

Economics from the Top Down

new ideas in economics and the social sciences

Have We Passed Peak Capitalism?

Blair Fix

July 12, 2022



Among leftists, predicting the end of capitalism is a favorite parlor game. For example, as a graduate student in the 2010s, I remember discovering the 1976 edition of Marx's *Capital* and being struck by the introduction. Written by the Belgian Marxist [Ernest Mandel](#), the foreword concluded that it was 'most unlikely' that capitalism would survive another half-century.

This prediction (and many like it) did not age well. What capitalism's critics often misunderstand is that social orders rarely 'die'. More often, they fade into irrelevance. Just as no one can point to the end-date of feudalism, it seems unlikely that capitalism will have a decisive 'finish'. But what it may have is a *peak*.

The goal of this post is to chart the rise (and potential peak) of 'capitalism' . . . as I understand it. This caveat is key. To study a social system, we must first define it. To many people, capitalism is a 'mode of production' (a definition inherited from Marx). The view that I take here, however, is that capitalism is primarily an ideology — or what Jonathan Nitzan and Shimshon Bichler call a '[mode of power](#)'. Capitalism is a set of ideas that justify the modern social order.

Although there are many ways to chart the rise of capitalism, what interests me here is that it was the first major ideology to have spread during the era of mass publication. That means capitalism's rise (and potential peak) should be visible in the word frequency of written language.

For example, as capitalism spread, we'd expect that capitalist jargon — words like 'market' and 'price' — should become more common. And feudal jargon — words like 'fief' and 'vassal' — should become less common. Now, I've

chosen these specific words as an illustration. But for my actual analysis, I do not ‘choose’ the jargon words. Instead, I choose a corpus of text that I believe encapsulates the ideology in question (capitalism or feudalism). And from there, I let the jargon of the text speak for itself. The basic idea is that jargon words are those that are both frequently used in a text corpus and *overused* relative to mainstream English.

The first step of the analysis, then, is to select a corpus of ideological texts. To capture feudal ideology, I use a sample of 22 modern English bibles. I use modern translations because I don’t want text that contains archaic words (like ‘thou’). And I use the Bible because christian theology formed the backbone of European feudalism.¹ To capture capitalist ideology, I use a sample of 43 introductory economics textbooks. My claim is that these textbooks deal mostly in capitalist metaphysics; they describe a fantasy world of self-equilibrating markets in which each person earns what they produce.²

¹One could argue that since the various books of the Bible were written over many centuries, the text advocates a hodgepodge of different ‘isms’. (For example, the Old Testament is remarkably ‘tribal’.) That said, I don’t want to worry too much about the specific ‘ism’ that the Bible decrees. The point is that christian ideology formed the foundation of class relations in medieval Europe. For example, here is an oath of fealty used by the Emperor Charlemagne in 802:

By this oath I promise to be faithful to the lord Charles, the most pious emperor, son of King Pepin and Bertha, as a vassal should rightfully be to his lord, for the preservation of his kingdom and of his rights. And I will keep and hope to keep this oath which I have sworn as I know it and understand it, from this day henceforward, with the help of God, the creator of heaven and earth, and of these sacred relics.’

(quoted in Ganshof, 1952)

Notice the use of biblical language, both in the explicit reference to God, but also in use of the word ‘lord’ — a biblical synonym for God. The hallmark of christian ideology is its use of a heavenly hierarchy to legitimize earthly ones.

²The ideology of economics evolved with capitalism. It started life as ‘political economy’, a project that was explicit about the values it advocated. With the neoclassical revolution of the late 19th century, however, these values were purged and what was left was called ‘economics’ — a supposedly scientific description of a market economy. It is this neoclassical theory that fills modern economics textbooks.

The problem is that there is very little ‘science’ in neoclassical economics. Generations of critics have shown that its foundations are either untenable or untestable. (For example, see Steve Keen’s classic book *Debunking Economics* and Paul Mirowski’s *More Heat Than Light*. For a shorter summary, see Bichler and Nitzan’s ‘[The 1-2-3 Toolbox of Mainstream Economics](#)’.)

Like the Bible, neoclassical economics deals almost completely with metaphysics. Thus, it seems fitting to call economics an ‘ideology’. Its central claim is that when left alone, competitive markets lead to an optimal society. However, like the God of the Bible, these competitive markets are nowhere to be found in the real world. Instead, we find a world filled with massive corporations who wield power in the name of the ‘free market’.

With my sample of biblical and economics text, I first isolate the jargon words of each corpus. Then I use the Google English corpus to measure how the frequency of this jargon has changed over time. (As a consistency check, I also analyze the text in paper titles on the Sci-Hub database and book titles in Library Genesis.)

I find that over the last several centuries, biblical jargon became less popular and was slowly replaced by economics jargon. I also find evidence that the popularity of economics language peaked during the 1980s, and has since declined. Ominously, this peak coincides with an uptick in the popularity of biblical language. In simple terms, it seems that we (anglophones) are in the midst of an ideological transition.

Why ideology matters

In a moment, I'm going to dive into my methods for analyzing ideology. But before doing that, it's worth emphasizing why I think the study of ideology is important.

Karl Marx was likely the first political economist to exhaustively study the transition from feudalism to capitalism. In Marx's eyes, this transition was to be understood in material terms. Ideology was not an explanation . . . it was something to be explained:

Just as our opinion of an individual is not based on what he thinks of himself, so we cannot judge . . . a period of transformation by its own consciousness. . . . [O]n the contrary, this consciousness must be explained from the contradictions of material life.

(Karl Marx, 1859)

This is Marx's doctrine of historical materialism, which I think is boldly wrong. Ideas matter, and they have a life of their own within our collective consciousness. In a sense, Marx proved so himself by writing an ideological tome about capitalism that went on to inspire many anti-capitalist revolutions.

In broader terms, humans are a cultural species, which means that our behavior is driven in large part by our collective ideas. How and why these ideas evolve is poorly understood. But what seems clear is that the evolution of our ideological milieu is an important topic of inquiry. To turn Marx on his head, we can and should judge a period of transformation by its own consciousness.

Dissecting an ideology

Given that we're interested in the transformation of ideology, how should we study it? One option is what I will call the 'philosophical approach'. To understand ideological change, we analyze the meaning of a belief-system. Then, we try to map this meaning onto the changing social zeitgeist.

Another option is what I will call the 'statistical approach'. Rather than engage with an ideology's broader meaning, we break the belief system into its smallest components — namely, words. The idea is that the features of an ideology can be discerned through the frequency of its vocabulary. Words that are frequent tell us what the ideology emphasizes. And words that are infrequent tell us what the ideology ignores. The main advantage of this statistical approach is that it allows us to parse massive quantities of text, and so gauge the written zeitgeist in a way that no philosopher ever could.³

To analyze the rise of capitalist ideology, I will use the statistical approach. The core of my method is illustrated in Figure 1, which demonstrates how I use word frequency to classify the vocabulary of a text.

On the horizontal axis in Figure 1, I plot word frequency within the text itself. As an example, in economics textbooks we expect that the word 'price' will be ubiquitous, while the word 'acetophenone' (an organic compound) will be rare.

Although it is tempting to judge a text by its word frequency alone, we need to account for a more general feature of language, which is that some words get used more than others. For example, in the broader English language, 'price' is a common word. And except for a small corner of organic chemistry, 'acetophenone' is extremely uncommon. So to judge economists' use of these words, we should also compare their frequency to what is found in 'mainstream English'. I will call this comparison 'relative frequency':

$$\text{relative frequency} = \frac{\text{text frequency}}{\text{frequency in mainstream English}}$$

³To give you a sense of the analytic scale we can achieve with the statistical approach, consider the size of the Google English database. It reports the frequency of about 800 billion words (written over more than four centuries) and can be parsed by a modern computer in a few minutes. (Compiling this database took far longer.) If a person attempted to read the same volume of text, they would die long before they finished. For example, supposing you read 300 words per minute for 8 hours a day, 365 days a year, you would need about 15,000 years to parse 800 billion words.

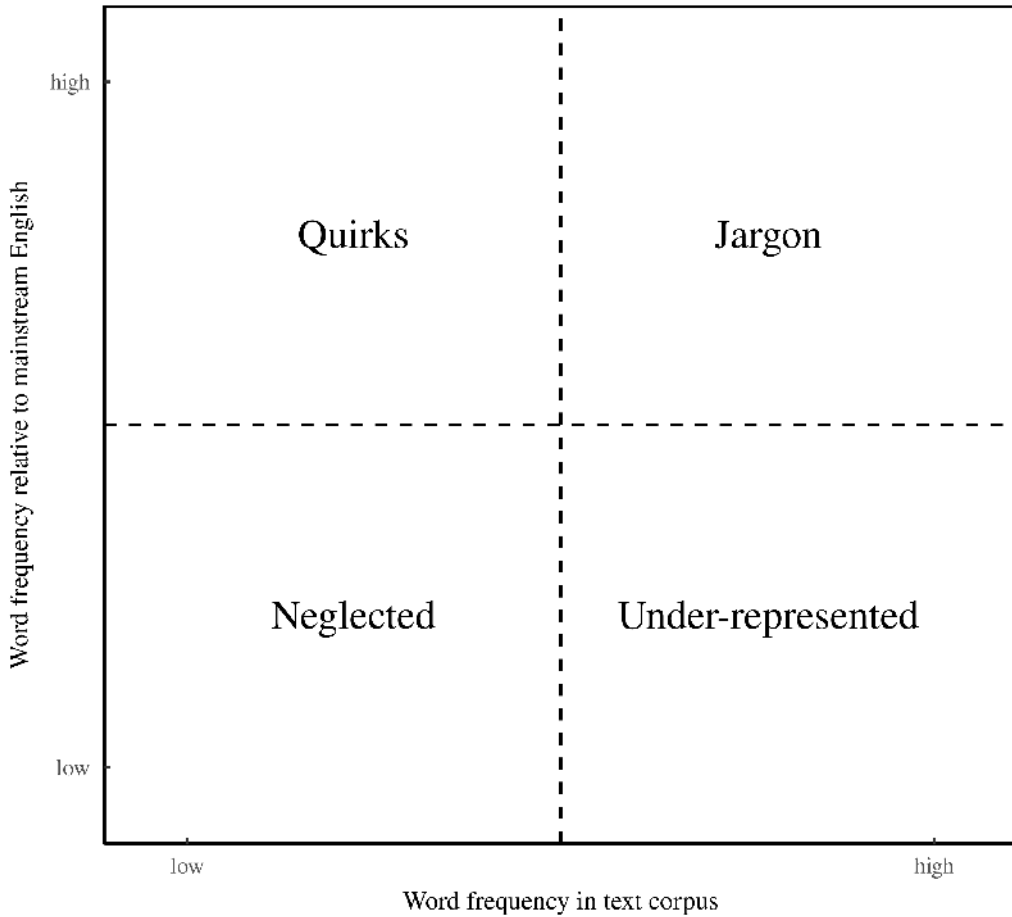


Figure 1: Categorizing the vocabulary of a corpus of text

This figure illustrates my method for classifying words within a sample of text. On the horizontal axis I measure word frequency within the text itself. On the vertical axis, I compare this frequency to that found in ‘mainstream English’. Based on these two measurements, I categorize words into four quadrants: ‘jargon’, ‘quirks’, ‘neglected’ or ‘under-represented’.

In Figure 1, I plot this ‘relative frequency’ on the vertical axis. Based on these two dimensions of frequency, we can classify the vocabulary of a text into four quadrants:

1. *Jargon*: words that are common in the text corpus and overused relative to mainstream English.
2. *Quirks*: words that are rare in the text corpus, yet overused relative to mainstream English.
3. *Under-represented*: words that are common in the text corpus but underused relative to mainstream English.

4. *Neglected*: words that are rare in the text corpus and underused relative to mainstream English.

Of these four quadrants, ‘jargon’ words and ‘neglected’ words are the most interesting. Jargon words tell us about the concepts that define a text. And neglected words tell us about the omissions that define a text.

I will use this classification scheme to measure the changing ideological landscape of the English language. In what follows, I assemble a corpus of economics textbooks and a corpus of bibles. Then I determine their ‘jargon’ using the quadrant system shown in Figure 1. The advantage of this approach is that we do not need to choose the jargon words ourselves. Instead, we let the text speak for itself.

That said, my classification scheme still requires some subjective decisions. Most importantly, to construct our quadrants, we need to compare word frequency in the text to the frequency found in ‘mainstream English’. And that means we need to decide on a representative sample of mainstream writing. However, given that there are many forms of writing (books, newspapers, websites, email, text messages), it is not clear what a ‘representative sample’ of English writing would/should be. To simplify things, I will focus only on the language found in books.

As my measure of ‘mainstream English’, I use data from the Google English corpus. This massive dataset, which is derived from Google’s trove of digitized books, contains a sample of roughly 800 billion words covering text written over the last four centuries. While I will call this sample ‘mainstream’ English, note that I do not mean ‘colloquial’ English. The Google corpus includes both popular books as well as technical monographs. So it does not tell us about the language heard in a typical English conversation. Instead, the Google corpus quantifies the average linguistic patterns found across the whole spectrum of anglophone books.⁴

The other major decision that goes into my word classification system is the choice of threshold that separates the left and right quadrants (in Figure 1). There is no linguistic feature that tells us exactly where this threshold

⁴Some scientists are skeptical of using the Google English corpus to gauge the English language. (For example, see [this paper](#) by Pechenick, Danforth and Dodds.) One problem is that the corpus contains one of each book, so it does not account for a book’s popularity. I personally do not think this is a problem, so long as we are clear that we are measuring what authors *write*, not what people read. Another criticism is that the Google books corpus contains a huge number of scientific books, so it is not a sample of ‘colloquial’ English. Again, I don’t see this fact as a problem, since I am concerned here with the state of ideas expressed in English, not the state of the ‘common’ tongue.

should lie. However, when we plot word frequency on a logarithmic scale (see the horizontal axes in Figures 2 and 5), I find that the value of 50 words per million lies roughly in the middle of the frequency range. So I use this number as my dividing line between left and right quadrants.

Analyzing the language of economics

Having described my word classification scheme, let's now apply it to the language found in economics textbooks. As a reminder, these books serve as my corpus of 'capitalist ideology'.

To analyze the language of economics, I have collected a sample of 43 undergraduate economics textbooks, shown in Table 1. Although not exhaustive, this sample contains most of the standard texts used in undergraduate economics courses (taught in English). When possible, I tried to get the 'micro', 'macro' and 'general' versions of each book. To isolate the textbook content, I trimmed each textbook of front matter, and then fed the resulting text into a word-counting algorithm. (For details about the text processing, see the Sources and methods.) The resulting language sample contains about 10.7 million words, with a vocabulary of roughly 35,000 unique words.

In Figure 2, I take the vocabulary in these economics textbooks and plot it on my quadrant system. In this chart, each point represents a word. The horizontal axis shows the word's frequency in economic textbooks. On the vertical axis, I plot relative frequency — the ratio of textbook frequency to the frequency in the Google English corpus. In each quadrant, the colored points represent the 1000 words that are most overused (for quirks and jargon) or most underused (for neglected and under-represented). Looking ahead, it is the red-colored 'jargon' words whose frequency I will track over time.

Some things to note about Figure 2. If the language in economics textbooks was identical to the Google English corpus (something we do not expect), then all the points would cluster around the horizontal axis. But that is not what happens. Instead, we find a large vertical spread in relative frequency. This spread indicates that economics textbooks overuse some words (those which appear above the horizontal axis) and underuse others (those which appear below the horizontal axis).

When interpreting the data in Figure 2, note that both axes use logarithmic scales. These scales allow us to visualize the entire vocabulary of our text corpus on one chart; however, they have the effect of compressing fre-

Table 1: A corpus of capitalist ideology from 43 economics textbooks

Author	Title	Year
Arnold	Economics	2008
Arnold	Macroeconomics	2008
Arnold	Microeconomics	2011
Blanchard & Johnson	Macroeconomics	2012
Case, Fair & Oster	Principles of Economics	2012
Case, Fair & Oster	Principles of Macroeconomics	2011
Case, Fair & Oster	Principles of Microeconomics	2008
Cowen & Tabarrok	Modern Principles of Economics	2011
Frank & Bernanke	Principles of Economics	2008
Frank & Bernanke	Principles of Macroeconomics	2008
Frank & Bernanke	Principles of Microeconomics	2008
Hubbard & O'Brien	Economics	2009
Hubbard & O'Brien	Macroeconomics	2011
Hubbard & O'Brien	Microeconomics	2013
Krugman & Wells	Economics	2009
Krugman & Wells	Macroeconomics	2005
Krugman & Wells	Microeconomics	2012
LeRoy Miller	Economics Today	2011
LeRoy Miller	Economics Today: The Macro View	2011
LeRoy Miller	Economics Today: The Micro View	2011
Mankiw	Principles of Economics	2008
Mankiw	Principles of Macroeconomics	2011
Mankiw	Principles of Microeconomics	2011
McConnell, Brue & Flynn	Economics	2008
McConnell, Brue & Flynn	Macroeconomics	2006
McConnell, Brue & Flynn	Microeconomics	2011
Nicholson & Snyder	Microeconomic Theory	2004
Nicholson & Snyder	Microeconomic Theory	2007
Nicholson & Snyder	Microeconomic Theory	2011
Parkin	Macroeconomics	2011
Parkin	Microeconomics	2011
Parkin, Powell & Matthews	Economics	2005
Perloff	Microeconomics	2011
Perloff	Microeconomics	2014
Pindyck & Rubinfeld	Microeconomics	2012
Pindyck & Rubinfeld	Microeconomics	2014
Rittenberg & Tregarthen	Principles of Economics	2009
Rittenberg & Tregarthen	Principles of Macroeconomics	2009
Rittenberg & Tregarthen	Principles of Microeconomics	2009
Samuelson & Nordhaus	Economics	2009
Varian	Intermediate Microeconomics	2005
Varian	Intermediate Microeconomics	2010
Varian	Intermediate Microeconomics	2014

quency variation, which is actually enormous. For example, in my sample of economics textbooks, the most common word ('price') outnumbers the least

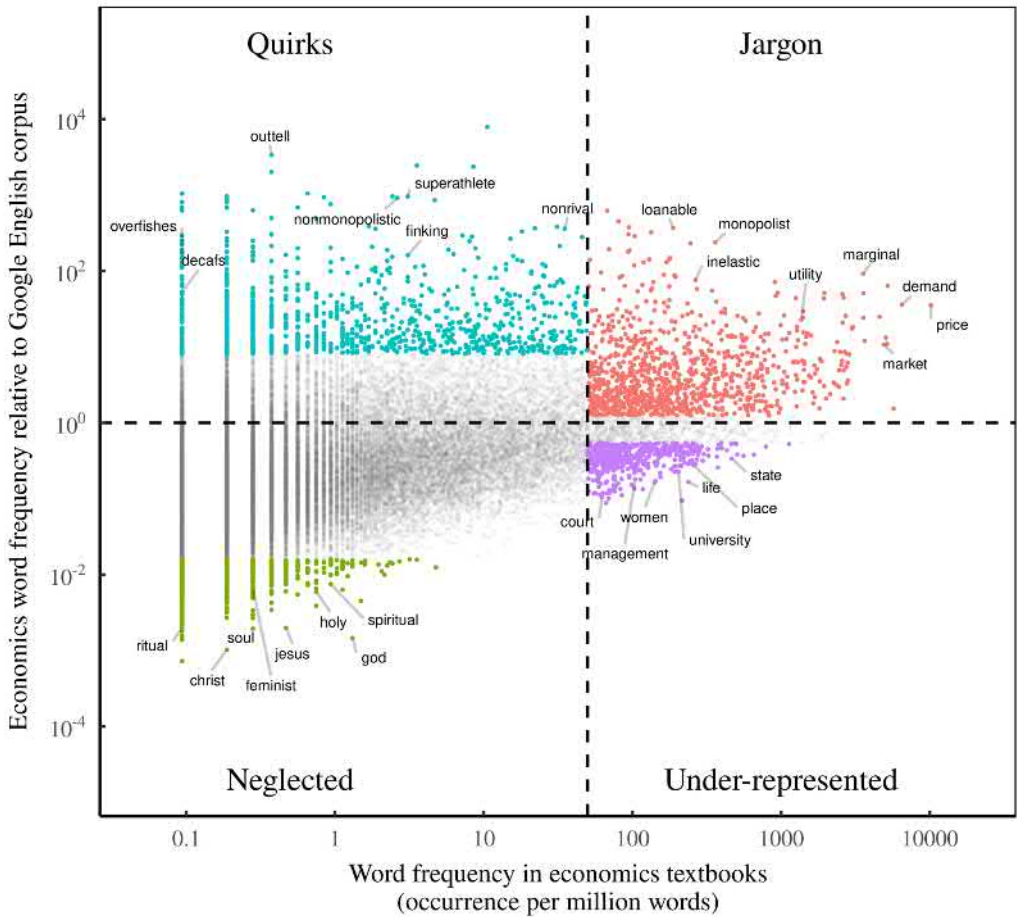


Figure 2: Dissecting the language found in economics textbooks

This figure analyzes word frequency in a sample of 43 economics textbooks. Each point represents a word, with its frequency in economics textbooks plotted on the horizontal axis. On the vertical axis, I compare this textbook frequency to word frequency in the Google English corpus (averaged over the years 2000–2019). In each quadrant, the colored points represent the 1000 words that are most overused (for quirks and jargon) or most underused (for neglected and under-represented). [[Sources and methods](#)]

common words (such as ‘ritual’) by a factor of 100,000. (Note that preposition words would be even more frequent. But since they are uninteresting for ideological analysis, I have removed them. For more details about the text processing, see the Sources and methods.)

Just as word frequency varies immensely, so does relative frequency. For instance, economics textbooks use the word ‘loanable’ about 370 times more frequently than the Google English corpus. And they use the word ‘christ’ about 1000 times less frequently than the Google English corpus.



Figure 3: Economics jargon

This cloud shows the most overused words from the economics jargon quadrant in Figure 2. These are words that are frequent in economics textbooks and overused relative to mainstream English.

Because there are about 35,000 words plotted in Figure 2, I cannot label all of them. However, the word cloud in Figure 3 gives you a sense for the words that define economics writing. Here I show several hundred words from the ‘jargon’ quadrant. When you read an economics textbook, these are the words that stand out.

In addition to its jargon, economics is also defined by the words it neglects. Figure 4 shows several hundred of the these neglected words. Interestingly, many of them relate to christian ideology. (For example, ‘lord’, ‘chapel’, ‘christ’, ‘monk’ and ‘prayer’, and ‘jesus’.)

In a sense, this exclusion of religious words is banal, since economics presents itself as a ‘science’, and all science is secular. My claim, however, is that this scientific appearance is mostly a veneer. Under the hood, economics textbooks care little for the scientific method. (For example, I challenge the reader to find a mainstream economics textbook that subjects its ideas to an empirical test. I doubt you will find one, for the simple reason that neoclassical doctrines amount to untestable metaphysics.)

Viewed in this light, economics ideology is remarkable for being the first widely accepted social myth that did away with gods. Commenting on this new secular belief system, Jonathan Nitzan and Shimshon Bichler [observe](#) that economics took the “hierarchical rule of God” and replaced it with a “flat law of nature, a secular–universal mechanism that regulated and equilibrated both heaven and earth.”

Analyzing the language of the Bible

Having analyzed the language of economics textbooks (my corpus of capitalist ideology), let’s do the same for the christian Bible, which will serve as my corpus of feudal ideology.

To analyze biblical language, I’ve selected a sample of 22 modern translations of the Bible, shown in Table 2. I use modern translations because I want to put these bibles on the same footing as economics textbooks. In other words, I don’t want a biblical text that is littered with archaic words (‘thou’) and tenses (‘sayeth’) — language that we already know has become uncommon. To isolate the content, I have trimmed each Bible of front matter. The resulting language sample contains about 13 million words, with a vocabulary of roughly 32,000 unique words.

In Figure 5, I analyze the vocabulary of this biblical corpus using my quadrant system. Each point represents a word. The horizontal axis shows word frequency in my sample of bibles. The vertical axis compares this biblical frequency to the word frequency found in the Google English corpus. In each quadrant, the colored points represent the 1000 words that are most overused (for quirks and jargon) or most underused (for neglected and under-represented). In the analysis that follows, I will track the frequency over time of the red jargon words.

As with economics, we find that biblical language is different than mainstream English. Diving into these differences, Figure 6 shows several hundred words from the jargon quadrant — words that the Bible most overuses. Here we find the names of many biblical characters and the places in which they lived. We also find a large helping of familiar christian terms — words like ‘tabernacle’, ‘sinners’, and ‘crucified’.

Table 2: A corpus of feudal ideology from 22 modern translations of the Bible

Version	Translation Date
American Standard Version	1901
Bible in Living English	1943
Borean Reader's Bible	2016
Catholic Public Domain Version	2009
Christian Standard Bible	2016
Common English Bible	2010
Easy English Bible	2001
English Standard Version	2001
Good News Bible	1966
JPS Tanakh	1917
NET Bible	2001
New American Bible	1970
New American Standard	1971
New Century Version	1983
New International Version	1973
New Life Version	1986
New Living Translation	1996
New Revised Standard Version	1993
Open English Version	2010
Revised New Jerusalem Bible	2019
The Message	1993
World English Bible	2000

When we turn to the words that the Bible neglects, shown in Figure 7, we find a collection of modern terminology. (For example ‘television’, ‘newspaper’, ‘nuclear’ and ‘matrix’.) Given that the Bible is an iron-age document, it is unsurprising that it should neglect these words — even when it is translated into modern English.

We also find among the neglected words a significant amount of economics jargon — words like ‘output’, ‘industry’, and ‘investors’. As we will see shortly, it turns out that the Bible systematically neglects the jargon of economics (and vice versa). This fact suggests that we are dealing with two mutually exclusive ideologies.

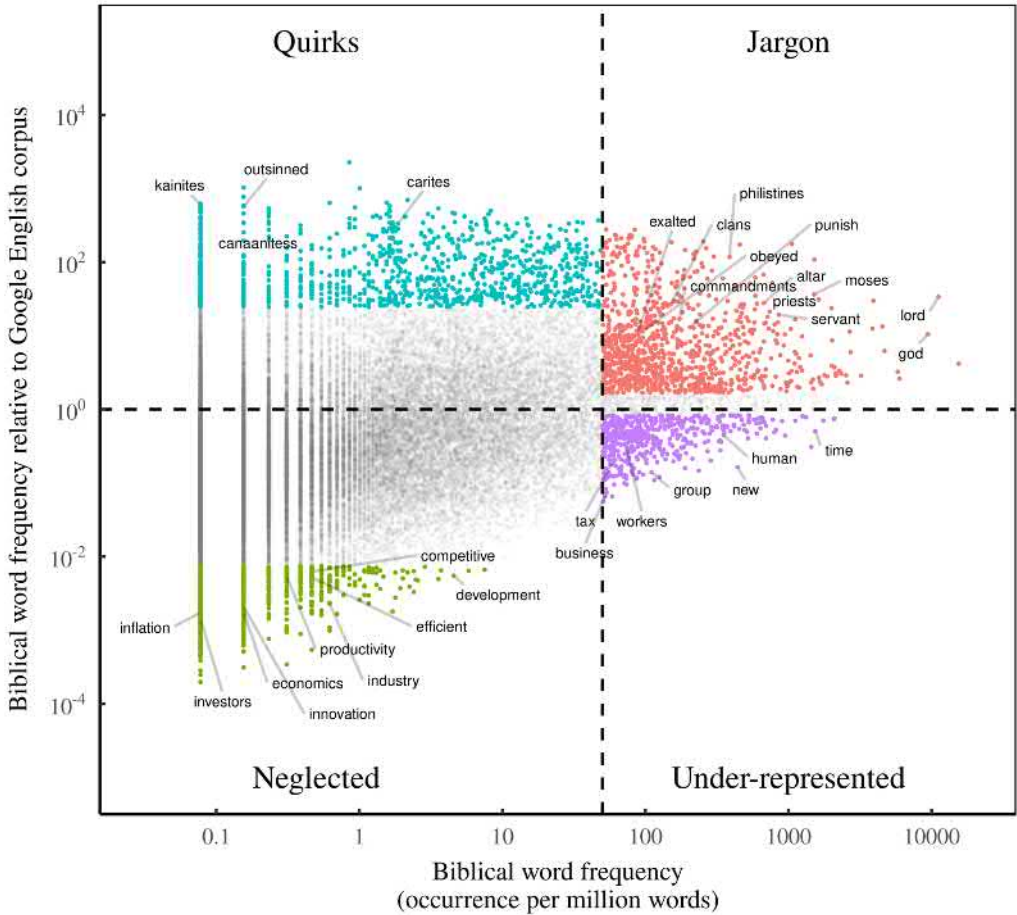


Figure 5: Dissecting biblical language

This figure analyzes word frequency in a sample of 22 modern translations of the Bible. Each point represents a word, with its biblical frequency plotted on the horizontal axis. On the vertical axis, I compare this biblical frequency to word frequency in the Google English corpus (averaged over the years 2000–2019). To characterize the language, I divide it into four quadrants. In each quadrant, the colored points represent the 1000 words that are most overused (for quirks and jargon) or most underused (for neglected and under-represented). [[Sources and methods](#)]

Ideological poles

Looking ahead, I'm going to take the jargon words from my corpus of bibles and economics textbooks and measure how their (respective) frequency has changed with time. Before we get to this analysis, though, it is worth making some predictions.

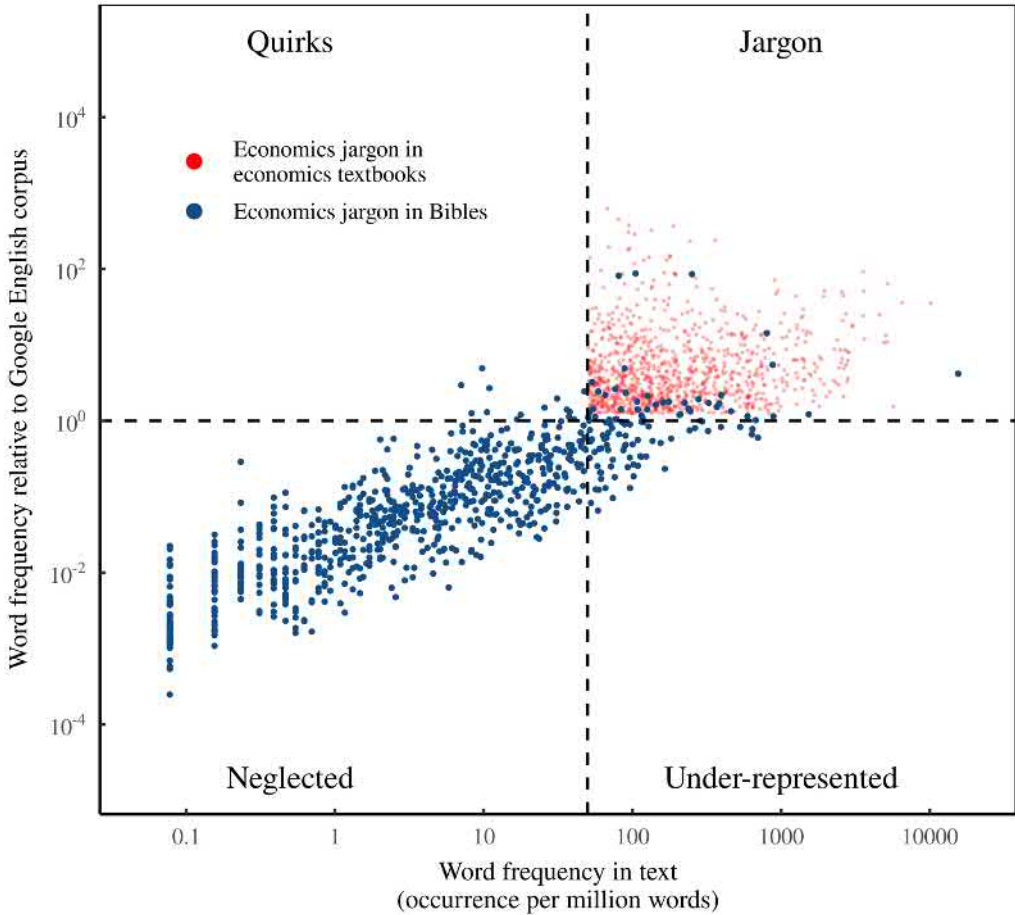


Figure 8: The location of economics jargon in economics textbooks and in the Bible

This figure shows the top 1000 jargon words selected from my sample of economics textbooks. The red points show the ‘position’ of these words in economics textbooks (by definition, in the ‘jargon’ quadrant’). The blue points indicate the ‘position’ of the same words in my sample of bibles. The vast majority of these words are found in the ‘neglected’ quadrant, indicating that the Bible ignores the jargon of economics.

The changing frequency of biblical and economics jargon in the Google English corpus

We’re now in a position to look at the changing ideological landscape that is written in the English language. To quantify ideological change, I measure how the frequency of biblical/economics jargon has changed with time in the Google English corpus.

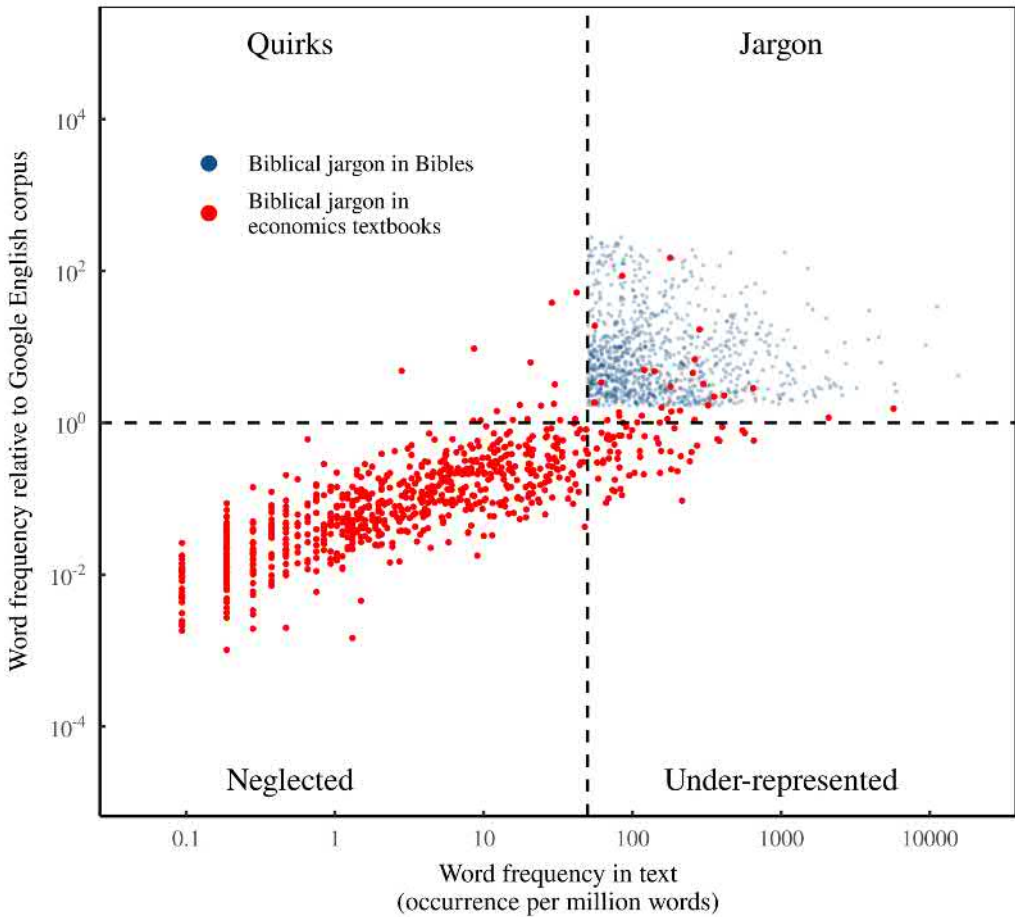


Figure 9: The location of biblical jargon in the Bible and in economics textbooks

This figure shows the top 1000 jargon words selected from my sample of bibles. The blue points show the 'position' of these words in my sample of bibles (by definition, in the 'jargon' quadrant'). The red points indicate the 'position' of the same words in my sample of economics textbooks. The vast majority of these words are found in the 'neglected' quadrant, indicating that economic textbooks ignore biblical jargon.

Figure 10 shows my results. Here the blue line shows the annual frequency of biblical jargon. The red line shows the annual frequency of economics jargon. Note that the frequency is expressed per thousand words, so you can interpret it like a batting average. For example, during the 1950s, for every 1000 words contained within the Google corpus, about 130 of them were economics jargon. (If our economics jargon 'batted' 1000, it would mean that English writing consisted entirely of economics jargon.)

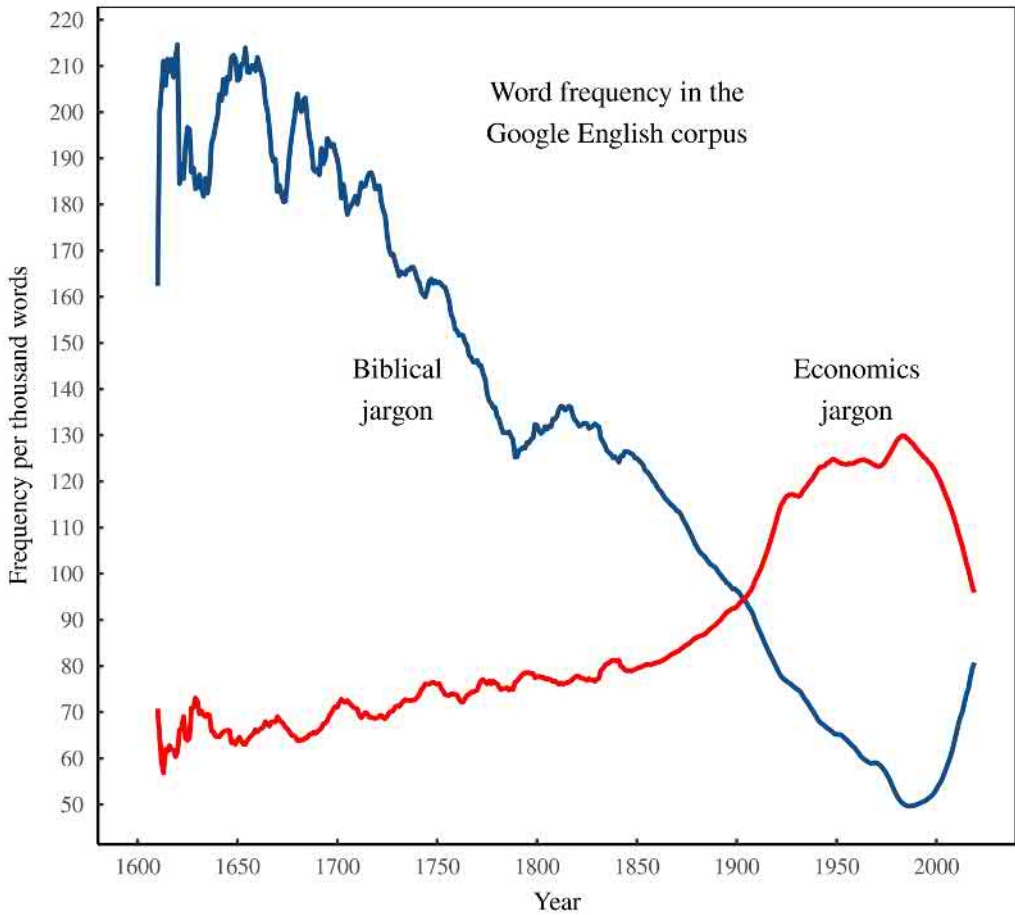


Figure 10: The changing frequency of biblical and economics jargon in the Google English corpus

Each line shows the frequency of biblical/economics jargon within the Google English corpus. I define ‘economics jargon’ as the 1000 most overused words from the jargon quadrant of Figure 2. And I define ‘biblical jargon’ as the 1000 most overused words from the jargon quadrant of Figure 6. In each year, I sum the frequency of these words in the Google corpus, and plot the pattern over time. To smooth the trends, each line shows the 10-year trailing average of jargon frequency.

Looking at Figure 10, let’s start with what is unsurprising. Over the last four centuries, biblical jargon fell out of favor while economics jargon became more popular. This is exactly what we expect for the transition from feudalism to capitalism.

To give you a sense for the scale of this linguistic transformation, note that in the 17th century, English writing was overwhelmingly religious. For every 1000 words written during that period (and captured by the Google English corpus), on average about 200 of them were biblical jargon. That may not

sound significant, until you realize that my sample of bibles themselves only ‘bat’ 300. In other words, for every 1000 words in these bibles, 300 of them are ‘jargon’. So it’s not an exaggeration to say that the English literature of the 17th century was dominated by theology.

From the 18th century onward, however, this christian hegemony steadily waned. At the same time, the jargon of economics grew more popular, becoming the dominant ideology around the turn of the 20th century.

While we should be cautious about interpreting these trends (because they depend in part on my analytic assumptions), they seem consistent with what we know both about the history of capitalism and about the history of economic thought. We know, for example, that the transition to capitalism has deep roots, but that the change accelerated during the late 19th century, as Western countries began to industrialize. We also know that the origin of modern-day economics dates back centuries, but that these ideas were not mainstream until after the marginal revolution of the late 19th century. So over all, the long-term pattern in Figure 10 meets our expectations.

That said, something conspicuous happened after 1980: the popularity of economics jargon began to wane, and the popularity of biblical jargon began to rise. What should we make of this reversal? Does it indicate that we (anglophones) have passed the peak of capitalist ideology? I think the answer is yes.

Reinforcing evidence

There is a saying in science that extraordinary claims require extraordinary evidence. I admit that I cannot muster ‘extraordinary’ evidence to back my claim that we’ve passed the peak of capitalist ideology. However, I can muster several independent lines of linguistic evidence that point to the same conclusion. In what follows, I look at different ways of tracking the popularity of biblical and economics language. Most of these measurements suggest that capitalist ideology (as measured by the language of economics) peaked in the 1980s.

The language similarity index

In Figure 10, I tracked the frequency of biblical and economic ‘jargon’ words. The advantage of this method is that it is (in my opinion) simple to understand. The disadvantage of the jargon approach is that it tracks only a small

portion (about 3%) of the vocabulary of each corpus. To address this shortcoming, I look now at a measure that I call the ‘language similarity index’. This index quantifies the similarity between two language samples using *all* of the words common to both vocabularies.

The idea behind the similarity index is that we can quantify language similarity by comparing the frequency of words in two samples of text. For example, if the word ‘market’ occurred at the same frequency in two text samples, this would indicate that the writing is similar. Of course, ‘market’ is just one word. So what we should do is compare the frequency of *every* word that is contained within both samples of text. The closer the frequency of each word, the more similar the texts.

In more technical terms, the similarity index is defined as follows. Let f_a^i be the frequency of word i in language sample a . And let f_b^i be the frequency of the same word in language sample b . For the intersection of all words found in both samples, the similarity index is defined as:

$$\text{similarity index} = \frac{100}{\text{mean}(\lfloor |\ln(f_a^i) - \ln(f_b^i)| + 1 \rfloor)}$$

Here $||$ indicates taking the absolute value, and $\lfloor \rfloor$ indicates rounding down to the nearest integer. The resulting similarity index varies from 0 (no similarity between language samples) to 100 (the language samples have identical word frequency). The advantage of the similarity index is that it measures all words in a text. The disadvantage is that it requires a large text sample to give accurate results. (See the Sources and methods for a discussion.)

In Figure 11, I apply the similarity index to measure the similarity between the Google English corpus and: (a) my sample of bibles (blue line); and (b) my sample of economics textbooks (red line). Because the similarity index requires a large text sample to be accurate, I restrict the analysis to the last two centuries.

Looking at Figure 11, we see that the trends are not identical to those found using jargon (in Figure 10). Using the similarity index, we find that the decline of biblical language is less pronounced, while the increase of economics language is more rapid. That said, the overall pattern is similar. Over the last two centuries, mainstream English became less similar to modern bibles and more similar to economics textbooks. And around 1980, this pattern reversed.

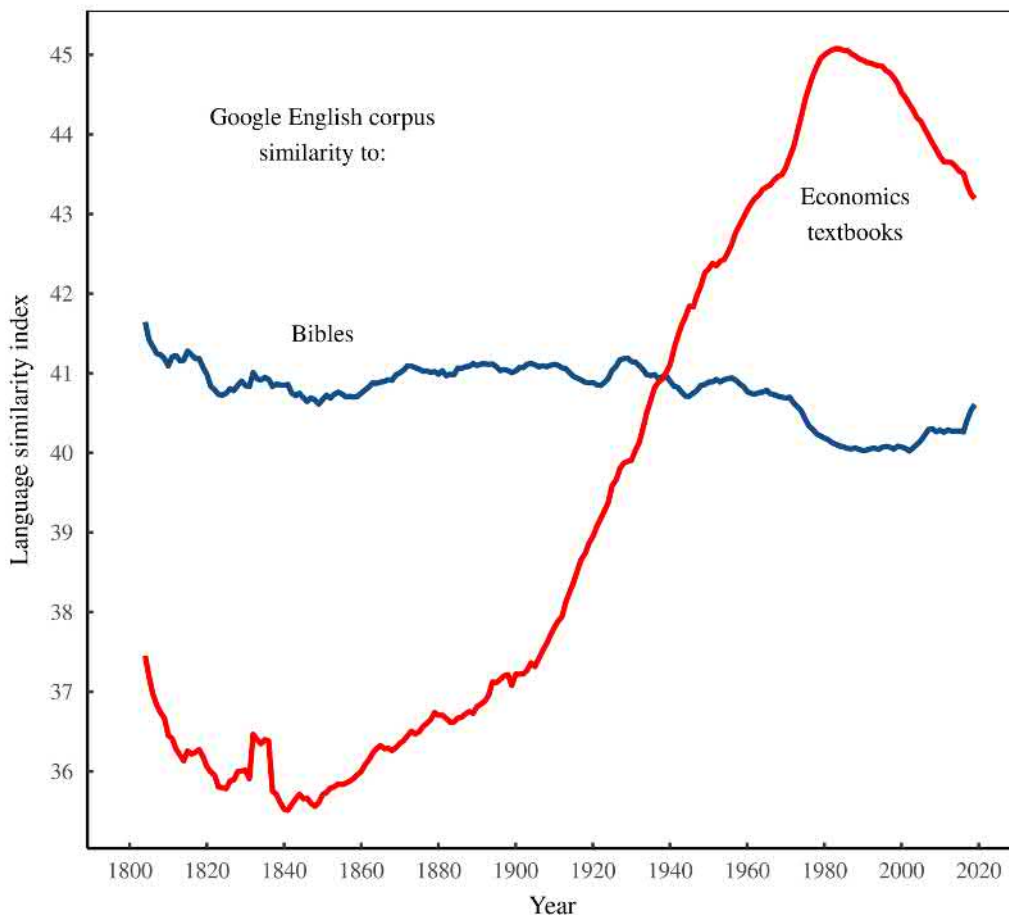


Figure 11: The similarity index between the Google English corpus and my sample of bibles and economics textbooks

The 'language similarity index' measures the extent to which the words in two samples of text have the same frequency. An index of 0 indicates no similarity. An index of 100 indicates that the word frequency is identical. The blue curve compares the similarity of the Google English corpus to my sample of bibles. The red curve compares the similarity of the Google English corpus to my sample of economics textbooks. To smooth the trends, each line shows the 5-year trailing average of the similarity index.

The frequency of biblical and economics jargon in other samples of English writing

Although the Google English corpus is the largest sample (by far) of historical English writing, it is wise to check that our results hold across multiple sets of data. With that in mind, I turn now to the word frequency found in two other sources of historical text:

1. The titles of academic articles on Sci-Hub

2. The titles of books on Library Genesis

For some context, Sci-Hub and Library Genesis are two databases that make copyrighted material available for free. (Publisher's call this act 'piracy'. The database maintainers see it as 'liberation'.) Sci-Hub has amassed a collection of some 90 million scientific papers. And Library Genesis holds a collection of over 20 million books.

Since these articles and books are available in full text, we could in principle analyze their word frequency as Google has done with its corpus of books. However, the scale of the task is enormous; it would involve downloading and processing many terabytes of data.⁶ While certainly possible, I leave this intensive task for the future. Here, I look only at the *titles* within these two databases.

At first glance, studying titles alone seems like a poor source of English text. Keep in mind, however, that Sci-Hub and Library Genesis host many millions of entries. So even when we look only at titles, we are still dealing with a substantial body of writing. (For details about the sample size, see Figure 17 in the Sources and methods.) The caveat is that the Sci-Hub and Library Genesis datasets are not large enough to apply the language similarity index. So I look only at the frequency of jargon.

Figure 12 shows the frequency of biblical and economics jargon within article titles on Sci-Hub. Consistent with the trends found in the Google corpus (Figure 10), it seems that over the last two centuries, biblical jargon fell out of favor, while economics jargon grew more popular. We also see a plateau in this pattern circa 1980. However, it's not clear from the Sci-Hub data that the popularity of economics jargon has actually peaked, since after 2010, we see an uptick in the trend.

Turning to the book titles in Library Genesis, Figure 13 shows the frequency of economics and biblical jargon over the last century. Again, we see that until 1980, biblical jargon waned in popularity, while economic jargon grew more popular. After 1980, the pattern reversed.

⁶For example, public domain advocate Carl Malamud recently released an [ngram database](#) which analyzes the full text of some 107 million academic articles (many of which are suspected to have come from Sci-Hub). The problem is that this dataset is an unwieldy 37.9 Terabytes. Crunching data of this scale is beyond the scope of the current analysis.

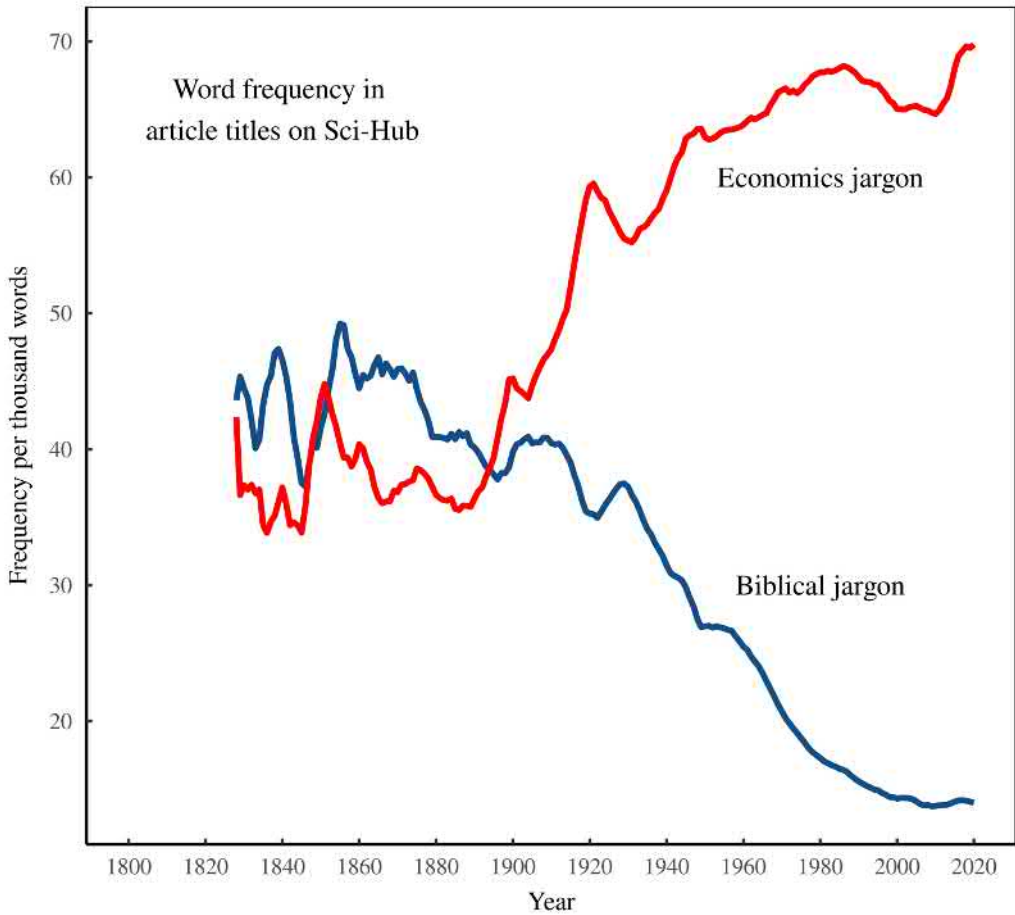


Figure 12: The frequency of economics jargon and biblical jargon in article titles on Sci-Hub

The red curve shows the frequency of economics jargon in the titles of articles in the Sci-Hub database. The blue curve shows the frequency of biblical jargon. The ‘jargon’ of each corpus is defined as the 1000 most overused words from the respective jargon quadrants in Figures 2 and 5. I show data over the period during which the Sci-Hub dataset contains more than 10,000 words per year. To smooth the trends, each line shows the 5-year trailing average of jargon frequency.

Looking at the results in Figures 10–13, the trends from the various measurements/datasets appear consistent with one another. Until about 1980, biblical language became less popular and economics language grew more popular. After 1980, these trends reversed. (For rigorous consistency checks, see Figures 19 and 20 in the Sources and methods.)

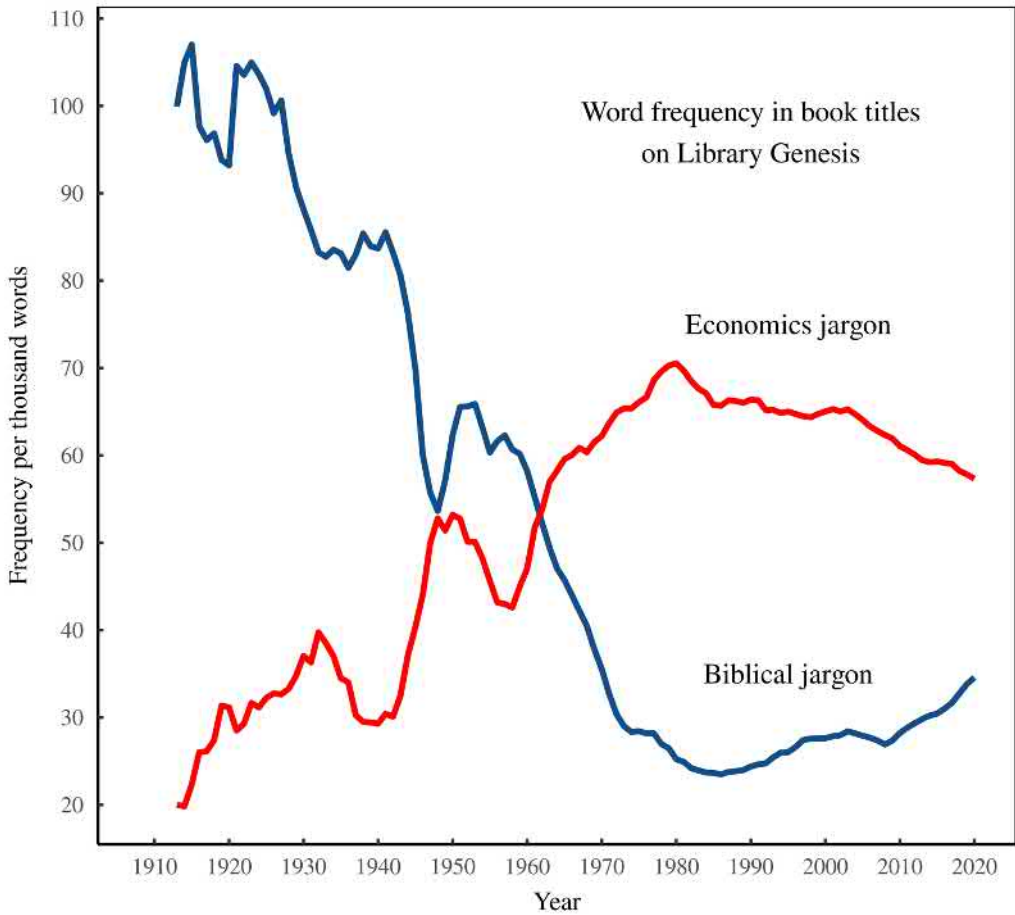


Figure 13: The frequency of economics jargon and biblical jargon in book titles on Library Genesis

The red curve shows the frequency of economics jargon in the titles of books in the Library Genesis database. The blue curve shows the frequency of biblical jargon. The ‘jargon’ of each corpus is defined as the 1000 most overused words from the respective jargon quadrants in Figures 2 and 5. I show data over the period during which the Library Genesis dataset contains more than 10,000 words per year. To smooth the trends, each line shows the 5-year trailing average of jargon frequency.

The peak of capitalist ideology

Having collected several independent measures of English word frequency, we’re now in a position to judge if and when capitalist ideology peaked (as measured by the frequency of economics language).

Figure 14 analyzes the timing of this peak. In panel A, points indicate the date when economics language peaked in popularity, as judged by the corresponding metric on the vertical axis. The error bars show the range for the

top ten years. Using the same convention, panel B shows the date when biblical language reached its minimum popularity. The evidence suggests that economics ideology peaked around 1980, with biblical ideology reaching a minimum around the same time (or shortly thereafter).

Before discussing the significance of this peak, a note of caution. Technically, the evidence demonstrates that there has been a peak in the popularity of economics language. Because the future is unwritten, we cannot be certain that we have observed *the* peak.⁷ However, assuming there is no reversal of current trends, we are left with a fascinating conclusion: capitalist ideology (as I measure it) seems to have peaked around 1980.

The date of this peak is interesting, because it coincides with the so-called ‘neoliberal’ turn in anglophone politics. In 1980, Ronald Reagan became US President, running on a platform that heralded the benefits of free-market capitalism. In Britain, Margaret Thatcher ran on a similar platform, and was elected Prime Minister in 1979.

The term ‘neoliberal’ refers to the fact that this free-market movement (which extended far beyond Reagan and Thatcher) hearkened back to the *laissez faire* policies of the 19th century. The linguistic evidence, however, suggests that this way of characterizing the neoliberal turn is not entirely accurate. Yes, free-market ideas existed in the 19th century and were promoted by neoclassical economists. But they were not the dominant ideology. Instead, it was during the ‘neoliberal’ era that the ideology of economics achieved its maximum appeal. And so perhaps a more appropriate way of thinking about neoliberalism is that it was not a return to the past; it was a statement of ideological *supremacy* — the moment when there really was no alternative to the doctrines of economics.

The neoliberal euphoria didn’t last long. From the 1990s onward, the popularity of economics ideology waned. I can think of many reasons for this decline, all of which relate to the fact that economics promises a free-market road to utopia. The more this myth does not pan out in reality, the less convincing it becomes.

⁷US oil production provides a good example of a peak that was not *the* peak. In 1970, US oil production peaked, and then proceeded to decline for many decades. Many analysts thought they had witnessed *the* peak of oil production. However, during the 2010s, there was an explosion of shale oil extraction that reversed the decline, and eventually pushed oil production to [new heights](#). The lesson: we can never be certain that a peak is permanent.

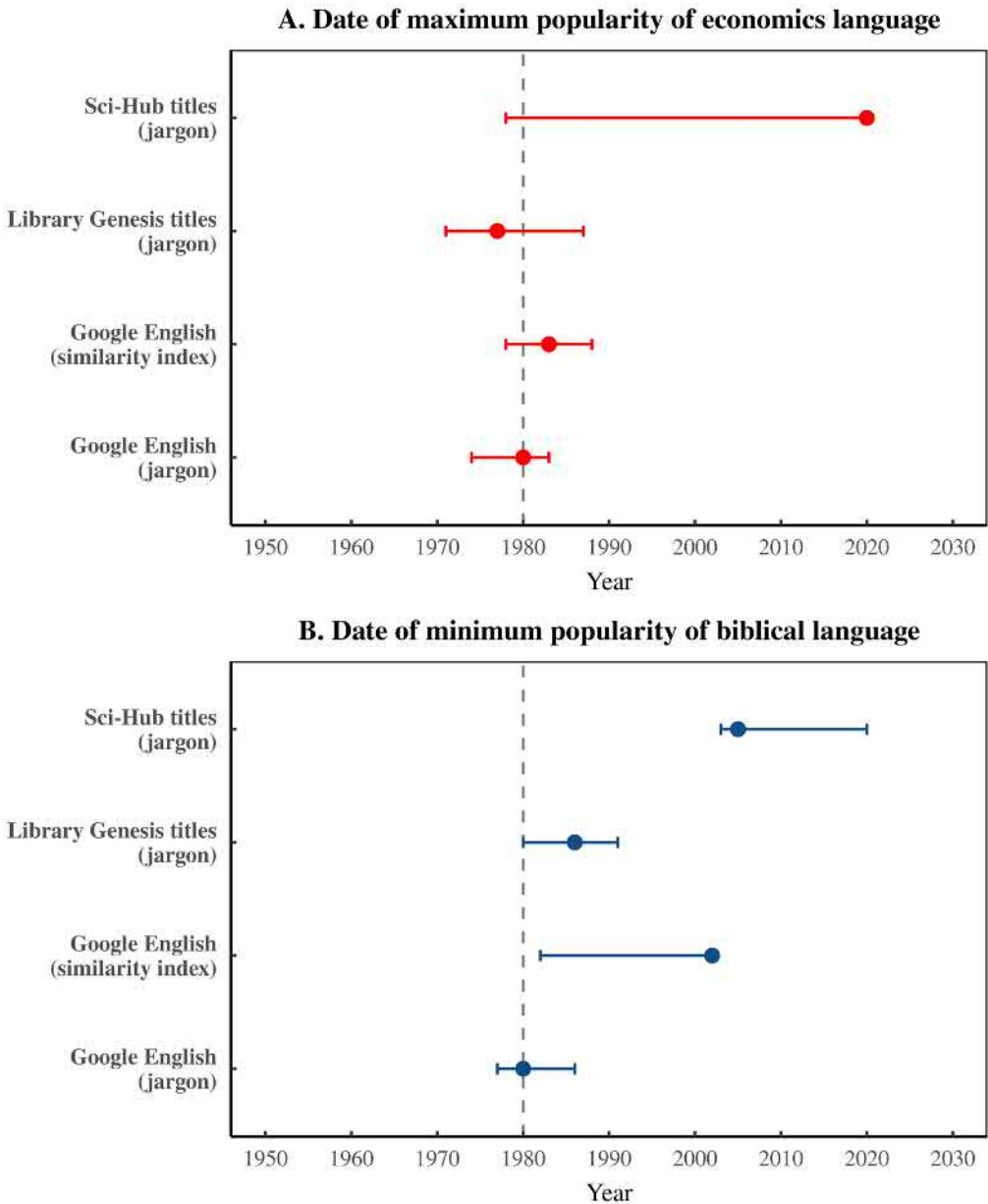


Figure 14: The timing of peak economics

Panel A analyzes the date at which economics language peaked in popularity in my samples of English writing. Panel B analyzes the date when the popularity of biblical language reached a minimum. Points indicate the date of maximum (for economics) or minimum (for biblical language) popularity. Error bars show the range for the 10 years with maximal/minimal popularity.

As an example, take the problem of inequality. Beginning in the 1980s, anglophone countries experienced an [explosion of income inequality](#), likely caused by the policies of neoliberal politicians. Apart from being socially corrosive, the problem with inequality is that it tends to undermine the economic doctrine that income stems from productivity. For example, if a CEO earns 5 times more than the average worker, many people can be convinced that this added income stems from the CEO's greater productivity. But if the CEO earns 500 times more than the average worker, attributing this windfall to 'productivity' seems laughable.

The 1980s also marked the period when climate change became a growing concern. People discovered that mainstream economists paid virtually no attention to the environment, other than to say that the market would solve our problems. But the market did not solve our ecological problems, which are today more severe than ever.

Overall, I suspect that since the 1980s, there has been a wave of hostility to neoclassical economics. Part of this hostility has come from the left, and from social scientists seeking to construct a more realistic theory of human interaction. How impactful this left-leaning movement has been, I do not know. (An interesting research project would be to use the linguistic methods developed here to study the extent of progressive ideology.)

What we do know is that a major source of economics hostility seems to have come from christian reactionaries. Indeed, this is the most plausible explanation for why biblical language is on the rise. And yet here we have a puzzle. Many evangelical christians proclaim a faith in both [God and free markets](#). So you would think that if evangelical ideology became more prominent, biblical and economics language might *both* become more popular. And yet there is little evidence for such a trend.

Without researching the language of evangelical christianity in more detail, it's difficult to resolve this puzzle. But I suspect that given its appeal to biblical literalism, the evangelical movement has less in common with the secular doctrines offered in economics textbooks, and more in common with the religious doctrines of the Bible.

An age of ideological discord

The linguistic evidence suggests that anglophones are in the middle of an ideological transition — a period when economics ideology is losing its dominance but has not yet been replaced by another hegemonic belief system. Assuming this inference is true, then there should be many signs of the ongoing ideological struggle.

In his book *Ages of Discord*, biologist-turned-historian Peter Turchin provides a rich set of evidence that over the past few centuries, the United States has experienced several ‘cycles of discord’. What Turchin means is that the US seems to oscillate between periods of social harmony and periods of conflict. Unsurprisingly, Turchin argues that the United States is now in an ‘age of discord’.

Might this discord relate to the ideological transition that is written in the linguistic data? To test this possibility, I propose something called the ‘ideological discord index’. To frame this index, let’s highlight an interesting (almost paradoxical) feature of ideology. Humans discern ideology much like we smell things. For example, when you walk into a bakery, you immediately smell the aroma. Yet the baker, who has been in the bakery all day, smells nothing. That’s because humans mostly smell *new* scents. When an aroma becomes ubiquitous, it quickly loses its odor.

The same is true of ideology. As a rule, a dominant ideology is an ‘odorless’ one. In other words, when an ideology is ubiquitous, people cannot sense its presence. Instead, they see a ‘natural order’. However, when ideologies are put in conflict, they suddenly have a ‘smell’. ‘Those people’ have the wrong ideology and it ‘stinks’.

The ‘ideological discord index’ builds on this analogy. It proposes that the world appears ‘discordant’ to the degree that no ideology dominates. Conversely, as one ideology becomes hegemonic, the world appears ‘concordant’.

Figure 15 illustrates this thinking. In the top panel, the red and blue lines plot the popularity of two opposing ideologies. Over time, ideology A falls out of favor and is replaced by ideology B. The bottom panel shows the corresponding ideological discord index. When ideology A is dominant, there is little discord. But as the two ideologies approach equal popularity, discord rises. Then, as ideology B becomes dominant, discord falls.

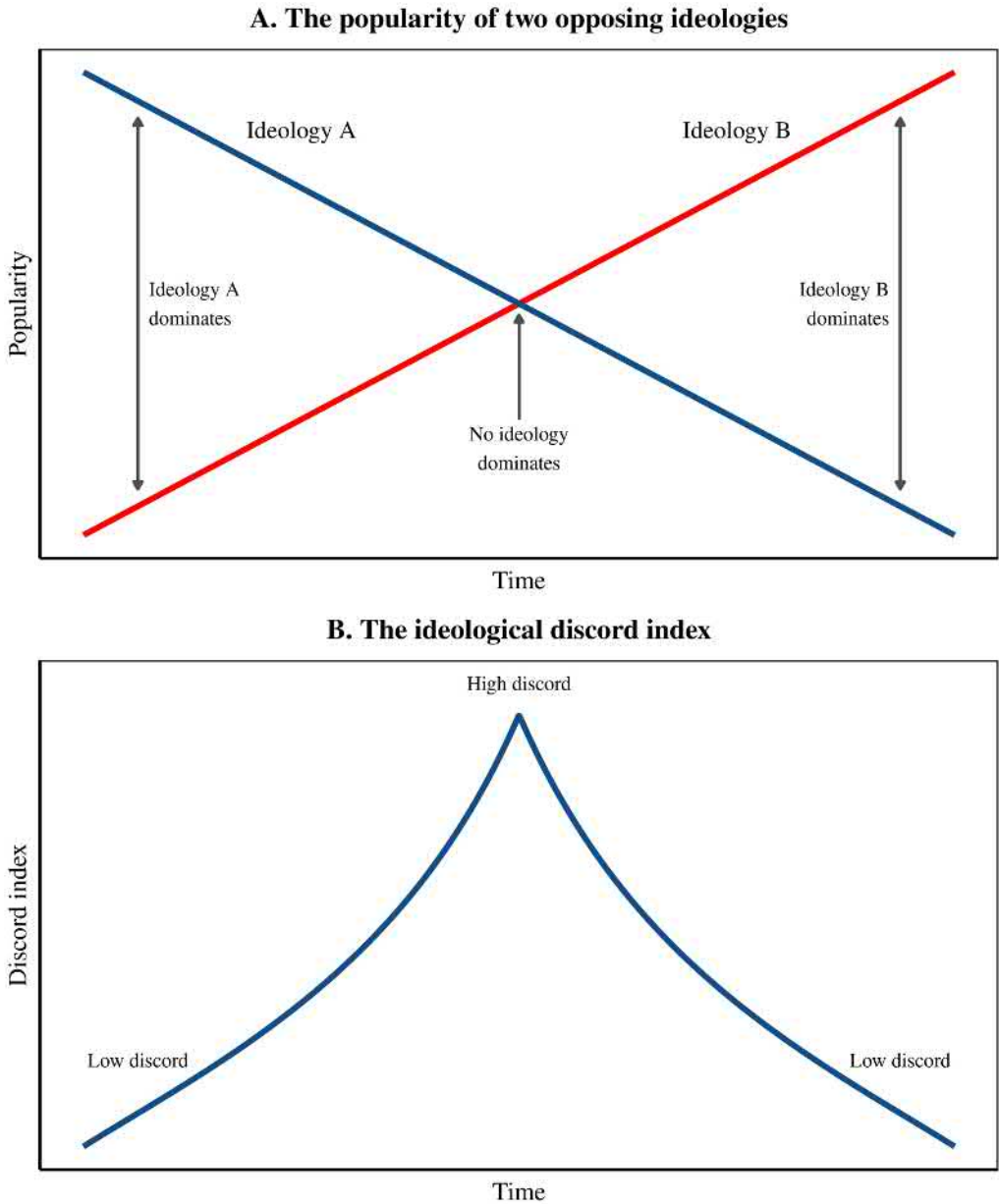


Figure 15: The ideological discord index

This figure illustrates the principle behind the ‘ideological discord index’, a measure which quantifies the degree of conflict between two opposing ideologies. Panel A shows a hypothetical ideological transition during which ideology A falls out of favor and is replaced by ideology B. The red and blue lines illustrate the popularity of each ideology. Panel B shows the corresponding ideological discord index. In visual terms, the discord index is inversely proportional to the popularity difference (in panel A) between the two ideologies. Thus, when ideology A dominates (left side) or when ideology B dominates (right side), the discord index is low. However, when no ideology dominates (center), ideological discord is high.

In more technical terms, I define the ideological discord index as follows. Let p_a and p_b measure the popularity of two ideologies, a and b . The ideological discord index is:

$$\text{discord index} = \frac{100}{|\ln(p_a) - \ln(p_b)| + 1}$$

The logic here is that when one ideology becomes dominant ($p_a \gg p_b$ or $p_a \ll p_b$), the discord index approaches 0. In contrast, when the two ideologies have equal popularity ($p_a = p_b$), the discord index is 100.

In Figure 16A, I apply the discord index to the struggle between biblical and economics ideology. I measure the popularity of each ideology using the frequency of its jargon in the Google English corpus (the data shown in Figure 10). Over the last two centuries, it seems there have been two cycles of discord. The first cycle peaked at the turn of the 20th century — the moment when economics ideology first surpassed the popularity of biblical ideology. The second discord cycle began around 1980 — corresponding to the peak of economics popularity. The current rise of discord is due to the declining popularity of economics ideology and the rising popularity of biblical ideology.

Assuming the discord index captures a real ideological struggle, the tension should be visible elsewhere — for example, among politicians. If you follow US politics, you know that today there is a bitter (almost pathological) divide between Democrats and Republicans. However, if you are older, you might remember a time when bipartisanship was the norm.

In *Ages of Discord*, Peter Turchin shows that over the last two centuries, the US federal government has gone through several cycles of polarization, and is currently in an era of extreme partisanship. Figure 16B shows these cycles. Here I've used data from voteview.com which reports, for every session of congress, the 'economic/redistributive' stance of federal politicians on a 0–1 scale. (The dataset tracks all members of the House, the Senate as well as the President.)

For each congress, c , I calculate polarization by taking the standard deviation of the set of ideological stances, I_c , of the constituent politicians:

$$\text{polarization index} = 200 \cdot SD(I_c)$$

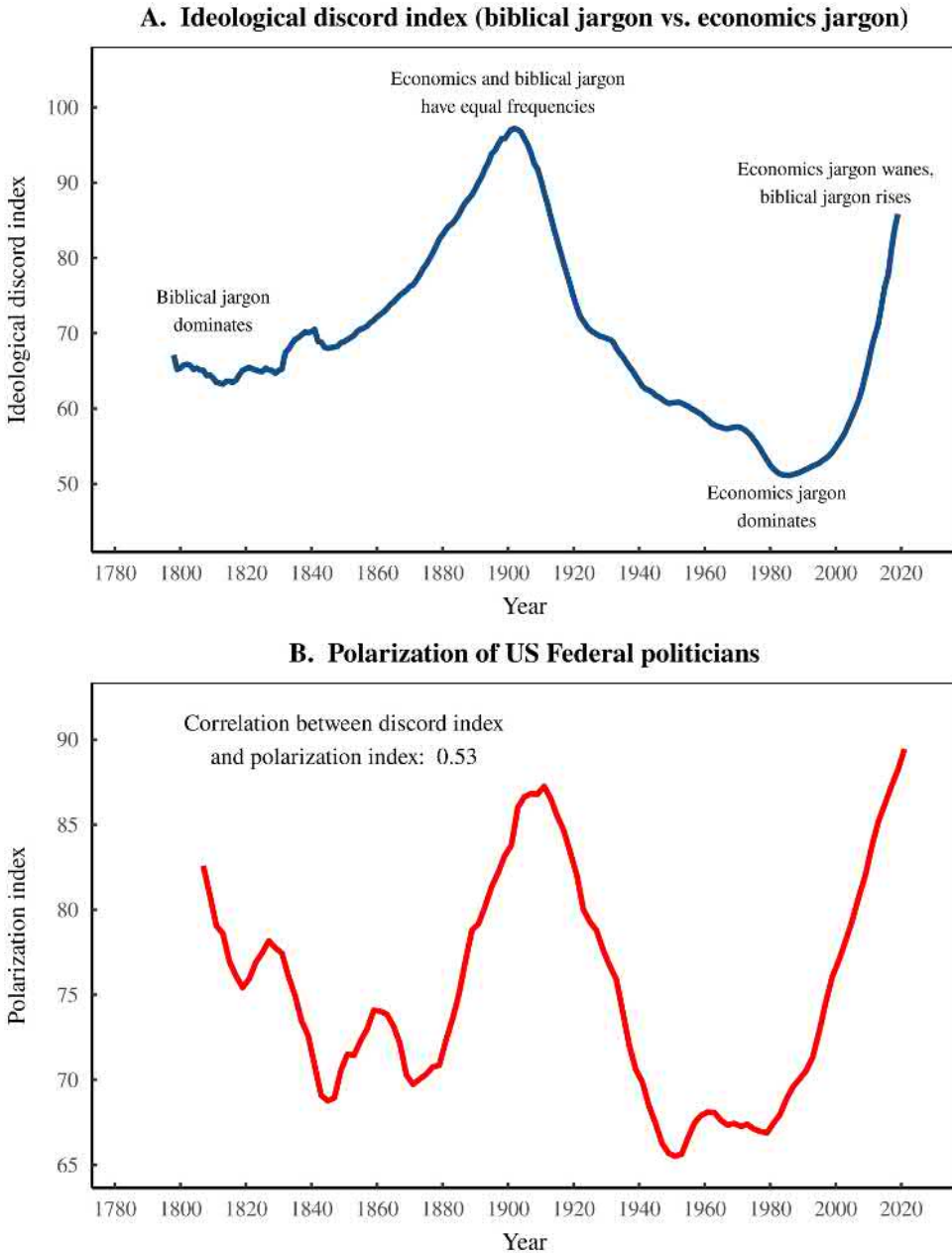


Figure 16: Ideological discord and the polarization of US Federal politicians

Panel A shows my ‘ideological discord index’, constructed from the frequency of biblical jargon and economics jargon in the Google English corpus. (See the data in Figure 10.) The discord index measures the lack of dominance between ideologies on a 100-point scale. Discord is at maximum (100) when biblical and economics jargon have the same frequency. Discord is at a minimum (0) when one type of jargon completely dominates the other. Panel B shows my measure of the polarization of US federal politicians. Constructed using data from Voteview.com, the index measures the polarization among all members of the House and Senate, plus the US President. A larger value indicates more extreme polarization. Note that both the discord index and the polarization index have been smoothed with a 10-year trailing average. For more details about these two metrics see the [Sources and Methods].

The polarization index varies from 0 (no polarization, meaning everyone holds the same ideology) to 100 (maximum polarization, meaning there are two equally sized groups that hold opposing ideologies.)

As Figure 16 demonstrates, the cycles of political polarization seem to match the oscillations of the ideological discord index. In other words, during periods of ideological conflict, US politicians become more polarized. Of course, that connection makes sense in abstract terms. After all, the US is supposed to be a democracy, so divisions amongst the population ought to induce polarization among politicians. However, it is not obvious (to me) that this political polarization relates to the specific conflict between biblical ideology and economics ideology. And yet the evidence suggests that it does.

For many Americans, the growing struggle between secular liberalism and reactionary christianity has been highlighted by recent Supreme Court decisions, especially the overturning of *Roe v. Wade*. Importantly, the linguistic evidence suggests that this secular-religious divide runs far deeper than the Supreme Court. On that front, journalist Amy Littlefield may not be exaggerating when she [describes](#) the rise of a ‘christian legal army’ that is fighting to install theocratic law.

Ideological futures

As a critic of neoclassical economics, I am happy that its doctrines are becoming less popular, and that in this sense, we have passed ‘peak capitalism’. But like many leftists, I had hoped that when capitalist ideology faded, it would be replaced by something better.

True, this ‘something better’ may still happen. Today, there are many vibrant movements seeking to build a world that is more equal and more sustainable. What remains unknown is the extent to which these alternatives are challenging the ideology of economics . . . or for that matter, the reactionary politics of the christian right.

On that front, Marx [called](#) religion the ‘opium of the people’, and he was right. Religion is a powerful ideological drug. But it is perhaps not the *most* powerful one. Give people a choice between living as a devout christian who is poor or as a secular liberal who is rich, my guess is that most people will chose the latter option. (This is one way to read 20th-century history.)

If you take away the opiate of wealth, however, people will quickly return to their religious drug of choice. It [happened](#) after the Soviet Union collapsed. And it seems to be happening within anglophone countries today. The growth engine has slowed and the gains that remain are going almost exclusively to the rich. If these trends continue (and the linguistic patterns that go with them), the future may look less like a post-capitalist utopia and more like [The Handmaid's Tale](#).

Support this blog

Economics from the Top Down is where I share my ideas for how to create a better economics. If you liked this post, consider becoming a patron. You'll help me continue my research, and continue to share it with readers like you.



Sources and methods

Data and code for the analysis are available at the Open Science Framework: <https://osf.io/ynp3t/>

English word list

To ensure that all samples of writing are restricted to the same base of words, I keep only the words that are within a common English 'dictionary'. To construct this dictionary, I combine the following word lists:

- The [Grady Augmented](#) word list from the R [lexicon package](#)
- The Project Gutenberg word list from [Moby Word II](#)
- An English word list from [Leah Alpert](#)

From the resulting list, I remove words that are uninteresting for ideological analysis:

- remove words with 2 or fewer characters
- remove prepositions (using R [lexicon](#) function `pos_preposition`)
- remove English numerals from 1 to 100 (i.e. one, two, three ...)

- remove acronyms (words containing ‘.’)
- remove contractions (words containing apostrophes)
- remove hyphenated words
- remove words containing numbers

To adjust for different spellings, I also replace all British spellings with American versions (i.e. change ‘labour’ to ‘labor’). The resulting dictionary contains about 500,000 words.

Samples of historical written English

I use three samples of historical written English:

1. [Google English corpus \(2019\)](#)
2. Titles of articles in the Sci-Hub database (dumped in August 2021, and downloaded from [Library Genesis](#))
3. Titles of books in the [Library Genesis](#) nonfiction database (dumped in June 2022)

After cleaning the data (keeping only the words that are in my English dictionary), I discard years in which the database has fewer than 10,000 words. Figure 17 shows the resulting sample size of each database over time. Google English is by far the largest sample, exceeding the others by several orders of magnitude.

The language similarity index

The idea behind the language similarity index is that it measures the average difference in word frequency between two samples of text. Letting f_a^i and f_b^i be the frequency of word i in two samples of text, a and b . A simple definition of the similarity index would be:

$$\text{similarity index} = \frac{100}{\text{mean}(|f_a^i - f_b^i| + 1)}$$

However, since word frequency varies over an enormous range, a better approach is to take the logarithm of frequency:

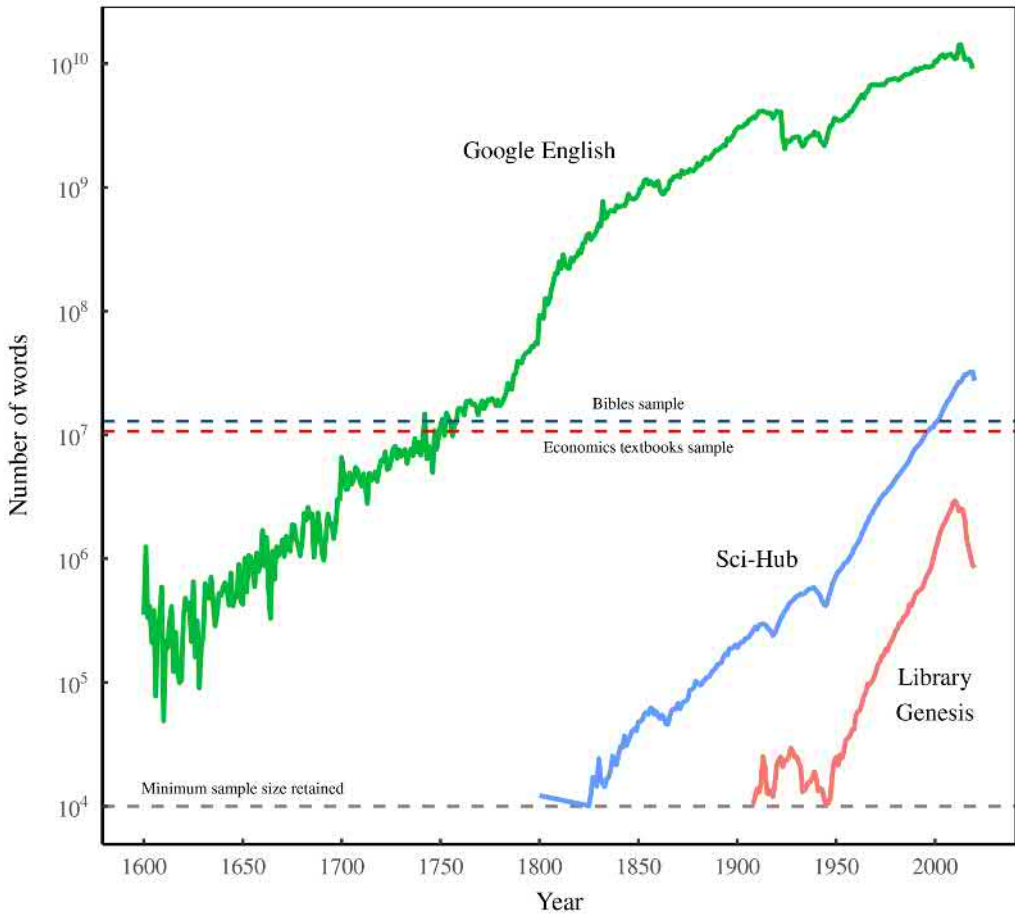


Figure 17: Number of words in the Google English, Sci-Hub and Library Genesis databases

Each line shows the number of words as a function of year in the given language database. Note the logarithmic scale on the vertical axis. I keep only the years in which each database includes more than 10,000 words. The blue and red dashed lines show the size of my sample of bibles and economics textbooks.

$$\text{similarity index} = \frac{100}{\text{mean}(|\ln(f_a^i) - \ln(f_b^i)| + 1)}$$

A problem with this definition, though, is that it requires an enormous sample to give accurate results. This is because as a rule, word frequency is distributed according to a power law, meaning the vast majority of words are quite rare. To capture the ‘true’ frequency of these rare words, we need an enormous sample. For example, if a word has a frequency of 1 per 10

million, we need to sample billions of words to recover this frequency reliably. The consequence is that in practice, the above definition of the similarity index is unworkable.

I fix the sample-size problem by taking the floor of the log differences, as follows:

$$\text{similarity index} = \frac{100}{\text{mean}(\lfloor |\ln(f_a^i) - \ln(f_b^i)| + 1 \rfloor)}$$

Taking the floor of the log frequency difference speeds up convergence of the similarity index to its correct value. Still, we need a fairly large sample for this metric to be accurate.

Figure 18 illustrates. Here I show what happens when we take a word sample from my economics textbook corpus and calculate the similarity between this sample and the full corpus. The resulting similarity index (vertical axis) depends on the size of the sample (horizontal axis). We recover the true similarity (100) as the sample size approaches the size of the economics corpus (about 10 million words). For smaller samples, the similarity index under-estimates the degree of language similarity.

Testing the consistency of the various measures of economics/biblical language

I have assembled four measures of the popularity of economics/biblical language:

1. Jargon frequency in the Google English corpus (Figure 10)
2. Language similarity index between bibles/economics textbooks and the Google English corpus (Figure 11)
3. Jargon frequency in Sci-Hub article titles (Figure 12)
4. Jargon frequency in Library Genesis book titles (Figure 13)

To verify the consistency of these different measurements, Figure 19 shows their ‘correlation matrix’. Each box shows the correlation between the corresponding time series labeled on vertical and horizontal axis. (Darker red indicates a stronger correlation.) Panel A correlates the various measure of

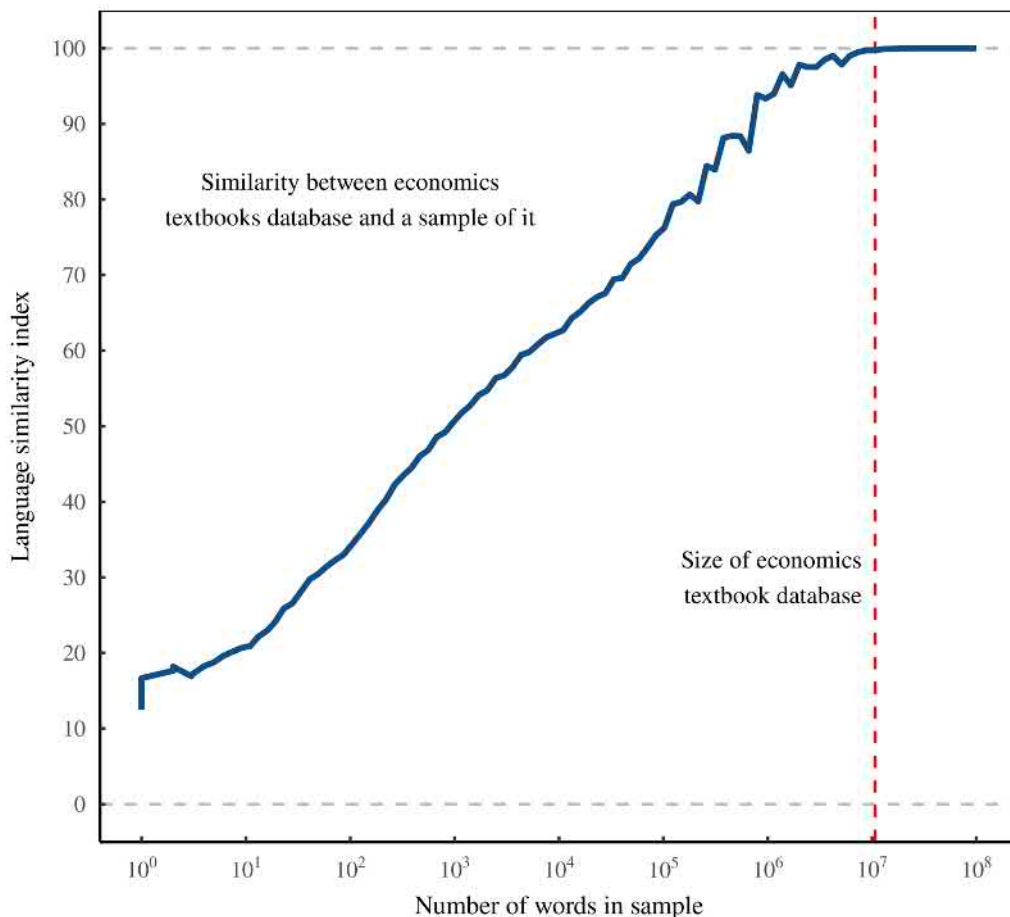


Figure 18: How the language similarity index is affected by sample size

This figure shows the result of an experiment where I sample words from the economics textbooks corpus and then calculate the similarity index between this sample and the full corpus. The horizontal axis shows the sample size. The vertical axis shows the corresponding similarity index. We reach the true similarity (100) only when the sample size is comparable to the database itself.

economics language. Panel B correlates the measure of biblical language. The correlations are generally high, indicating that the measurements are consistent with each other.

Figure 20 illustrates that for each of these measurements, the popularity of biblical language correlates negatively with the popularity of economics language.

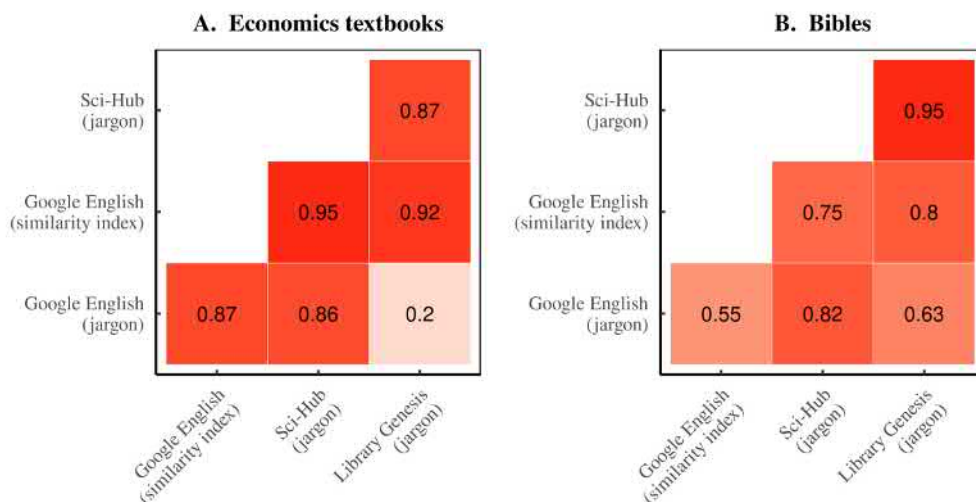


Figure 19: Cross correlating the four metrics of economics/biblical language

Each box in the ‘correlation matrix’ indicates the time-series correlation between the corresponding metrics on the horizontal and vertical axes. For example, in panel A, the top right box shows the time-series correlation between economics jargon in Library Genesis and in Sci-Hub. The top right box in panel B shows the equivalent correlation for biblical jargon. A deeper shade of red indicates greater correlation.

Polarization of US Federal Government

Data for the ideological stance of Federal politicians comes from voteview.com, series `nokken_poole_dim1`. It quantifies the ideological stance of each US president, senator and member of congress on an ‘economic/redistributive’ scale ranging from 0 (extremely conservative) to 1 (extremely liberal). For details about this data, see Adam Boche and colleagues’ discussion paper ‘[The new Voteview.com](https://voteview.com)’.

Further reading

Bichler, S., & Nitzan, J. (2021). The 1-2-3 toolbox of mainstream economics: Promising everything, delivering nothing. *Real-World Economics Review*, 98, 23–48.

Boche, A., Lewis, J. B., Rudkin, A., & Sonnet, L. (2018). The new vote-view.com: Preserving and continuing Keith Poole’s infrastructure for scholars, students and observers of congress. *Public Choice*, 176(1), 17–32.

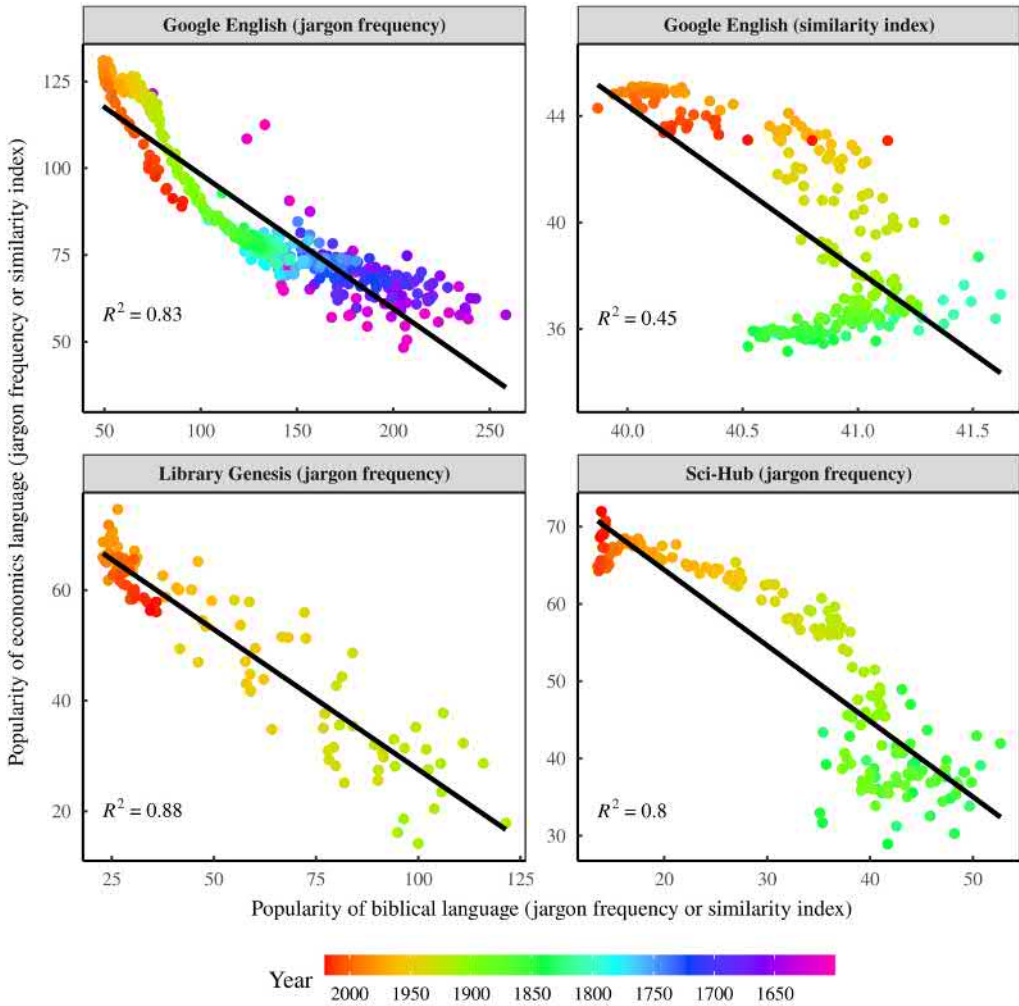


Figure 20: The popularity of economics language correlates negatively with the popularity of biblical language

Each panel shows a different measurement of word frequency in English, with the popularity of biblical language plotted on the horizontal axis and the popularity of economics language plotted on the vertical axis. Color indicates the year of the data. The correlations are strong and negative. When biblical language becomes less popular, economics language becomes more popular, and vice versa.

Ganshof, F. L. (1952). *Feudalism*. London: Longmans Green.

Keen, S. (2001). *Debunking economics: The naked emperor of the social sciences*. New York: Zed Books.

Mirowski, P. (1991). *More heat than light: Economics as social physics, physics as nature's economics*. Cambridge Univ Pr.

Nitzan, J., & Bichler, S. (2009). *Capital as power: A study of order and creorder*. New York: Routledge.

Turchin, P. (2016). *Ages of discord: A structural demographic analysis of American history*. Chaplin, CT: Beresta Books.