
# Mining for Meaning: Investigating Conceptions of Democracy Across Cultures

Authors: Isabella Duan*

Pranathi Iyer*

Emily Yeh*

Val Alvern Cueco Ligo*

*Each of the authors equally contributed to the paper

# Introduction

There is no single 'global' story about democracy (Ulbricht, 2018). How are differing conceptions of democracy constructed in different cultures? How do these constructions emerge, maintain, and compete for dominance? Leveraging word-embedding models trained on Google's Ngram corpus, we present an analysis of how conceptions of democracy evolve and engage with change in political institutions as well as with developments in economic and material life.

This paper makes three contributions to the existing literature. The first is methodological. Previous historical and sociological studies of democracy as political ideology are severely constrained in scope and scale: scholars typically engage with one or two cultures, attend to a small sample of texts, and focus on particular historical periods (Zhai, 2018; Crow, 2010; Miller et al., 1997). We use word embedding models trained on text from millions of books published over 100 years in American English, British English, French, German, Italian, Chinese and Russian, therefore offering more comprehensive analyses and inferences across social time and space with a year-level temporal resolution.

Our methodological innovation enables two theoretical contributions. Previous scholarship on the relationship between political culture and political institutions treats culture as stagnant (Almond, G. A., & Verba, S., 2015). We challenge this assumption of political culture and ask: does the change in conceptions of democracy precede regime changes and social reforms, lag behind, or do they intertwine with each other in a more nuanced way? We propose a measure for "conceptions of democracy" by extracting words that are substitutes and complements to the word "democracy." We then measure the change in a nation-state's conceptions of democracy and analyze how such change relates to regime changes. To further our investigation into the richness of conceptions of democracy, we also project the word "democracy" on the cultural dimensions of class and morality, interpreting how democracy becomes moral or immoral, rich or poor in contexts of political and social transformations over 100 years.

Furthermore, we also speak to the literature on the relationship between democracy and economic development. Since Lipset (1959), whose modernization theory famously argues that economic development is a prerequisite for a stable democracy, scholars of political science have debated empirical evidence for and against the theory (Przeworski, 2006). Moreover, there are also multiple causal mechanisms through which democratic institutions and cultures can pose positive, negative, or no overall effects on economic growth: democracy may bring political stability and extensive protection of property rights that are conducive to economic growth, or it may limit the growth by breaking down an integrated national market, imposing barriers against new market entrants, and enlarging costs to the implementation of monetary and fiscal policies. (Schultz, 1999) We provide a new perspective on the complex relationship between democracy and economic development at the nation-state level by measuring the cosine similarity of the word "democracy" represented as a vector in the high-dimensional vector space and six vectors representing exogenous word lists we compiled based on patterns from our exploration of complements and substitutes: economic growth, economic recession, economic liberalism (friedman), socialism (marx), centralized government, and decentralized government. We find little support for the hypothesis that cosine similarities predict GDP growths in democratic regimes but not in authoritarian regimes. However, we find some support for the hypothesis that GDP per capita growth predicts cosine similarities in authoritarian regimes but to a lesser extent in democracies.

# Data

We use samples from Google's Ngram corpus in American English, British English, French, German, Italian, and Russian from 1900 and 2009. For the Chinese language, data is only available from 1955 to

2009.[1] Our analysis is based on pre-trained word-embedding models[2] on a small sample of this corpus: we have one model for each year and each language, totaling 715 models. These models were trained on default settings–future research will explore tuning these hyperparameters for closer measurement. We consider our data and models to be reliable representations of how nation-states talk about and make sense of democracy. We do acknowledge, however, that our data do not distinguish between political groups within a nation-state and, because they come from published books, over-represent elite culture and ruling authorities as opposed to popular culture and suppressed political groups.

# Section 1: Tracing Conceptions of Democracy across Time

We model a nation-state's conception of democracy in two ways. First, we model the conception of democracy as a word frequency distribution of most similar words to the word "democracy." Word2Vec models project each word as a vector in a high dimensional space that tries to capture the semantic relations between words (Mikolov, 2013). We first locate the top 500 words that have the highest cosine similarity to the word "democracy," that is, the cosine of the angle between those word vectors and the "democracy" vector is the smallest. Then, we remove words that appear less than 100 times from this word list because we want to capture the broad trends in political discourse. After, we build a word frequency distribution based on those words and their frequency for each year. Lastly, we calculate the Kullback-Leibler divergence between a given year's distribution and the previous year's.

Second, we model the conception of democracy as complements and substitutes to the word "democracy." A consumer's preference for a commodity $X$ may be defined by what other commodities become more attractive when $X$ is purchased, or $X$'s complements, and what other commodities

[1] We would like to thank Shilin Jia for providing us with trained models.

[2] Model parameters used are sg=1, hs= 1, size=300, window=5, min_count=10, iter = 3. The size of the sample differs across cultures but is consistent within a culture: the sample size is a percentage of the smallest yearly corpus size within the culture's timeframe; this percentage is 5% for British English, US English, French; 1% for German and Russian; 50% for Chinese; unknown for Itlatian.

become less attractive when *X* is purchased, or *X*'s substitutes (Ruiz et al., 2020). Similarly, a nation-state's conception of democracy may be defined as what words are likely to be mentioned when "democracy" is mentioned, or democracy's complements, and what words are unlikely to be mentioned when "democracy" is mentioned, or democracy's substitutes. We acknowledge that a consumer's preference for commodity *X* cannot be fully reduced to a function of their preferences for other commodities conditional on their purchase of commodity *X*: one may have innate like for a commodity that has nothing to do with other commodities. Therefore, we make no claim that conception of democracy is nothing but complements and substitutes to democracy; rather, we hope to capture an important aspect of how people think and talk about democracy that is indicative of changes across time. We access context vectors of each word on our "most similar words to democracy" word lists and define words with the top 100 cosine similarities with the context vector of "democracy", as complements. Conversely, we define words with the lowets 100 cosine similarities as substitutes. We then count the total number of words that are added or removed from substitutes or complements. Figure 1 reports the change in conceptions of democracy as measured by these two measures. We qualitatively inspected the lists of most similar words, compliments, and substitutes. Table 1 showcases the Top 5 items on our wordlists in the German language between 1910 and 1960.

| **Year** | **Most Similar Words** | **Compliments** | **Substitutes** |
|---|---|---|---|
| 1910 | Zivilisation (civilization)<br>Weltanschauung (belief)<br>Religion (religion)<br>Arbeiterbewegung (labour movement)<br>Ethik (ethics) | Sozialismus (socialism)<br>**Judentums (judaism)**<br>Wirtschaft (economy)<br>Politik (politics)<br>Proletariats (proletariat) | Rasse (Race)<br>Protestantismus (protestantism)<br>Renaissance (renaissance)<br>Einzelstaaten ('individual states)<br>Sogen (so-called) |
| 1920 | Liberalismus (liberalism)<br>Antiken (antique) | Liberalismus (liberalism)<br>Sittlichkeit (morality) | **Diktatur (dictatorship)**<br>Arbeiterbewegung (labour movement) |

| | | | |
|---|---|---|---|
| | Sozialismus (socialism)<br>Kapitalismus (capitalism)<br>Proletariats (proletariat) | Fortschritts (progress)<br>Tradition (tradition)<br>Kampfes (struggle) | Demokratischen (democratic)<br>Revolutionären (revolutionary)<br>Weltanschauung (worldview) |
| 1930 | Liberalismus (liberalism)<br>Zivilisation (civilization)<br>Kapitalismus (capitalism)<br>Sozialismus (socialism)<br>Wirtschaft (economy) | Religion (religion)<br>Politik (politics)<br>Proletarischen (proletarian)<br>Idee (idea)<br>Versöhnung (reconciliation) | **Judentums (judaism)**<br>Freiheit (freedom)<br>Sozialismus (socialism)<br>Kapitalismus (capitalism)<br>Arbeiterschaft (labourers) |
| 1940 | Humanität (humanity)<br>**Diktatur (dictatorship)**<br>Zivilisation (civilization)<br>Sozialismus (socialism)<br>Revolution (revolution)<br>Intelligenz (intelligence) | **Diktatur (dictatorship)**<br>Idee (idea)<br>Selbständigkeit (independence)<br>Subjektiven (subjective)<br>Humanität (humanity) | Religion (religion)<br>Kultur (culture)<br>Weltpolitik (world politics)<br>Ausgestaltung (organize)<br>Proletariats (proletariat) |
| 1950 | Sozialismus (socialism)<br>Nationalismus (nationalism)<br>Kommunismus (communism)<br>Liberalismus (liberalism)<br>Weiterentwicklung (further development) | Verwirklichung (realization)<br>Demokratisierung (democratization)<br>Klassenkampfes (class struggle)<br>Progress (progress)<br>Menschheit ( | Einheit (unity)<br>Sogenannte (so-called)<br>**Diktatur (dictatorship)**<br>Nationalismus (nationalism)<br>Wiedergeburt (rebirth) |
| 1960 | Freiheit (freedom)<br>Nationalismus (nationalism)<br>Revolution (revolution)<br>Ideologie (ideology)<br>Katholizismus (catholicism) | **Judentums (judaism)**<br>Restauration (restoration)<br>Friedens (peace)<br>Allseitigen (universal)<br>Gemeinschaftsarbeit (community work) | Proletarischen (proletarian)<br>Marktwirtschaft (market economy)<br>Ideen (ideas)<br>Revolution (revolution)<br>Frieden (peace) |

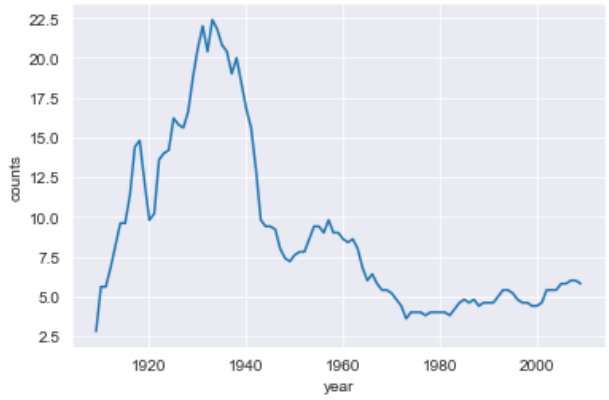
**Table 1: Top 5 Words Most Similar, Compliments, or Substitutes to the Word "Democracy" (Demokratie) in Germany, 1910 - 1960**

Our measurement reflects the rise and fall of Nazism: from 1920 to 1940, "dictatorship" moves from being a substitute of democracy to its complement; "Judaism" moves from being a compliment in 1910 to a substitute in 1930 and again back to a top complement in 1960. This could possibly reflect that, as Nazism took power, the way Germans talked about democracy excluded Jewish people, and a strong leader is compatible with the ideal of democracy; when Nazism was defeated, including Jewish people became a marker of democracy, and a dictator became the opposite of democracy. Our

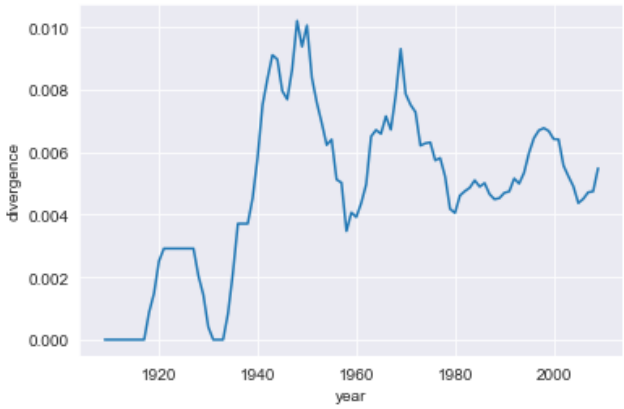
divergence measurement of conception change seems to predict regime change in Germany, whereas our word change count measurement seems less stable.

We acknowledge four limitations in our measurement. First, we did not remove synonymous words of "democracy" from our wordlists so that the word "democratic" ironically appears as a substitute to democracy: unsurprisingly the contexts by which people use the word "democracy," which is a noun, are very different from those surrounding "democratic," which is an adjective. Second, we adopted the same threshold (top 500 for most similar word vectors, top 100 for most similar context vectors, 100 as minimum word counts) for all languages. Our results thus cannot be compared and interpreted across languages. Further research may consider adopting customized thresholds for each language and, to facilitate cross-national comparisons, normalizing the measurement by the frequency of the word "democracy" appearing in each language. Third, because our models use 5-grams, our context vectors only capture 5-gram contexts. Future research may increase the window size of word embedding models to account for richer contexts. Finally, except for English and Chinese languages, we use Google Translate to interpret and validate our results. In our future work, we hope to recruit native speakers to translate results and facilitate interpretations.

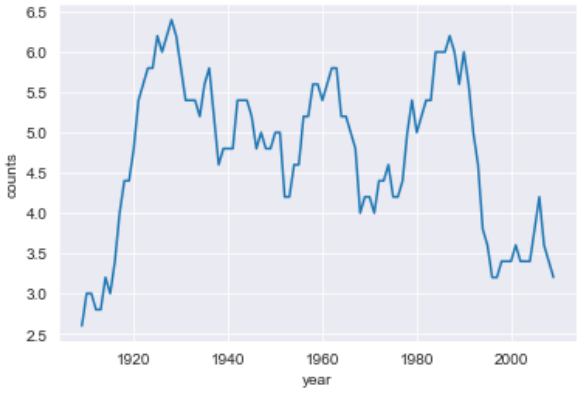
enggb: Counts of Word Change in Complements and Substitutes to Democracy



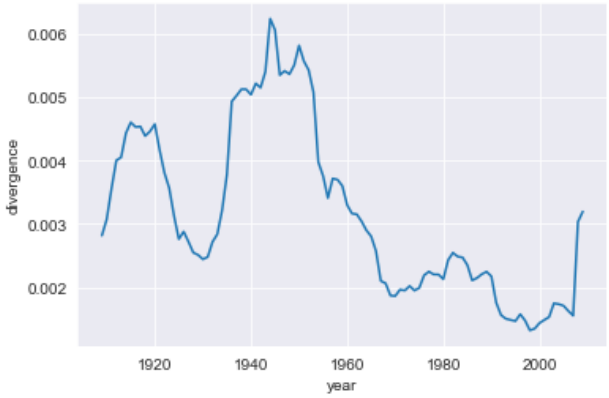

enggb: KL Divergence in Word Distribution of Most Similar Words to Democracy



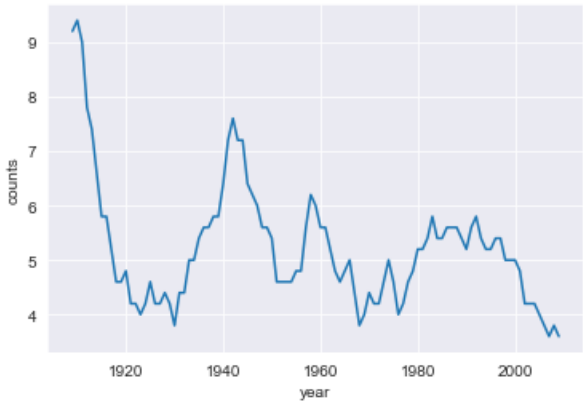

engus: Counts of Word Change in Complements and Substitutes to Democracy



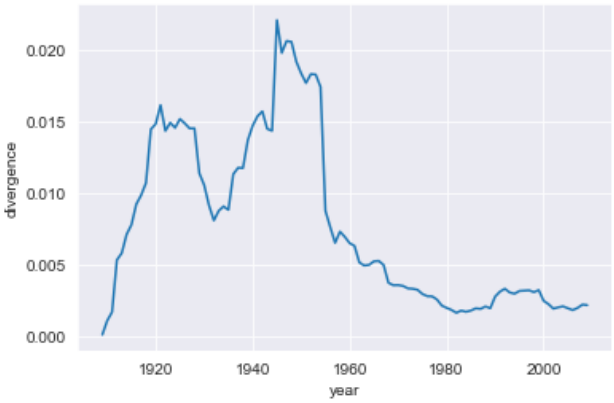

engus: KL Divergence in Word Distribution of Most Similar Words to Democracy



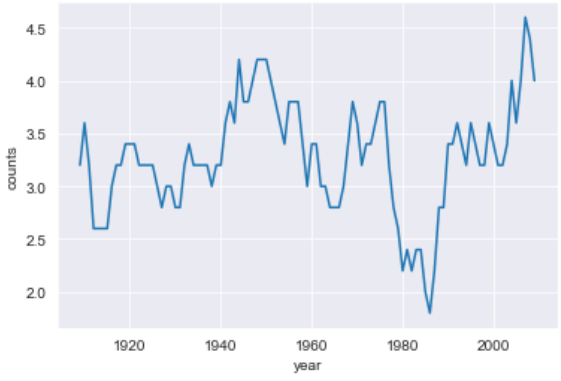

german: Counts of Word Change in Complements and Substitutes to Democracy



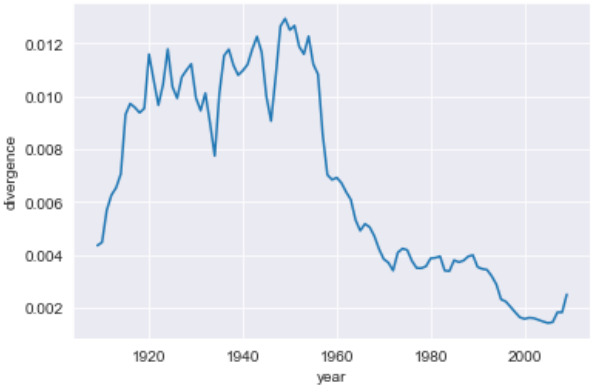

german: KL Divergence in Word Distribution of Most Similar Words to Democracy




italian: Counts of Word Change in Complements and Substitutes to Democracy




italian: KL Divergence in Word Distribution of Most Similar Words to Democracy



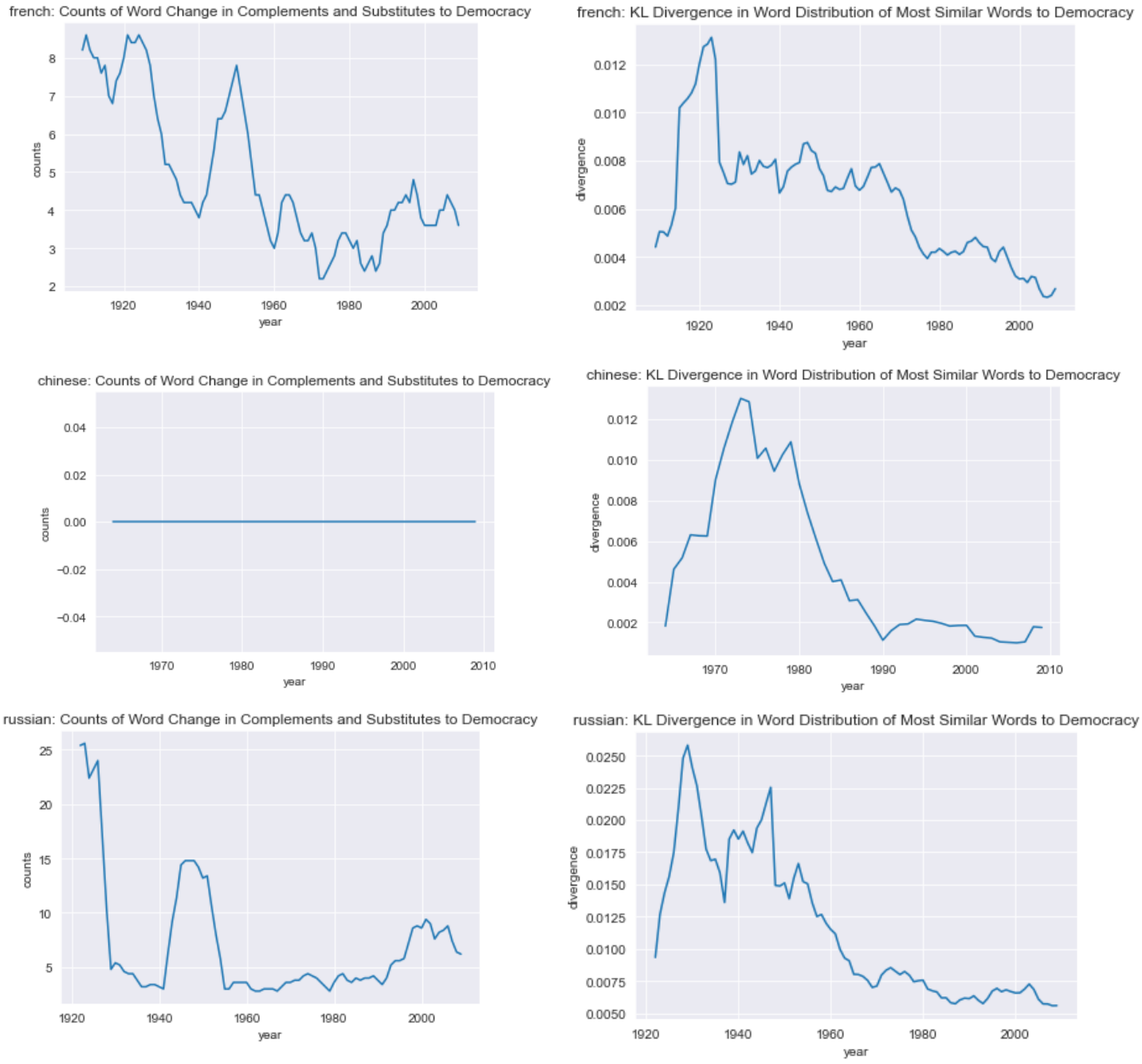
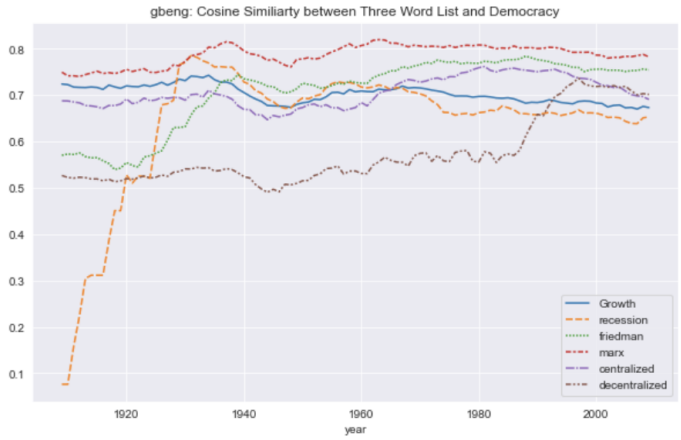



**Figure 1: Change in Conceptions of Democracy in Seven Nation-States in Two Measures**

Lastly, we observe that most words which appear as complements and substitutes in the German language between 1900 and 2000 (Table 1), can be bracketed into three broad categories of governance, economics, and ideology. For instance, words such as “socialism,” “liberalism,” “capitalism,” and “nationalism” can be classified as ideologies. Additionally, while “economy,” “market economy,” “laborers,” and “class struggle” can be categorized as economic concepts, words such as “politics,” “progress,” “world politics,” and “dictatorship” can be associated with governance. More importantly, this trend is apparent across all cultures and languages (Appendix 1). Therefore, in the following sections, we explore the relationship between democracy and each of these aspects. It is important to note that there were several other abstract and miscellaneous words that cannot be classified under the umbrella of economics, ideologies, or governance or any other category for that matter. We thus choose to discard them for the purpose of our research. However, further studies could delve deeper into their meaningfulness and associations with democracy.

# Section 2: Tracking Democracy Against Economics, Ideology, & Governance

From the section above, we arrived at three broad conceptions–economics, governance, and ideology. To measure how they changed relative to democracy, we first created word lists which were handcrafted in reference to exogenous, open-source repositories from the internet, such as Wordnet and Wikipedia. We then aggregated the word vector of each word in its respective list to a single representative vector for that concept. Lastly, we measured this vector's cosine similarity to the word "democracy" and tracked their changes over time. However, we observed that these associations were not significant since these wordlists contained words which are similar to democracy but opposite in meaning (for instance, socialism and capitalism) and canceled out each other's effect. We thus decided to split each of these concepts into two dichotomous sub-concepts: growth vs. recession (economics); economic liberalism/friedman vs. socialism/marx (economic ideology); centralized government vs. decentralized government (governance). We then plotted their similarity against democracy and Figure 2 reflects these changes for the US, UK, and France.



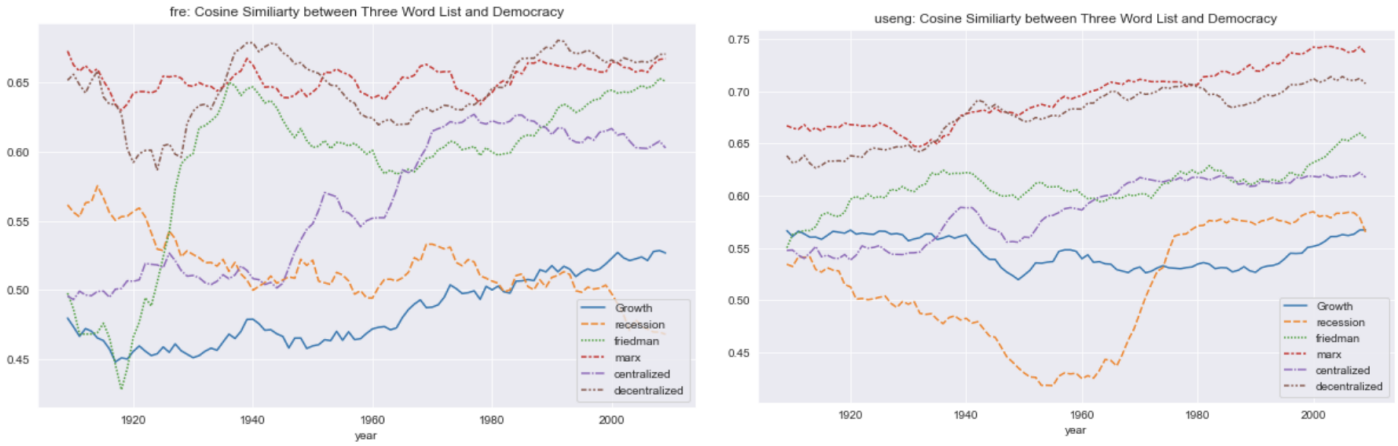



**Figure 2: Cosine similarity graphs for the US, UK, & France**

These graphs validate how conceptions of democracy differ across different countries. More importantly, they highlight how these conceptions respond to exogenous events. For example, across the graphs, these conceptions had dips in their similarity to democracy, starting from the beginning of World War II in 1939. From this, we sought to compare these changes to objective economic measures. Specifically, we tested how changes in conceptions of democracy affect changes in GDP per capita growth and vice versa, framing the rise and fall in each country as an event and hypothesizing that the variation in conceptions will be reflective of GDP growths.

We adopted the abduction process (Timmermans & Tavory, 2012) for theory-generation and made two hypotheses:

- Hypothesis 1: Change in conceptions of democracy predicts GDP growths in democratic regimes but not in authoritarian regimes.
- Hypothesis 2: Change in GDP growth predicts change in conceptions of democracy in authoritarian regimes but not in democratic regimes.

Our intuition comes from the observation that public discourses in democratic regimes about democracy in relation to economic development or governance generate ideas and actual policy changes that are conducive for economic growth. In contrast, the discourse about democracy is tightly manipulated and controlled in authoritarian regimes. We suspect that authoritarian regimes' strategically exploit economic growth to build rapport for regime legitimacy. Therefore, we expect conceptions of democracy to shift in an ad hoc manner in response to GDP growths in those regimes.

# Section 3: Testing the Relationship between Conceptions of Democracy and GDP Growth

In order to test our hypothesis, we conducted a time series analysis for the cosine similarities between the six conceptions and the word "democracy," and GDP per capita growth (World Bank, 2022) using Granger causality and Pearson correlation tests. Table 2 reflects Granger's causality results on whether past values of cosine similarities have a statistically significant effect on the current value of GDP per capita growth, taking past values of cosine similarities into account as regressors. We find that the relationship is heterogeneous across authoritarian and democratic regimes, suggesting that our first hypothesis is likely to be false. How discourse about democracy affects GDP per capita growth depends on nation-state contexts: in China, cosine similarities between the word "democracy" and economic liberalism reliably track GDP per capita growth, as market liberalization since the 1980s stimulates economic growth. Table 3 reports Granger's causality results on whether past values of GDP per capita growth have a statistically significant effect on the current value of cosine similarities. We find some support for our second hypothesis: GDP per capita growth to great extent predicts cosine similarities in authoritarian regimes like China and Russia but to a lesser extent in democracies.

Table 4 reflects the Pearson correlation test between cosine similarities and GDP per capita growth. We find that the relationship is largely insignificant, with only China reflecting consistently significant

results across the different conceptions. This is unsurprising as the Pearson test does not account for the same lead-lag considerations that the Granger causality test considers.

For a visual representation, we plot changes in GDP per capita growth and cosine similarities of the aspects of growth and recession with democracy as depicted in Figure 3. The plots across countries showcase synchronous movements in GDP per capita growth and cosine similarities. Moreover, they capture significant events with respect to the economic climates of countries. For instance, for the US, despite being synchronous with GDP per capita growth, recession similarity is fairly stable prior to 2008. However, we see that after 2008, there is a significant drop in recession cosine similarity and GDP per capita growth. While the drop in GDP growth itself can be attributed to the 2008 financial crisis, the downward slide in recession cosine similarity can be associated with the fact that people utilized words analogous with economic conditions which are more severe than a recession to describe the economic and political situation in the country at the time.

| | **Growth** | **Recession** | **Friedman** | **Marx** | **Centralized** | **Decentralized** |
|---|---|---|---|---|---|---|
| USA | 0.015 (Lag4) | 0.029 (Lag4) | 0.041 (Lag5) | 0.048 (Lag 4) | | |
| Great Britain | | | 0.0 (Lag4) | | | |
| Italy | | | | | | |
| France | 0.064 (Lag5) | | | 0.039 (Lag5) | 0.016 (Lag1) | 0.041 (Lag1) |
| Germany | 0.022 (Lag1) | 0.033 (Lag3) | 0.041 (Lag5) | 0.032 (Lag2) | 0.057 (Lag5) | 0 .044 (Lag3) |
| Russia | 0.009 (Lag3) | 0.012 (Lag5) | 0.073 (Lag1) | 0.006 (Lag2) | 0.002 (Lag3) | 0.01 (Lag3) |
| China | 0.029 (Lag 2) | 0.029 (Lag1) | 0.019 (Lag4) | | | |

**Table 2: Granger's causality test results examining whether cosine similarities between the word democracy and six wordlists predict GDP per capita growth**

| | **Growth** | **Recession** | **Friedman** | **Marx** | **Centralized** | **Decentralized** |
|---|---|---|---|---|---|---|
| USA | .069 (Lag 4) | | | | | .051 (Lag2) |
| Great Britain | .002 (Lag4) | | | | .047 (Lag3) | |
| Italy | | .03 (Lag3) | | .058 (Lag1) | | |
| France | | | | | .011 (Lag1) | |
| Germany | | .006 (Lag2) | | | .022 (Lag1) | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Russia | | 0.0 (Lag5) | 0.0 (Lag4) | | .0.0 (Lag5) | .029 (Lag5) |
| China | .042 (Lag4) | 0.0 (Lag4) | .008 (Lag5) | 0.0 (Lag4) | .007 (Lag1) | 0.0 (Lag5) |

**Table 3: Granger's causality test results examining whether GDP per capita growth predicts cosine similarities between the word democracy and six wordlists**

| | **Growth** | **Recession** | **Friedman** | **Marx** | **Centralized** | **Decentralized** |
|---|---|---|---|---|---|---|
| USA | | | | | | |
| Great Britain | | | | | | |
| Italy | | | Corr: −0.326<br>p: .042 | | Corr: −0.418<br>p: .008 | Corr: −0.419<br>p: .007 |
| France | | | | | | |
| Germany | | | | | | |
| Russia | | | Corr: .568<br>p: .027 | | | Corr: .539<br>p: .037 |
| China | Corr: −0.326<br>p: .042 | | Corr: .526<br>p: 0.0 | Corr: .472<br>p: 0.0 | Corr: .483<br>p: .001 | Corr: .346<br>p: .03 |

**Table 4: Pearson Correlation results between GDP per capita growth and cosine similarities between the word democracy and six wordlists**

We acknowledge three limitations of the proposed measure. First, these tests were only measured on GDP data from 1970 onwards. Second, the cosine similarities only measure whether the word "democracy" and words from our wordlists are talked about together and in similar ways. It compresses the specific contexts and sentiments with which words are talked about: for example, does a higher association between the word "democracy" and words from the economic liberalism word list reveal a rejection of the ideology, or an affinity, or neither? Further research may analyze the concordance of democracy and specific words, such as "free market," and generate more detailed historiography of how exactly democracy and free-market are talked about together. Third, we do not have causal mechanisms on the nation-level to fully validate and explain the results. The lack of mechanisms also entails that we are uncertain of which maximum lag value to employ for granger's causality test.


**Figure 3: Representation of time series of GDP per capita growth and cosine similarities of growth and recession with democracy**

While this section explores the relationship between conceptions of democracy and GDP per capita growth—a common objective measure of a country's economic performance, we would like to gather subjective and qualitative measures that can shed enriched information about the perceptions of democracy across cultures and time. We do this in the following section by projecting the word "democracy" on the cultural dimensions of affluence and morality.

# Section 4: Projecting Democracy on Dimensions of Affluence and Morality

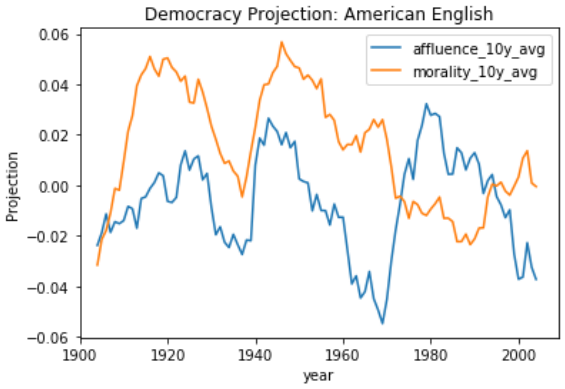
In the political science literature, studying public attitudes toward democracy often requires empirically acquired data such as the World Value Survey to understand temporal shifts in attitudes. Kozlowski et al. applied word embedding models to the sociological culture analysis and showed that cultural dimensions derived from the vector spaces correspond to the individuals' perceptions of affluence, gender, and status in the real world (2019). We adopted the affluence and morality word lists from this work in an attempt to project the word "democracy" onto the rich-poor and good-bad value spectrums of each year (See Appendix 3 for the full translation). We argue that the temporal fluctuations of the respective positions of the word "democracy" can reflect how people of a certain nation-state feel about democracy differently through time.

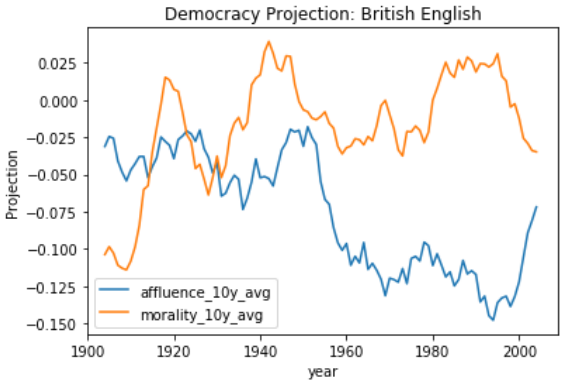
To construct the spectrums for projection, the affluence and morality word lists are translated into the six target languages; the normalized vector differences between rich words and poor words in the affluence wordlist, and the normalized vector differences between moral words and immoral words constitute two dimensions. The keyword "democracy" in the corresponding languages is projected on the spectrums of each year. If democracy leans toward the rich or moral side, then the cosine similarities to the wordlists would be positive values, and negative values otherwise. The results are presented in Figure 4, with the blue and orange lines representing the rolling averages of the projection

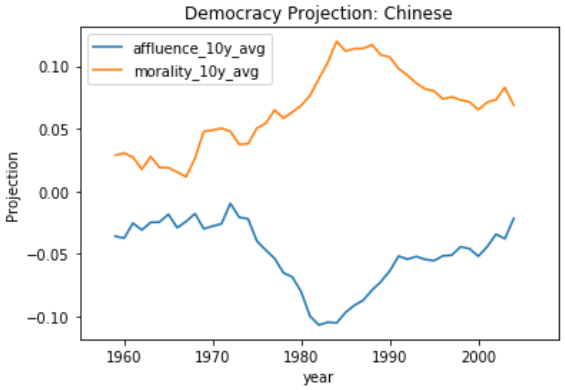
values of affluence and morality spectrums. The line plots show that their peaks and valleys correspond to historical events of the nation-states. Major political events such as world wars, regime changes, social movements, and financial crises match the turning points or drastic shifting of the relative position of democracy.

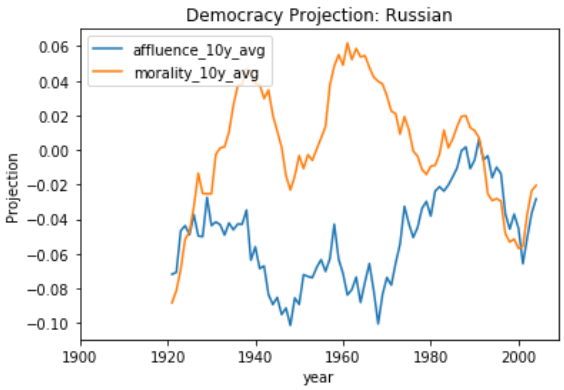
For example, in both American and British English corpuses, democracy was projected on the 'richer' and 'better' sides of the spectrums during the two World Wars. The trend corresponds to the theoretical argument of democracy as a justification to fight against counter-ideologies in wars. The countries who value democracy and consider themselves as authentic democracies hold the responsibility for peace-making, and thus the capability and strength were emphasized in general wars.

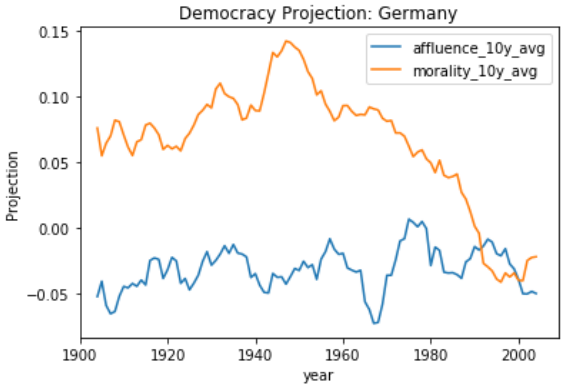
In the Chinese projection, the morality line peaked and the affluence line plunged in the mid-1980s. The trend could possibly resonate with the prelude of the 1989 Tiananmen Square Incident. Since the start of the economic reform in 1978, the incumbent political leader Deng Xiaoping coined "the Socialism of Chinese Characteristics," claiming to develop socialist democracy under the Communist Party, as part of the ideological justification of embracing market economy and derailing from the Soviet-styled planned economy. The Chinese government has since created an official narrative of "democracy" distinctive from Western constitutionalism, which emphasizes eliminating poverty. The economic reform and ideological collection could explain why the censored publishing sector presents the rebound of affluence association to democracy. Conversely, in Russia, the "richness" and the "morality" of democracy reached a small peak in the late 1980s to 1990, which corresponds to the collapse of the Soviet Union.

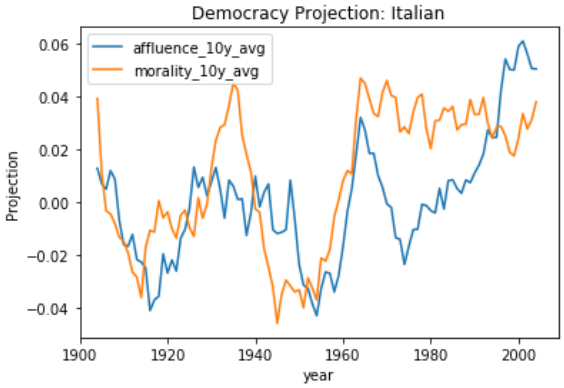

Democracy Projection: American English
affluence_10y_avg
morality_10y_avg
Projection
year


Democracy Projection: British English
affluence_10y_avg
morality_10y_avg
Projection
year


Democracy Projection: Chinese
affluence_10y_avg
morality_10y_avg
Projection
year


Democracy Projection: Russian
affluence_10y_avg
morality_10y_avg
Projection
year


Democracy Projection: Germany
affluence_10y_avg
morality_10y_avg
Projection
year


Democracy Projection: Italian
affluence_10y_avg
morality_10y_avg
Projection
year

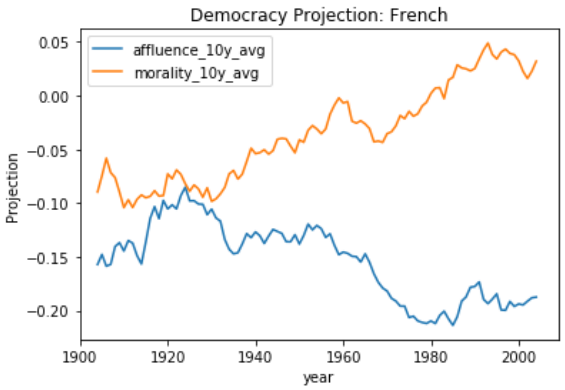

Democracy Projection: French
affluence_10y_avg
morality_10y_avg
Projection
year

**Figure 4: Projection of democracy on affluence and morality**

We acknowledge three limitations in our method. First, the time lagging effect is uncertain, which means that changes in perception can happen and be represented immediately after a certain incident, or the books can be discussing a historical event decades later. Since the model is pre-trained, we have no access to the concordances and contexts, which refrains our scope of interpretation. Second, the different sample sizes and linguistic properties prevent direct comparison among languages. Thus, we could not examine the different effects of a global shock (e.g. World Wars) on democracies and authorities. Third, there are a variety of possible explanations for the high mortality and low affluence periods commonly observed in different languages. One plausible explanation of this pattern in the 1970s could be the increasing discussion on waves of decolonization and democratization; the prevalence of the Modernization Theory gave rise to the discourse of political transformation and economic development in developing countries.

# Conclusion

This paper provided a detailed analysis of conceptions of democracy in seven nation-states. We demonstrated various ways in which word-embedding models trained on Google's Ngram corpus can be used to render a historical investigation into how different nation-states construct political ideologies with a high temporal resolution—what story they narrate about democracy. In our future work, we hope to extend the scope of this paper, employ our abduction process to other aspects of "democracy" and other political ideologies, and adopt a more robust methodology and validation, which includes interrogating these quantitative findings with specific excerpts from the corpus to capture a richer understanding of what these numbers reflect.

# References


Gerring, J., Bond, P., Barndt, W. T., & Moreno, C. (2005). Democracy and economic growth: A historical perspective. World politics, *57*(3), 323-364.

Crow, D. (2010). The party's over: citizen conceptions of democracy and political dissatisfaction in Mexico. *Comparative Politics*, *43*(1), 41-61.

Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy1. *American political science review*, *53*(1), 69-105.

Mikolov, T., Yih, W. T., & Zweig, G. (2013, June). Linguistic regularities in continuous space word representations. In Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies (pp. 746-751).

Miller, A., Hesli, V., & Reisinger W. (1997). Conceptions of Democracy Among Mass and Elite in Post-Soviet Societies. *British Journal of Political Science, 27*(2), 157-190. doi:10.1017/S0007123497000100

Przeworski, A., Alvarez, R. M., Alvarez, M. E., Cheibub, J. A., Limongi, F., & Neto, F. P. L. (2000). *Democracy and development: Political institutions and well-being in the world, 1950-1990* (No. 3). Cambridge University Press.

Ruiz, F. J., Athey, S., & Blei, D. M. (2020). Shopper: A probabilistic model of consumer choice with substitutes and complements. *The Annals of Applied Statistics*, *14*(1), 1-27.

Schultz, K. A. (1999). Do democratic institutions constrain or inform? Contrasting two institutional perspectives on democracy and war. International Organization, 53(2), 233-266.

Ulbricht, T. (2018). Perceptions and conceptions of democracy: Applying thick concepts of democracy to reassess desires for democracy. *Comparative Political Studies*, *51*(11), 1387-1440.

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. American Sociological Review, *84*(5), 905-949.

Zhai, Y. (2019). Popular conceptions of democracy and democratic satisfaction in China. *International Political Science Review*, *40*(2), 246-262.

Almond, G. A., & Verba, S. (2015). The civic culture. Princeton university press.

Timmermans, Stefan and Iddo Tavory. 2012. "Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis." Sociological Theory 30(3) 167–186.

The World Bank. (2022). World Development Indicators. Accessed February 25th, 2022, https://databank.worldbank.org/source/world-development-indicators#.

# Appendix

All codes and data used in this project can be found in the GitHub repository: https://github.com/cultural-proj/democracy

1. Most similar words, frequency distributions, complements substitutes for 7 nation-states, 6 languages: https://github.com/cultural-proj/democracy/tree/main/comsub.

2. Exogenous, open-source, theoretically driven word lists pertaining to economic development, economic ideology, and governance:

| | | |
|---|---|---|
| Economic Development Wordlist | Growth | development, employment, prosperity, boom, bullish |
| | Recession | underdevelopment, unemployment, stagnation, recession, bearish |
| Economic Ideology Wordlist | Friedman | industrialization, modernization, globalization, capitalism, market |
| | Marx | socialism, equality, environmentalism, tax, welfare |
| Governance Wordlist | Centralization | bureaucracy, authority, sovereignty, administration, dictatorship, authoritarian, communism |
| | Decentralization | anarchism, libertarianism, liberalism, federalism, individualism, pluralism, public |

3. Word lists used to construct cultural dimensions of affluence and morality:

| | | |
|---|---|---|
| Affluence wordlist | Rich | richer, richest, affluence, advantaged, affluent, classy, costly, exorbitant, expensive, exquisite, extravagant, flush, invaluable, lavish, luxuriant, luxurious, luxury, moneyed, opulent, plush, posh, precious, priceless, privileged, propertied, prosperous, developed, solvency, successful, sumptuous, swanky, opulent, upscale, valuable, wealthy, ritzy, opulence, solvent, moneyed, rich, affluence |
| | Poor | poorer, poorest, poverty, disadvantaged, destitute, beggarly, economical, impecunious, inexpensive, ruined, necessitous, skint, cheap, economical, penurious, threadbare, cheap, unmonied, indigent, threadbare, plain, cheap, worthless, underprivileged, bankrupt, unprosperous, underdeveloped, insolvency, unsuccessful, plain, basic, needy, squalid, valueless, impoverished, ramshackle, indigence, insolvent, moneyless, penniless, penury |
| Morality wordlist | Moral | good, moral, honest, virtuous, virtue, righteous, chaste, principled, unquestionable, noble, uncorrupt, scrupulous, altruistic, chivalrous, honest, commendable, pure, dignified, holy, valiant, upstanding, guiltless, decent, chaste, righteous, ethical |
| | Immoral | evil, immoral, bad, dishonest, sinful, vice, wicked, transgressive, unprincipled, questionable, nefarious, corrupt, unscrupulous, selfish, knavish, crooked, reprehensible, impure, undignified, unholy, fiendish, villainous, guilty, indecent, unsavory, odious, unethical |