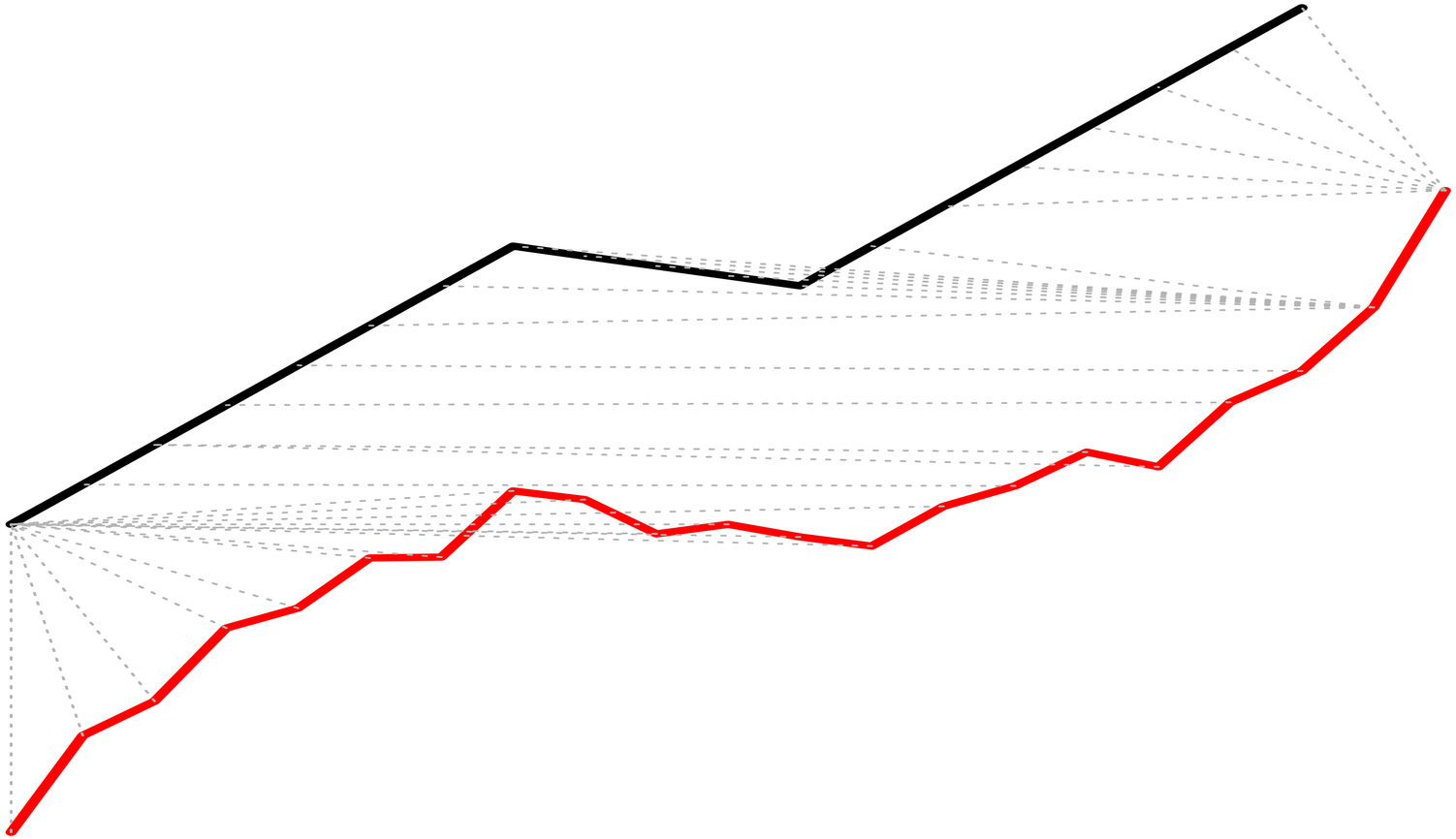
The Patterns of Conflict Emergence (PaCE)

Thomas Chadefaux
Trinity College Dublin

30 August 2022

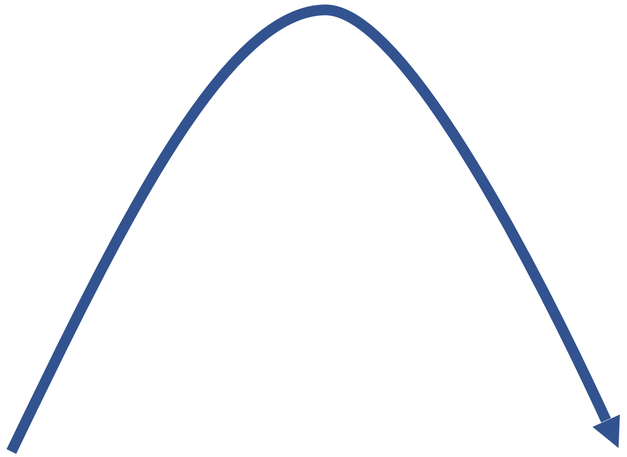
PaCE

- ERC 2022-27
- Trinity College Dublin, Department of Political Science
- Team:
  - Political Science - Hannah Frank
  - Economics - Jian Cao
  - Physics - Gareth Lomax
  - Computer science - Thomas Schincariol
  - Machine Learning - Adalberto Quiros

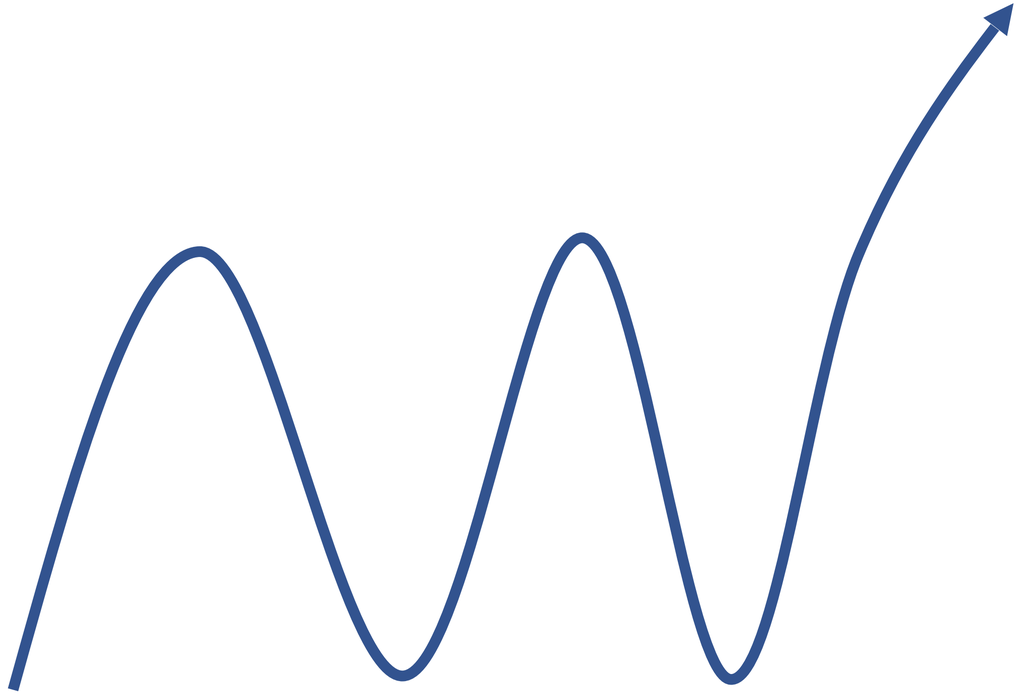


PaCE: Project overview

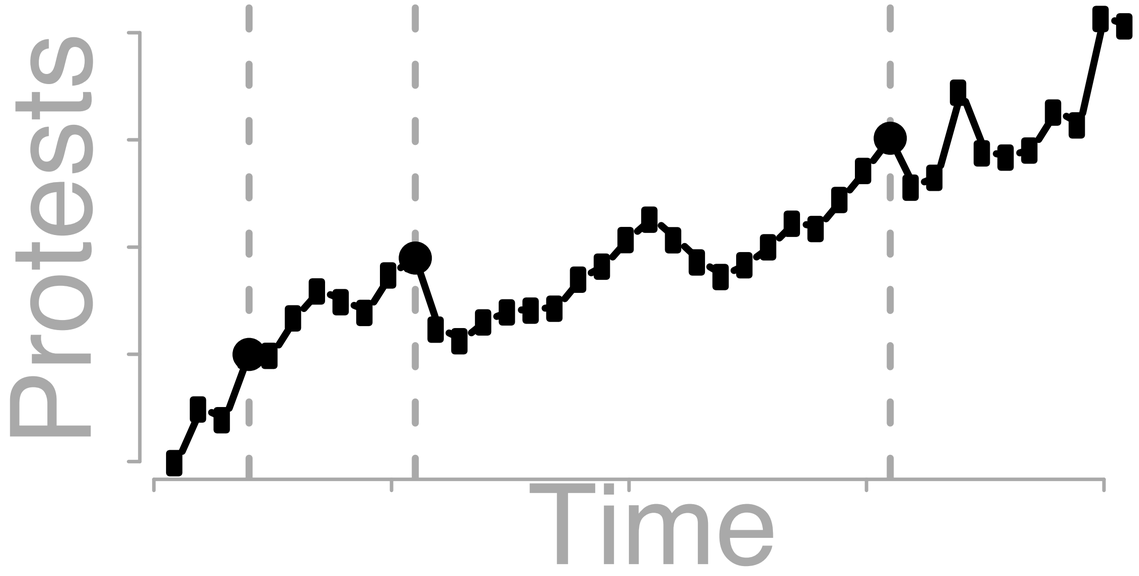
Stories



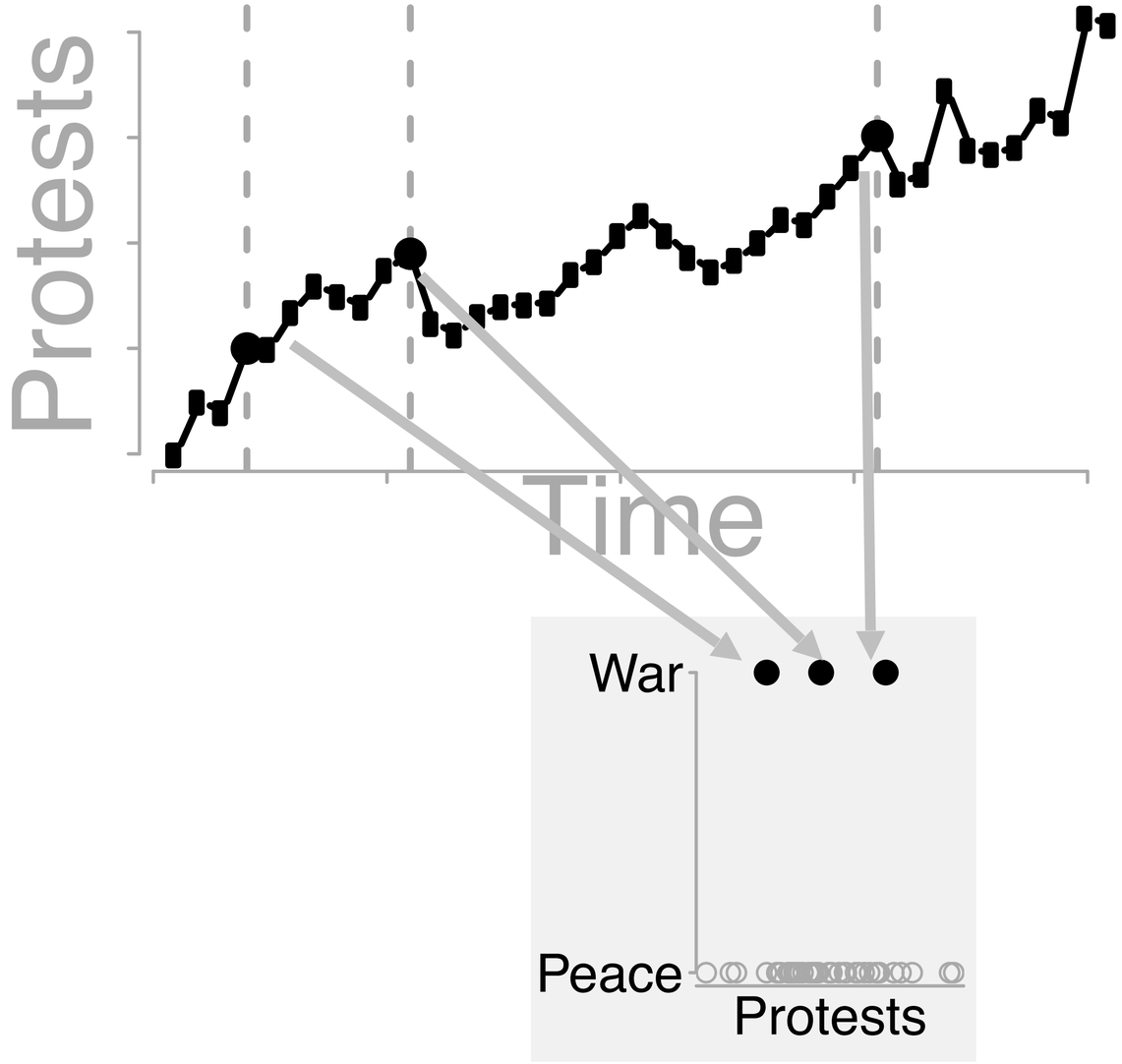
Tragedy



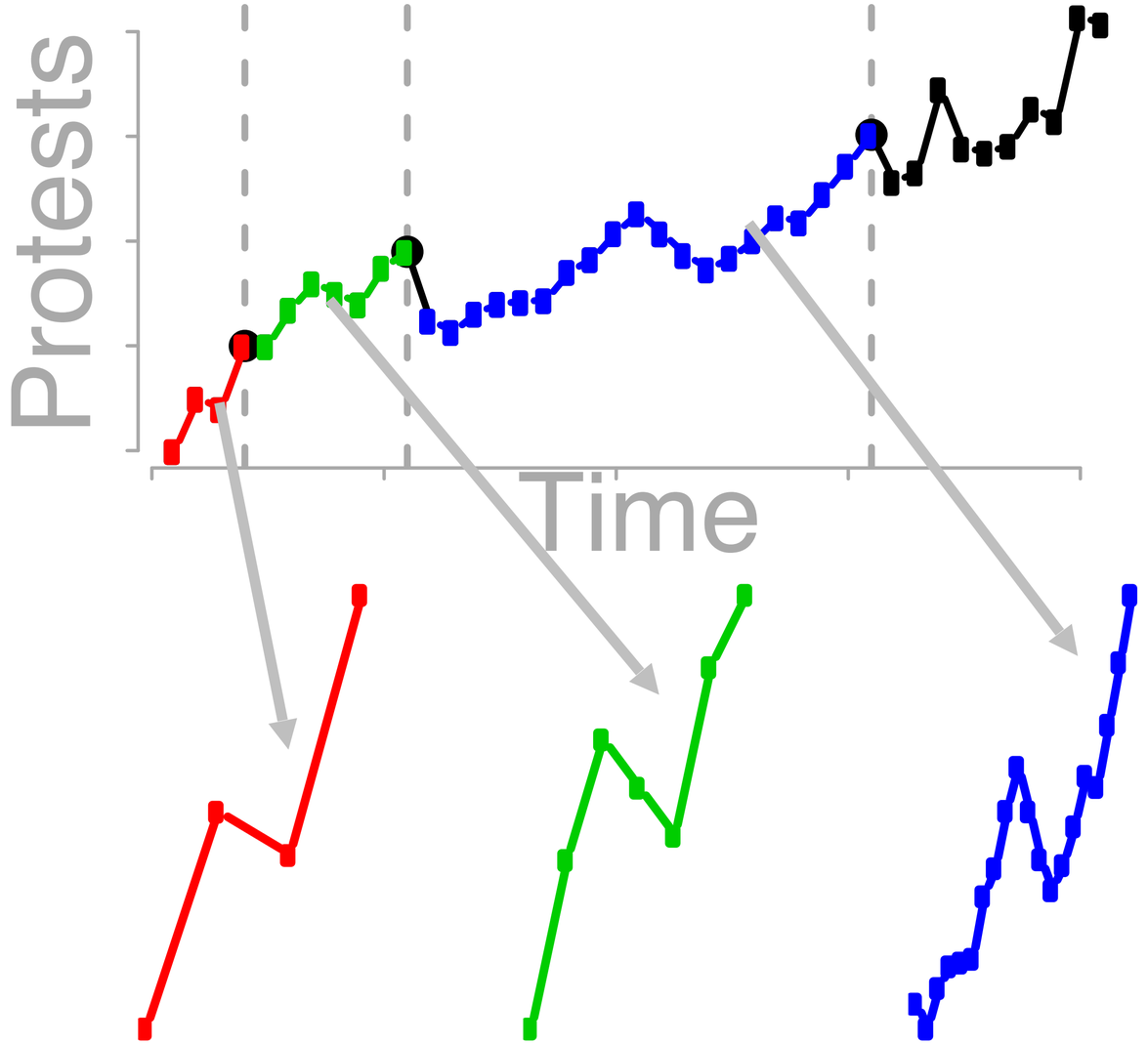
Seven point story


Protests
Time

Observations-
as units of analysis



Sequences
as units of analysis

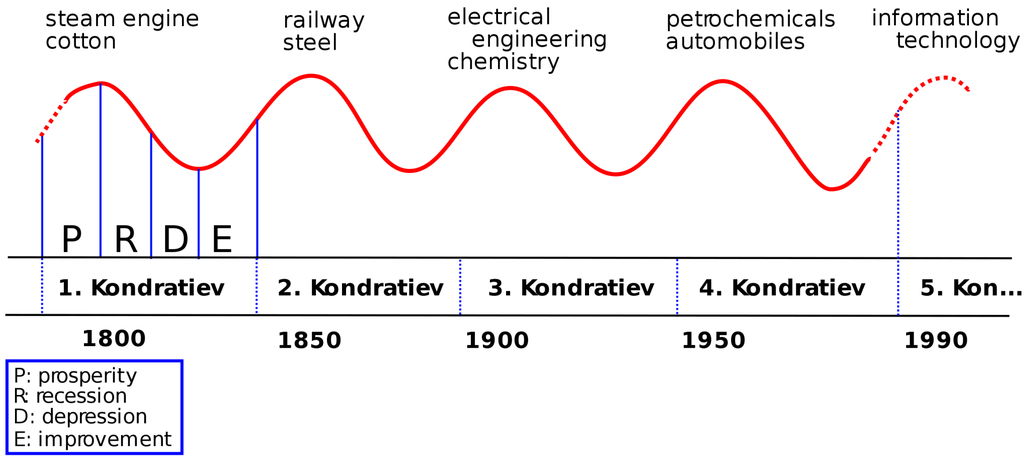


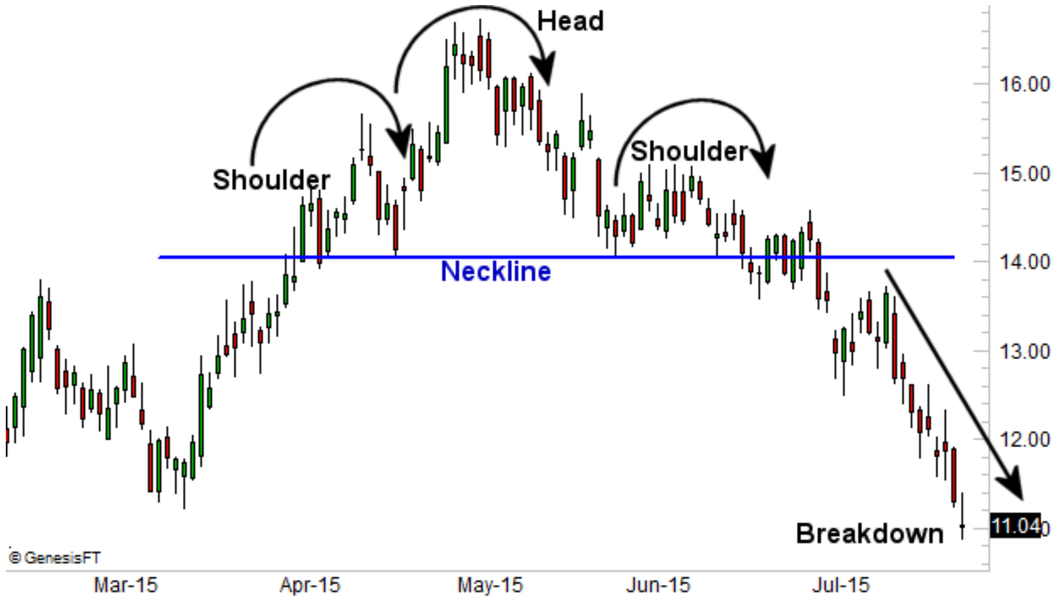
Sequences in other fields

In social sciences

- Patterns
  - Marx: Repeating motif over time
  - Kondratiev waves: alternating intervals of high sectoral growth and intervals of relatively slow growth

- No patterns
  - Efficient market hypothesis
  - Richardson





PaCE: One Big Question

Does history repeat itself?

PaCE: Three questions

1. Are there patterns in the temporal sequences that lead to conflict?

2. Can we use these motifs to predict conflict?

3. Can we cluster these patterns to create new / inform existing theories of conflict

How?

Predict…

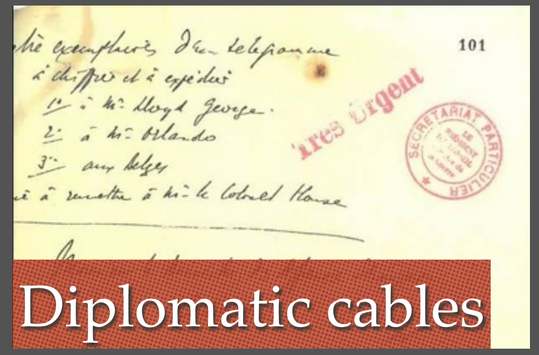
- Interstate and Civil wars / conflict
- Migration flows
- Protests
- Violence
- and more

… using

Events

Financial markets

News



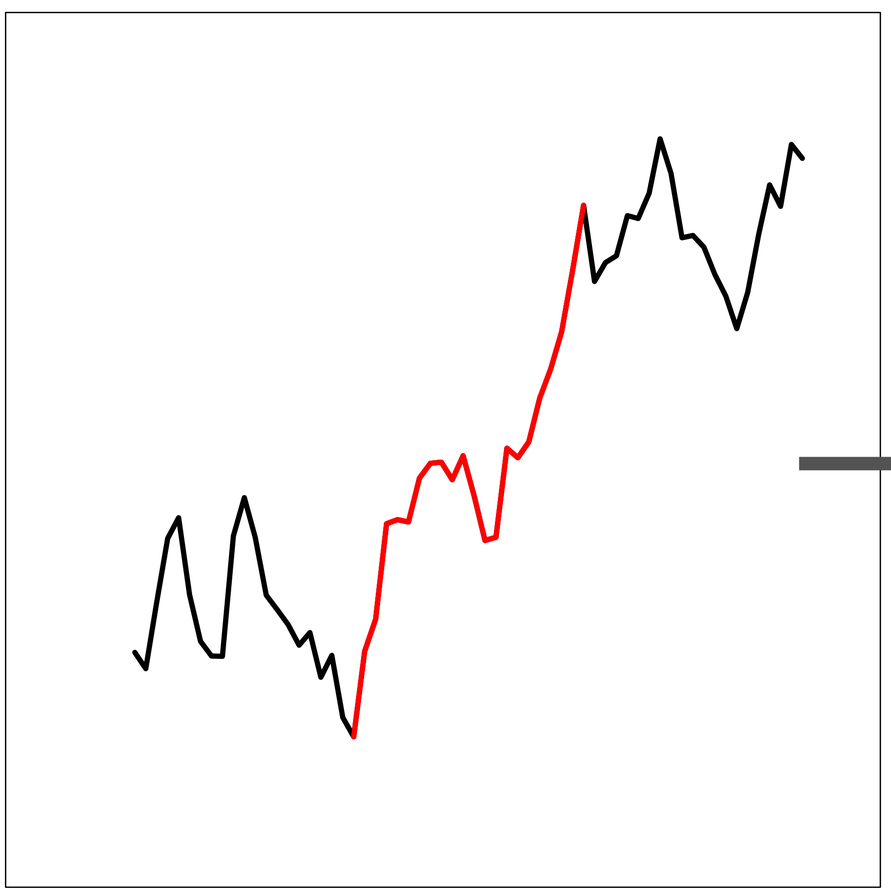
Diplomatic cables

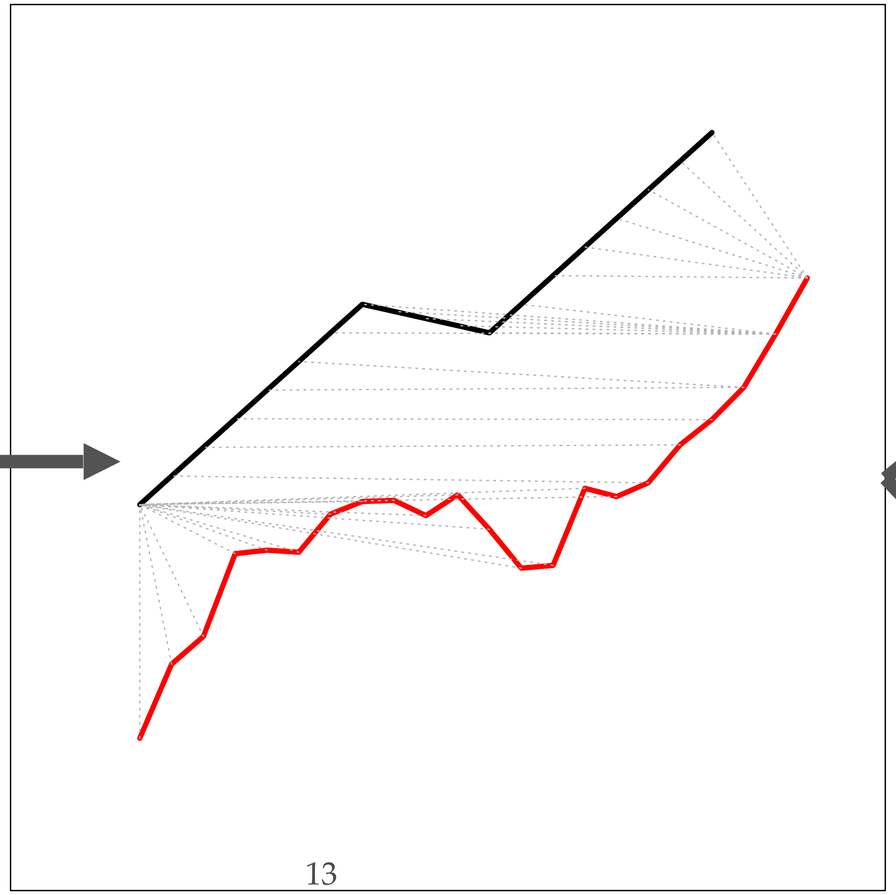
Satellite imagery

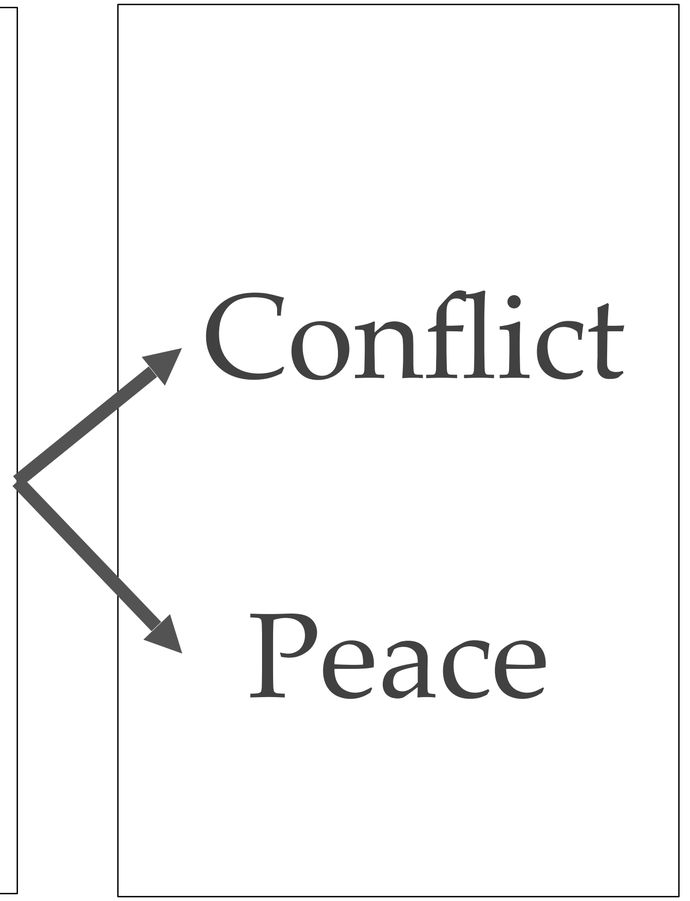
Data acquisition & processing

Distance measures

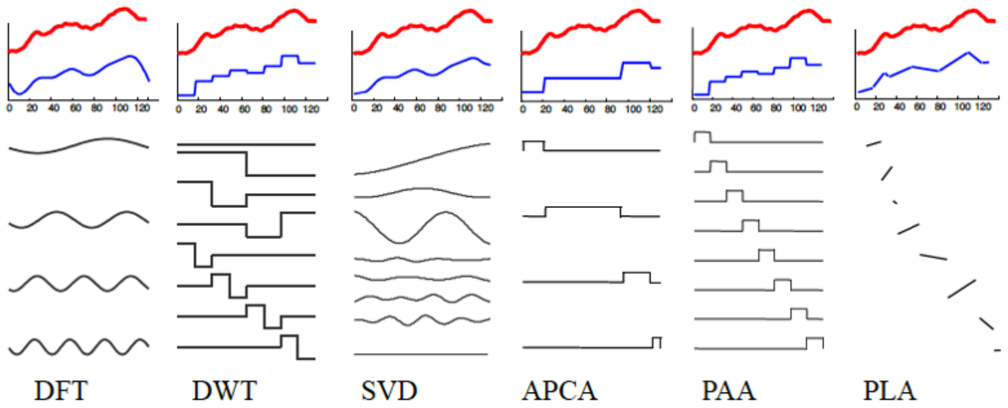
Classification



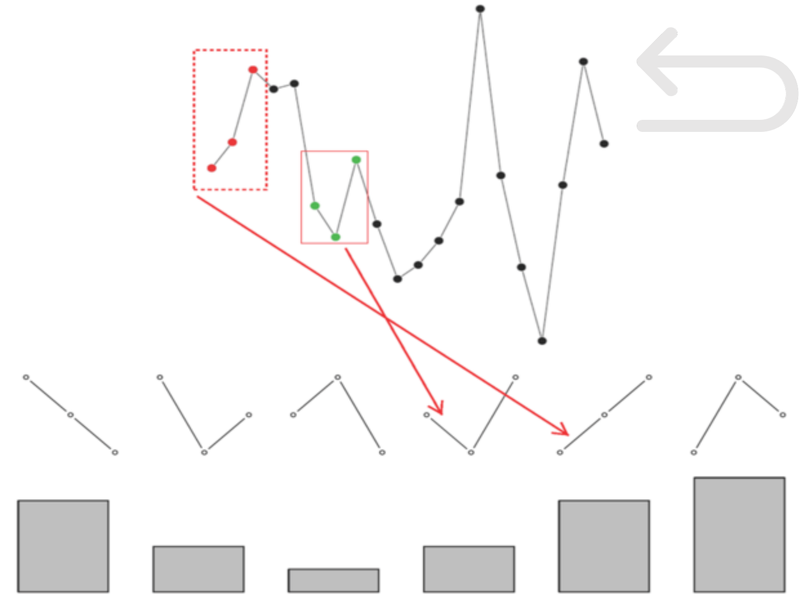




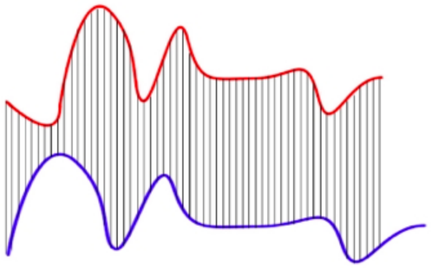
Methodology: Distance Measures

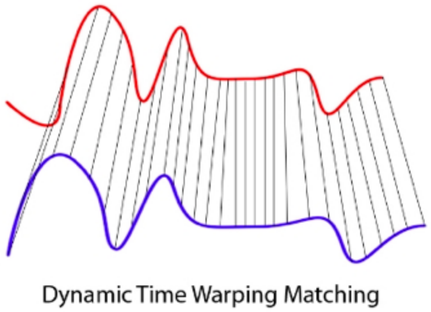


Dimensionality Reduction



Permutation Distribution





Dynamic Time Warping

Clustering

Clustering

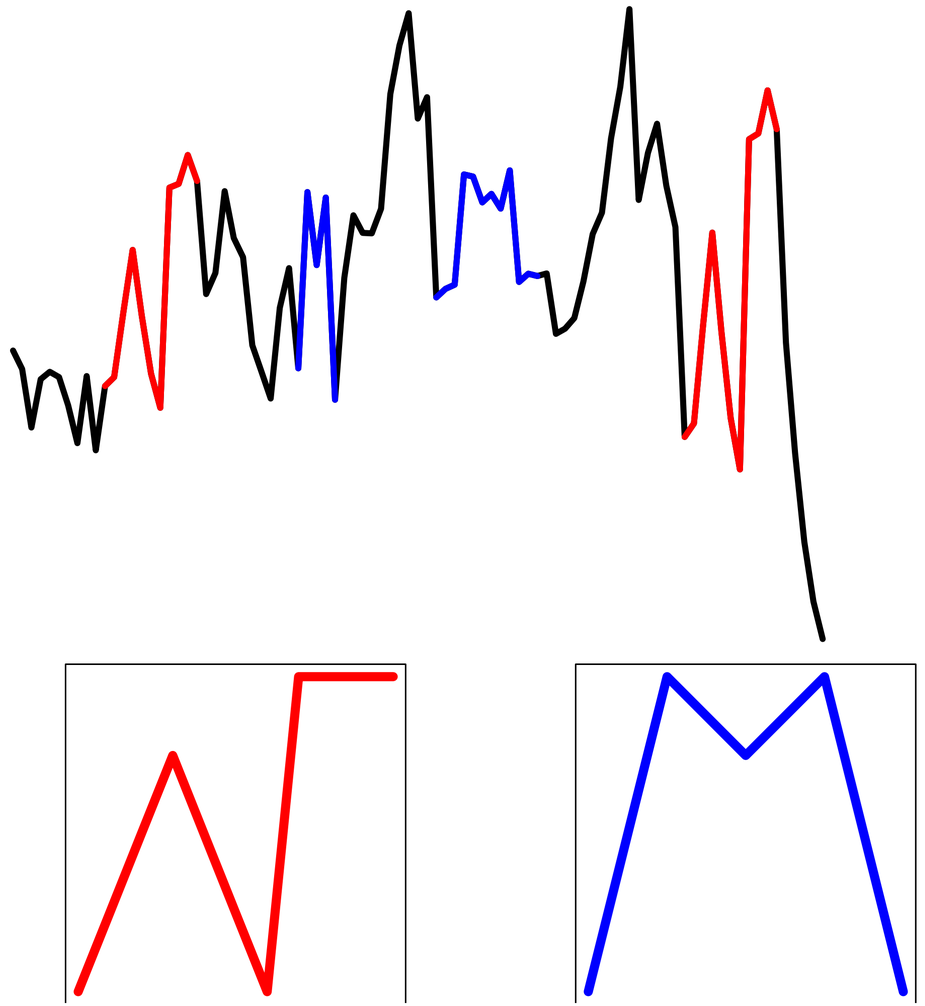
Predictions

Forecasting

- **Backward forecasting**
- **Live forecasting**

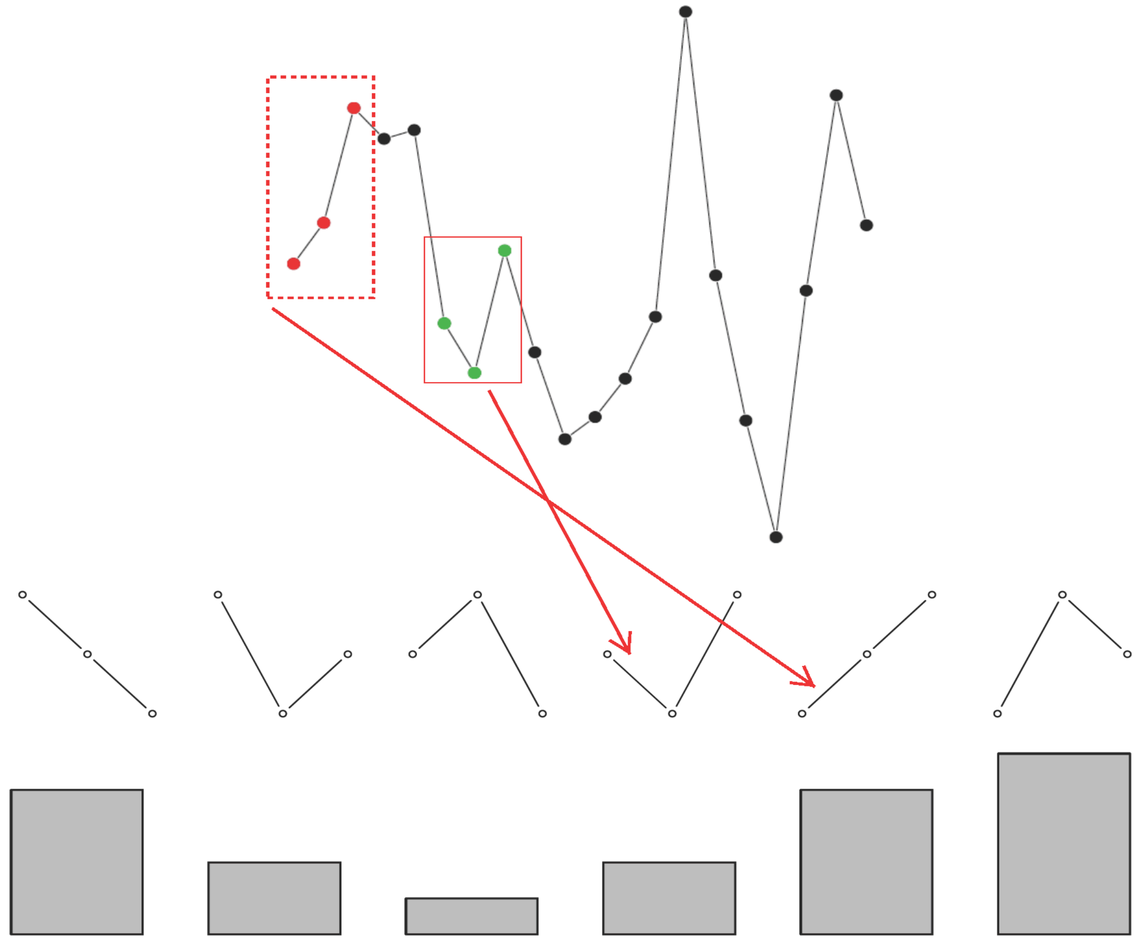
Theoretical implications 1

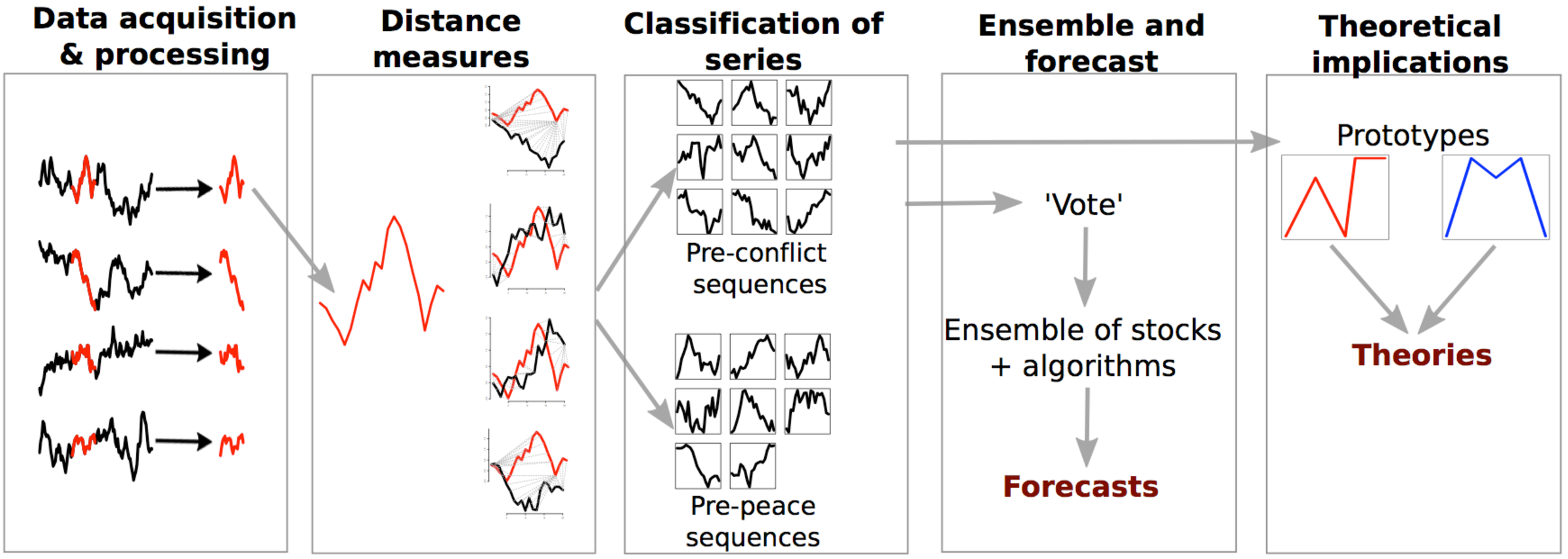
Prototypes for theory building



Theoretical implications 2

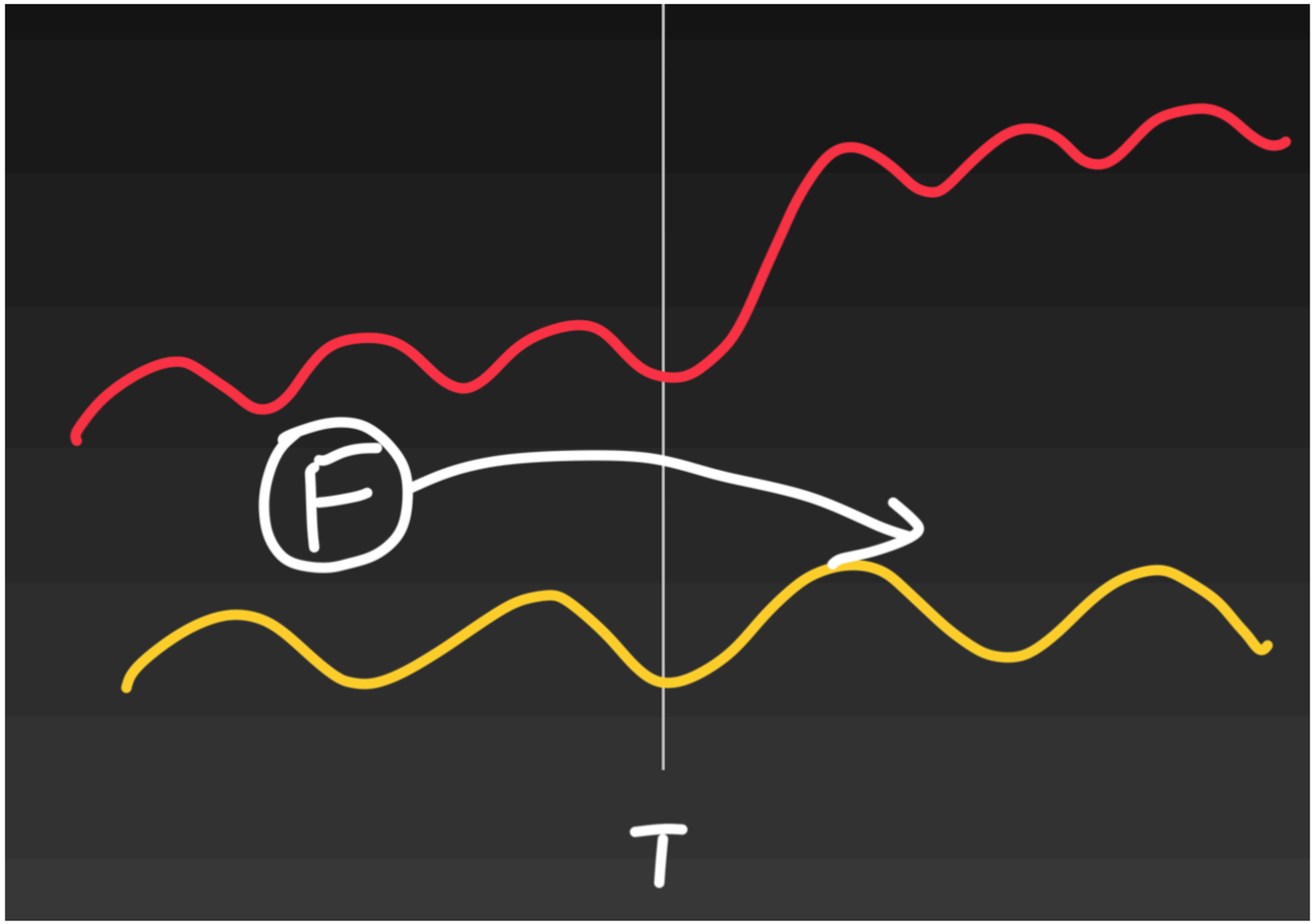
Fundamental limits to the predictability of conflict events?



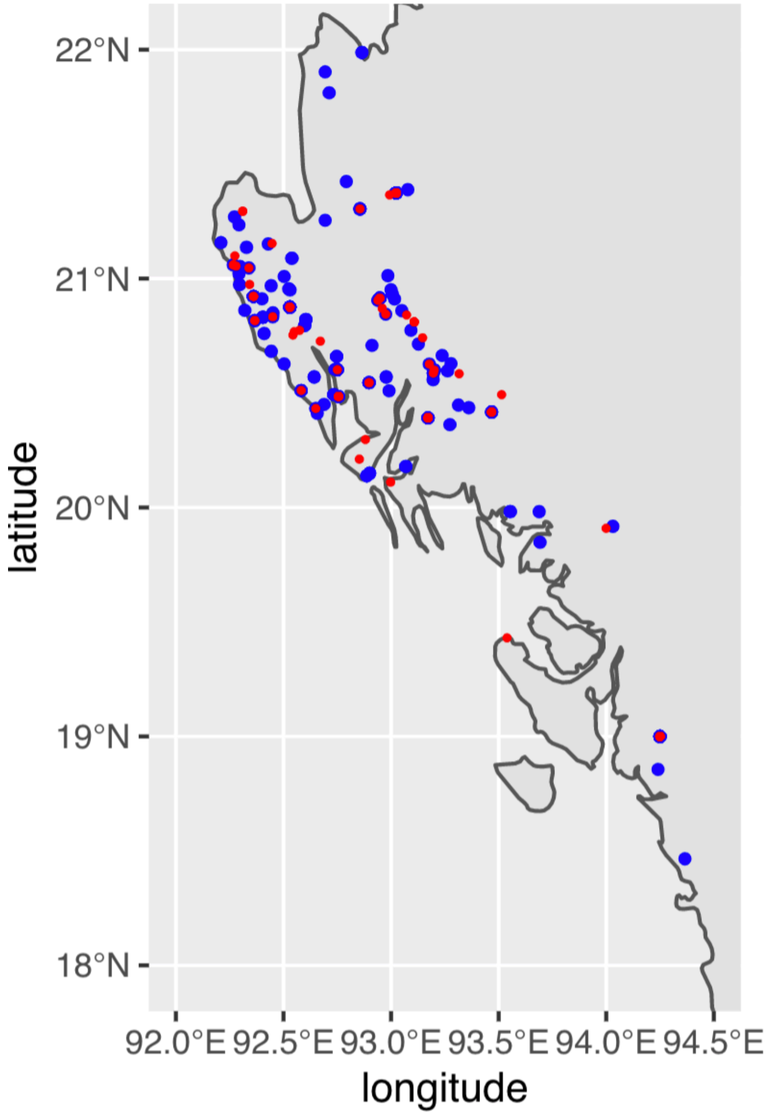

Data acquisition & processing
Distance measures
Classification of series
Ensemble and forecast
Theoretical implications
Pre-conflict sequences
Pre-peace sequences
'Vote'
Ensemble of stocks + algorithms
Forecasts
Prototypes
Theories

Current Projects

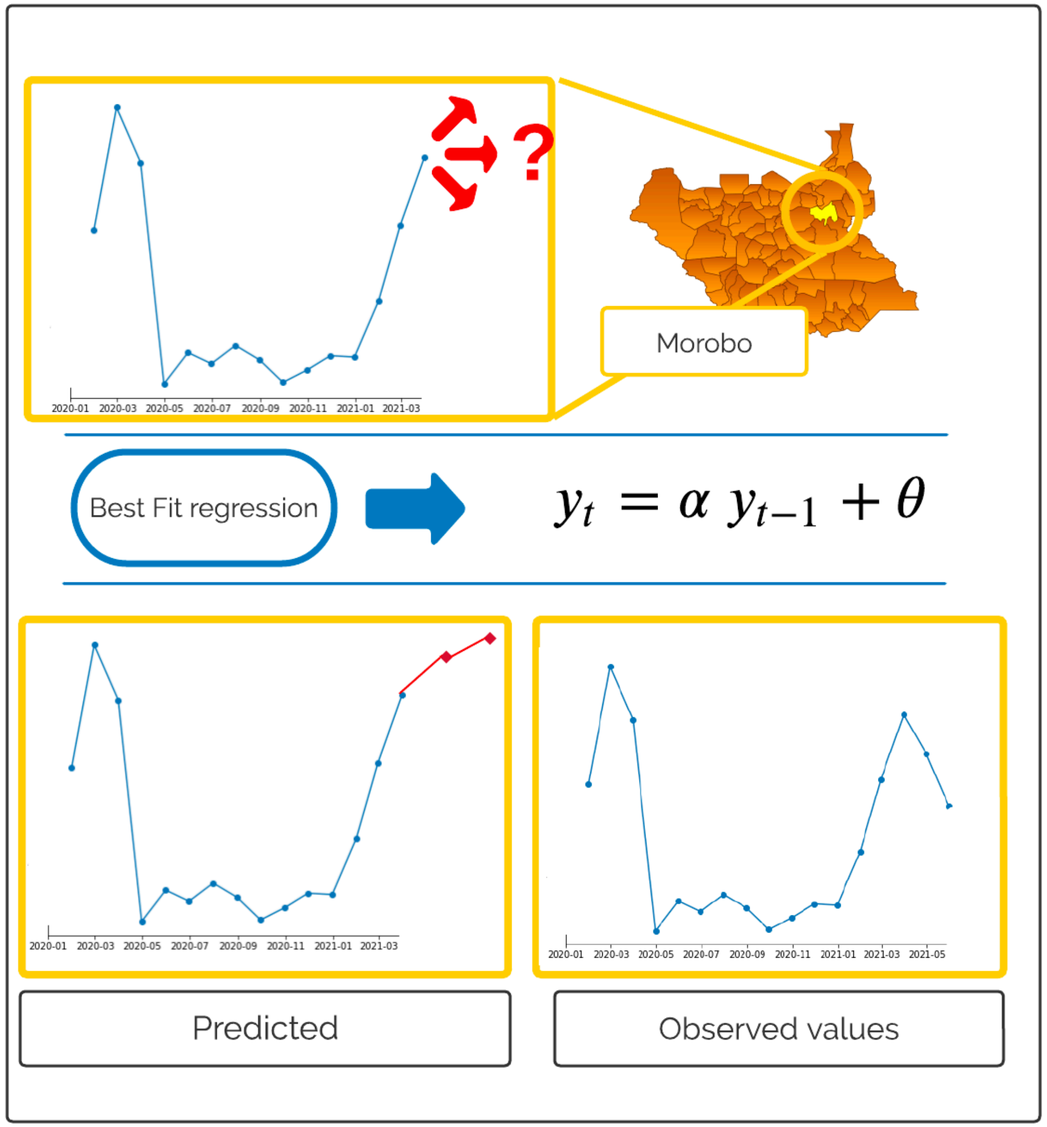
Causal Inference: Dynamic Synthetic Controls

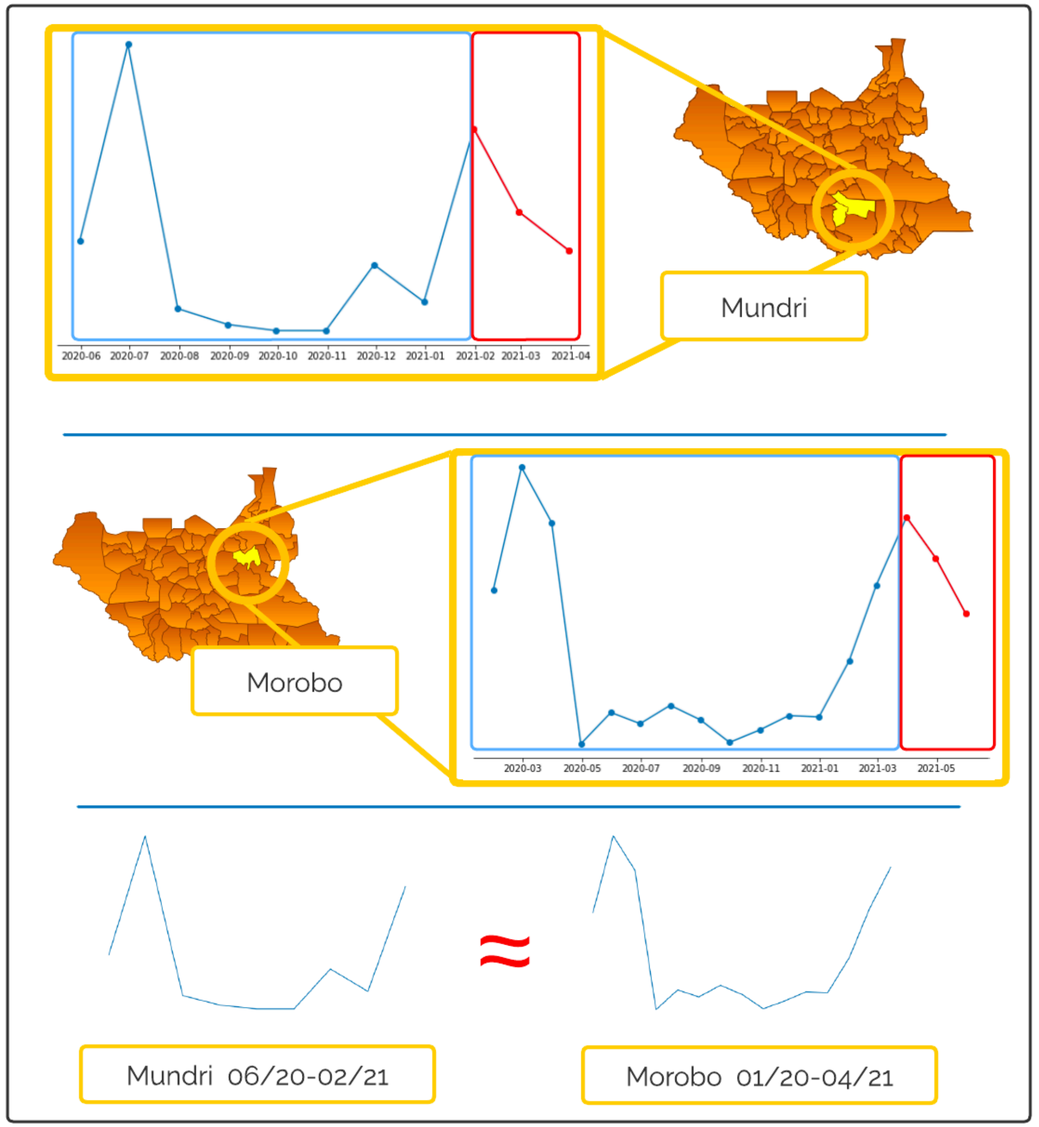


Improving data: temporal accuracy



Example : Migration flow in Morobo - South Sudan

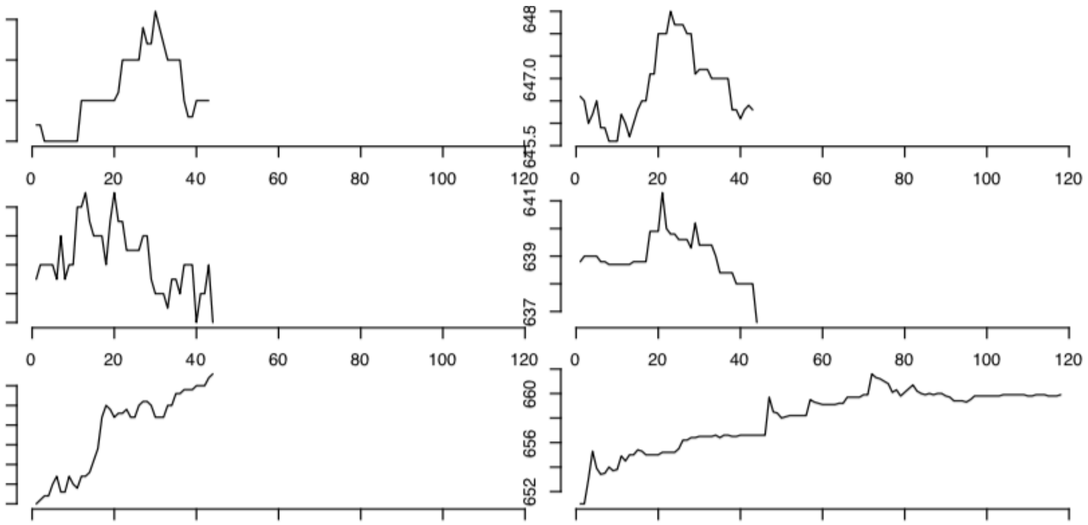


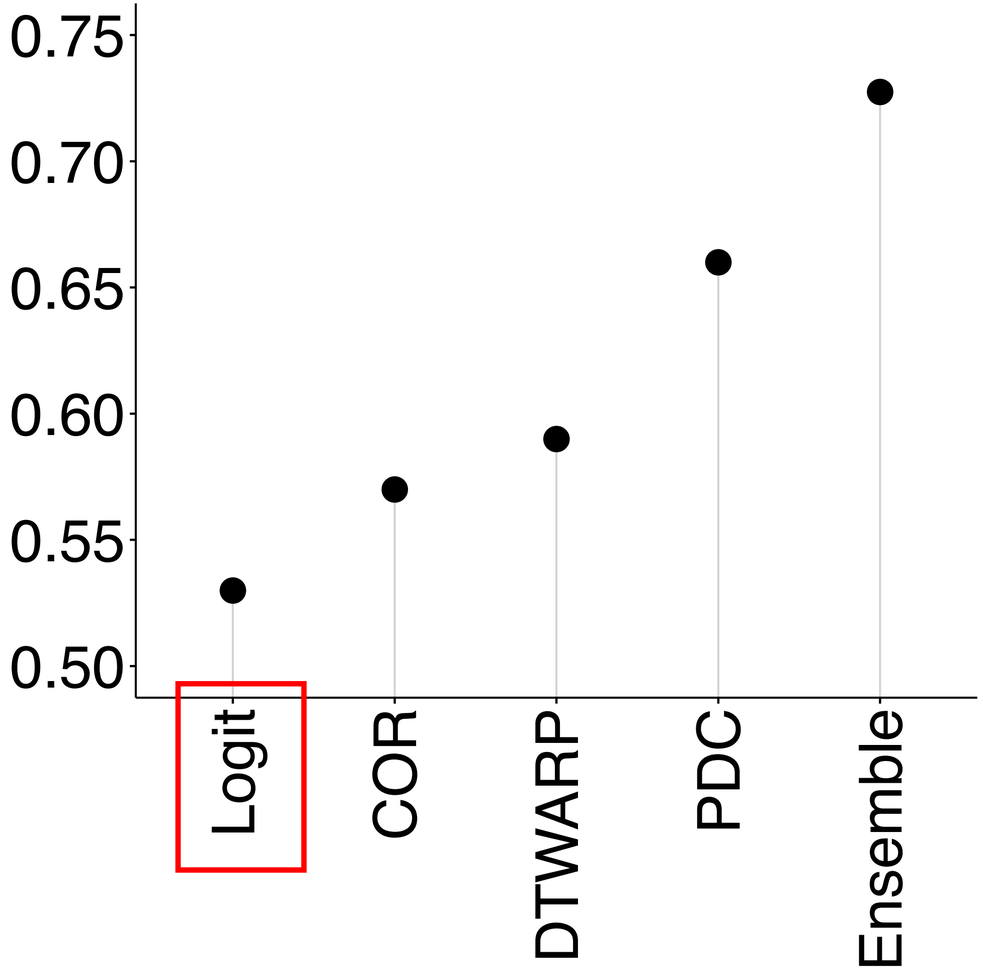


Forecasting: Climate change and conflict

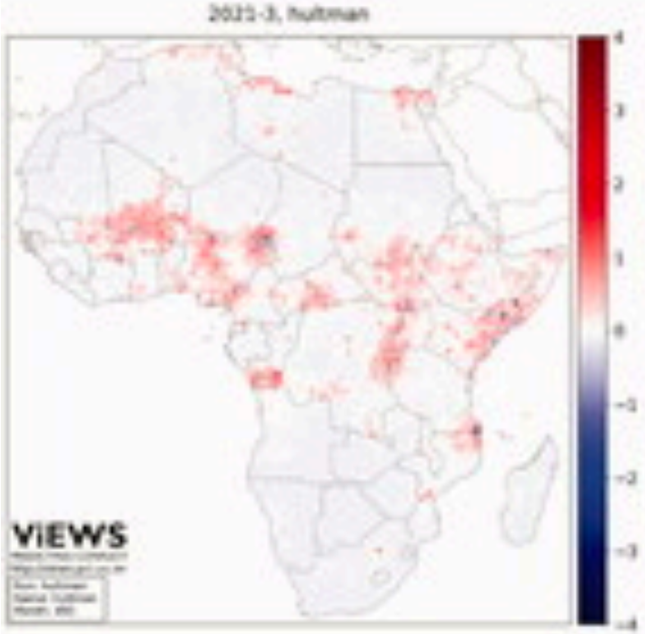
Forecasting: Protests

Forecasting: Palestinian Attacks

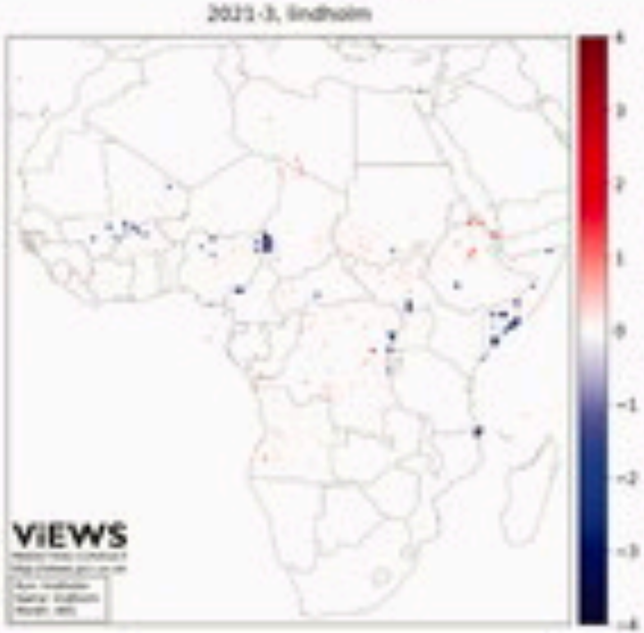




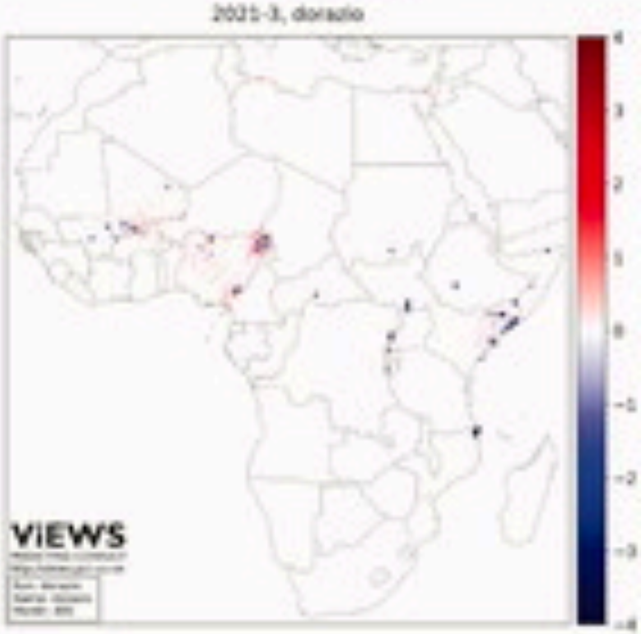
Brandt



Chadefaux



Fritz et al.

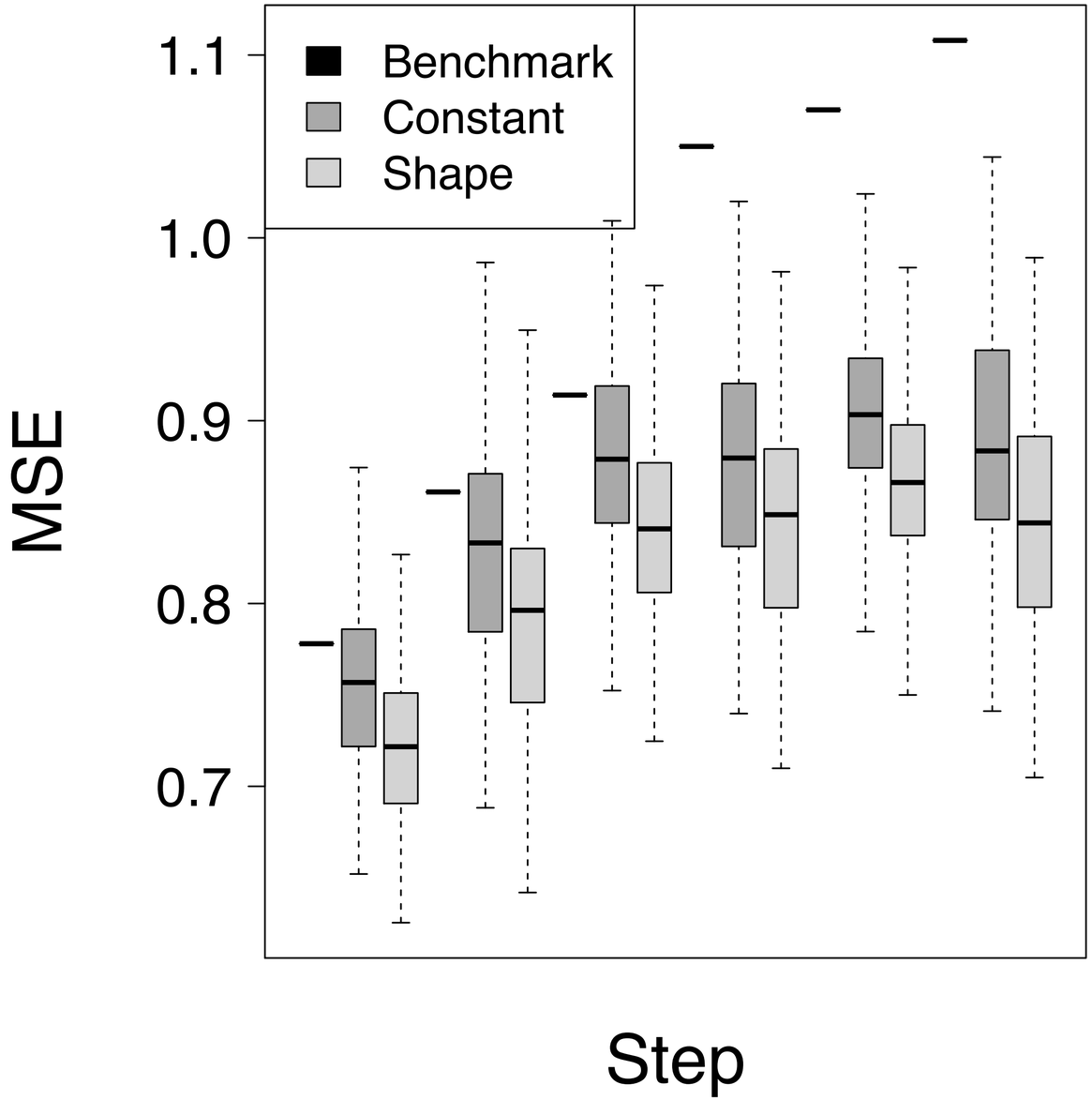


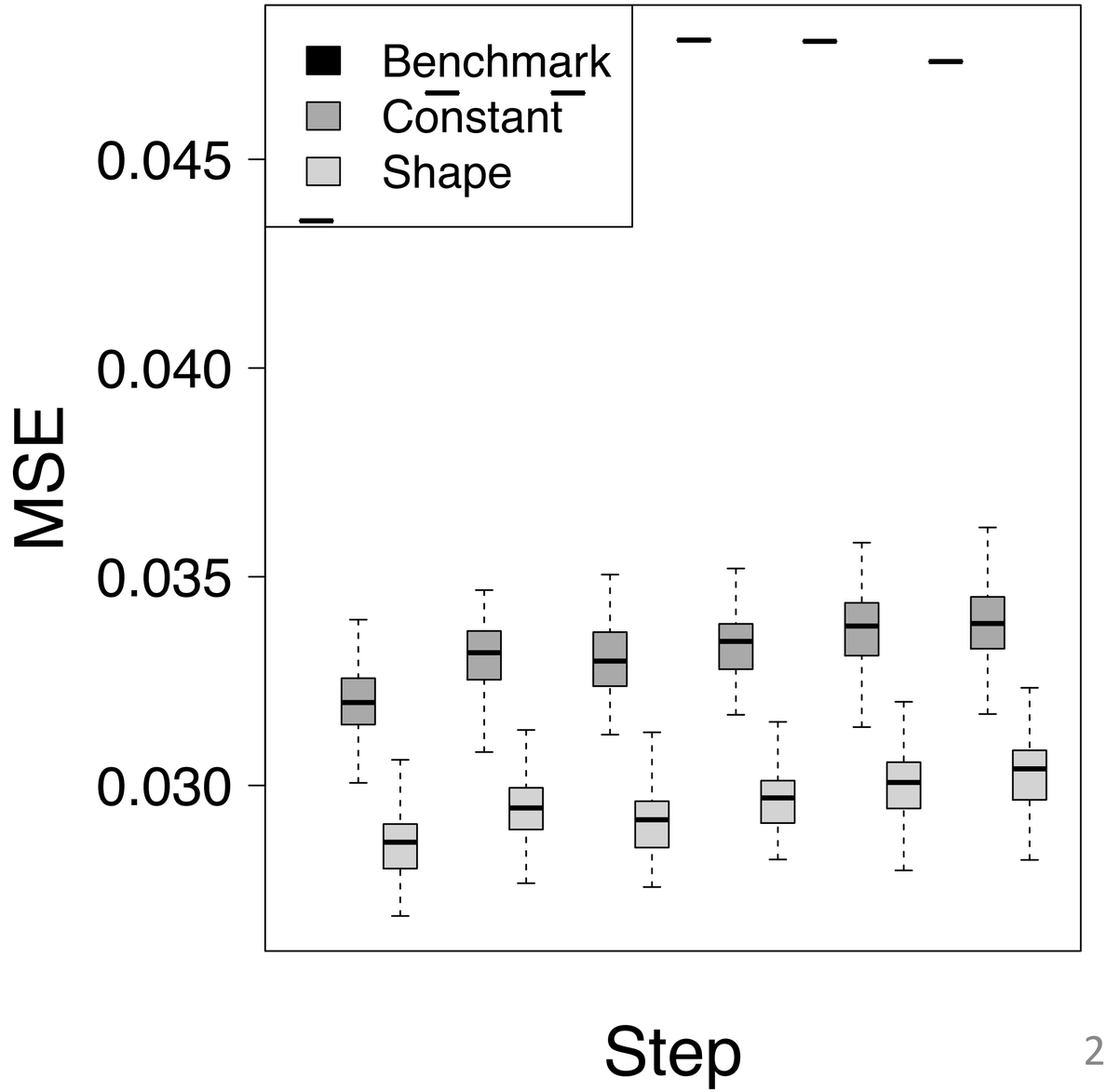
Hultman et al.

Lindholm et al.

D'Orazio & Lin

Forecasting conflict





Diplomatic Correspondence and Conflict Anticipation

Constantine Boussalis Thomas Chadefaux Silvia Decadri

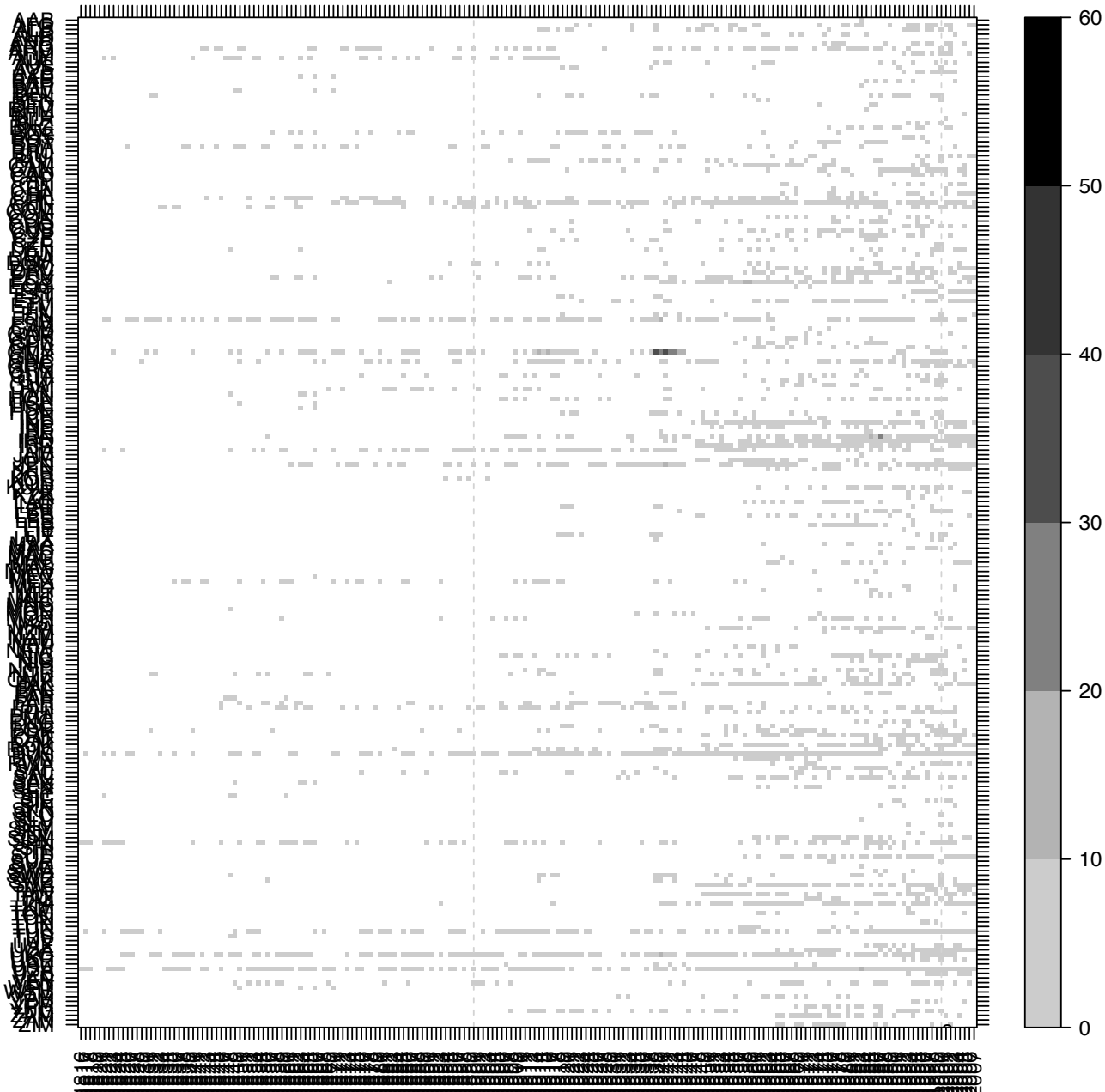
Trinity College Dublin

Workshop on Computational Diplomacy
Crans-Montana, August 30, 2022

Sleepwalkers?

"There will be no European general war. The six great powers cannot afford a clash of arms. [They] will hesitate at the last moment and endeavor to adjust matters." — Los Angeles Times, July 27, 1914

Conflict Anticipation

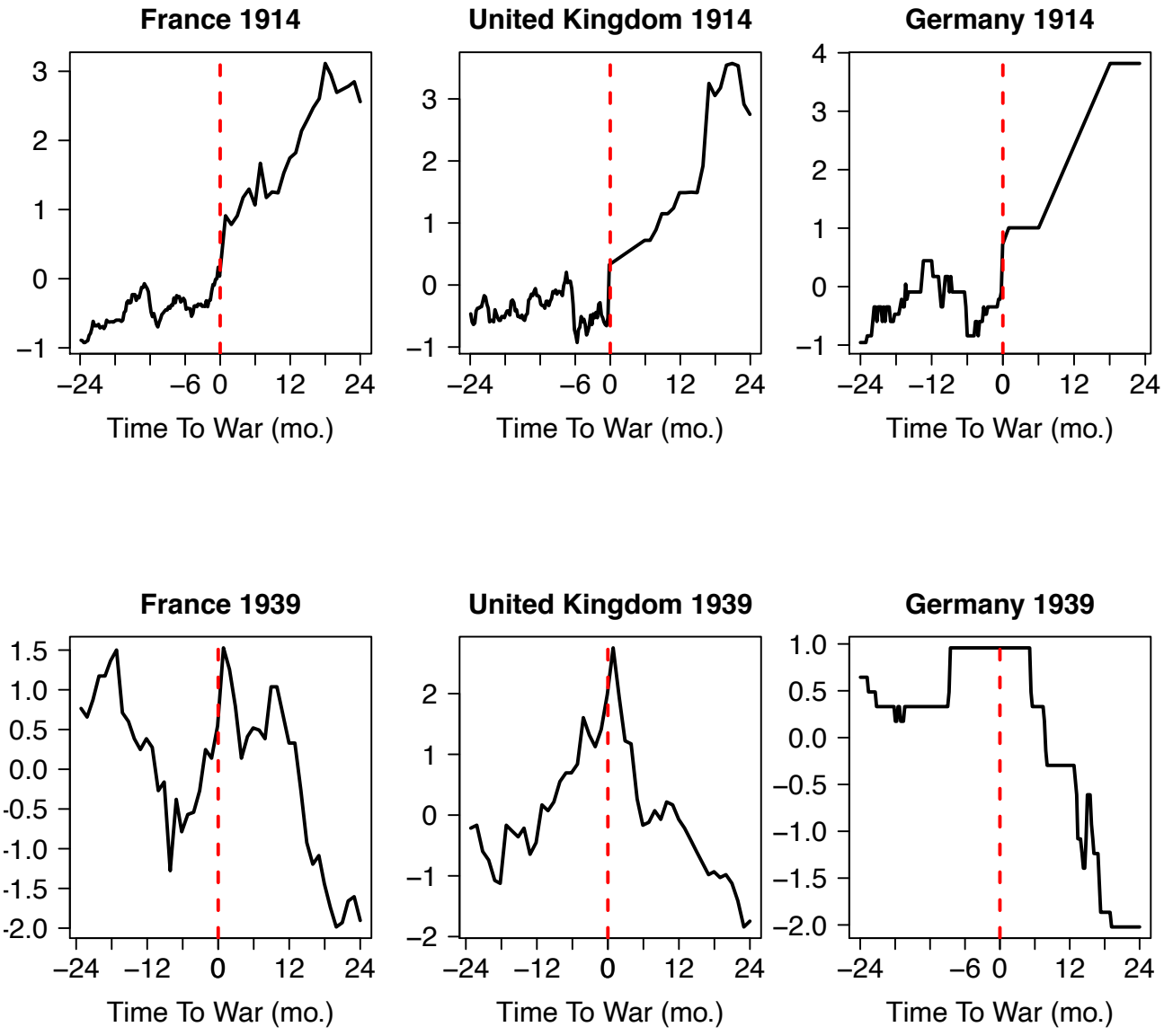


Why do we keep fighting?

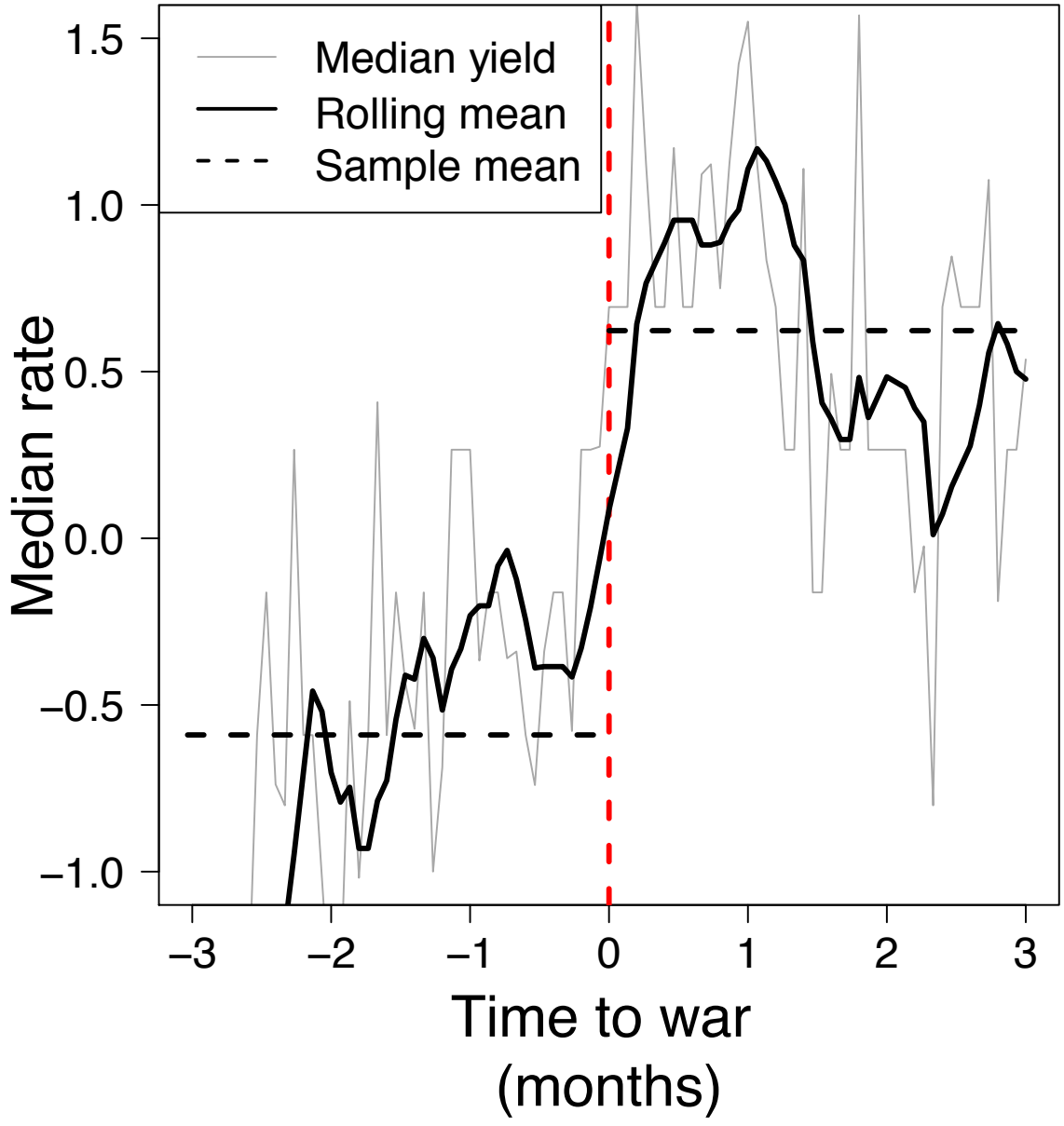
Three possible options:

1. We anticipate correctly but choose to fight
2. We anticipate correctly but cannot prevent war
   e.g., Bargaining failures
3. **We fail to predict the risks and costs of war**

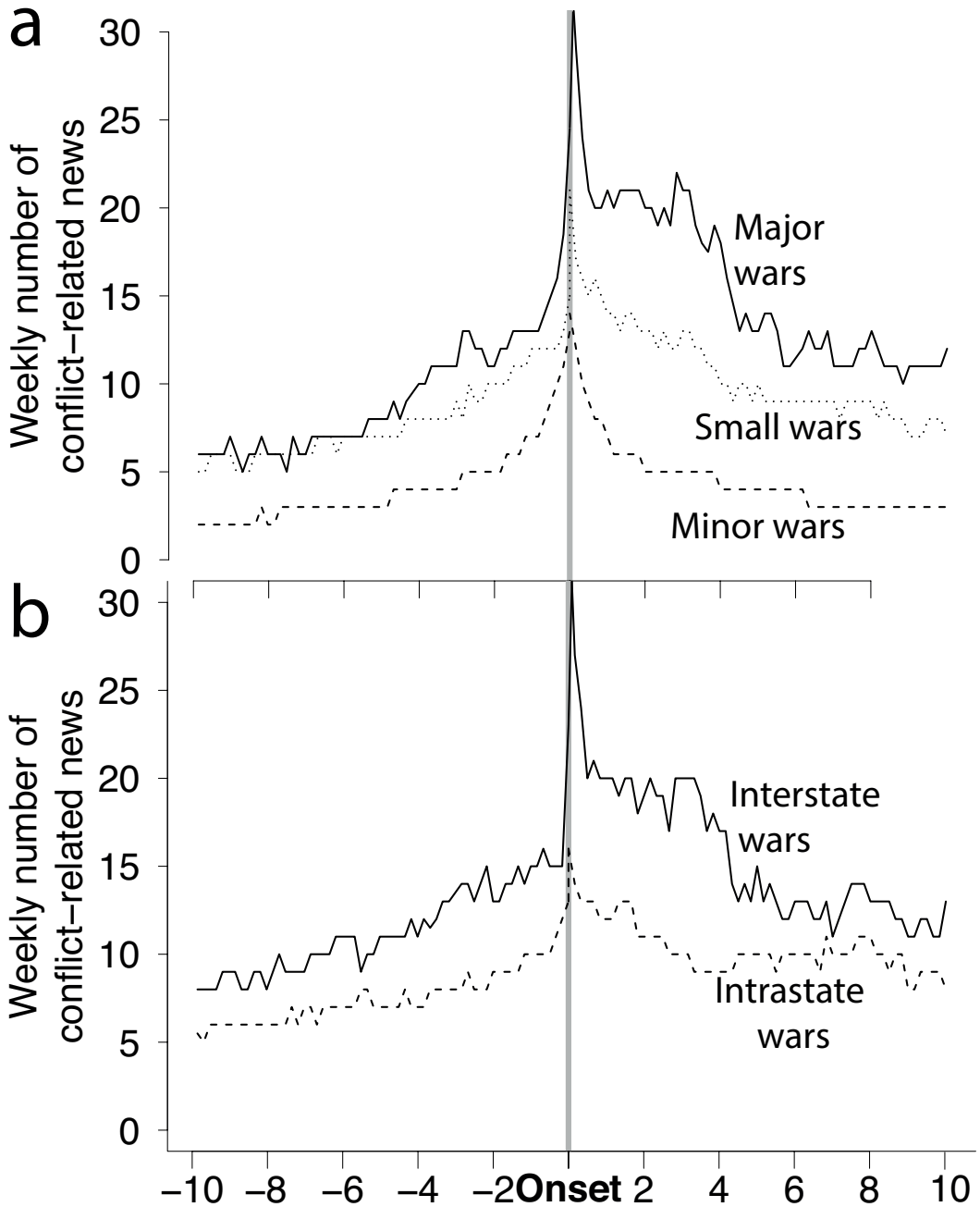
Who predicts? Markets



Who predicts? Bond markets



Who predicts? News



Who predicts?

- Journalists (Chadefaux 2014, Mueller & Rauh 2018). News help, but limits
- Financial Markets (Schneider 2006, Schneider 2017, Chadefaux 2017). Underestimate the risk
- Leaders: no available data.

Motivating question: do insiders know better?

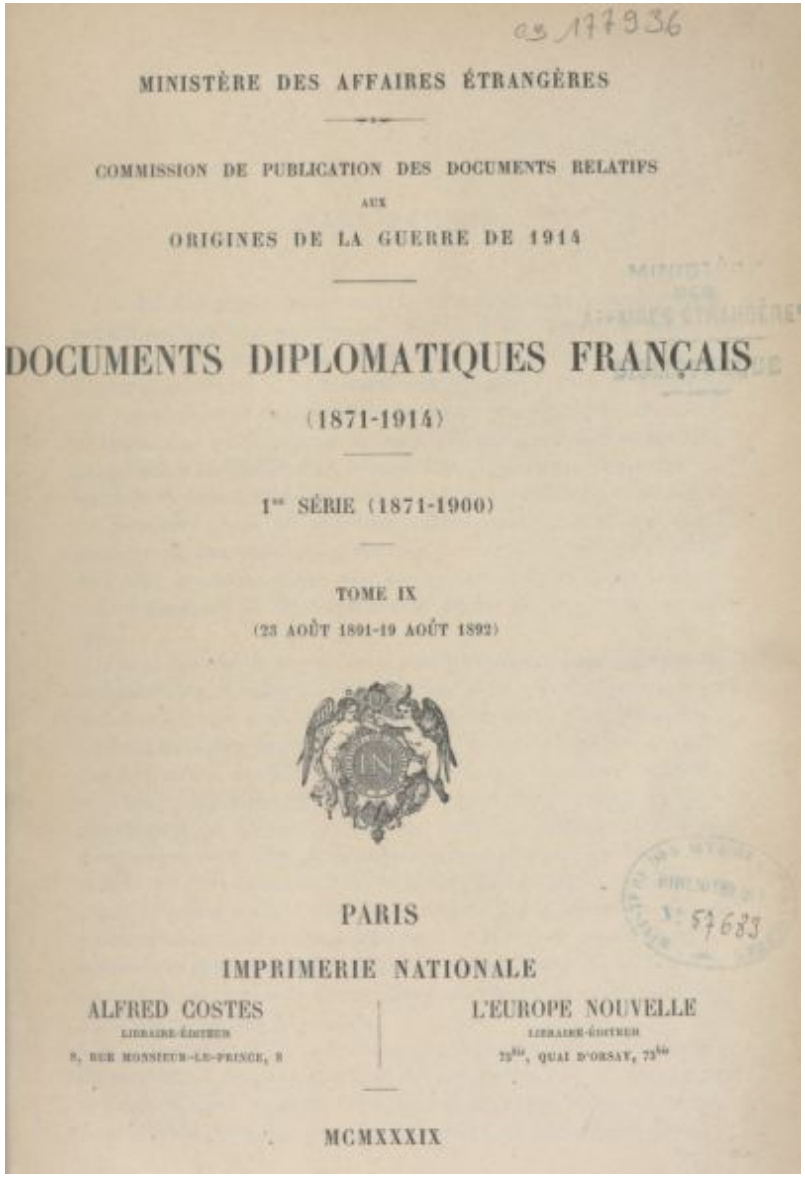
Data

Diplomatic cables

Country	Digitized resources	Online link
France	1871—1973	https://www.diplomatie.gouv.fr/fr/archives-diplomatiques/
Germany	1918-1945	http://www.ifz-muenchen.de/aktuelles/themen/akten-zur-auswaertigen-politik/open-access/
Great Britain	1898-1939	http://dbpo.chadwyck.com/marketing/index.jsp
Austria	NA	https://www.bmeia.gv.at/en/the-ministry/publications/
Canada	1909-1945	http://www.international.gc.ca/history-histoire/documents-documents.aspx
Ireland	1919-1951	http://www.difp.ie/browse-volumes/
Switzerland	1848-1989	http://db.dodis.ch/?lang=fr
USA	1861- 1980	https://history.state.gov/historicaldocuments/pre-truman https://history.state.gov/historicaldocuments

Data

A corpus of French diplomatic cables (1871-1914)



MINISTÈRE DES AFFAIRES ÉTRANGÈRES

COMMISSION DE PUBLICATION DES DOCUMENTS RELATIFS AUX ORIGINES DE LA GUERRE DE 1914

DOCUMENTS DIPLOMATIQUES FRANÇAIS

(1871-1914)

1re SÉRIE (1871-1900)

TOME IX

(23 AOÛT 1891-19 AOÛT 1892)

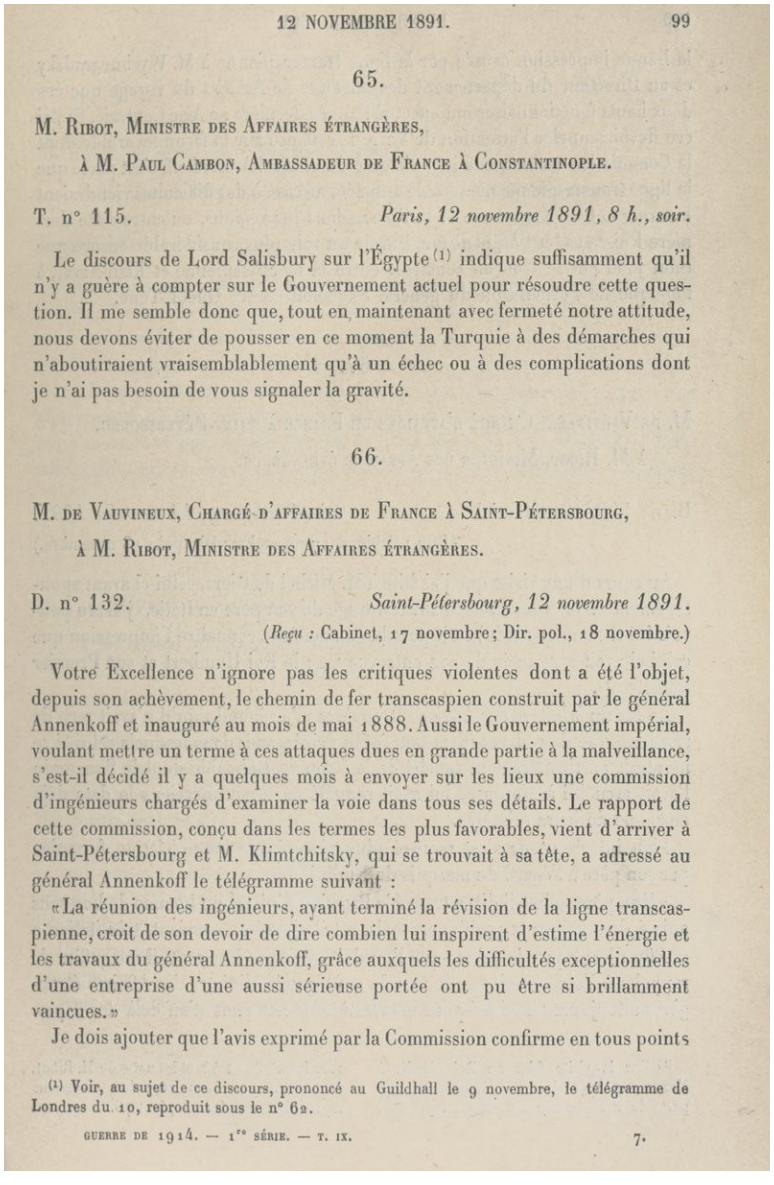
PARIS

IMPRIMERIE NATIONALE

ALFRED COSTES, LIBRAIRE-ÉDITEUR, 8, RUE MONSIEUR-LE-PRINCE, 8

L'EUROPE NOUVELLE, LIBRAIRE-ÉDITEUR, 73bis, QUAI D'ORSAY, 73bis

MCMXXXIX



12 NOVEMBRE 1891. 99

65.

M. RIBOT, MINISTRE DES AFFAIRES ÉTRANGÈRES,

À M. PAUL CAMBON, AMBASSADEUR DE FRANCE À CONSTANTINOPLE.

T. n° 115. *Paris, 12 novembre 1891, 8 h., soir.*

Le discours de Lord Salisbury sur l'Égypte (1) indique suffisamment qu'il n'y a guère à compter sur le Gouvernement actuel pour résoudre cette question. Il me semble donc que, tout en maintenant avec fermeté notre attitude, nous devons éviter de pousser en ce moment la Turquie à des démarches qui n'aboutiraient vraisemblablement qu'à un échec ou à des complications dont je n'ai pas besoin de vous signaler la gravité.

66.

M. DE VAUVINEUX, CHARGÉ D'AFFAIRES DE FRANCE À SAINT-PÉTERSBOURG,

À M. RIBOT, MINISTRE DES AFFAIRES ÉTRANGÈRES.

D. n° 132. *Saint-Pétersbourg, 12 novembre 1891.*

(*Reçu :* Cabinet, 17 novembre; Dir. pol., 18 novembre.)

Votre Excellence n'ignore pas les critiques violentes dont a été l'objet, depuis son achèvement, le chemin de fer transcaspien construit par le général Annenkoff et inauguré au mois de mai 1888. Aussi le Gouvernement impérial, voulant mettre un terme à ces attaques dues en grande partie à la malveillance, s'est-il décidé il y a quelques mois à envoyer sur les lieux une commission d'ingénieurs chargés d'examiner la voie dans tous ses détails. Le rapport de cette commission, conçu dans les termes les plus favorables, vient d'arriver à Saint-Pétersbourg et M. Klimtchitsky, qui se trouvait à sa tête, a adressé au général Annenkoff le télégramme suivant :

« La réunion des ingénieurs, ayant terminé la révision de la ligne transcaspienne, croit de son devoir de dire combien lui inspirent d'estime l'énergie et les travaux du général Annenkoff, grâce auxquels les difficultés exceptionnelles d'une entreprise d'une aussi sérieuse portée ont pu être si brillamment vaincues. »

Je dois ajouter que l'avis exprimé par la Commission confirme en tous points

(1) Voir, au sujet de ce discours, prononcé au Guildhall le 9 novembre, le télégramme de Londres du 10, reproduit sous le n° 62.

GUERRE DE 1914. — 1re SÉRIE. — T. IX. 7.

Data

A corpus of French diplomatic cables (1871-1914)

- Commission de publication des documents relatifs aux origines de la Guerre de 1914-1918.
- 'Documents diplomatiques français relatifs aux origines de la guerre de 1914'
  From End of the Franco-Prussian War (1871) $\rightarrow$ onset of the First World War

Data

A corpus of French diplomatic cables (1871-1914)

"Count Munster came to find me to deny the rumor, circulated by the Parisian correspondent of the Times, about a conversation that would have taken place in Berlin between the Italian King and Prince Bismarck, according to which the Chancellor of the Empire would have pressed Victor-Emmanuel to arrange the return of Savoy and Nice to Italy. The German Ambassador explicitly told me that his communication was official. I rush to transmit it to your Excellency, with the character that Count Munster intended to give it and without adding anything to this guarantee."

Ambassador in London to Minister of Foreign Affairs, May 13 1874.

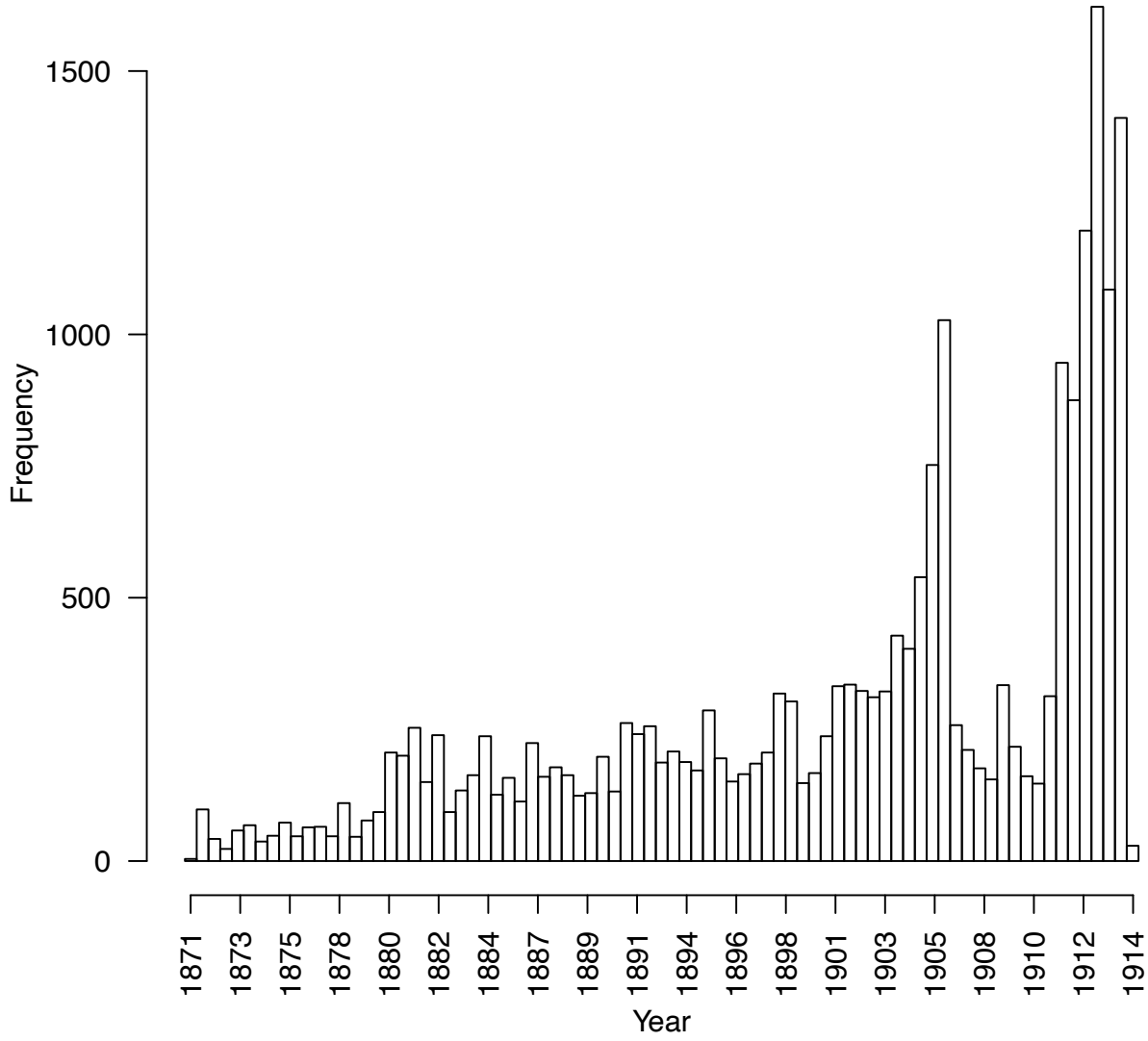
Data

A corpus of French diplomatic cables (1871-1914)

"Should you deem it appropriate, and should other Powers do the same, you are authorized to bring or have someone bring to prince Bismarck our condolences for the heinous assassination attempt against him. It is good in any case that our feelings of indignation are known."

Minister of Foreign Affairs to Ambassador in Berlin, July 14 1874

Data


1500
1000
500
0
Frequency
1871
1873
1875
1878
1880
1882
1884
1887
1889
1891
1894
1896
1898
1901
1903
1905
1908
1910
1912
1914
Year

Data

A corpus of French diplomatic cables (1871-1914)

Some challenges:

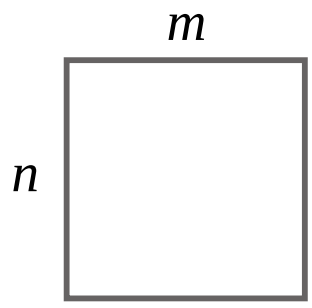
- extract date with regexp.
- 3,849 unmatched $\rightarrow$ Crowdflower
- 553 remaining imputed
- Quality of OCR'd texts; corruption in texts, particularly a problem in older documents.
- French sentiment and emotion detection. Although many advances in recent years for French NLP.
- Cable selection process

Quantitative Text Analysis Pipeline

21,747
diplomatic cables

15,657
issues of Le Figaro

Text pre-processing
Collect meta-data



Document-term
frequency matrix

(1) Prediction of **sentiment** using CamemBERT language model (RoBERTa architecture)

Topic modeling with Dynamic Non-Negative Matrix Factorization (2)

+

(3) Lexicon-based measurement of **emotion** and **CAMEO verbs**

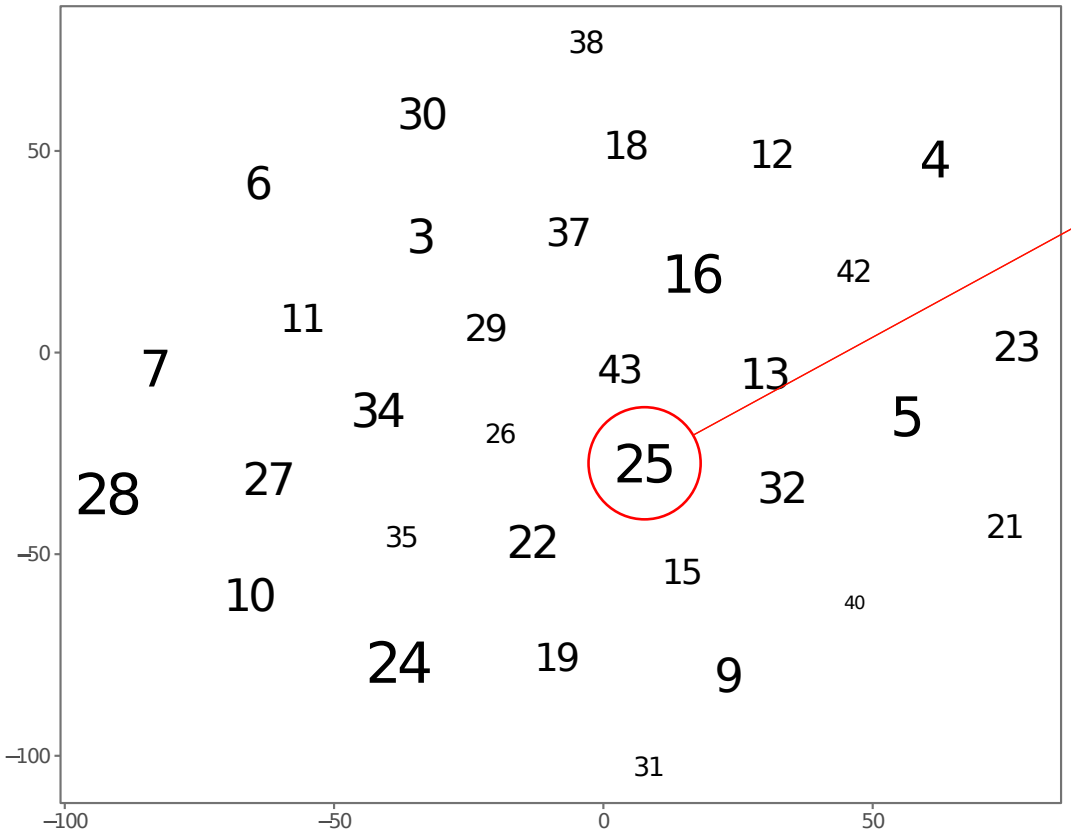
Unsupervised learning of latent topics

Overview of Dynamic NMF

- Rely on *unsupervised* machine learning approach to uncover though the corpus of diplomatic cables
- We use a *dynamic* topic model, which allows topics to evolve over time
- Algorithm takes a tf-idf document-term matrix as input
- Per-document normalized topic weights are the output. We use these data as predictors in our forecasting models.

Topic modeling diplomatic cables

Semantic validity

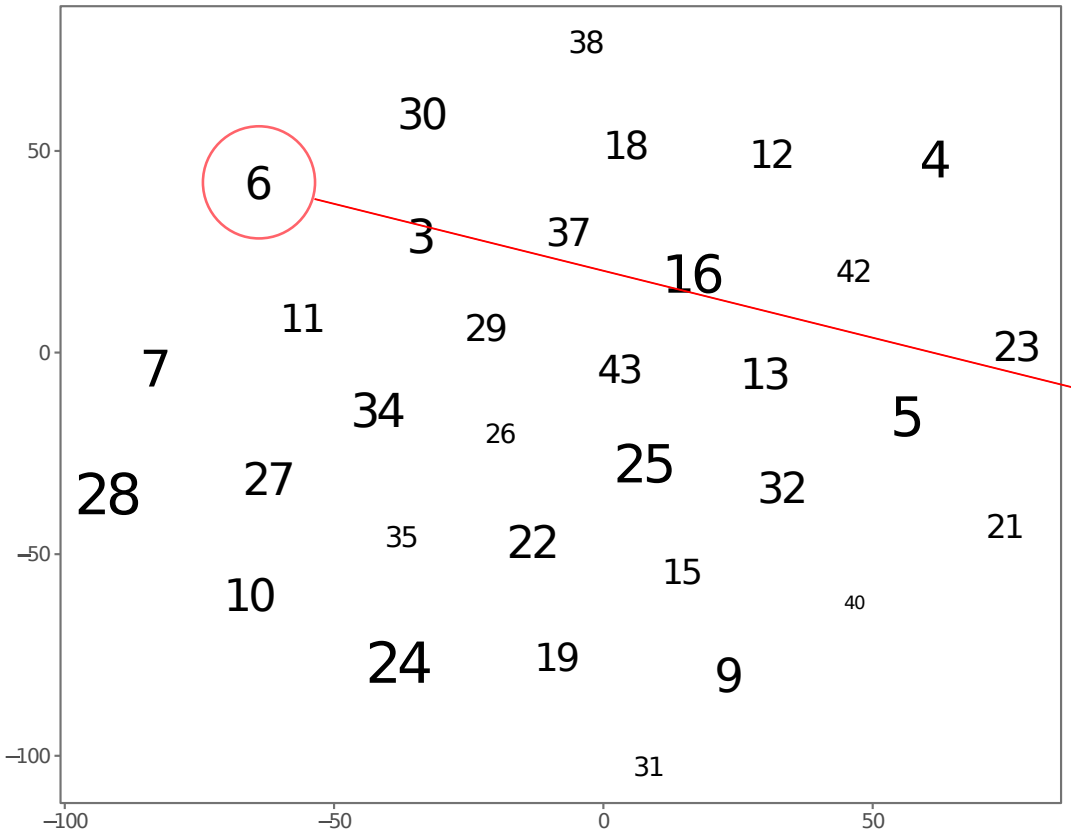


Topic 25: "Military"
troupe general matin guerre mobilisation frontiere ...

3: Germany/Emperor
4: Russia
5: Italy
6: Turkey
7: Germany
9: Egypt/Britain
10: Spain/Morocco
11: China
12: Bulgaria
13: Greece
15: Egypt/Sudan
16: Royalty
18: Serbia/Austria
19: UK Governemnt
21: Japan/Korea
22: Int'l Negotiations
23: Vatican/Pope
24: Germany/Bismark
25: Military
26: Russian Government 2
27: Mahreb
28: Franco-German Conflict
29: Russian Government
30: Turkish Government
31: Portugal
32: Tunisia
34: Finance
35: United States
37: Austria-Hungary
38: Montenegro/Albania
40: Indochina
42: Romania
43: Persia

Topic modeling diplomatic cables

Semantic validity



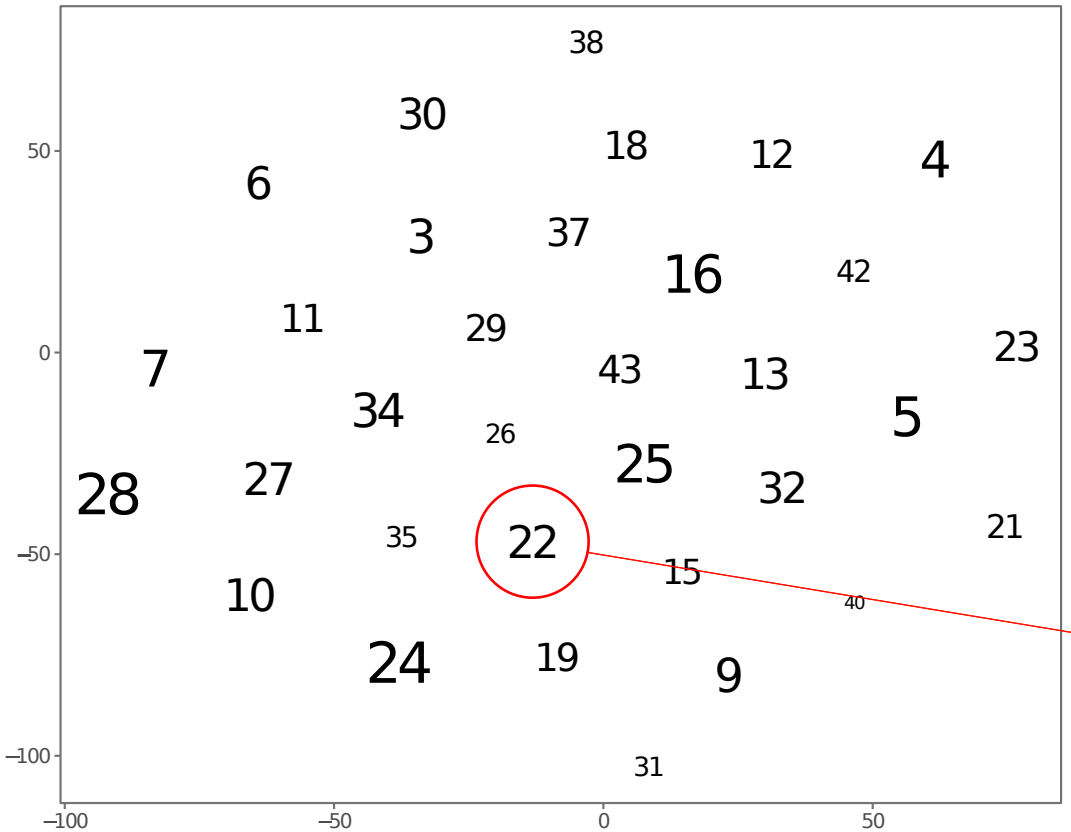
3: Germany/Emperor
4: Russia
5: Italy
6: Turkey
7: Germany
9: Egypt/Britain
10: Spain/Morocco
11: China
12: Bulgaria
13: Greece
15: Egypt/Sudan
16: Royalty
18: Serbia/Austria
19: UK Governemnt
21: Japan/Korea
22: Int'l Negotiations
23: Vatican/Pope
24: Germany/Bismark
25: Military
26: Russian Government 2
27: Mahreb
28: Franco-German Conflict
29: Russian Government
30: Turkish Government
31: Portugal
32: Tunisia
34: Finance
35: United States
37: Austria-Hungary
38: Montenegro/Albania
40: Indochina
42: Romania
43: Persia

Topic 25: "Military"
troupe general matin guerre mobilisation frontiere ...

Topic 6: "Turkey"
sultan majeste constantinople maghzen ...

Topic modeling diplomatic cables

Semantic validity



3: Germany/Emperor
4: Russia
5: Italy
6: Turkey
7: Germany
9: Egypt/Britain
10: Spain/Morocco
11: China
12: Bulgaria
13: Greece
15: Egypt/Sudan
16: Royalty
18: Serbia/Austria
19: UK Governemnt
21: Japan/Korea
22: Int'l Negotiations
23: Vatican/Pope
24: Germany/Bismark
25: Military
26: Russian Government 2
27: Mahreb
28: Franco-German Conflict
29: Russian Government
30: Turkish Government
31: Portugal
32: Tunisia
34: Finance
35: United States
37: Austria-Hungary
38: Montenegro/Albania
40: Indochina
42: Romania
43: Persia

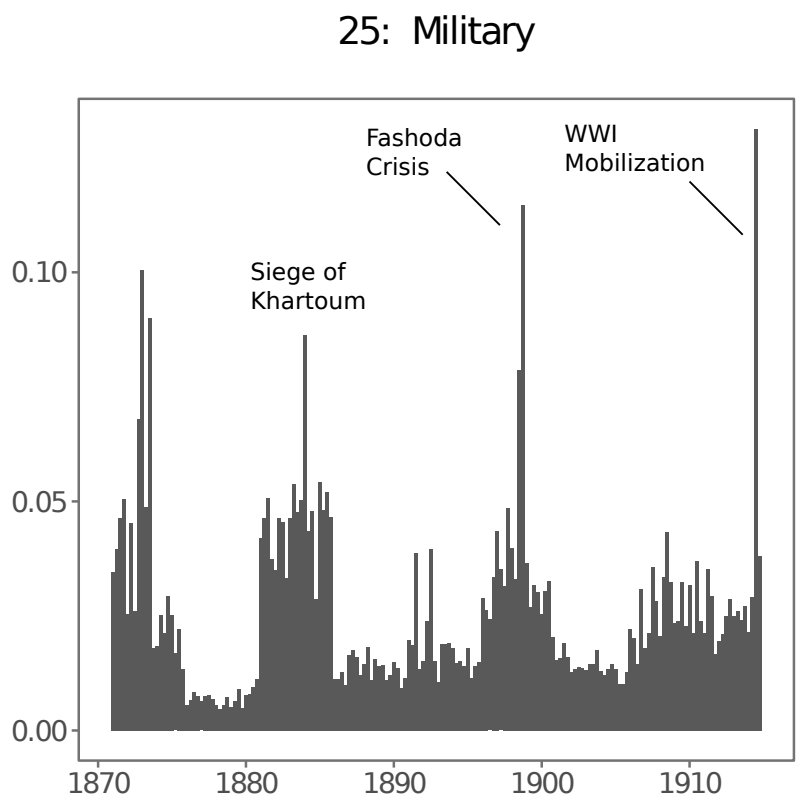
Topic 25: "Military"
troupe general matin guerre mobilisation frontiere ...

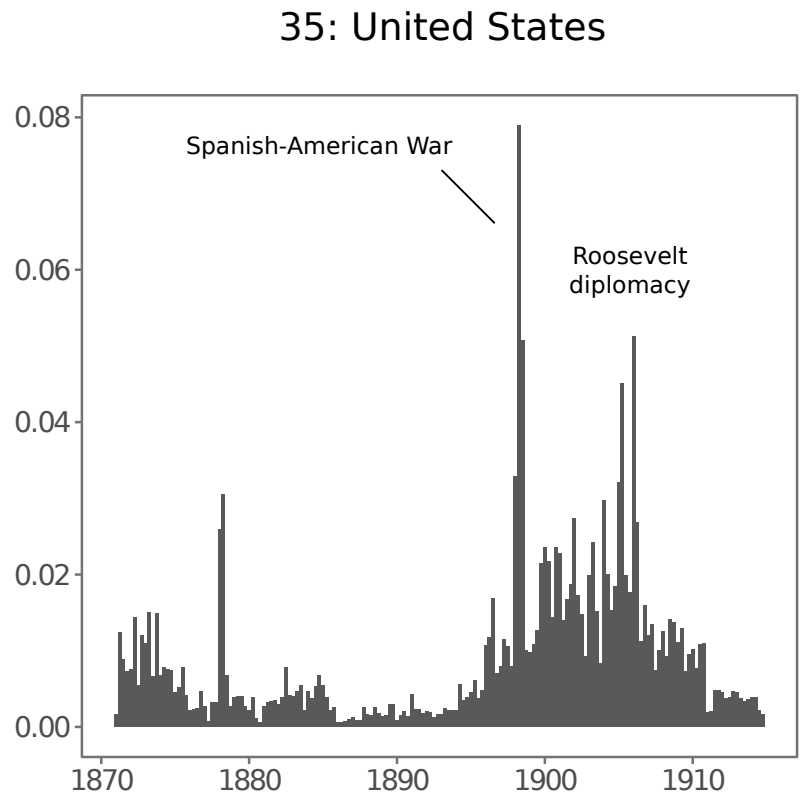
Topic 6: "Turkey"
sultan majeste constantinople maghzen ...

Topic 22: "Int'l Negot."
conference proposition puissance devoir question ...

Topic modeling diplomatic cables

Predictive validity of a selection of topics





Data

Newspaper articles from *Le Figaro* (1871-1914)

Forecasting MIDs with French diplomatic themes

- 480 months, 36 of which have a MIDs involving France
- On average 1 year and 8 months between MIDs (median = 1 year)

Sample Disputes

Larger disputes:

- Franco-Siamese war 1893
- Boxer rebellion 1900
- Sino-French war 1883
- WWI

Also smaller ones:

- Nov. 1885 to Dec. 1885, France–Germany over Togoland. Germany threatened.

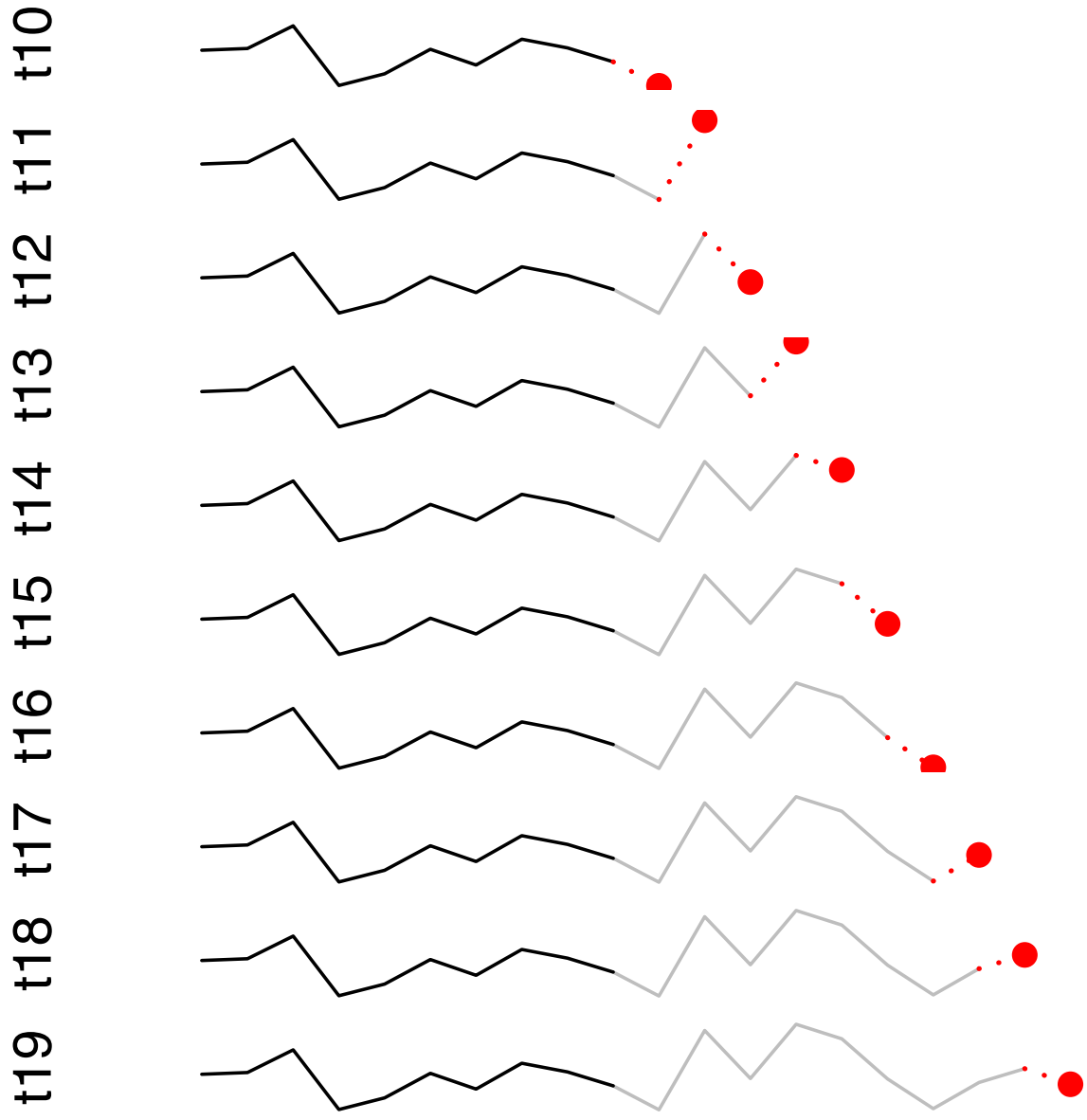
Forecasting MIDs with French diplomatic themes

Overview of statistical model

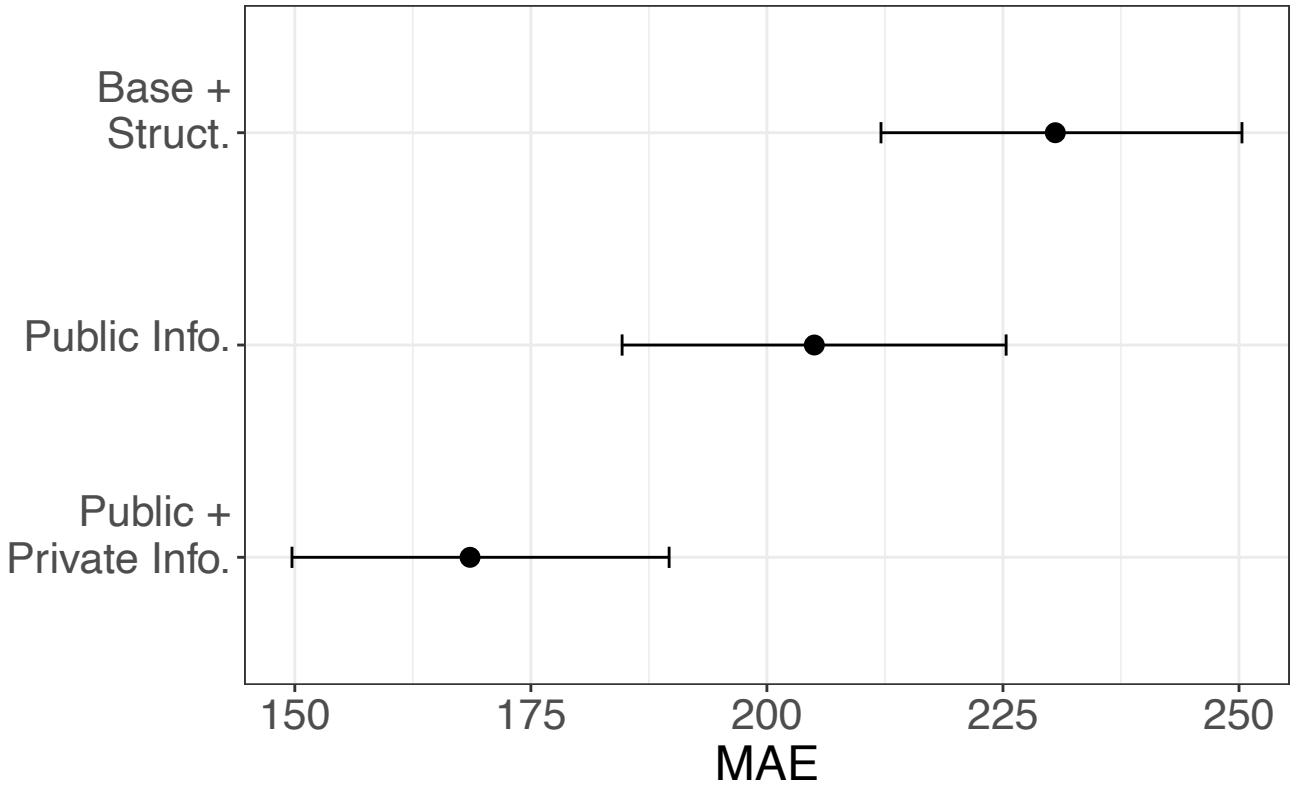
- We forecast the *Time to the next MID*.
  e.g. on July 1st, 1914, the time to the next MID is 31 days.
- OLS with:
  - Baseline (time dependence variables);
  - Structural variables (e.g., military exp., trade);
  - Our 45 topics (each one is a time series, one observation per month)
- We calculate the error as $e_i = |\hat{time}_i - time_i|$
- Repeat this for all 300 forecast months and calculate the Median Absolute Error (MAE)

Forecasting MIDs with French diplomatic themes

Out of Sample Forecasting



Forecasting MIDs with French diplomatic themes



Boussalis C., Chadefaux T., Decadri S., Salvi A.. Forthcoming" Public and Private Information in International Crises: Diplomatic Correspondence and Conflict Anticipation". *International Studies Quarterly*.

Does Sentiment help?

- We measure the sentiment (positive and negative speech) of sentences in the diplomatic cables and newspaper corpora.
- For each document (cable or newspaper issue), we calculate the proportion of sentences that are predicted to have a 0.6, 0.7, 0.8, or 0.9 confidence score for positive (negative) sentiment.

Does Sentiment help?

Example of cable with high **negative** sentiment

T. Paris, 5 mars 1883, 7 h. 45, soir. L'arrangement dont vous avez pris l'initiative au sujet du Tonkin ne tient pas un compte suffisant de notre situation conventionnelle, et consacrerait au profit de la Chine des concessions auxquelles il ne nous appartient pas de souscrire (1). Pour bien marquer notre sentiment à cet égard, je me vois à regret dans la nécessité de mettre fin à votre mission. En attendant vos lettres de rappel, vous vous appliquerez à établir que notre décision n'implique aucune pensée d'hostilité contre la Chine, notre seul but étant d'assurer l'ordre au Tonkin dans les conditions du traité de 1874 (2). 7.

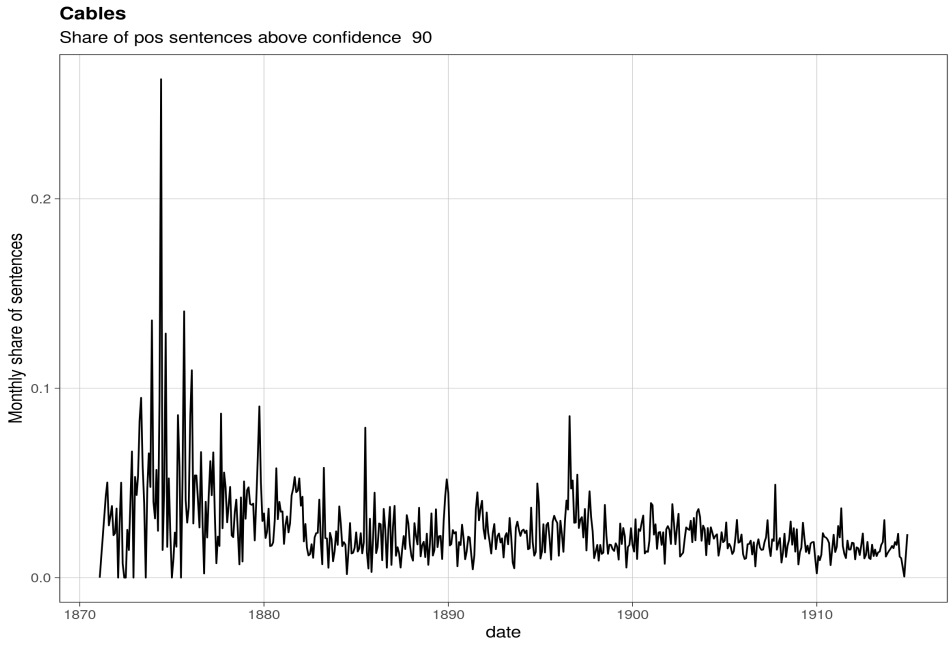
Does Sentiment help?

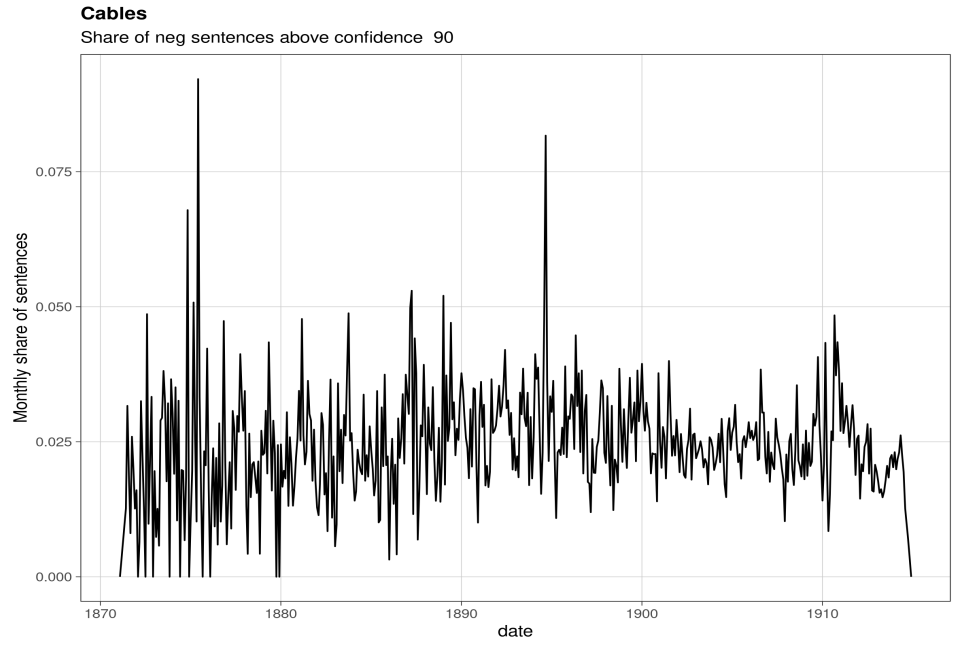
Example of cable with high **positive** sentiment

T. Versailles, 15 mars 1873. J'espère que, lorsque ce télégramme vous arrivera, le traité qui nous occupe depuis si longtemps sera signé et qu'en recevant, ce soir, Sa Majesté l'Empereur et Roi, vous pourrez La remercier à la fois, et de sa présence, et du concours qu'Elle nous a prêté dans une œuvre longue et difficile. Remerciez La, pour ma part, d'avoir par le traité d'aujourd'hui contribué à cicatriser les plaies de la guerre, et facilité, autant qu'il était en Elle, l'apaisement des passions qui ont récemment désolé le monde. C'est la paix définitive que nous venons de signer, et cette paix entre deux grandes nations sera un sûr garant de la paix universelle dont l'Europe ne pourra manquer de se réjouir(1). 190.

Does Sentiment help?

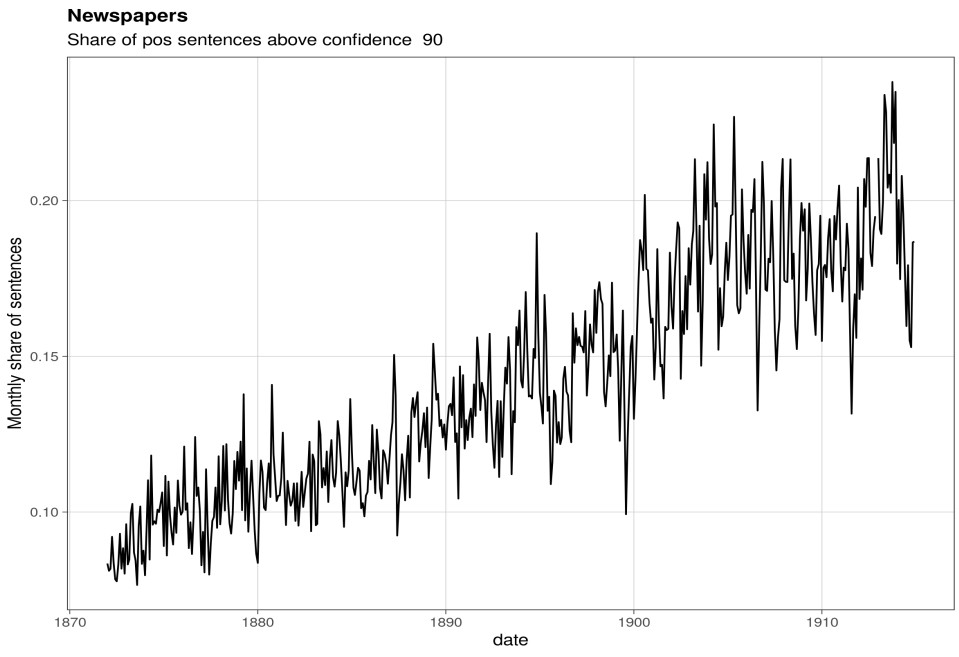
Diplomatic cable sentiment over time

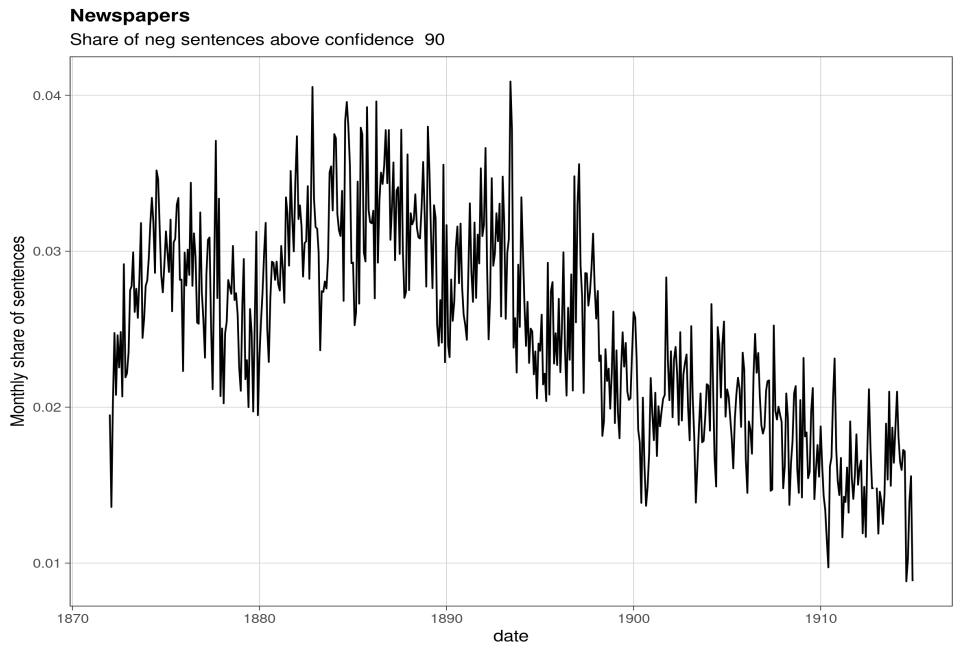




Does Sentiment help?

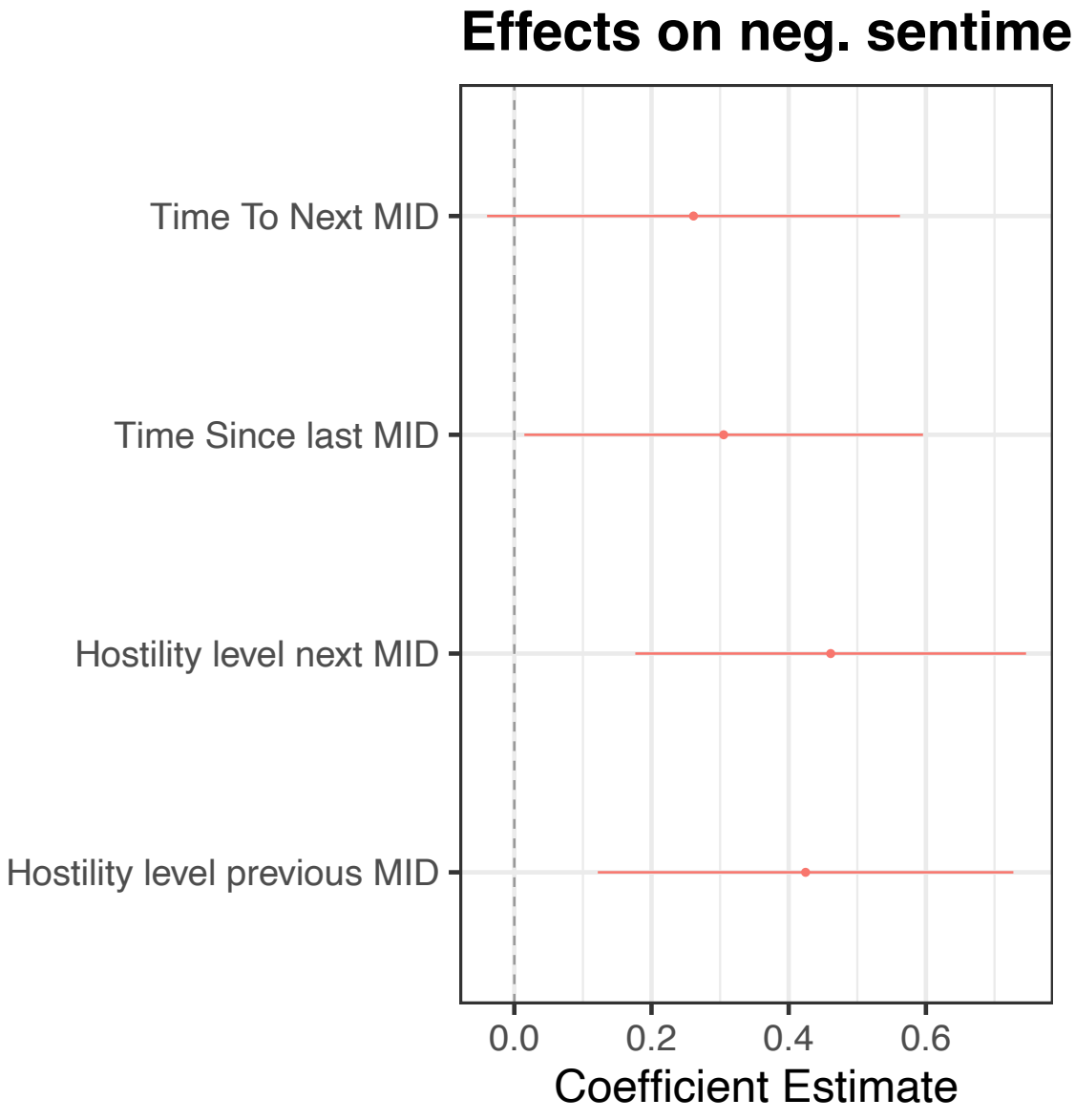
Le Figaro issue sentiment over time





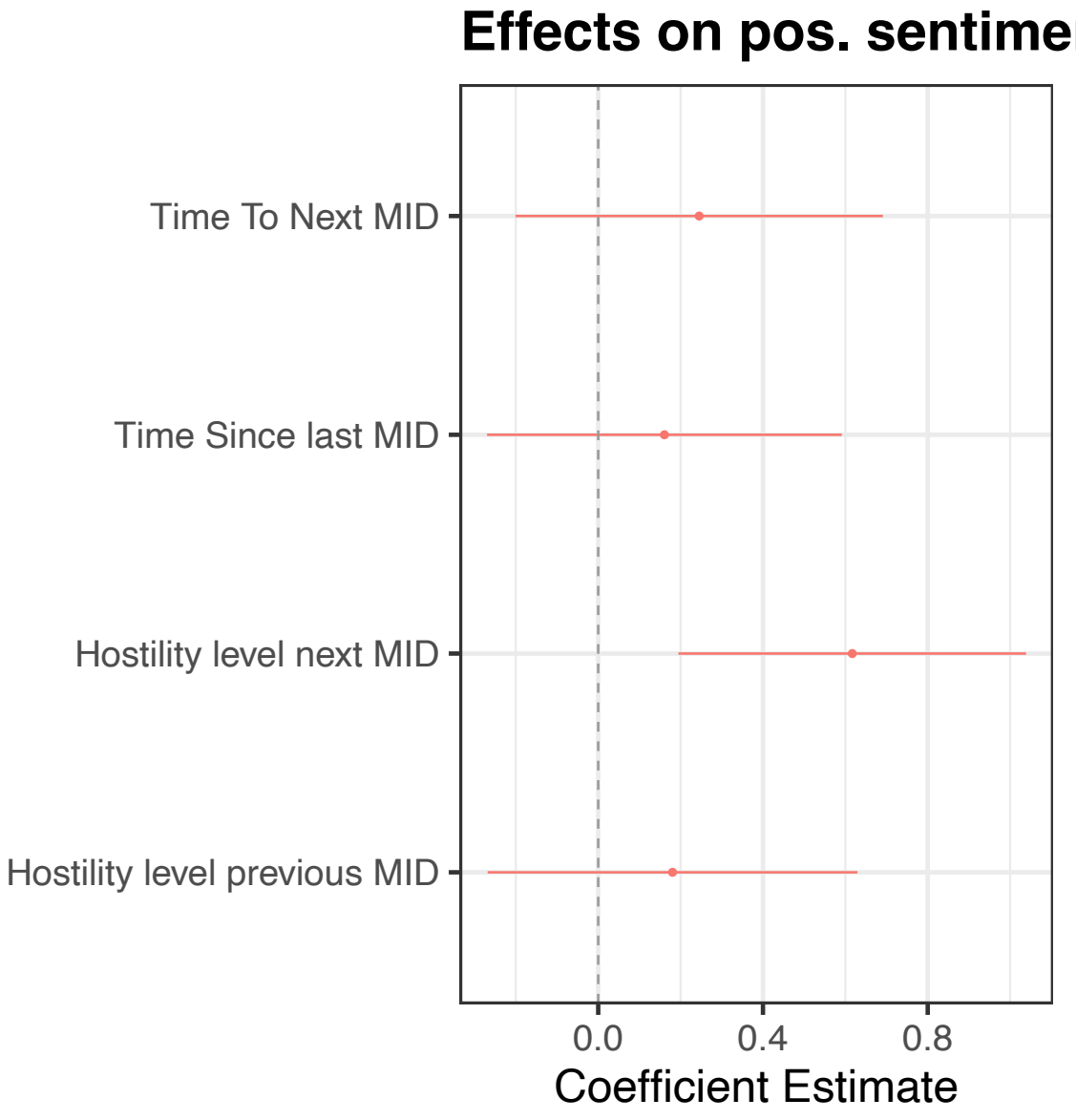
Does Sentiment help?

What affects sentiments (negative)?



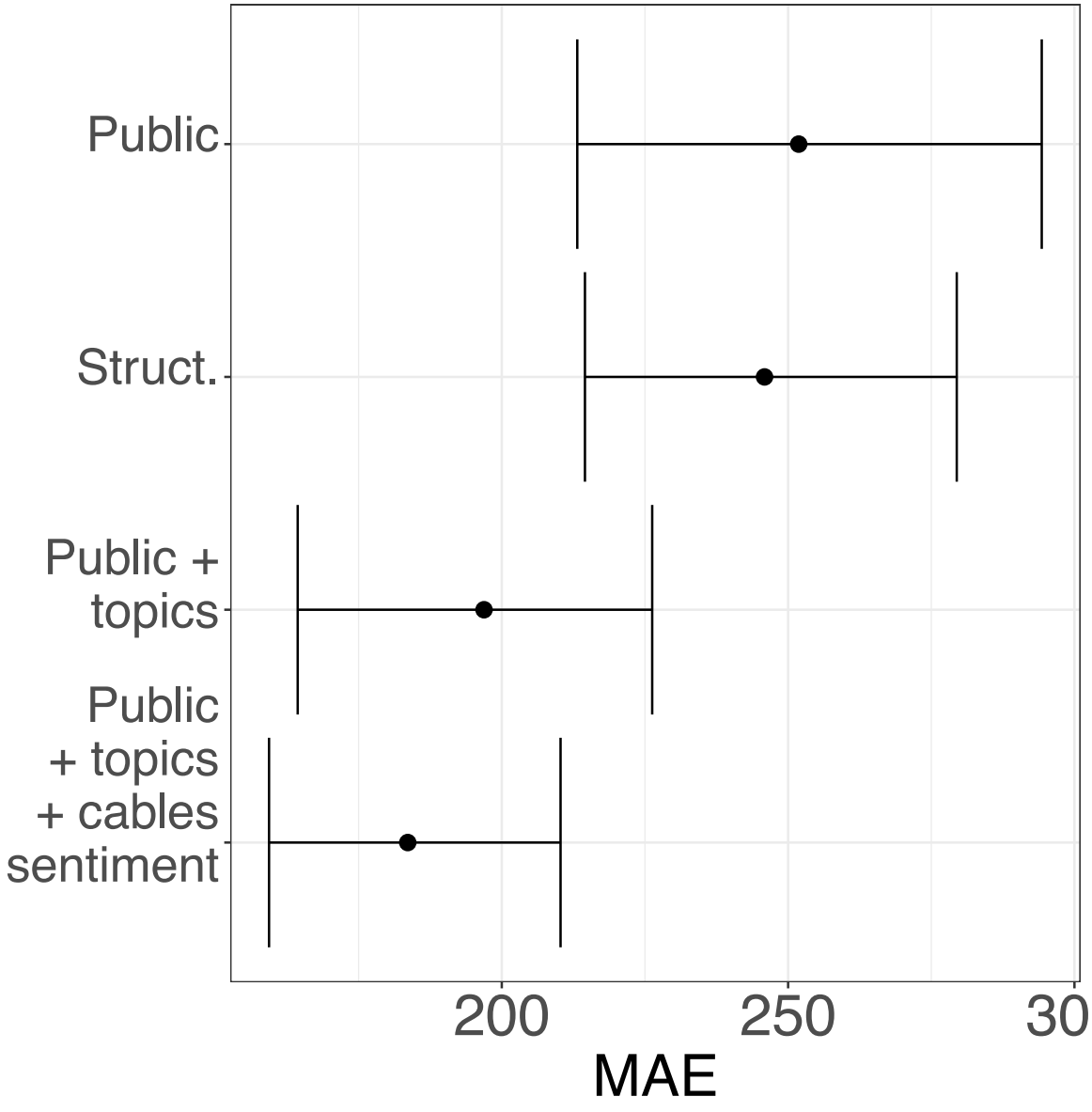
Does Sentiment help?

What affects sentiments (positive)?



Does Sentiment help?

Out-of-sample results



Conclusion

- New data: “insiders’ information”
- Findings: significantly reduces error in estimating time to war, but still very imperfect.

Thank you

Thomas Chadefaux
Department of Political Science
Email: thomas.chadefaux@tcd.ie
URL: www.forecastlab.org