From Pen to Print: Tracing the Evolution of Poverty Narratives in The Times Over Centuries

Dr. Ran Tao

Dr. Jin Zheng

Prof. Adrian Bell

Our aim & research questions

Explore the media narratives regarding the poverty issue in the newspapers.

- What topics discussed in the articles, and are there any time-varying changes over the time?
- Is there any association between “economic”-related topics and the financial market index?
- What is the overall political preference of the selected newspaper articles towards the different parties?

https://pixabay.com/illustrations/question-mark-question-response-1019820/

Data Source

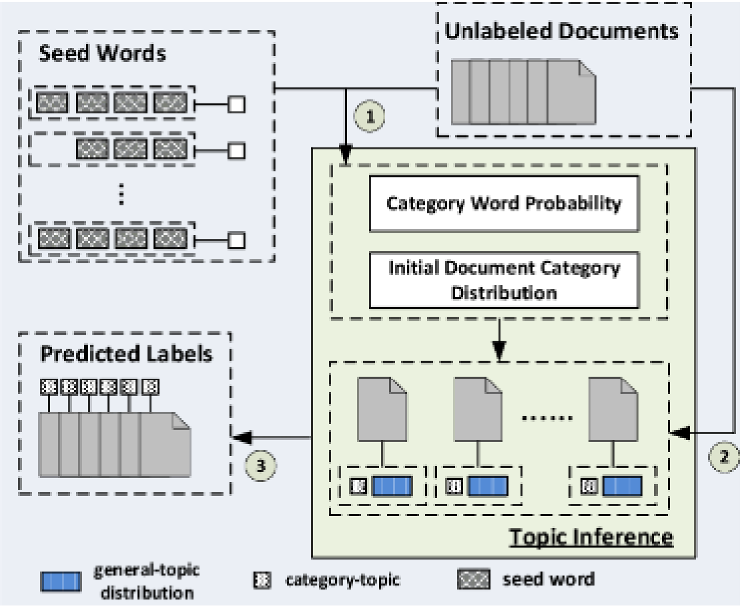
- *The Times* newspapers discussing the poverty issue between 1785 to 2012,
- The University Library automatically converts PDF files into editable txt formats.
    - OCR: a value between 0 to 100

https://www.historic-newspapers.co.uk/old-newspapers/the-times/

Methods

1. Seeded LDA
   a. Topic modelling
   b. Semi-supervised MLs

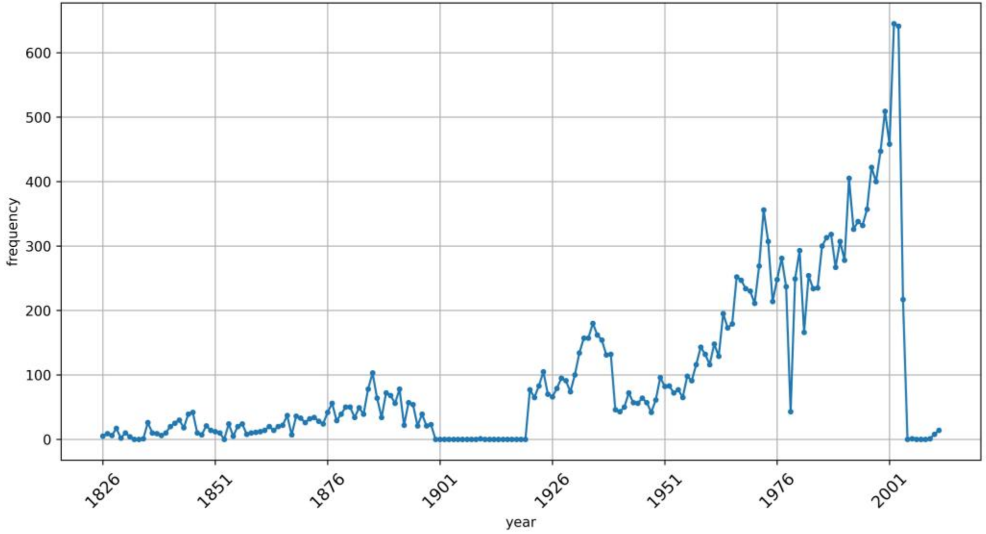
2. Advantages compared to LDA:
   a. Control topic numbers
   b. Incorporate target keywords
   c. Filter “noises”



Topics

1. Poor Relief
2. Workhouse Conditions
3. Unemployment
4. Child Poverty
5. Education and Poverty
6. Government Policies/Social Reform Movements
7. Poverty and Migration/War
8. Poverty and Crime/Drug/Alcohol

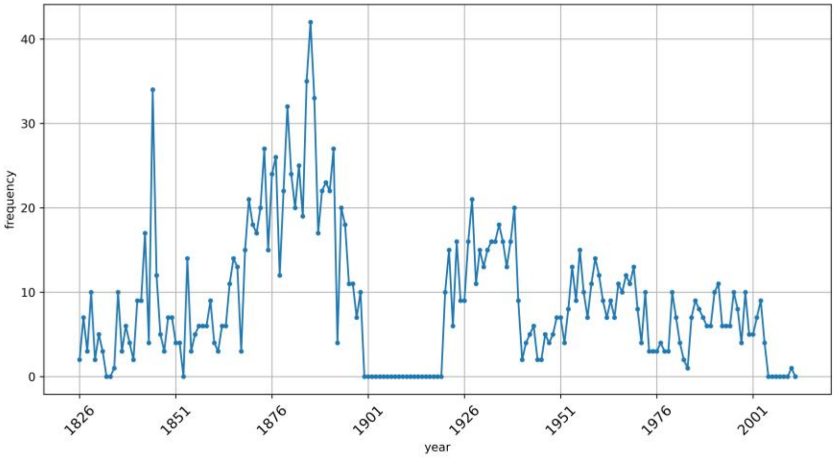
Time-Series Trends -- Overall



Note: articles are filtered with the OCR value above 80.
18,521 articles remain

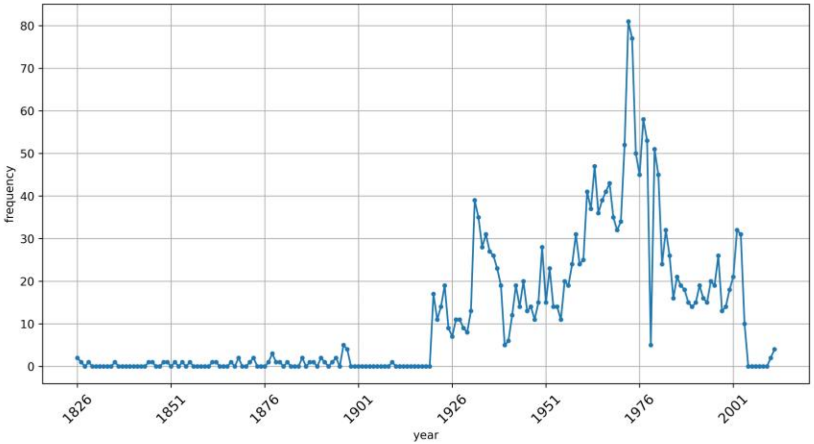
Year	Mean	Stdev	Min	0.25	0.5	0.75	Max
Number_of_Articles	116	130	1	23	64	168	645

Workhouse Conditions



Note: articles are filtered with the OCR value above 80.
unrecognised words are also filtered out.

Unemployment



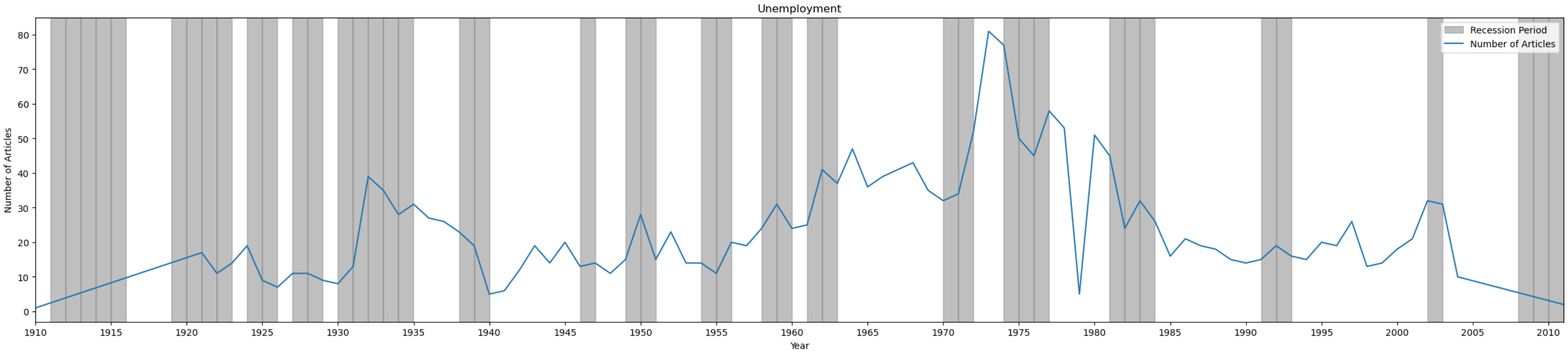
Note: articles are filtered with the OCR value above 80.
unrecognised words are also filtered out.

Proportion of topics in different periods

Topic / Percentage (%)	Total	19th	20th
Education and Poverty	18.3	1.3	17.3
Child Poverty	18.3	0.4	20.6
Poverty and War/Migration	15.5	#N/A	17.0
Poverty and Crime/Drug/Alcohol	12.3	36.8	10.5
Unemployment	11.6	2.3	14.1
Workhouse Conditions	8.9	46.0	5.0
Government Policies/Social Reform Movements	8.1	12.9	8.7
Poor Relief/Poverty Alleviation	7	0.1	6.9

- "Workhouse Conditions" articles were the top-ranked topics in 19 century, but were down to the bottom one in 20 century.
- "Unemployment" articles were covered less in the early period whereas they were frequently appeared in the 20 century.
  - it was recognized as one of the main structural factors linking to poverty.
- Other topics such as "Education and Poverty", "Child Poverty", as well as "Poor Relief" also exhibit an increasing trend over the time.

Unemployment and Recession



- There is a virtually high association between NBER-dated recession period and the number of “unemployment” news articles.
- For example, the Great Depression (1929–1939) was overlapped with the high media counts.

Determinants of "unemployment" narratives

$$News_t = \alpha + \beta_1 Recession_t + \beta_2 Ret_t + \beta_3 Ret_{t-1,t-12} + \beta_4 Vol_t + Time + \varepsilon_t$$

- Where $News_t$ is the number of newspaper coverage regarding "unemployment" in year-month t.
- $Recession_t$is a dummy variable of 1 if the year-month t is a NBER-dated recession period.
- Ret_t is the monthly return in month t.
- $Ret_{t-1,t-12}$is the past 12-month cumulative return.
- Vol_t is the past 12-month return volatility.
- Time is the year fixed effects.

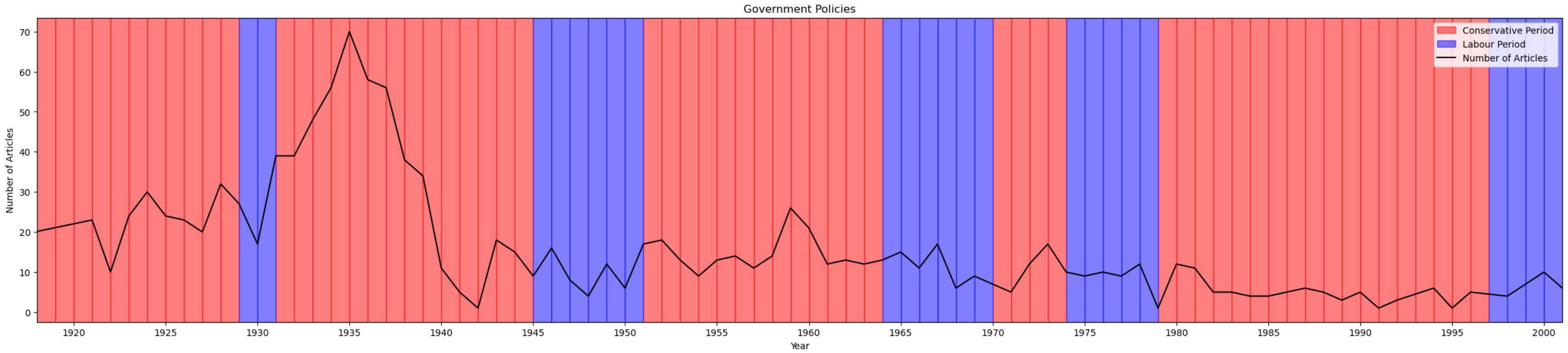
Determinants of “unemployment” narratives

	(1)	(2)
	Unemply news	Total news
Ret_t	0.903	1.014
	(1.14)	(0.25)
Ret_t-1,t-12	0.9200**	1.9902
	(2.31)	(0.97)
Vol	1.918	-6.6648
	(1.35)	(-0.92)
Recession	0.1641	-0.1642
	(1.38)	(-0.27)
_cons	0.7949***	20.6284***
	(4.3)	(21.7)
R-squared	0.5832	0.8281
N. of Obs.	1799	1799
Year Fixed Effects	Yes	Yes

t statistics in parentheses ="* p<0.10; ** p<0.05

- The number of “unemployment” articles is positively associated with the past 12-month cumulative return of UK equity index.
- In contrast, we do not observe such an association between the total Times newspapers and the past return performance of the equity index.
- This suggests that journalists’ coverage of economic-related topics largely depends on macroeconomic conditions.

“Government policies” and Political Parities



- We do not observe any significant difference between Labour and Conservative Party in terms of “government policies” news articles.
- The spike of media counts in the early sample period appears to be related to the Great Depression (1929–1939).

bristol.ac.uk

“Government Policies” and Political Parities

$$SentScore_t = \alpha + \beta_1 Conservative_t + \beta_2 Recession_t + \varepsilon_t$$

- Where $SentScore_t$ is the aggregated sentiment score at a year-month t.
- $Conservative_t$ is a dummy variable of 1 if the Conservative party represents the government.
- $Recession_t$ is a dummy variable of 1 if the year-month t is a NBER-dated recession period.

	(1)	(2)
	HIV4	LM2011
Recession	0.0409***	0.0185
	(2.62)	(0.72)
Conservative	0.0149	0.0172
	(0.97)	(0.68)
_cons	0.2608***	-0.4308***
	(19.46)	(-19.55)
R-squared	0.0151	0.002
N. of Obs.	546	546
Year Fixed Effects	No	No

Follow-ups

- Due to poor printing quality, the early sample texts tend to contain many unrecognised English words even though the OCR value indicates a high conversion rate.

- An example:

  I wouldnotba iqwita him to conne to my house

- ChatGPT predicts the original sentence is likely to be:

  I wouldn't invite him to come to my house.

- A large number of such recognised English words have critically affected the accuracy of textual analysis. Many keywords which should be in the NLP corpus have to be filtered out.

- This, as a result, leads to an inaccurate topic modelling as well as sentiment analysis outcomes.

- Large language models are likely to correct such texts. This will need further fundings to support massive usage.