Does a Movement Emerge? An Analysis Of Egyptian and Palestinian Prison Let

Yousuf Abdelfatah

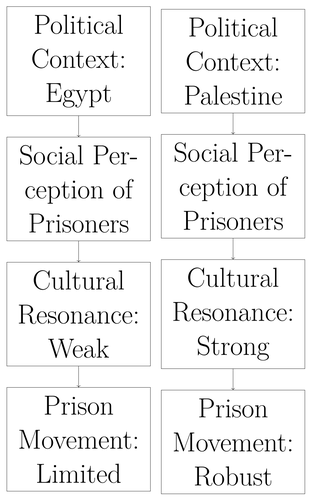
Department of Politics, Princeton University

Summary

- This paper explores how and why prisoners' movements emerge in some contexts but not in others.
- Employs a novel empirical strategy, using Natural Language Processing techniques to analyze Egyptian and Palestinian prison letters.
- Findings suggest that variation in movement emergence are related to the interplay between external political dynamics and internal resistance narratives, with Palestinian detainees more likely to connect their detention with the broader national movement.
- Contributes to the literature on prisons as sites of contentious politics, while adding to our understanding of the structural factors that cause movements to emerge more generally.
- Serves as a demonstration of how we can use machine learning techniques to take advantage of unstructured data in low-information environments.

Theory

I theorize that the emergence of prisoners' movements is contingent upon the social and political opportunity structure; particularly the way prisoners are perceived and their roles are framed within their respective movements and societies. With respect to Egypt and Palestine, the Palestinian case sees a robust prison movement while the Egyptian one does not due to the differing political context that exists outside of the prison walls. These



Analysis Pipeline

Optical Character Recognition

Two distinct corpuses are used for this project, the first containing 471 letters from Egyptian political prisoners, and the second containing 162 letters from Palestinian political prisoners. Egyptian political prisoner letters are sourced from archives of letters publish by the El Nadeem Center for the Rehabilitation of Victims of Torture in a series of reports titled "The Oppression Archive." Palestine prisoner letters are sourced from The Palestinian Museum's Digital Archive and Palestine Memory. The letters were converted into a machine-readable format using the Google Vision API.

Keyword Extraction/Frequency Analysis

I implement a simple keyword analysis both as a sanity check and to get a basic understanding of what the letters in each corpus seems to refer to

Topic Modeling...

topic modeling, using Latent Dirichlet Allocation (LDA) topic modeling, a statistical method used in natural language processing to discover abstract topics within a collection of documents. Experimented with topic modeling based on Arabert embeddings, but LDA provided more interpretable results

Embedding Regression...

Finally, I compare the embeddings of keywords in each corpus, using the method outlined by Rodriguez, Spirling and Stewart (2022). I will provide evidence that the embeddings are substantively dierent by comparing the nearest neighbors to the embeddings of interest, and show that these dierences are statistically significant by employng an embedding regression.

Keywords/Frequencies

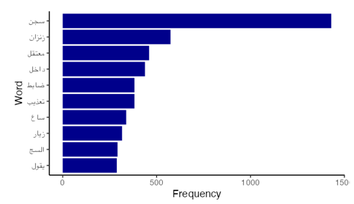


Figure: Egypt WC

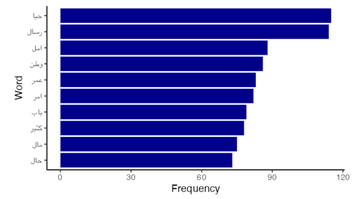
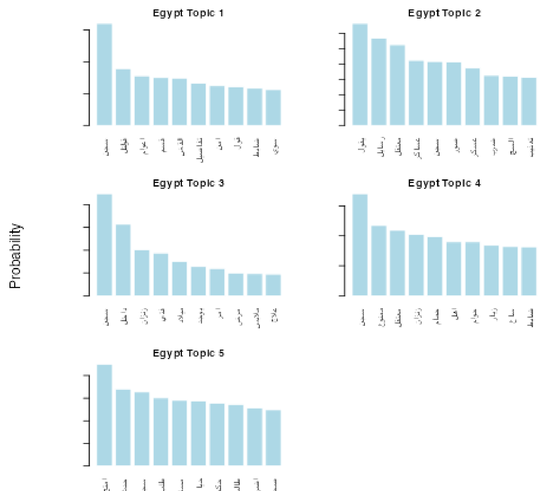


Figure: Palestine WC

- "Aseer" (captive) appears as an identifier for prisoners in the Palestinian case. The Egyptian letters tend to use "detainee" or "prisoner"
- Most frequent Palestinian words include "watan" (nation), "ard" (land), and 'tareekh' (history)
- Egyptian letters make use of more place-based words such as "jail", "cell", and "the inside"
- Egyptians also seem to talk more about prison conditions and treatment, using words such as "visitation" and "torture".
- Palestinian corpus employs agency words more frequently, such as "astatee'" ("I am able")

Topic Modeling



Embedding Regressions

Given the relatively small sizes of the corpuses being used in this analysis, they are an ideal candidate for the embedding approach advanced by Rodriguez, Spirling and Stewart (2022). In this approach, pretrained embeddings are used to calculate the corpus specific embeddings for specified words of interest by taking an average of neighboring words. Three keywords are selected for the analysis here: prison, detainee, and nation. A comparison of nearest neighbors shows some preliminary dierences. The embeddings closest to nation in the Palestinian corpus is much more tied to a sense of homeland and movement, while the nearest neighbors in the Egypt case don't give us much information. Similarly the nearest neighbors for prison in the Palestinian case appear to refer to collective suering, while the Egyptian embeddings seem to refer to more personal conditions. The same is true for the embeddings around detainee. However, in all three cases there is significant overlap between the words in the embeddings and so we require additional analysis to be confident in our results. This additional analysis takes the former of an embedding regression. Regressing the embeddings for out terms of interest on their respective corpus confirms our suspicions, with all three terms providing us with statistically significant dierences, and p-values near zero as displayed in Table 1.

variable	coefficient	normed.estimate	p.value
Detainee	Egypt	3.72	0.00
Nation	Egypt	3.40	0.00
Prison	Egypt	2.98	0.00

Table: Embedding Regreesion Results

References

[1] Chen, J. et al. (2016) Cereb Cortex.
[2] Norman, K. (2010) Hippocampus.
[2] Baldassano, C., et al. (2017) Neuron.
[3] Ben-Yakov A., & Henson, N. (2018). J Neuro.

Download the poster:
`https://tinyurl.com/CCN19-QL`