# THE UNPOPULAR FUNCTION
# TESTING THE ECONOMIZATION HYPOTHESIS AGAINST A GOOGLE NGRAM VIEW OF TRENDS IN FUNCTIONAL DIFFERENTIATION (1800-2000)

## Steffen Roth

**Abstract**

The present article considers the economization of society a hypothesis rather than a fact. The hypothesis is tested against the results of a *Google ngram viewer* analysis of the most frequent function system references in the Google Books corpus for the year 1800 through 2000. Despite the remarkable growth figures the economic word frequency shares feature in the English, French, and German language corpora, the results suggest rejecting the economization hypothesis. In fact, the growth trends of economic word frequencies are stopped in all of the three language areas, in none of which the economy ever reached a dominant position throughout the entire 200 years. The results give reason to assume that the idea of an economized society is an intellectual artifact rather than a fact. This fact is emphasized not to prove the marginal relevance of research in economic risks and benefits, but rather in terms of a suggestion to consider re-focusing research foci and drawing increased attention to function systems beyond the politico-economic double stars of social science. Maybe even the solution to the present "economic" "crises" is not in more, but rather in less attention to the economy.

**Keywords:** functional differentiation; function systems; social systems; Google ngram; culturomics; economization.
**JEL Code:** Z10, Z13.

## Introduction

The distinction between autonomous function systems is considered a core concept of modern societies. Functional differentiation makes the difference between truth and money, hospitals and

banks, and between a political and a clerical economy. In fact, the most important definitions and criticisms of modern society are based on functional differentiation, since all discussion on capitalist or economized societies implicitly refer to an underlying distinction and ranking of function systems. This idea, however, is not understood without ambiguity. On the one hand, function systems are considered fundamentally equivalent. Accordingly, there is no way of arguing that the political system is essentially more important than the economy, religion or sport. On the other hand, there seems to be plenty of empirical evidence for such rankings in modern societies. In the following, we assume that it is not despite, but because of their basic equivalence that function systems can be ranked at all.[1]

Looking at prominent rankings, we find that there is little consensus on the question of whether capitalism results either from the primacy of a particular form of politics or from the primacy of the economy. Still, most people would agree on the idea that present societies are subject to economization. This "increasing influence of economic factors and values on the political agenda and other areas of society" (Blumler & Kavanagh, 1999: 210) includes the economization of health, art, science, religion, and, as a matter of course, society as a whole (Habermas & McCarthy, 1985; Polanyi, 1957). As a result of this "economic turn" (Smart, 2003) or fethishization of the economy (Foucault, 2008), the "economization of every sphere of existence" (Kane, 2010: 81) emerges so omnipresent and dominant that even the proliferation of economics is taken for an indicator of economization (Çalışkan & Callon, 2010) or a need of de-economization (Latour, 2004), respectively. The problem with the corresponding marginalization of the "other areas of society" (Blumler & Kavanagh, 1999: 210) is not only in the fact that "(o)ther social institutions are seen (once again) as mere puppets in the hands of powerful economic trends and actors" (Stehr, 2002: 4) or as subjects to the respective political counter-performances, but also in the empirical basis of these statements. So far, the existence of an economization of society is supposed rather than studied. In the following, however, the economization of society will be considered a hypothesis that still needs to be defended.

Subsequent to a theory statement, the present article test the hypothesis of an economization of society against the results of a Google Ngram Viewer analysis of the most frequent function

---

[1] If they were essentially unequal, they would already be ranked and, therefore, could no longer be ranked. Function systems can therefore be treated as nominal data that feature a skewed distribution whenever it comes to the analysis of concrete segments of society.

system references in the Google Books corpus for the year 1800 through 2000. The results give reason to assume that the idea of an economized society is an intellectual artifact rather than a fact.

## Theory statement

Differentiation refers to an intrasystem process of subsystem formation (Luhmann, 1977, 1997). The first known forms of subsystems within society were families and tribes. Families coexisted with other families of the same tribe in the same manner as tribes coexisted with other tribes. As a result, early societies are said to be differentiated into identical and co-equal segments of society. However, in the course of the Neolithic revolution, the early forms of differentiation were soon complemented and superposed by stratification, which is commonly associated with the formation of hierarchical social orders, such as the Indian cast system or the Occidental Estates of the realm. However, "(m)odern society is no longer characterized by a stratification of lineage, clans and families, but by a differentiation of function systems" (Vanderstraeten, 2005: 476). Despite their importance to modern society, there is still little consensus on necessary and sufficient conditions for the definition of function systems. Looking at existing working definitions and non-exhaustive lists of function systems, however, a hardcore of 10 function systems can be extracted: Policy, the economy, science, art, religion, law, health, sport, education, and the mass media (Roth, 2013).

Testing the economization hypothesis, in the following, we will focus on the importance of the economy as compared to the other nine function systems.

## Hypotheses

The basic assumption of the functional equivalent and mutually exclusive nature of function systems represents an excellent groundwork for the null hypothesis, which the present article proposes to test. Representing coequal nominal data, function systems can be assumed equally relevant to social systems. The null hypothesis is, therefore, as follows: *(H0) Function systems relevancies exhibit a uniform distribution in social system.* On the other hand, empirical research seems to indicate an unequal distribution of function system relevancies. The alternative hypothesis reads as follows: *(H1) Function systems relevancies exhibit unequal distributions in*

*general and a to the economy in particular in the Anglophone (H1.1), Francophone (H1.2) and Germanophone world (H1.3).* The hypothesis will be tested against both weak and strong definitions of economization as a process and a state description. Economization will, consequently, be analyzed with regard to their incidence in one or several cultural areas as well as to their individual course and their relative importance.

## Method

The key assumptions proposed in this article are the temporal change of and the intercultural differences in the importance of particular function systems to particular societies. The major problem becomes how to measure this importance in a context that can be expected to allow for somewhat representative information on societies as a whole.

The importance of concepts is often defined in terms of the frequency of their occurrence in given corpora, which is considered "the simplest and most impartial gauge of word importance" (Kloumann et al., 2012:1) or the popularity of objects, ideas, and persons (Ophir, 2010; Bohannon, 2011)[2]. Given the scope of the present research question and the corresponding scale of the research field, research utilizing the Internet as the largest possible corpus would make sense. Despite the fact that web search engines are said to return word frequency estimates that are highly consistent with established methods (Blair et al., 2002), the problem with Internet word frequency analyses is that the analysis of this most recent media hardly allows for long-term trend analyses. Hence, the present discussion of popular trends in functional differentiation analyses changes the frequency distributions of function system references in the second largest available text data collection.

Since its start in 2004, the Google Books project has digitalized some 15 million of the estimated 115 million books ever published. In 2007, a Harvard research team (Michel et al., 2011) recognized the research potential of the Google Books corpus, performed considerable quality checks, and finally reported the compilation of a representative corpus of more than five million books or 500 billion words covering seven language areas and a time span of 600 years.

---

[2] The present interpretation of word importance does not refer to where word importance is inversely related to word frequency. Such an indexical approach to word importance would only make sense if the present article was interested in comparing the discriminatory abilities of the concepts involved. All function systems, however, are on the same level of analysis and, therefore, feature the same degree of discriminatory power.

The development of this enormous data soon raised hopes of a golden age of digital humanities, which would open up new types of historical knowledge (Ophir, 2010), as it has already given birth to the discipline of culturomics as "the application of high-throughput data collection and analysis to the study of human culture" (Michel et al., 2011: 181). The access to the Google Books corpus is facilitated by the Google Ngram Viewer - an open-access interface that allows for trending in terms of the production of customized time-series plots for entered search terms.

Testing the economization hypotheses, the present article will focus on the analysis of trends in the English, French, and German Google Books 1-gram corpora 1800-2000.[3] Due to the sheer size of the databases concerned, further in-depth analyses will be conducted using the Google Million, an English-language corpus of only one million words designed to fulfill the same representativeness requirements as the larger English version(s), however, with improved handling characteristics.

In this setting, the null hypothesis will be rejected in favor of the alternative hypothesis (H1) when the time-series plots produced by the Google Ngram Viewer feature not only unequal distributions in general, but also a bias to the economy in particular in any or all of the analyzed language areas (H1.1-1.3). In concrete terms, the hypotheses will be weakly confirmed if the output of the Google Ngram Viewer query supports the respective trend assumptions in at least two out of three language areas. In concrete terms, the plots should display an increase in the importance the economy. A strong confirmation would, moreover, call for a dominant position in the texture of function system references for all language areas.
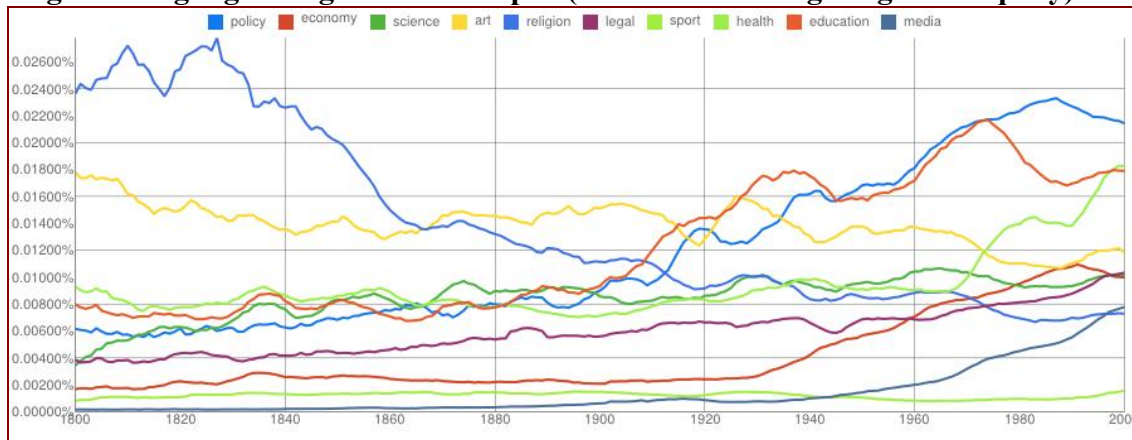
## Results

The ten function systems exhibit unequal occurrence frequencies that vary within and across all the language areas:

In the English case (cf. Fig. 1), at the beginning of the sample period, law is the most dominant function system followed by religion and art. At the end of the period, policy gained on law, which ranks second and is followed by health and education. Currently, religion ranks ninth,

---

[3] The data from 1500-1800 is likely to feature biases due to insufficient sample sizes: "The oldest works were published in the 1500s. The early decades are represented by only a few books per year, comprising several hundred thousand words. By 1800, the corpus grows to 98 million words per year; by 1900, 1.8 billion; and by 2000, 11 billion" (Michel et al., 2011: 176).

outperformed even by the formerly marginal economy and mass media system. A closer examination of the function systems calls attention to the word frequency shares of the economy as compared to the other function systems. Rather than an economization, the decline of religion ($\hat{f} = 0.25$)[4] and the rise of policy ($\hat{f} = 3.7$) are the most striking trends. Religion is outperformed by policy during World War I and by science in the early 1930s, with the latter being a result of slow, but steady growth.

**Figure 1: Word Frequency Shares of the Names of the Function Systems in the Overall English Language Google Books Corpus (source: own Google Ngram enquiry)**



At about the same point in time, the formerly marginal economy takes off ($\hat{f} = 5.0$). In 2000, the economy and science share about the same numbers of mentions in the Google Books corpus (100ppm)[5] and are, therefore, both lower-middle ranking function systems: Law and policy (240-220ppm), health and education (180ppm), and art (120ppm) relegate the first five ranks. On the verge of the millennium, religion was finally outpaced by the mass media. Together with religion, mass media (80ppm) shares the penultimate ranks. Sport is last (>10ppm).
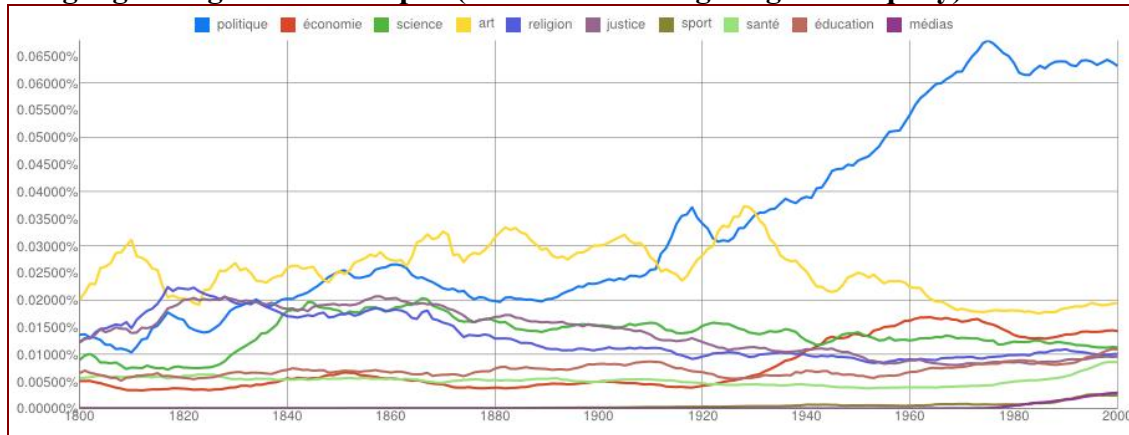
In the French case (cf. Fig. 2), art initially ranks first in front of religion, justice, and policy. Between World War I and World War II, it was overtaken by the political system, whose increase ($\hat{f} = 4.9$) to a remarkable frequency of 680ppm experienced a smaller counter trend

---

[4] Fold change of the relative word frequency calculated in terms of the ratio of the visually assessed word frequency of 2000 and 1800: $\hat{f} = e(2000)/e(1800)$.

[5] The abbreviation ppm represents "parts per million", i.e. the relative word frequency per million words in the Google Book corpus in 2000. The unit *per million* is used to avoid longer chains of digits after the decimal point. The most common English words account for 45,000ppm (4.5%: *the*), 30,000ppm (3%: *of*) and 23,000ppm (2.3%: *and*). *Time*, the most frequent noun in the Google Million corpus, has an appearance of 1,300ppm. *Law*, for example, appears about four times less than *time* and about two times less than *life*, which is still about as important as *men*, *English* or *family*.

around 1980. Uncontested since the last confrontation with law, which occurred between 1945-1960, policy is ranked first. In 2000, policy occurred twice as often as law and three times more often than art.
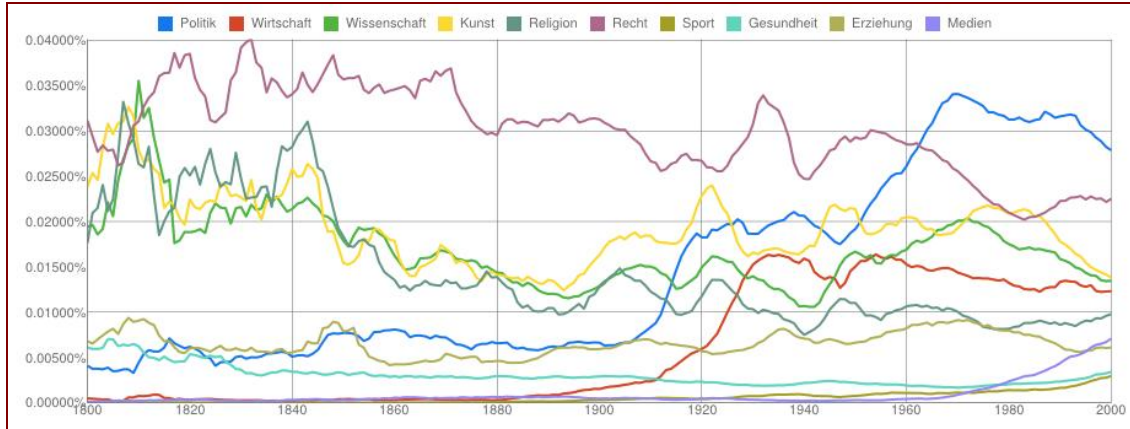
**Figure 2: Word Frequency Shares of the Names of the Function Systems in the French Language Google Books Corpus (source: own Google Ngram enquiry)**



As of the late 1920s, after an increase of economic communication ($\hat{r}$ = 3.0; 150ppm), the economy ranks third, closely followed by science, education, justice, religion, and health. Mass media and sport come last.

A comparison of the English and the French data uncovers the higher relevance the French function systems have in their corpus. The word politique, however, must be considered an exception to this rule because it is both a noun and an adjective. Complemented by the adjective(s), the English and the German political performance would be about the same as the French. Still, it is notable that politique is more important to its corpus than words like man, life or work are to the English, given that the first refers to both gender and genus and the latter is both simultaneously a noun and a verb.

**Figure 3: Word Frequency Shares of the Names of the Function Systems in the German Language Google Books Corpus (source: own Google Ngram enquiry)**

Additionally, despite a considerable decline (f̂ = 0.67 as of 1800, 0.53 as of 1930), second ranked art (200ppm) is still more important to the French corpus than second and third ranked health and education are to the English.

The German data (cf. Fig. 3) displays the largest variances and the broadest scope of changes within a language area. In the early 19th century, Germany seems 'governed' by law as well as science, art, and religion. The latter triad appears to share a common destiny in terms of their collective decline up until the 1960s, when science and art finally separated from the downtrend of religion, which appears to have stopped since the 1980s. After a most notable increase (f̂ = 7.0), since the early 1960s, Politik (280ppm) has ranked first, followed by the legal system (220ppm) as well as art and science (130-140ppm). The economy (120ppm) is ranked fifth due to an increase in importance dating back to the late 19th century rather than to the late 1920s, like in the case of both the English and the French.

## Discussion

The first finding of the presented analyses is that the Google Ngram Views of functional differentiation actually react to history, which is true with regard to both the interregional and the regional level. On the one hand, all language areas display a decline of religion as well as clear evidence of the impact of Black Friday on the takeoff of policy and the economy. On the other hand, the results also reflect particularities related to the histories of the individual languages areas: The German time-series plot (cf. Fig. 3) displays a veritable program change for an entire language area from a realm of priests, poets, and philosophers to a national economy. This politico-economic campaign started soon after the death of Karl Marx, which is considerably

earlier then in the other two cases. In the English and French language area, it was not until Black Friday that the economy became at least moderately relevant.

Despite the remarkable growth figures the economic word frequency shares feature in all of the three corpora, the results suggest rejecting the economization hypothesis in both the strong and the weak variant. In fact, in two cases, the period of the potential trend was rather short (E: 1930-1990; F: 1925-1975). The trends are stopped in all of the three language areas, in none of which the economy ever reached a dominant position throughout the entire 200 years.

## Limitations and Future research questions

One of the most serious limitations of the Google Ngram Viewer is that it allows for case-sensitive queries in ngram only. Ngrams can, consequently, not be bundled into one single graph, just as the queries cannot be for lemmas.[6] Further critique may come into play with the fact that studying the key media of the Gutenberg Galaxy might not be the key to the emerging Internet society. However, there is evidence that online and offline contents do not differ all too much. At the same time, the advantages of a books corpus are evident: Book content can be considered subject to stricter selection and, therefore, a better indicator for importance. Moreover, the Internet corpus does allow for research within the sample period of several hundred years. Further research limits are, as a matter of course, related to restricted language competences and the corpus data size. In fact, it would be best to compare all language areas and analyze 2- or more-grams, e.g. with regard to word co-occurrences, in order to get much deeper context-information. Future research on "Big Data" in general and the present topic in particular could, therefore, call upon international cooperation and access to more powerful computer resources.

Future expeditions in the corpus might then not only open up further and more snapshots of trends in functional differentiation, but also allow for an answer to the question of whether or not functional differentiation is indeed a master trend in present societies. In this sense, an extended Google Ngram View of modern societies could inform on the actual state and trend of modernity itself.

---

[6] A search engine for the Google Books corpus designed by Marc Davies allows for these and further options, however, for the American English corpus only.

## Conclusion

Our analysis of the English, French, and German language sets of the largest available text corpus does not corroborate the perhaps most prominent state description and trend prediction related to functional differentiation. Starting as a trend in line with the rising popularity of the early socialist movement, the "economization of society" never resulted in high weights of the economy and stopped rather early (German: 1950ies, French: 1970ies, English: 1990ies).

If it is true that economists "perform" economies, then the presented results also support the idea that economization critics perform economization and that both parties jointly perform an economy bias in social sciences that is incongruent with the average to marginal relevance of the research object. In this respect, it is congruent that the different dictions and notations of economization or commodification are as marginal entries in the English language corpus as misspellings or exotic forenames are. This fact is emphasized not to prove the marginal relevance of research in economic risks and benefits, but rather in terms of a suggestion to consider re-focusing research foci and drawing increased attention to function systems beyond the politico-economic double stars of social science. Maybe even the solution to the present "economic" "crises" is not in more, but rather in less attention to the economy

## References

Blair, I., Urland, G., & Ma, J. (2002). Using Internet search engines to estimate word frequency. *Behavior Research Methods* 34(2), 286-290.

Blumler, J. G., & Kavanagh, D. (1999). The Third Age of Political Communication: Influences and Features. *Political Communication* 16(3), 209-230.

Bohannon, J. (2011). The Science Hall of Fame. *Science* 331(6014), 143.

Habermas, J., & McCarthy, T. (1985). *The Theory of Communicative Action, Vol. 2: Lifeworld and System: A Critique of Functionalist Reason.* Beacon Press: Boston.

Kane, C. L. (2010). Programming the Beautiful, *Theory, Culture & Society* 27(1), 73-93.

Kloumann, I. M., Danforth, C. M, & Harris, K. D., et al. (2012). Positivity of the English language. *PloS one* 7(1), e29484.

Latour, B. (2004). *Politics of Nature: How to Bring the Sciences Into Democracy.* Cambridge: Harvard University Press.

Luhmann, N. (1977). Differentiation of Society. *The Canadian Journal of Sociology / Cahiers canadiens de sociologie* 2(1), 29-53.

Luhmann, N. (1997). Globalization or World society: How to conceive of modern society? *International Review of Sociology* 7(1), 67-79.

Michel, J.-B., Shen, Y. K., & Aiden, A. P., et al. 2011. Quantitative Analysis of Culture Using Millions of Digitized Books. *Science* 331(6014), 176-182.

Ophir, S. (2010). A New Type of Historical Knowledge. *The Information Society* 26(2), 144-150.

Polanyi, K. (1957). *The great transformation.* Beacon Press: New York.

Roth, S. (2013). Die Zehn Systeme. Ein Beitrag Zur Kanonisierung Der Funktionssysteme. *SSRN eLibrary* (http://ssrn.com/abstract=2148935).

Smart, B. (2003). An Economic Turn. *Journal of Classical Sociology* 3(1), 47-66.

Stehr, N. (2002). *Knowledge and Economic Conduct: The Social Foundations of the Modern Economy.* University of Toronto Press: Toronto.

Vanderstraeten, R. (2005). System and Environment: Notes on the Autopoiesis of Modern Society. *Systems Research and Behavioral Science* 22(6), 471-481.

**Contact**

Steffen Roth

ESC Rennes School of Business

2 Rue Robert d'Abrissel

35000 Rennes

steffen.roth@esc-rennes.fr