Machine Learning and the Rise of Consumerism

Albert Sun[1], Donald Pepka[1], Daisy Zhan[1], Andrew Scofield[2]

Faculty Leads: Astrid Giugni, Ph.D[1], Jessica Hines, Ph.D[2]   Project Manager: Chris Huebner[1]

Duke University[1], Birmingham-Southern College[2]

Introduction

Word embeddings have been effective in understanding stereotypes and cultural trends, but they haven't been used extensively on corpora prior to the 18th century[1,2]. **Our research uses word embeddings to analyze words related to consumerism within the Early Modern period when the market economy began to dramatically expand in England.**
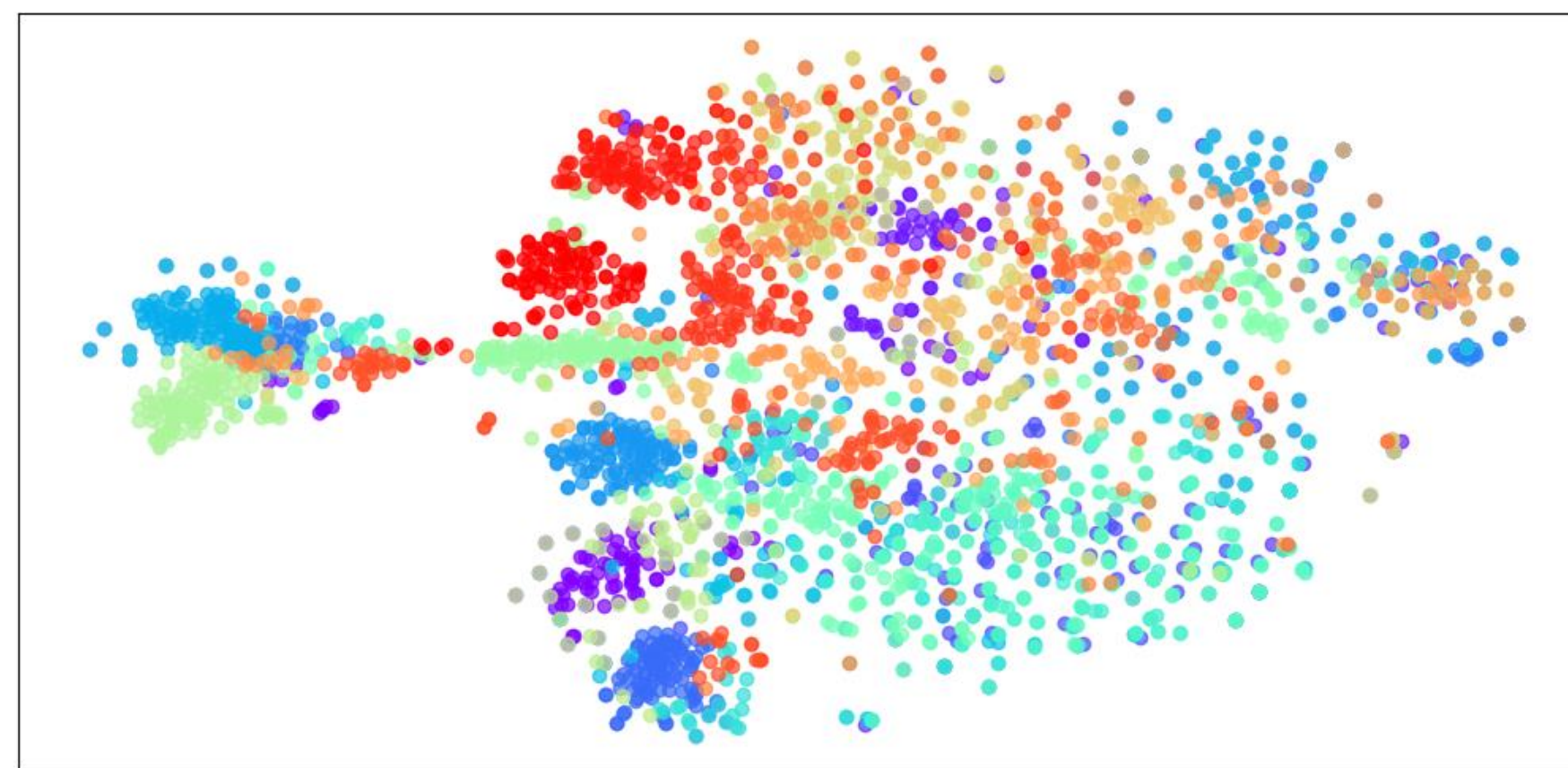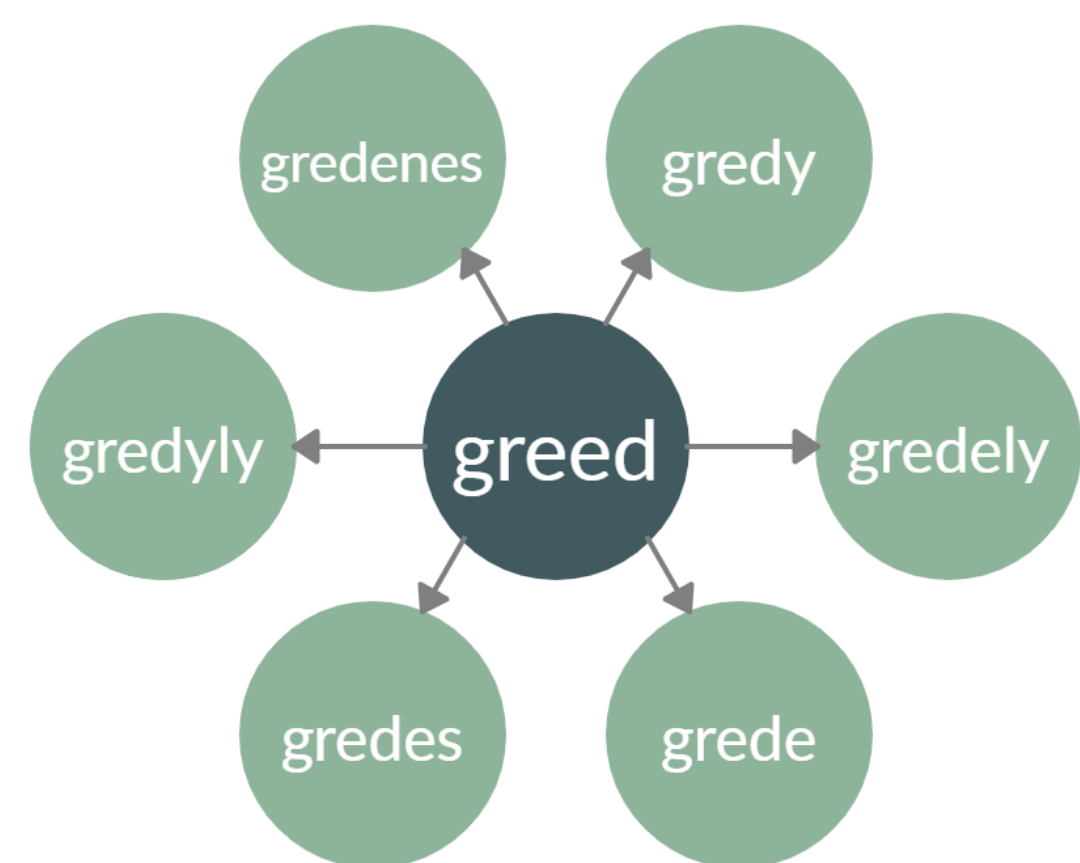

Fig. 1. Word Embedding Model for consumerism-related words in 1670-1700. TSNE representation.

Methodology

1. Data Cleaning

- Scrape data from EEBO TCP Phase I.
- Merge spelling variations with some help from the University of Michigan's Middle English Compendium[3].

2. Word Embedding Models

- Generate Word2Vec Word Embedding Models[4]. Find cosine similarity between word vectors of interest over time.

$$\cos(\theta) = \frac{x \cdot y}{\|x\|\|y\|}$$

- Validate Models via analogy tests and bootstrapping resampling to evaluate accuracy and consistency.

$$\overrightarrow{Queen} - \overrightarrow{King} + \overrightarrow{Man} = \overrightarrow{Woman}$$

- Create artificial "cultural axes" to track words' semantic relationships with gender, race, and social class.

$$\overrightarrow{W} = \frac{(\overrightarrow{Queen} - \overrightarrow{King}) + (\overrightarrow{Woman} - \overrightarrow{Man}) + \cdots}{N}$$

Birth of 'Economic' Consumption

Consume transforms from a "medical" to an "economic" idea in the beginning of the 17th century. Consumption is seen as increasingly "luxurious", and its association with "disease" continuously decreases.
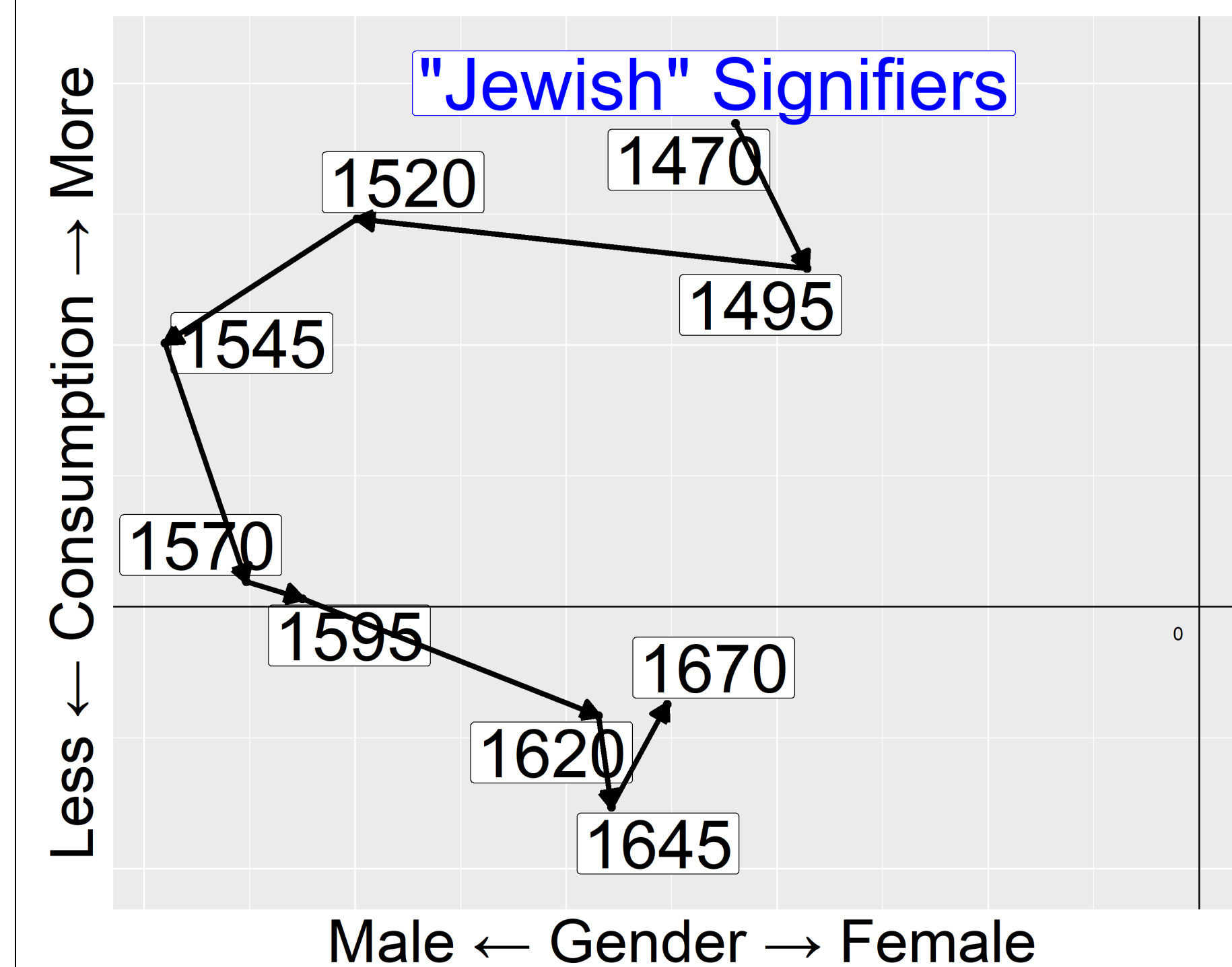

"Consume" develops economic connotations
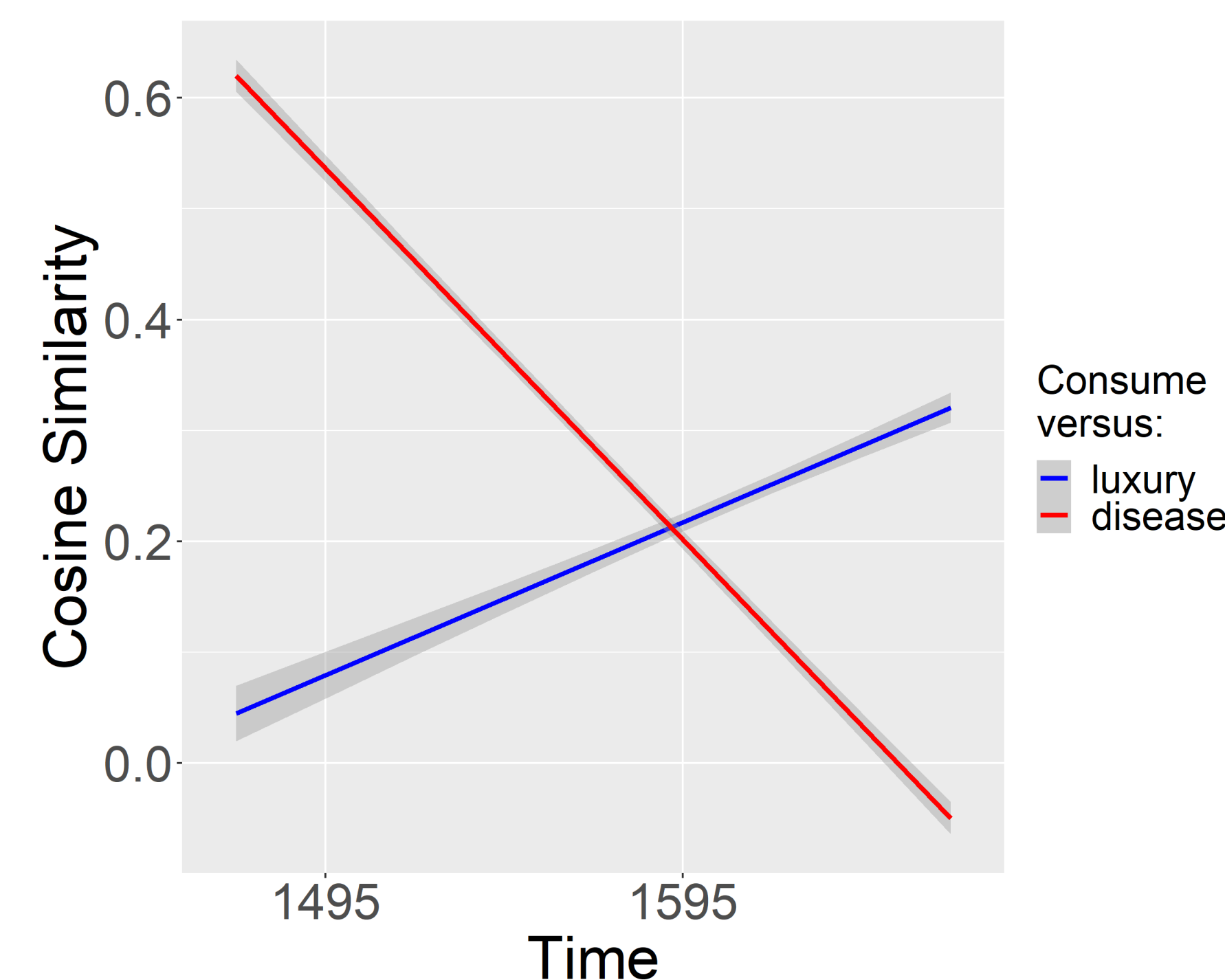
Fig. 2. Consumption transforms from a "medical" to an "economic" idea around the beginning of the 17th century supporting Withington's argument.[5]

Gender Axis Insights

1. **"Consume," "consumption," and "greed" are mostly neutral overall but become more gynocentric over time,** possibly suggesting that women during this period are increasingly being seen as primary consumers in comparison to men.

2. **The word "possession" remains consistently androcentric over time,** possibly due to dominantly male land-ownership.

Consumption-related words and their gender contexts
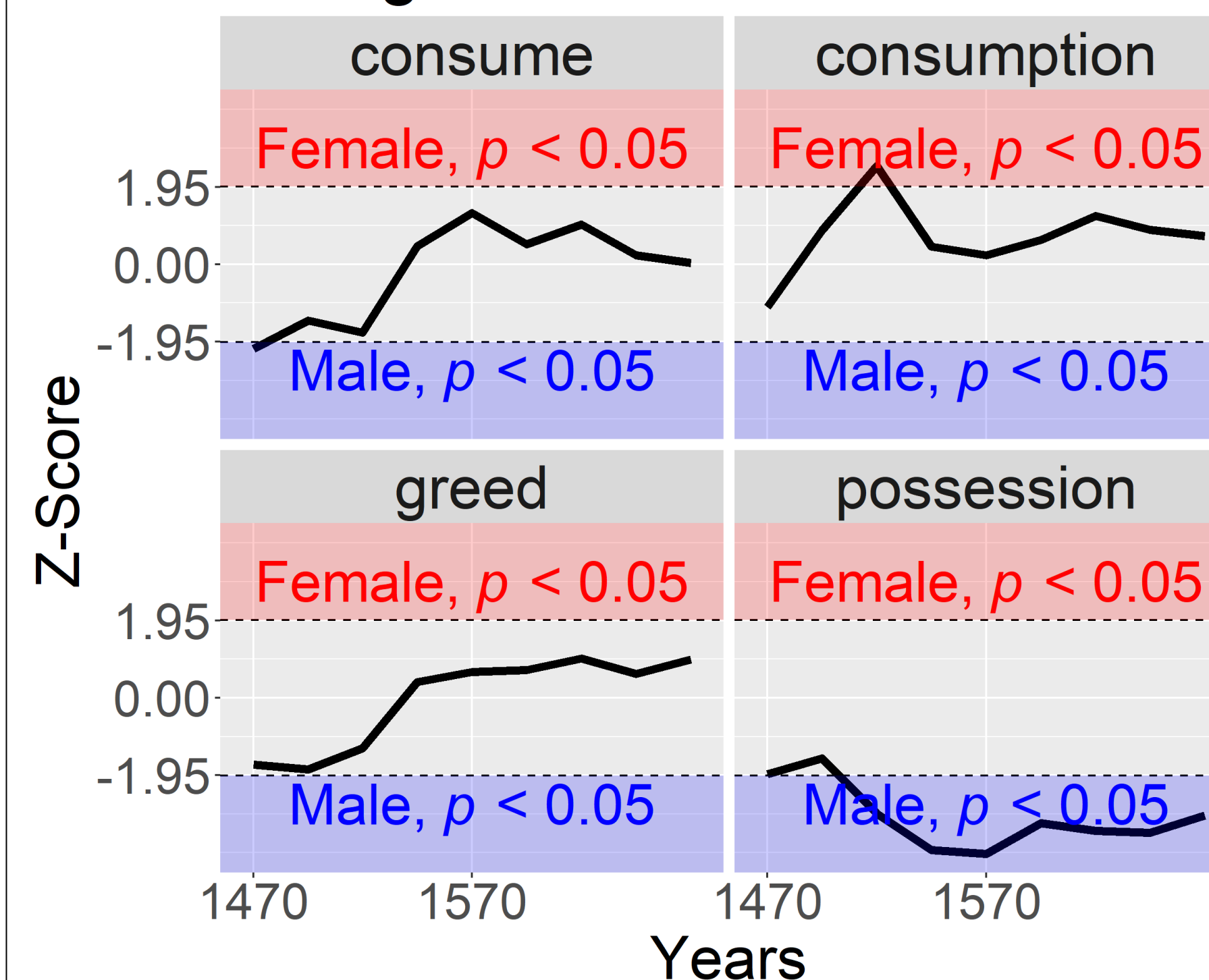
Fig. 3. Comparison of consumption-related words over time to gender culture vector.

Z-Score values are calculated from relative cosine similarity.

Intersectional Axis Insights

1. **Jewish-signifying words become more related to masculinity and less to consumption** during 1495-1544.

2. **There is a sudden spike in the relationship between Jewish-signifying words and words related to consumption** from 1645-1700, overlapping with the resettlement of Jews in England.

Why did we look at these axes? Antisemitism is often expressed in terms of gender and ideas often related to consumerism.


Fig. 4. Jewish-signifying words mapped in context of consumption-related and gender-related lexicon over time.

Next Steps

1. **Expand** our analysis to EEBO TCP II, a set of 30,000+ more texts being released on Jan 1, 2021.

2. **Enhance** our word embedding's accuracy by improving our text-cleaning process, increasing bootstrap resample sizes, and possibly using stop words to reduce clutter.

3. **Compare** the efficacy of different Word Embedding structures, like GloVe, BERT, and fastText

References & Acknowledgements

Find our code, references, and more visualizations at:
https://github.com/albertyusun/For-the-Love-of-Greed