense:** military alliances, arms sales, border conflicts, cyber operations
- **D. Legal / territorial:** maritime claims, visa, extradition, international disputes
- **E. Multilateral:** UN votes, regional organizations, climate, health, human rights
- **F. Other significant:** humanitarian crises, sports diplomacy, cultural exchanges

CAMEO classification (Schrodt 2012): Four quadrant classes (Verbal/Material × Cooperation/Conflict), with 20 two-digit root codes and three-digit event codes within them.

Goldstein scale (−10 to +10): Anchored across event types: −10 military attack · −7 economic conflict · 0 neutral · +6 major agreement · +10 extreme cooperation.

Coding discipline: CAMEO and Goldstein assigned after the event is identified within one of the six dimensions; LLM applies bilateral context within fixed scoring conventions.