The AI-GPR Index: Measuring Geopolitical Risk using Artificial Intelligence

Matteo Iacoviello
Jonathan Tong

Federal Reserve Board and University of Wisconsin-Madison

Financial Risks International Forum
Paris - March 30, 2026

Measuring geopolitical risk: A new tool

- Economists have long tried to quantify hard-to-measure phenomena by counting specific words or combinations of words in newspapers — fast, transparent, and powerful. E.g. the GPR Index of Caldara-Iacoviello (2022)
- Limitation of these approaches: dictionary-based approach cannot easily tell whether "war" in an article means active conflict or retrospective on WWII; cannot say who is threatening whom; cannot tell whether crisis threatens oil supplies.
- Large language models change this: LLM reads each article and understands its meaning — just as a research assistant would, but at a scale that is impossible for humans.

What this opens up

Better measurement of the same concept (less noise, fewer false positives) and new measures that word-counting simply cannot produce: oil-specific risk, bilateral directed stress, actor-role networks

The AI-GPR index

What we do

We build a daily measure of geopolitical risk from 5 million newspaper articles spanning 1960 to 2025. Each article is read and scored by an AI model (GPT-4o-mini).

We extend the same approach with additional classification layers to produce new measures that keyword-based methods cannot produce.

1. The AI-GPR Index

- Same newspapers and same definition of GPR as Caldara and Iacoviello (2022)
- LLM scores every article 0–1 for geopolitical risk intensity
- Available daily from 1960 onwards

2. New Measures

- **Oil-GPR**: does this article discuss a geopolitical oil supply disruption, and if so where?
- **Bilateral GPR**: who is acting on whom: initiator, respondent, spillover countries
- **Event types**: military conflict, sanctions, terrorism, ...

Building the index: Step 1

Collecting Articles

Cast a **very** wide net to collect every article *possibly* discussing geopolitical risk

Article query:

```
airstrike* OR alliance OR annex* OR attack* OR
blockade* OR bomb* OR cease-fire OR combat OR
conflict* OR coup OR crisis OR diplomat* OR embargo
OR enem* OR hostage* OR hostil* OR invade* OR
invasion* OR militia OR military OR missile*
OR nuclear OR refugee* OR riot OR sanction* OR
sovereign OR terror* OR treaty OR troop* OR truce
OR unrest OR violence OR war OR weapon* ...
```

Numbers:

- 4.6 million pass the article query filter
- ∼10 million total articles (NYT, WaPo, Chicago Tribune, 1960–2025)
- Filter is very conservative by design; false positives handled in Step 2

Step 2: Full classification prompt (simplified)

Pass the articles to LLM model, and evaluate its GPR content

```
You will be given [the first 2,000 characters of] a news article.

Classify the article's assessment of geopolitical risk based only on what the article states or
strongly implies.

Geopolitical risk is defined as the threat, realization, and escalation of adverse events associated
with:  wars (...); escalation of existing wars (...); major terrorist attacks (...); and tensions
among states and political actors (...).

Assign a geopolitical risk score from 0.0 to 1.0 based on the following scale:
0.0--0.2:  No mention of geopolitical risks.
0.2--0.4:  Mentions of minor tensions, diplomatic disputes, or isolated incidents....
0.4--0.6:  Discussion of significant tensions...
0.6--0.8:  Substantial discussion of major war initiation/escalation risks....
0.8--1.0:  Extensive coverage of imminent or new war, major war escalation, severe terrorism threats,
or critical threat to international stability.

Movies, books, anniversaries of old events, and obituaries should receive a score of 0.0 ...
Purely domestic political events (elections, protests, internal policy debates) should score 0.0 ...
```

The scoring scale in practice: Results

Score: 0.2 (minor tensions)

Soldier wounded in suicide bombing wants right to sue military contractor

The Supreme Court is considering the legality of a suit filed by a U.S. soldier who was injured in an explosion at an air base in Afghanistan.

November 3, 2025

Score: 1.0 (imminent or active war / severe threat)

The Legality, or Illegality, of Killing a Foreign Leader, Explained

The United States and Israel started a war with Iran by killing its supreme leader, Ayatollah Ali Khamenei.

Listen · 8:48 min

Share full article

Calculating the AI-GPR index

Daily index formula

$$\text{AI-GPR}_t = \frac{1}{\bar{S}} \times \frac{1}{A_t} \sum_{i=1}^{N_t} S_{it}$$

- S_{it} = GPR score assigned by LLM to article i on date t
- N_t = number of scored articles on date t (those passing the filter)
- A_t = Total articles published on date t (denominator, not just filtered)
- $\bar{S}$ = normalization constant

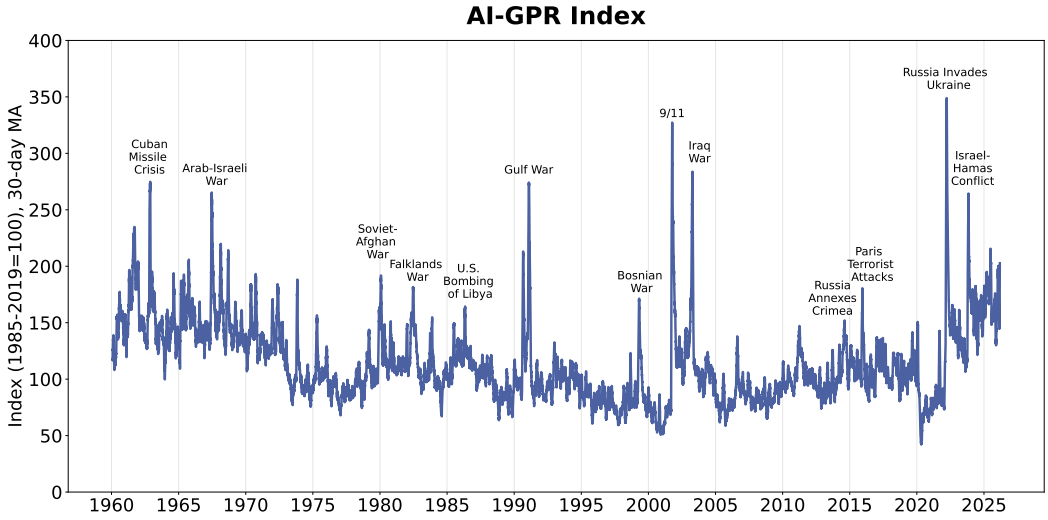
Result of Steps 1+2:

- 4.6M articles filtered
- 1.2M (26%) receive score > 0
- 15% of articles classified as GPR-relevant

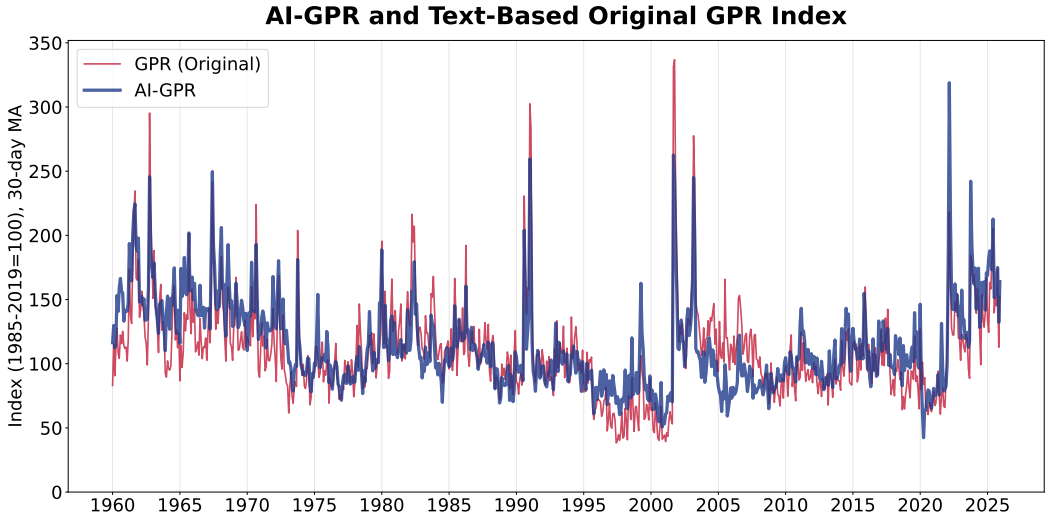
Key properties:

- Daily frequency, 1960–2025
- Continuous intensity (not binary)
- Corr. 0.69 with original GPR; higher persistence (autocorr. 0.73 vs. 0.62)

The AI-GPR index, 1960–2025



AI-GPR vs. original GPR: Monthly comparison (1960-2025)



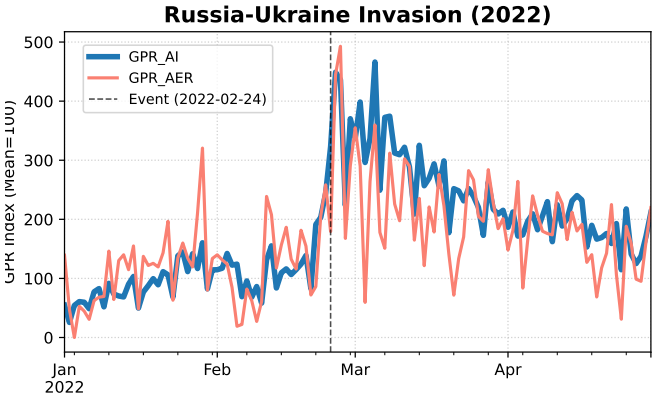
Measurement validation

Concern	Test	Finding
Stochasticity	Re-run 3x on 1,050 articles	Corr. 0.99; 98% identical scores
Text truncation	2,000-char vs. full text	Corr. 0.93; 86% identical scores
Bias vs. humans	Human audit (1,050 articles)	Corr. 0.81; type-I error 13% vs. 21% for keywords
Model sensitivity	GPT-4o, GPT-5-mini, GPT-5	Cross-model corr. 0.86–0.94
Type-II error	LLM on excluded articles	$<1\%$ receive positive score

Type-II errors and Carlson and Dell (2025)

Carlson and Dell show keyword classifiers systematically undercount geopolitically relevant articles (type-II error) and propose a statistical post-hoc correction. Our LLM approach addresses this directly: we classify 15% of all articles as geopolitically relevant vs. 3.3% under keywords, with a false-negative rate below 1% on excluded articles — reducing type-II errors at the classification stage rather than correcting for them afterwards.

How accurate is the AI-GPR? Russia's invasion of Ukraine (2022)



Daily AI-GPR (blue) and keyword-based GPR (red) around Russia's invasion of Ukraine. The AI-GPR spikes sharply on invasion day with less day-to-day noise.

Application 1: Geopolitical risk and stock returns

	(1) OLS	(2) Decomp.
ΔGPR_t	-0.136^{**}	
	(0.057)	
Persistent $\widehat{\Delta \text{GPR}}_t$		-0.271^{**}
		(0.134)
Shock $\hat{u}_t$		-0.119^{*}
		(0.062)
Obs.	3,452	3,448
Sample	1960–2026	

Weekly excess returns (%). Pairs-bootstrap SE in col. (2).

GPR and stock returns (col.1)

- Analyze relationship between changes in GPR and changes in weekly stock returns
- OLS coefficient is negative and significant
- Same exercise with keyword-based GPR: OLS coefficient imprecisely estimated $\Rightarrow$ AI-GPR provides a sharper signal

Decomposing stock returns (col.2)

- Decompose ΔGPR into:
  - persistent: AR(4) fitted values
  - shock (residuals) component
- Persistent component coefficient (-0.27) is twice aggregate OLS (-0.14):
- Markets price "long tail" of geopolitical risk, not just short-term movements

New measures: Geopolitical oil supply disruptions

Two-stage LLM classification

Stage 1: every article receives a GPR score 0–1 (AI-GPR)
Stage 2: articles with GPR > 0.5 are re-classified: does this article discuss an *oil supply disruption caused by a geopolitical event*? If yes, which region?

Coverage (of articles with GPR > 0.5)

- 12.6% contain oil/energy keywords
- 9.3% classified as oil disruption

Regional breakdown

- Middle East: 63%
- Russia: 14% (surges after 2014)
- North Africa: 11%
- Venezuela, W. Africa, SE Asia: 5–6% each

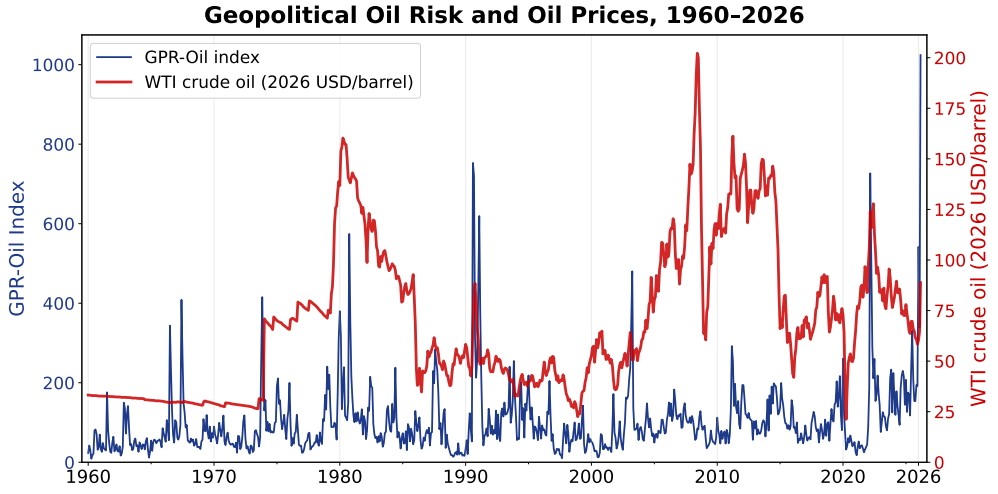
$\text{Oil-GPR}_t = \sum_{i \in \text{disruptions}} S_{it} \;/\; A_t$ (weighted by GPR intensity, normalized by total articles)

OIL-GPR: The second-stage classification prompt

```
TASK 1: Determine if this article discusses events that could disrupt oil or energy supplies.

Oil/Energy disruption includes:
- Military conflicts in oil-producing regions
- Sanctions on oil-producing countries
- Attacks on oil infrastructure, pipelines, refineries, or shipping routes
- Political instability in major oil producers
- Blockades of oil shipping lanes (Strait of Hormuz, Suez Canal, etc.)
- Oil embargoes or export restrictions
- Terrorist attacks on energy infrastructure

Answer ‘‘yes’’ if the article discusses events that could affect oil/energy production,
transportation, or supply. Answer ‘‘no’’ otherwise.

TASK 2: If ‘‘yes’’, identify which oil/energy producing region(s) are affected.
Select from: Middle East, Russia, USA, Venezuela, North Africa, West Africa, Central Asia, North
Sea, Canada, Mexico, Latin America, Southeast Asia, China, Other.
```

Applied to all articles with GPR score > 0.5 that contain oil/energy keywords.

OIL-GPR index and oil prices



Oil-GPR index (left axis, 30-day MA) and WTI oil price (monthly, right axis). Sample ends March 27, 2026

Application 2: A proxy SVAR with geopolitical oil price shocks

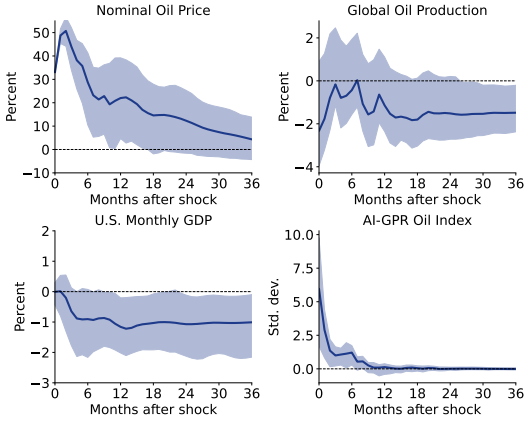
External instrument

- We use cumulative monthly WTI log-price changes on days when Oil-GPR spikes > 2 s.d. above its mean to construct an external instrument for a VAR.
- These are days with a clear geopolitical trigger—not OPEC quota decisions, demand fluctuations, or weather events.
- Oil-GPR spikes provide a valid instrument: correlated with geopolitical supply shocks, uncorrelated with demand shocks.

VAR setup

We embed this variable in a proxy SVAR with: real oil price (WTI), global oil production, U.S. real GDP, and the Oil-GPR index. 12 lags, monthly, 1975–2024.

Proxy SVAR: Impulse responses to a geopolitical oil shock



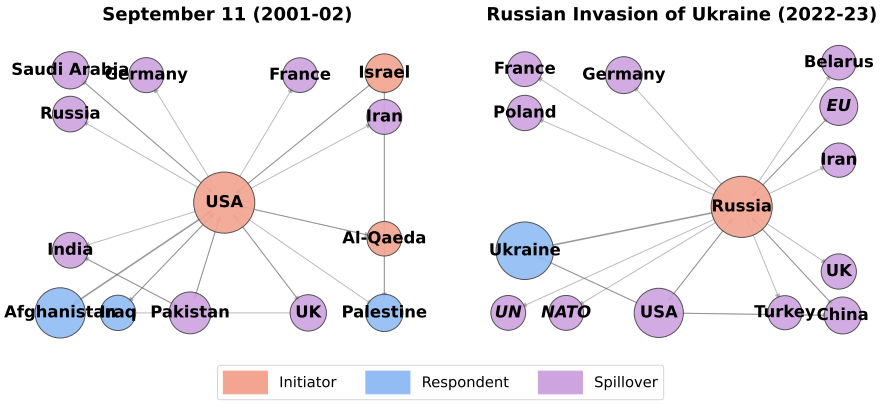
Impulse responses to a one-standard-deviation geopolitical oil supply shock (normalized to 33% oil price rise on impact). Shaded bands: 80% bootstrap CI. 12 lags. Sample: 1975–2024.

Application 3: Geopolitical risk networks

```
TASK: Identify the key country-level actors in this geopolitical event and assign each a role.

Roles are defined as follows:
- initiator:    the country that launched or triggered the hostile action (military attack, invasion,
sanctions, blockade, etc.)
- respondent:  the country that is the direct object of the hostile action
- spillover:     a country not directly involved but significantly affected (energy shock, refugee
flows, trade disruption, etc.)

Rules:
- List up to 5 actors total; include only countries with a meaningful role in THIS article
- Use standard country names (e.g.  ‘‘United States’’, ‘‘Russia’’, ‘‘Saudi Arabia’’)
- Assign exactly one role per actor
- When initiation is genuinely contested (civil wars, mutual escalations), assign both parties the
role of initiator
- If the article is too vague to identify actors, return an empty actors list[3pt]
```

Applied to all articles with GPR score > 0.5. Enables directed networks (who acts on whom) and bilateral time series for 120 country pairs.

Networks and bilateral GPR: Two examples



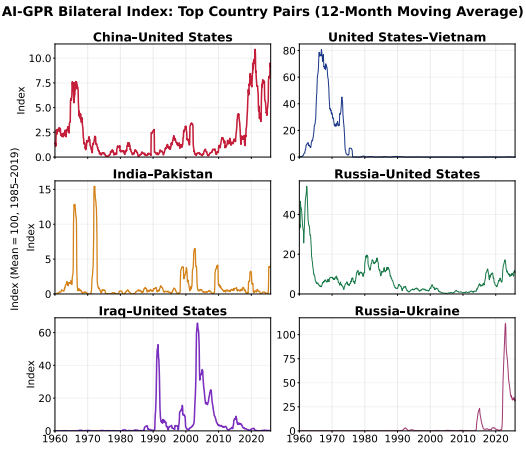
Directed networks for 9/11 and Russian Invasion of Ukraine. Node size $\propto$ GPR-weighted involvement. Color: initiator / respondent / spillover.

Bilateral geopolitical risk: A directed measure

What it Measures

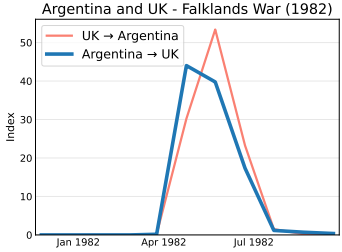
For each ordered country pair (i, j): the share of articles in month t where country i is the *initiator* and country j is the *respondent*. Directed — Russia acting on Ukraine $\neq$ Ukraine acting on Russia.

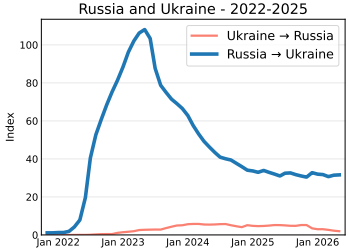
- Country pairs at monthly frequency, 1960–2025
- Extends Caldara and Iacoviello (2022) country sub-indices to a full directed bilateral matrix

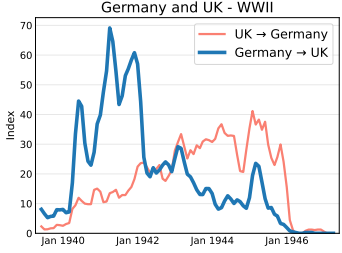


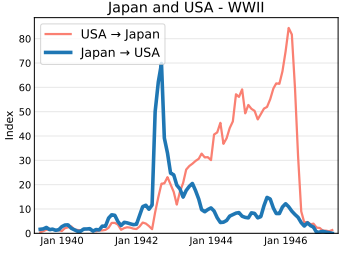
Selected pairs by bilateral GPR: sum of both directions.

Selected bilateral GPR indexes









Application 4. Does bilateral GPR depress bilateral trade?

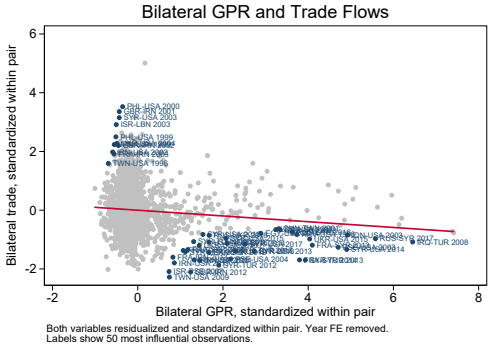
GPR and Trade

$$X_{ijt} = \beta\, \mathrm{GPR}_{ij,t-1} + \alpha_{ij} + \gamma_t + \varepsilon_{ijt}$$

- X_{ijt}: bilateral trade standardized.
- $\mathrm{GPR}_{ij,t-1}$: bilateral GPR, standardized.
- Controls: pair FE, year FE, log GDP.

Result

A 1-s.d. increase in bilateral GPR (lagged 1 year) $\Rightarrow$ 0.10 s.d. decline in bilateral trade, robust to year FE, GDP controls, and country$\times$year FE.



Coefficient estimates (and 90% CI) from gravity regressions of bilateral trade on bilateral GPR, by lag. Estimated on bilateral trade data, 1962–2023.

Conclusions

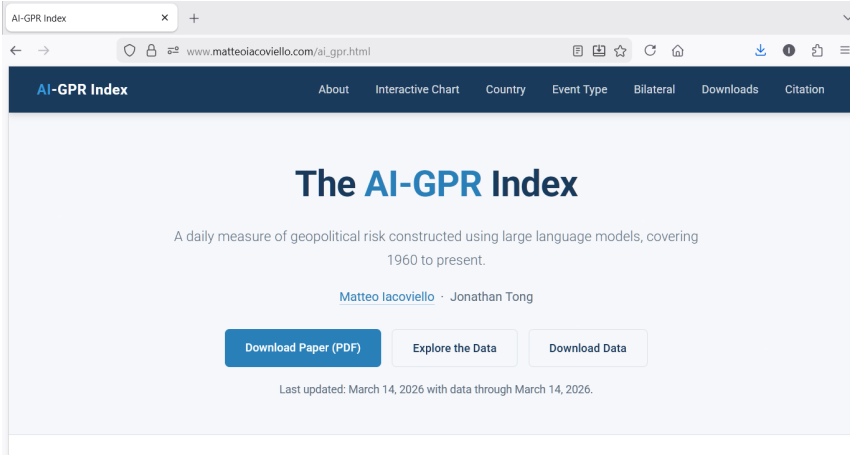
The AI-GPR Index

We construct an LLM-based measure of geopolitical risk with greater precision than keyword methods, and can measure GPR-related phenomena that cannot be easily measured otherwise

Four Applications

- **Stock returns**: AI-GPR deliv1ers a sharper negative effect; markets are especially sensitive to the persistent component of GPR
- **Oil shocks**: two-stage LLM identifies 65 years of geopolitical oil disruptions; proxy SVAR shows adverse effects of GPR-related oil price shocks
- **Networks**: LLM extracts directed actor roles (initiator / respondent / spillover) across major historical episodes
- **Decompositions**: event type, country, and bilateral — bilateral GPR predicts bilateral trade declines

All data available at matteoiacoviello.com/ai_gpr.html



Daily and monthly indices, country decompositions, bilateral pairs, event types — all downloadable. Interactive charts with smoothing controls.