

SPEAKING IN TONGUES, A TEXT ANALYSIS OF ECONOMIC OPINION AT NEWSWEEK, 1975-2007

TIAGO MATA
CLAIRE LEMERCIER

HOPE Center Working Paper No. 2011-02

January 2011



Duke University
Center for the History of Political Economy

Speaking in tongues, a text analysis of economic opinion at *Newsweek*, 1975-2007

Tiago Mata

t.mata@duke.edu

Center for the History of Political Economy
Duke University, USA

Claire Lemerrier

claire.lemerrier@sciences-po.org

CNRS researcher, Center for the Sociology of Organizations
Sciences-Po, France

Acknowledgments:

Tiago Mata thanks financial support from a 2006-08 postdoctoral fellowship grant from the Fundação para a Ciência e Tecnologia, Portugal.

Abstract of

Speaking in tongues, a text analysis of economic opinion at *Newsweek*, 1975-2007

by

Tiago Mata

t.mata@duke.edu

Center for the History of Political Economy

Duke University, USA

Claire Lemerrier

claire.lemerrier@sciences-po.org

CNRS researcher, Center for the Sociology of Organizations

Sciences-Po, France

Among American news magazines *Newsweek* holds the distinction of having hosted some of the most authoritative interpretation of economic events. Its cast of columnists included two of the most acclaimed academic economists and some of the most widely read business journalists of the late twentieth century. The purpose of this paper is to examine their writings as examples of “economic journalism”. We used statistical methods to study word co-occurrences and identify vocabulary classes for bodies of texts for 1975-1990 and 1991-2007. We are thus able to distinguish two domains of discourse: political economy and the causal economy. We find traces of changes in economic discourse resulting from the end of the Cold War as well as point out differences in the language of academics and journalists.

JEL Codes: A11, B20, N01

Keywords: economic opinion, policy, economic journalism, text analysis, economic ontology

Wordcount: 9490

Version: 12/16/10

1. The economy of words

The news magazine is an invention of the 1910s and 1920s. Novel technologies for photographic reproduction, halftone and telephotography, allowed for cheaper and timely printing of visuals. In taking up the magazines' pages, visuals challenged the text. Although a feature of journalism globally, magazine writing required authors to make difficult choices to achieve concision, intelligibility and liveliness. The formula that developed was successful. Magazines captured into the 1950s the imagination of readers, and set trends in popular and consumer culture. At the middle of the twentieth century, they were the most profitable items in the portfolios of publishing companies in the Western World.

Politics and entertainment are the best known foci of the news magazines, yet alongside these was also coverage of economic and social issues. Writers were asked to examine America's economic life: to describe its state, to sketch its structure, to reduce uncertainty about its future and to line up policy options for public scrutiny. And to do it in half a page, in 800 words or less. The subject of this essay is economic opinion writing at one news magazine. We want to observe how the language changed in the course of three decades and how authors with different backgrounds and intellectual commitments approached the task.¹

¹ We aim to study the phenomena of economic journalism and opinion in its own

For scholars intent on characterizing economic culture and discourse, magazine writing poses significant challenges. There is too much of it, to the extent that economic opinion extends over decades, if not centuries.² But there is also too little, since each piece carries insufficient content to sustain an extended analysis. It is therefore impossible to analyze this discourse with the “close reading” or more or less contextualized exegesis generally used in intellectual history. It is equally difficult to perform semi-automated content analysis that require the analyst to code parts of the corpus according to themes or categories that he or she has previously defined. At least as a first approach, and especially if one wants to capture language patterns that can be as discreet as they are common, it is appropriate to analyze a large corpus through statistical methods.³

The most common strategy to probe economic newsprint has been to query words and their valence. One study of this kind counted occurrences of the word “competitive”

right, as economic culture. The extent to which this script influenced the views of its readers is not of our immediate concern. As it will become clear throughout the paper we believe economic journalism does significant ontological work that may not be exactly reproduced in popular opinion, but that is plausibly a resource in its formation.

2 We take for granted that the learned magazines/reviews of the 19th century are a different species from the news magazines of the 20th. On the former see Fetter, F.W. (1965) “Economic Controversy in the British Reviews, 1802-1850.” *Economica*, 32(128): 424-437.

3 On the use of tools that both compensate for the cognitive difficulties in examining very large corpuses and allow for the identification of patterns that would be neglected as they make up a subtle “background noise” of discourse, see e.g. Mayaffre, Damon. (2001) “History and information technology : the French are way behind”, *Lexicometrica*, 3. Accessed from: <http://www.cavi.univ-paris3.fr/lexicometrica/article/numero3/dm2001.htm>, December 2010. Contrary to this author's title such techniques are often of French origin.

over 10 year periods at the *New York Times*. Constructing ratios of counts of the word with a variety of related words such as “monopoly” or “cooperative”, the author sought to disambiguate and track its changing meaning.⁴ Yet another approach has been to use (human) coding to identify the context in which words occur. In a study of the media construction of the phenomenon of “downsizing”, over 4000 articles, where the word appeared in the first 300 words, were coded by two scholars as either announcements, company-focused, economy or worker's perspective.⁵ Our approach bears some kinship with this work, to the extent that we examine collocation of words. However, we deploy a statistical algorithm which allows us to look beyond single words or pairs and to look across the whole corpus. Furthermore, we do not rely on human coding for the statistical analysis: the vocabulary patterns that we will discuss are not theory motivated or expert attributed but are elicited “bottom up” from the patterns in the text (more details in section 3). Despite our emphasis on the automatism of our methods, we do not seek to do away with the important judgment and interpretative work of the historian or social scientist. The results of the statistical analysis pose significant challenges of interpretation and contextualization.

Our efforts go to identify distinctive kinds in economic discourse. We look for

4 George, David. (2008) “On being “competitive”: the evolution of a word.” in *Real-World Economics Review*. n.48.

5 Hollister, Matissa. (undated) “Speaking of Downsizing. The use of the term “downsizing” in American news media 1975-2007” working paper. Accessed from: <http://www.dartmouth.edu/~socy/pdfs/HollisterSpeakingDownsizing.pdf> , November 2010.

distinctions *within* economic opinion to describe salient topics. We want to both observe change in these topics and their configuration across time (section 4) and differences in the vocabulary and style of authors (section 5). The dimensions of time and author, unlocks a matrix to discuss the history of economic opinion and journalism. We are able to discuss the extent to which the backgrounds, professional and generational of authors is expressed in their editorial discretion, and we can also examine how world events, economic and political history, have set the agenda and the language of economic magazine writing.

2. *Newsweek*

News-week saw its first issue in February 17, 1933. It was founded by Thomas J. Cardell Martyn, a reporter at both *Time* magazine and, from 1925, at the *New York Times*. The new publication's prospectus addressed the former employer, stating that "*Time* is too inaccurate, too superficial, too flippant and imitative" and promised a magazine "written in simple, unaffected English [in] a more significant format [with] a fundamentally sober attitude on all matters involving taste and ethics." Born in the worst year of the Great Depression, it was a respectable 32-page digest of weekly news.⁶ As its own self-promoting advert explained, it was a publication for the "busy

⁶ The original sections were: Front Page, Home News, Foreign News, Headlines, Sport, Business Science, Aviation, Entertainment, Books, Law, Religion, Forth Estate, Education.

individual caught up in the turmoil of a bewildered civilization.”⁷ In its original format articles were numerous, short, lively and unsigned.

In 1937 *News-week* underwent its first major change. Martyn's starting fund had been depleted and he sold his stake as the magazine merged with *Today*, another failing publication of the 1930s. Samuel T. Williamson was replaced as editor-in-chief by Malcolm Muir, a former McGraw-Hill executive. The hyphen was dropped from the magazine's title in the October 4, 1937 issue, and new features were added.⁸ Muir held a long tenure at the magazine from 1937 to 1961. *Newsweek* benefitted from the general rise in distribution of all magazines, its circulation rose between 1950 to 1962 in 80%, to a total of 1,5 million copies. In the view of several commentators, the magazine succeeded only as a follower of others. Its sectional organization paralleled that of *Time* and it offered little more coverage and analysis than that found in the *New York Times*.⁹

One of the features added in 1937 was a business opinion column, titled “Business Tides”. Its first author was Raph Robey, an economist with a doctorate from Columbia University, who had a career between banking, teaching and the publishing industry. He had been financial editor at the *New York Post* and editor at the *Washington Post*

7 Watkins, Julian L. (1933) “A letter to News Minded People” *News-Week*, March 18, p. 33.

8 Sumner, David E. (2003) “A History of Time, Newsweek, and U.S. News & World Report” in *Encyclopedia of International Media and Communications*.

9 Elliott, O. (1980) *The World of Oz* 1st ed., Viking Adult. Tebbel, J. & Zuckerman, M.E. (1991) *The Magazine in America, 1741-1990* 1st ed., Oxford University Press, USA.

before joining *Newsweek* as associate editor in 1937. He continued in the magazine as columnist between 1938 and 1946. Robey finally left the magazine to become chief economist at the National Association of Manufacturers.¹⁰ The replacement was Henry Hazlitt, a former staffer at the *Wall Street Journal*, and writer at the *New York Evening Mail* and the *New York Sun*. As essayist and for 12 years editorial page writer at the *New York Times*, Hazlitt developed a voice for commenting on the economy, often as a contrarian speaking for anti-statism in an age of growth in federal government.

In 1961 *Newsweek* was purchased by the Washington Post group. Osborn Elliott was placed as editor-in-chief and introduced major changes to the magazine in the following years. He set out to compete with *Time*, and even if the publication never reached its competitor's readership share, it became the most profitable of the two publications. *Newsweek's* success rested on its ability to stay current, tracking the transformations in American culture. It had access to the JF Kennedy and Lyndon Johnson administrations, and offered sympathetic and extensive coverage of the anti-war movement and the new wave of feminism. It also changed its economic opinion.

In 1963-5, in the wake of the Kennedy-Johnson's tax cuts and the rapid economic expansion that followed, economists made the covers of the magazines. Walter Heller, chairman of the Council of Economic Advisers, was on the cover of *Time* in issues in

10 "Dr. Ralph Robey Economist, Dies." *The New York Times*, July 7, 1971, p. 35. For the summer of 1946, a few columns were penned by John Willis Love.

March 1961 and June 1962 and the economy headlined other issues like that of March 31, 1963, on "The Rising US Economy." *Newsweek* responded by adding from 11 January 1965, Henry Wallich as an alternate to 71 year old Hazlitt. Wallich had credentials that Hazlitt, an opponent to the "new economics" of the White House, could not match.¹¹ He had been a member of the Council of Economic Advisers during the Eisenhower Presidency and was a Professor at Yale University. In September, Hazlitt left the publication. His parting words were gracious even if preoccupied in responding to the criticisms of "liberals".¹² Wallich in the meantime had convinced the publisher to add two other academic economists to the business pages, Paul Samuelson and Milton Friedman. Samuelson was the first John Bates Clark medallist, an award (bi-annual until 2009) given to the economist under the age of 40 that has made the "most significant contribution to economic thought and knowledge." Friedman won the medal 4 years later. Samuelson was the first American to receive the Nobel in Economics in 1970. Friedman received this accolade in 1976. Both were former advisors to Presidential candidates: Samuelson to J. F. Kennedy in 1960, Friedman to Barry Goldwater in 1964. The magazine introduced, in 12 September 1965, Friedman as "a leading conservative economic thinker" and Samuelson as "one of the nation's leading liberal economists." With both extremes of the political spectrum spoken for, Wallich was left with the center. As he put it: "I'm in the middle." In the year the trio began their writing, *Time* responded by resurrecting John Maynard Keynes to its cover in

11 Tobin, J. (1974) *The new economics, one decade older*, Princeton University Press.

12 Hazlitt, Henry (1966) "Parting Words" *Newsweek*, September 12.

December 31, and quoting Friedman announced “We are all Keynesians Now!”¹³

During the years that followed there was no lack of economic drama: a string of budgetary and monetary crises, oil shocks, rising inflation. The magazine was also, and unrelated to the economic climate, thrown into turmoil going through five changes in its editorial leadership.¹⁴ Its economic columns however did not change much. Wallich left the publication in 1974 to join the Federal Reserve Board. Samuelson and Friedman wrote alternating columns until 1982, when Lester Thurow replaced Samuelson. A Professor at MIT, where he was Samuelson's colleague, Thurow had a Harvard PhD degree in Economics and was the author of a popular book of 1981, *The Zero-Sum Society*.

The age of the economists ended when Richard Smith took the helm of the publication in 1984, a post he would keep for 15 years. One of his original changes was to replace the academics by journalists. The economic column was entrusted to Robert Samuelson.¹⁵ Jane Bryant Quinn was paired with him to write on personal finance and

13 *Time* did not have prominent economists writing its opinion. Instead it created a “Board of Economists” that it ostensibly advertised as a brain trust for the economic reporting in the magazine. Although they were not active in the reporting they lent credibility and credentials to *Time*'s economics.

14 The *Newsweek* editors were: Samuel T. Williamson (1933 – 1937), Malcolm Muir (1937-1961), Osborne Elliot (1961-1969 and 1972-1975), Ed Kosner (1969 – 1972), Lester Bernstein (1975 -1982), William Broyles (1982-1984), Richard Smith (1984-1998), Mark Whitaker (1998 – 2006), Jon Meacham (2006-2010).

15 “Joining NEWSWEEK as a contributing editor is Robert J. Samuelson” *Newsweek*, February 20, 1984, p. 64

consumer affairs. R. Samuelson was a 1967 graduate of Harvard University with a B.A. in government, who began his journalism career as a reporter on *The Washington Post's* Business Desk. After four years he left the paper to freelance, and was hired by *The National Journal* as an economics correspondent in 1976, writing its "Economic Focus" column. He joined *Newsweek* in 1984 with already a vast experience in economic writing.¹⁶

The corpus of texts analyzed in this essay covers all the economic columns published in the 1975-2007 period, after the departure of Henry Wallich, and before Daniel Gross joined the magazine and became R. Samuelson's alternate. The columnists in our study are: Paul Samuelson, Milton Friedman, Lester Thurow, and Robert Samuelson.¹⁷

3. Method: Co-occurrence analysis as a topic model

We accessed the set of texts by querying the Lexis Nexis database for *Newsweek*, for the

16 Robert Samuelson has earned many journalism awards, including in 1993 the John Hancock Award for Best Business and Financial Columnist and the Gerald Loeb Award for Best Commentary. Samuelson earned the National Headliner Award for Feature Column on a Single Subject in both 1992 and 1993. He received another National Headliner Award in 1987 for Best Special Interest Column, and the Gerald Loeb Award for Commentary in 1986. He was named a Loeb finalist in 1988 for his columns on the October 1987 Wall Street crash. Before coming to *Newsweek*, Samuelson also won a 1981 National Magazine Award and a 1983 Loeb Award.

17 Later versions of this paper will include analysis from earlier periods, with the goal of ultimately having the full corpus of *Newsweek's* economic opinion.

section “columnists”, and for the names of the already mentioned columnists.

Sometimes, author content at Lexis Nexis was not only to be found under the section category of “Column” but also under “Business”, and “National”: we kept that distinction in our metadata. Our corpus thus included 783 different texts, 352 for 1975-90 and 431 for 1991-2007, each economic column was coded individually. Each was characterized by the section attributed in Lexis Nexis, their date of publication (month and year), their author, and the number of visuals (i.e. charts or tables) that they included.

To analyze this corpus, we followed a co-occurrence analysis, or “bags of words” approach, which is the statistical analysis of the joint appearance of words in the same sentences. More specifically, we used the software Alceste, that is one of the products of an approach sometimes described as “French automatic textual data analysis”¹⁸, and that is finding increasing use in English psychology and political science literatures.¹⁹

18 Dargentas, Magda & Le Roux, Dominique (2005), “Potentials and Limits of Secondary Analysis in a Specific Applied Context: The Case of EDF—Verbatim”, *Forum: Qualitative Social Research/Forum Qualitative Sozialforschung*, 6-1, art. 40. Accessed from: <http://nbn-resolving.de/urn:nbn:de:0114-fqs0501404>, December 2010. On French software generally and on Alceste, see the review by Guérin-Pace, France (1998), “Textual Statistics. An Exploratory Tool for the Social Sciences”, *Population: An English Selection*, 10-1, 73-95.

19 See especially Bailey, Andrew & Schonhardt-Bailey, Sheryl. (2008), “ Does Deliberation Matter in FOMC Monetary Policymaking? The Volcker Revolution of 1979,” *Political Analysis*, 16-4, 404-427; Schonhardt-Bailey, Sheryl. (2008), “The Congressional Debate On Partial-Birth Abortion: Constitutional Gravitas And Moral Passion,” *British Journal of Political Science*, 38, 383-410; Schonhardt-Bailey, Sheryl. (2005), “Measuring Ideas More Effectively: An Analysis of Bush and Kerry’s National Security Speeches,” *PS: Political Science and Politics*, XXXVIII-4, 701-711

Alceste processes large corpuses of texts through the following steps:

1. Words are categorized by an internal dictionary that (a) separates “function words” (articles, auxiliary verbs, pronouns, etc.) from all other words that will be the object of statistical analysis and (b) it groups forms that are variants of the same word around a root or “reduced” form (e.g. industry, industries and industrial are grouped, as well as woman and women).

2. A table is constructed that indicates, for each sentence of the text, if it does or does not include each of the analyzed words (e.g. industr-, woman/en, etc.). The software segments texts into quasi-sentences which it labels as “elementary context units” (ECUs): these units include ca. 200-250 characters each and the software iterates to have them end with a punctuation mark as often as possible. Very long sentences are thus cut and very short ones joined.

3. The binary table is submitted to a clustering process of hierarchical classification. This process aims at separating the entire set of sentences (UCEs) of the corpus into two distinct classes so that these two classes share as little vocabulary as possible (in the ideal case, they would have no word in common, “function words” are excluded at this stage of the process). Each class is then split according to the same principle, and the process repeats until a robust classification is found.

At each stage of the process, the best possible partition of sentences in classes is chosen according to a chi-square criterion²⁰. Out of all the possible partitions, Alceste chooses the one that maximizes criterion:

$$\chi^2 = k_1 k_2 \sum_{j \in J} \left[\left(\frac{k_{1j}}{k_1} - \frac{k_{2j}}{k_2} \right)^2 / k_j \right]$$

where

$$k_{1j} = \sum_{i \in I_1} k_{ij}; \quad k_1 = \sum_{j \in I_t} k_{1j}; \quad k_j = k_{1j} + k_{2j}$$

The chi-square statistic is employed not, as is standard, as a test, but as a metric of dissimilarity between classes, where j denotes a word, and k_{1j} and k_{2j} denote frequency of contextual units in classes 1 and 2 that contain the word j .

Although our terms “out of all the possible partitions” and “until a robust classification is found” hide somewhat complex statistical processes which we do not report here²¹,

20 This table and the formula are adapted from Kronberger, Nicole and Wolfgang Wagner (2000). “Keywords in Context: Statistical Analysis of Text Features”. In *Qualitative Researching with Text, Image and Sound: A Practical Handbook*. M. W. Bauer and G. Gaskell (eds.). London, Sage Publications: 309. This textbook discusses Alceste on pages 299-317.

21 More details are given in Kronberger, Nicole and Wolfgang Wagner (2000). art.

what the software's approach is intuitive. It employs a clustering method often used on very sparse tables: ours is especially sparse, as most of the words of the whole corpus do not appear in each given sentence (generally 95% in Alceste analyses)²². In our case, employing a hierarchical clustering that joined sentences with a very *similar* vocabulary, as opposed to dividing the corpus into classes with a very *different* vocabulary as we do, would produce artifacts such as a large number of very small classes, small clusters of extremely similar sentences, and one very large class encompassing most of the corpus. The contrasting approach we employ allows (a) to distinguish classes (typically 3 to 9 classes) of sentences that are characterized by vocabulary that they do not share with the other classes, but also, as a consequence, by words shared inside the class and (b) to compute the distance between these classes, or “semantic contexts”, and to represent them in a tree diagram.

We would like to stress that these classes are produced by an inductive, “bottom-up” process. They are not influenced by any human coding or by research hypotheses, and hence have an essentially descriptive character²³. Their meaning requires interpretation provided by the historian, in the same way as the latent dimensions arising from factor analysis (a technique akin to Alceste in many ways). Alceste is eminently not a toolkit

cit. Some sentences cannot easily be classified by Alceste and form a residual “non-classified” class that is omitted in the results presented below. Specialists generally advise not to use a classification with more than 30% of unclassified sentences. In our case, this proportion is 24% in the first period and only 15% in the second one.

22 Guérin-Pace, France (1998). art. cit.

23 On the drawbacks of human coding, see esp. Guérin-Pace, France. (1998), art. cit. and Schonhardt-Bailey, Sheryl. (2005), art. cit.

for *a priori* hypothesis testing (although some of its additional results can be used for this purpose), but one for exploration and description. It does not require the analyst to hypothesize a topic model; to the contrary, its emergent vocabulary classes provide one possible, statistically-based definition of the main topics in the corpus and of the associated language patterns.

4. The classes of sentences produced during the third step are characterized by the words (and grammatical forms, and phrases) that are over- and under-represented in them, as well as by the characteristics (here the metadata of: author, section, date, charts) of the texts in which they are over- and under-represented and finally, by their most typical sentences, those composed by most over-represented words in the class. All this is again done by using a chi-square criterion. Alceste computes the number of times a word, sentence or grammatical form appears in a class and compares it to the expected value of its appearance across the corpus; and it computes the number of times a sentence by, e.g. Friedman, is included in a class and compares it to chance frequencies. Although the classification is made solely on the basis of "content words" (mainly nouns, verbs, adjectives and adverbs), it is possible, at this stage, to check if each class of sentences is also characterized by the over- or under-representation of "function words" (such as personal pronouns) or, for instance, of texts written by one author, or published in one given year. "Function words" and attributes of the texts are used as "illustrative", or passive variables in the analysis. Although Alceste doesn't analyze style *per se* (e.g. metaphors or the tense of verbs), some of these features can be

brought after the analysis to characterize the style associated to each vocabulary class. The automatic selection of typical sentences is also expedient in terms of interpretation, as it offers us a core set of sentences (and not only isolated words) that represent the class. It will allow us to extract quotations from the large corpus that are statistically meaningful, and hence addressing a common criticism that faces authors dealing with large texts, i.e. that they can “cherry pick” quotations that best fit their arguments.

Finally, Alceste provides a battery of additional analyses , visualizations and ways to explore the corpus (e.g. “concordances”, or lists of all the sentences including one word). Drawing on these, we have mainly used one such option for this paper: the *tri croisé* or “cross-tabulation”. This is a statistical procedure that, works in near reverse to the above described procedure since it starts with an hypothesis, a criterion that splits the text into two or more segments. From this meta criteria it determines the words that are most typical (the most over-, but also under-represented) of each portion of the corpus, once again on the basis of a chi-square test. *Tri croisé* can thus be used to contrast the vocabulary of one author's (or one section's) sentences to that of the other authors' sentences; in this case too, Alceste gives additional results such as the most typical sentences by this author. *Tri croisé* can also be used to explore the vicinity of each word by contrasting, e.g., the part of the corpus made of sentences including “Europe” to the part made of sentences not including “Europe”.

4. Ontologies of the economic (1975-1990 vs 1991-2007)

The events of 1989-1991 in Eastern Europe were of epochal significance. In 1992, there was a Presidential election in the USA, ending 12 years of Republican White House. This overlap of international and national events sets a credible cut point for our corpus. The choice is all the more felicitous because it splits Robert Samuelson's writings for the first period (roughly 6 years) to amounts comparable to P. Samuelson (7 years) and Milton Friedman (9 years).

At the first stage of our analysis, we examine changes from two periods of the corpus. Our contention is the evolution of Alceste's sentence classes and of statistical patterns of over- (and under-) representations of words in them provides us a map of an economic ontology. In this first analysis we assume that authorship does not explain these patterns, an assertion that will be discussed in the following section.

The first result we obtain is a frequency count of the most used "content words" in both corpora, which is a good proxy on the the topics of economic columns in *Newsweek* from period one to period two. Taking the 15 most common words in each period (with ca. 600 to ca. 1800 occurrences), we do not find much change. A third of these words are the same: new, govern-, spend-, tax-, year, rate(s), high-, American, price(s), work-, make. Their frequency might arise from a focus on domestic questions and economic policy, and forever a stress on novelty that primes urgency and excitement. The

changing words are: econom-, inflation-, percent-, market-, then say-, bill, people and it.²⁴ The simple word count analysis suggests stability, but as we shall see, by asking the question of what differences emerge in the formation of topics between the two periods, the language structure becomes both more complex and interesting. Alceste's classifications take into account that columnists discussing “economic” questions in fact discuss different subjects in different ways. It thus allows us to construct a model(s) of topics and through them discuss changes to economic ontologies in the press.

The topic model for 1975-90 exhibited seven classes of words that roughly evenly divided the corpus. Four classes were separated from three others, each cluster of classes including half of the sentences. The first cluster we will call the domain of “political economy”, the other the domain of the “causal economy” (see illustration 1). Although nothing in the inductive statistical method requires conformity of topics between the two period corpora (the independent analyses are independent), we observe the same clustering in the 1991-2007 corpus, although with revealing differences once we have a closer look.

Class numbered 1, over-represented in the “Business” section, had as top three most over-represented “content words”: families, security, care. The class demarcates text

24 What we know about the changes in the cast of writers from the two periods might suggest a correlation between these changes and the departure of professional economists, who we will see were more likely to repeat words such as percent, inflation or market and less likely to personalize their writing.

addressing matters of personal and family wealth, often describing social problems from the point of view of the people concerned. The connection between these subjects and the policies of governments is expressed in that class's relative proximity (see the tree diagram on top of the illustration) with class numbered 2, generally present in the same "Business" section and headlined by the words: republic- (generally Republicans), Senate, democrat- (generally Democrats), as the most over-represented. Sentences in this class often reported on the events in Congress and more generally what could be labeled "politics", frequently with more tenuous connection to economic problems conventionally conceived. One example, the sentence of the class that was most statistically representative, is

"Were the Democratic Party to nominate a Fred Harris, or some other serious populist who would scare the bourgeois voters the way George McGovern rightly or wrongly did in 1972, the Republicans could probably win the White House with ease in November." - Paul Samuelson, February 1976

Oddly, the terms "wife" and "diapers" were also salient in this class. Their presence denotes a trope of Robert Samuelson. In 1990, he gave a personal twist to the political debate about environmentalism with references to his family use of disposable diapers; the item ran for some weeks. An example excerpt is,

“Worthless badge: What my critics really resent is that I've denied their moral superiority. Using cloth diapers is an environmental badge, and I've said the badge isn't worth much. Their response is to accuse me of rationalizing my bad behavior.” - Robert Samuelson, April 1990

However even this quirky tangent is loaded with the characteristic political and moral controversy, following the agenda of Congress, its protagonists and discourse.

Class numbered 3 had as top words: world, Europe, Soviet, and addressed geopolitical events slanted by the ideological Cold war. In this class you can also find: war, freedom and Adam Smith. It encompasses many sentences naming countries, often in the context of critiquing communist countries along political lines (freedom, not only economic freedom) or related to the military threat that they represented. The sentences of this class appear principally in 1989 and 1990, when analyses about Eastern Europe were the most salient – and, quite naturally, in the “International”, but also “Reports” sections. This class was the closest to the set of sentences, numbered 7, reporting on: companies, firms and invest- (generally investment). A typical column addressed the structure of financial products available to consumers, stating that

“I knew that the wealthy investor or the corporate pension plan could utilize the Wells Fargo Bank of California, or the American National Bank and Trust Co. of Chicago, or the Batterymarch Financial Management

Corp. of Boston to engage in prudent across-the-market index investing. But I had then to report, sadly, that: "As yet there exists no convenient . . . fund that apes the whole market, requires no load, and that keeps commission turnover and management fees to the feasible minimum."" - Paul Samuelson, August 1976

Another example focused on the much discussed impact of technology in the management of human resources, and stated that,

"What firm does not now have a bigger legal staff? What firm has really used computers to get that much prophesied office revolution under way? What private managers are trying to improve decision-making so that they can fire managers? The answer: almost none. Those 5.5 million new white-collar workers prove it." - Lester Thurow, August 1981

Taken together, these four classes speak of private, of national, and of global affairs, hence very diverse scales of economic phenomena, but they all do so with reference to human choice and collective process. If we look for the most significant absences, words that are under-represented in this cluster, these are: inflation, rate, price, market and growth. Three of the classes (all but the more international class) are also characterized by an under-representation of the use of charts. We will label these classes as referring to "political economy."

The other cluster of classes, generally found in the “Columnists” section, is distinctive for imagining a world of cause and effect that seems to preclude choice. Some actors make an appearance in these classes – countries, “cartels”, presidents, the “chairman”, “consumers”, but do so in distinctively impersonal ways –, the core subject of these classes are the relationships between phenomena such as inflation, growth, deficits or consumption. Generally speaking, the “causal economy” classes were over-represented in *Newsweek* in the 1970s, while the “political economy” ones gained weight in the 1980s, and especially in 1988-90. Two factors might explain this pattern, one is a change, to be discussed in the next section, in the authorship of opinion from economists to journalists. The other factor are the dynamics of national and world events, where the concerns over inflation that took up the attention of public debate in the late 1970s and early 1980s, gave way to reporting and discussion about the political and economic revolutions that were bringing an end to the Cold War.

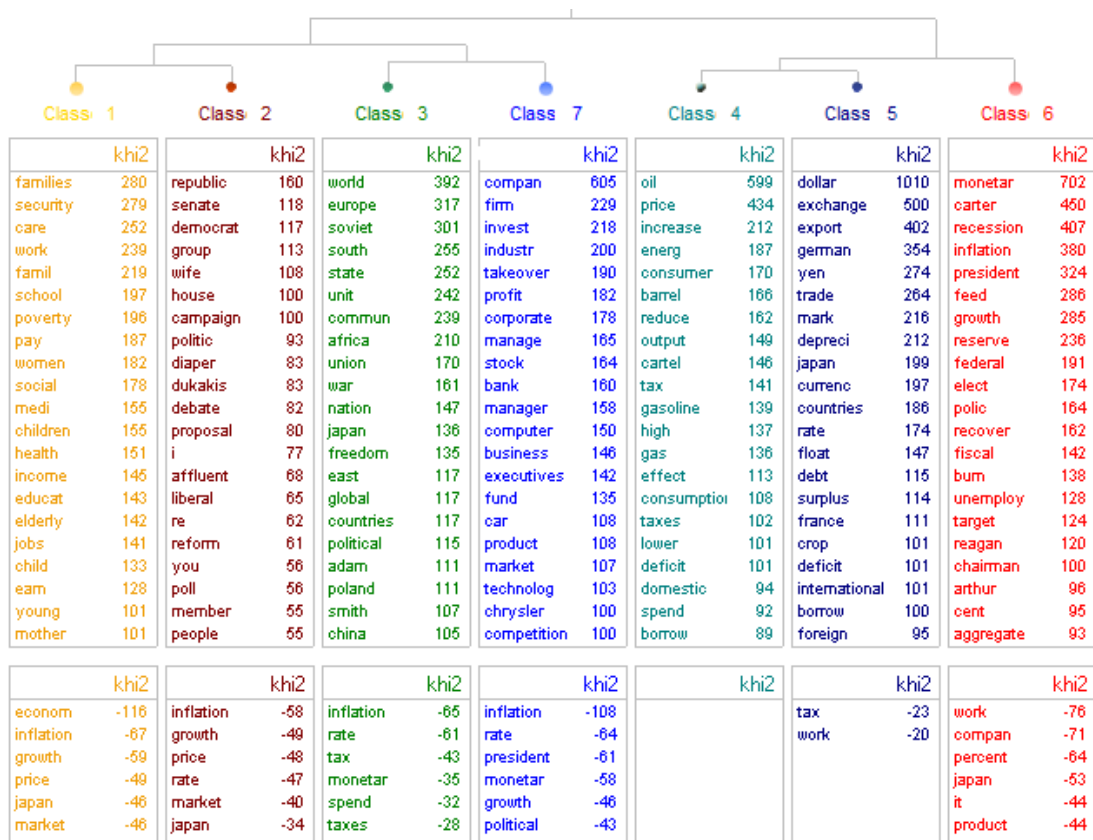


Illustration 1: Corpus 1975-1990. Over- and under-representations of “content words” in the 7 classes. The sentences split in the following shares: class 1: 17%; class 2: 11%; class 3: 13%; class 7: 11%; class 4: 8%; class 5: 20%; class 6: 20%

The second cluster included class numbered 4 with top words: oil, price, increase. This is one of the most time bound of all classes as it refers to the events and consequences of the 1970s oil shocks. Additional analysis reveals that the years 1975 and (less strongly) 1979 are distinctly over-represented in this class. Another feature of note, is that while “demand” and “supply” were terms used in all classes, they were only over-represented in this class. Close to that class is the one numbered 5 having most represented words: dollar, exchange, export. With class 6, it is the one that is most

concentrated in columns and inclusive of charts and tables. The focus of that class was trade with currencies as the centerpiece and crucial regulator of international affairs. Finally comes a class where there is an over-occurrence of words such as: monetary, Carter, recession, and also inflation, federal reserve, and Reagan. The class denotes the macroeconomic debate of the late 1970s over monetary management and efforts to contain inflation. It is striking that inflation was in this period firmly discussed as a monetary phenomenon, close but substantively distinct from (oil) cost shocks and consumer prices more generally.

For the second period 1991-2007, the topic model evidenced seven classes of words. Again we find a cluster of three classes and another of four (see representation of classes in illustration 2). The two clusters are not so distinct from the ones found in 1975-1990, and the second one is now slightly smaller than the first (40% vs. 60% of the sentences), but the distinction is maintained. Notably, in both periods, identification of grammatical categories reveals that words indicating the person (pronouns, possessives...) are consistently over-represented in classes of the first cluster and under-represented in classes of the second.

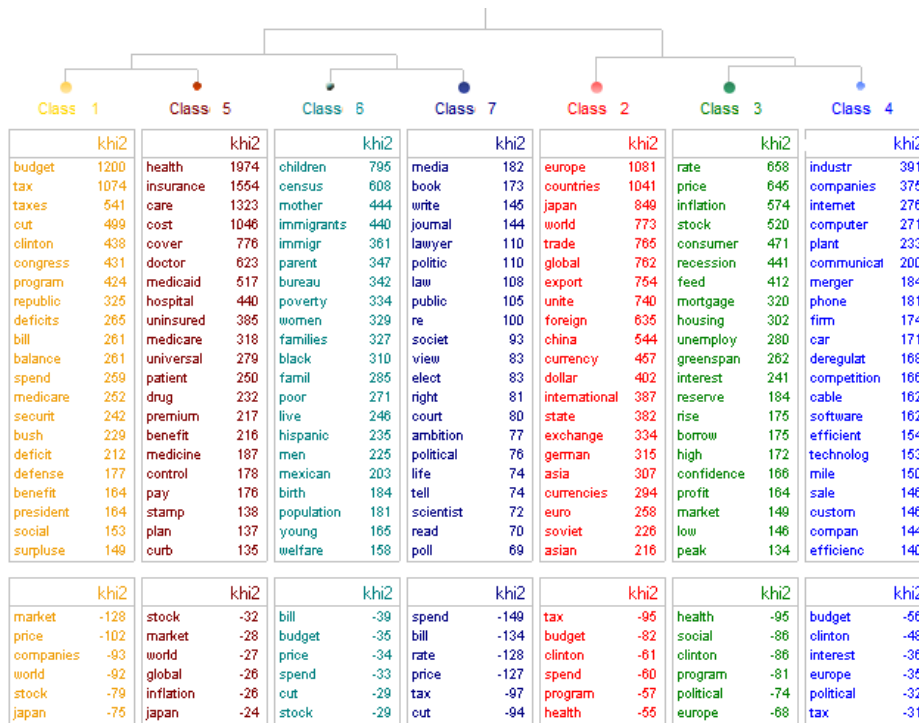


Illustration 2: Corpus 1991-2007. Over- and under-representations of “content words” in the 7 classes. The sentences split in the following shares: class 1: 19%; class 5: 7%; class 6: 6%; class 7: 20%; class 2: 17%; class 3: 21%; class 4: 11%

The first class identified in the second period had as most salient words: budget, tax, taxes. The sentences of this class were often published in the “National” section and rarely in the “Business” section of *Newsweek*. Notably, “budget” was not a salient word in our first period/analysis. Although it was very much present (ca. 300 times), it was evenly used in all of the semantic contexts found in the analysis, hence not significantly bound to any specific vocabulary class. “Fiscal” was in the class preoccupied by anti-inflationary policy, “taxes” was in a class together with “oil” and addressing consumers' purchasing power. In the second period these words co-occur in a class that is distinctive for denoting the dealings of Congress. One example reads:

“Nor would the benefits of a new tax system eliminate the political choices posed by the budget debate.

Even if, optimistically, Clinton and the Republicans agree to balance the budget by 2002, the issue won't be settled. Under the various budget plans, about half the required spending cuts would have to be made by future Congresses. And even if these cuts occur, Congress hasn't faced the broader issues raised by the retirement of baby boomers in the next century.” - Robert Samuelson, January 1996

Related to this class was a smaller one, numbered 5 also under-represented in the “Business” section, but present in the “National”, “Focus” and “Report” ones, with words such as health, insurance, and care heading the list of the most over-represented. The latter was closest to the class on Congress and not to class 6 with the headings of: children, census, mother. Our contention is that the class on health, insurance, and care, addressed the controversial status of the welfare state and national insurance, and discussions of its reform in the recent period. References to women and children, now appear grouped with matters of race and immigration, in a demographic frame. Closing this cluster of four classes was one, numbered 7, with top words: media, book, write. This is an emerging class that speaks of humanities subjects, which might include the art of writing, justice, and politics in a broader sense. It also connects to the

appearance of book reviews as a separate genre of writing, the class appears frequently in the "Business" section as well as in "Arts". This genre however also captured part of the vocabulary of class 2 from the first period, namely politics *per se*, with words such as election or polls, thus becoming more sharply separated from economic subjects, especially budgetary discussions. One typical example of class 7 was

"Even now, Hughes's elegantly written book is a wonderful read. But historians generally dissent from its central conclusion. They increasingly think that Eisenhower was more politically deft than people believed and that, in his basic judgments and policies, he conducted a successful presidency." - Robert Samuelson, January 2004

Another representative instance was

"Actually, it's only a brief comment on a windy treatise by a genuine law professor. As far as I can tell, no one ever read my short reply. Still, it plugged a good idea: the 'English rule.'"

I am a lawyer-basher and proud of it. Most American lawyers abhor the English rule. It requires the losing side in a civil suit to pay the winning side's attorneys' costs." - Robert Samuelson, April 1994

A cluster of three classes completes the 1991-2007 analysis. All of them were significantly under-represented in the “National” section of the magazine. Class numbered 2 was topped by words: Europe, countries, Japan, as it addressed the subject of international trade and competition. Within this cluster this was a fairly distinct class. The final two were more closely related. Class numbered 3, over-represented in the “Columnists” section, had: rate, price and inflation as most over-represented words and its themes are self-identified. It significantly under-represented personal pronouns and possessives. Class numbered 4, often in the “Business” section, was headed by: industr- (generally industries), companies, and internet and addressed business reporting. The changes to the vocabulary assigned to economic opinion on industry is one of our more significant findings. The references to “industries” doubled their weight in the corpus, from 11 to 21%, and moved from the realm of “political economy”, tied to geopolitical calculations, to that of the “causal economy”. Late twentieth century debates on globalization are accompanied by changes in vocabulary: the interest in innovation and communication industries, and a shift in the discourse about firms and management from the political to the impersonal. In addition, this movement happened at a time when the single economic writer favored a quite personal and familiar tone in his writing (more below). One example of this class was

“Since 1997, for instance, paper imports have tripled from 2.6 percent of U.S. consumption to 7.6 percent. And what about all the new information technologies? Yes, they also contributed to the weeding-out process and

the resulting productivity gains. But their role was secondary.

Sometimes the fittest firms are technology leaders--Dell and Wal-Mart, for instance. Both companies excel at using information systems to regulate inventories and orders. Well-deployed new technologies almost always enable companies to cut costs and expand markets." - Robert Samuelson, December 2003

Another instance of this class reads

"First, the PC explosion -- and everything associated with it, from software to ISDN lines -- has been a driving force of the U.S. economy. It has, for example, propelled strong business investment, as companies overhauled offices, factories and customer networks. In 1990, company purchases of high-tech equipment (computers, communications gear, instruments) was 20 percent of all business investment, which includes everything from office buildings to industrial machinery, reports economist Richard Rippe of Prudential Securities." - Robert Samuelson, April 1999

Comparing the results from the two periods, some classes have remained relatively stable. Our interpretation of the correspondence between periods is provided in

illustration 3. Of most interest are movements whereby classes of words, nearly wholesale, move between what we call the domains of “political economy” and the “causal economy.” And secondly, how some classes of the first period are broken or recombined into classes of the second period.

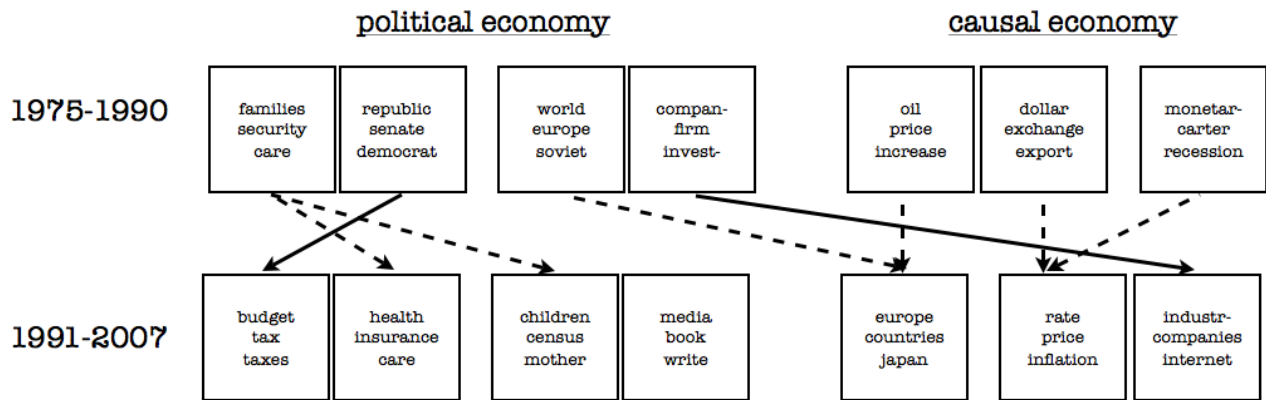


Illustration 3: Changes in the classes of vocabulary from 1975-1990 to 1991-2007. Solid lines refer to relatively stable classes. Dotted lines refer to classes that are broken or combined in the second period

Of the first kind the significant movements were that of classes on geopolitical and business affairs away from the domain of the political and the personal into the domain of the factual and impersonal. This, we suggest, accompanies a change in context, from the final years of the Cold War towards a new enthusiasm for a global order of free trade and enterprise. The most under-represented words for this pair of classes were: Clinton, budget, political. So while Europe in 1975-1990 period evoked the following analysis,

“Just because communism is crumbling in the Soviet Union does not mean Washington can ignore the military threat to Europe. A noncommunist Russia would still have 60 divisions west of the Urals, and a nationalist right-wing regime might be tempted to use them to protect the Rodina

(Homeland) -- and distract from economic problems at home." - Robert Samuelson, January 1990

In 1991-2007, Europe was not longer a piece of Washington's foreign policy, as in the example,

"Decreased capital flows to Latin America would curtail world trade, whose growth in 2001 is already projected by the World Bank to drop by more than half from last year. Latin countries need foreign investment (in dollars, euros or yen) to pay for some of their imports. Less investment would cause them to cut their purchases from each other and from the United States, Europe and Asia." - Robert Samuelson, April 2001

Here lay the domain of economic life that was independent of civic action and where players were the nation states' currencies, capital, and businesses. This domain, with classes 2, 3 and 4 of the second period was most salient in the 2000s, as compared to the 1990s, ie. as we move further away from the shadow of the Cold War.

Some other classes are reconfigured, while not moving in their domain of discourse.

The budget process emerges as the centerpiece of reporting on Congress (class 2 in 75-90 to class 1 in 91-07) that had in the earlier period been discussed in relation to electoral campaigns, polling and party drama, the latter topics pooled into class 7. Discussions on

families and care (class 1 in 91-07) offers us the most interesting transformation of this kind. In the earlier period references to families were akin to the following example,

“Since the young pay for their medical care while the elderly enjoy social security's medicare, when health costs are considered the average elderly family enjoys a standard of living that is essentially equal to that of the nonelderly. If social security and medicare were to be abolished, the standard of living of the elderly would be more than cut in half. Many elderly families depend almost entirely on social security for their standard of living.” - Lester Thurow, May 1982.

In the latter period, such topics can appear in two classes. Class 5 refers, quite specifically to the “Clintons” plans for welfare reform in 1993-4 (and evoked again as lesson in later years) as in the example,

“These assessments, though partially true, miss the larger truth. The real undoing of "reform" was its utopianism. The Clintons tried to create a perfect health-care system and, in the process, missed an opportunity to improve a system that is inevitably imperfect.

We would all prefer a system that provides "universal" insurance for every conceivable ailment (from diet disorders to long-term care), remains "affordable" and allows complete "choice" for patients and doctors alike.” -

Robert Samuelson, October 1994

While in class 5 the tone could be fierce, in class 6, family becomes subject of factual exposition, speaking of trends in demographical change. This class is distinctive for an over-representation of numbers. One illustration is,

“Indeed, the material well-being of poor children has generally improved, report Mayer and sociologist Christopher Jencks of Harvard in another study. In 1970, about 26 percent of the poorest fifth of children hadn't visited a doctor in the past year; by 1989, the figure was only 14 percent. In 1973, about 71 percent of these children lived in homes without air conditioning; by 1991, only 45 percent did. Unfortunately, these material improvements didn't translate into better social conditions. Crime rose; so did out-of-wedlock birthrates.” - Robert Samuelson, May 1997

The subjects: family, old age, childhood, are pulled apart from one period to the next. In presence of the debate about public insurance, the subjects are talked in the committed language of Congress and politics. When the reporting bears no connection with that controversy, subjects become the near opposite, in the language that reviews books, or analyzes public opinion.

Brought together, these changes suggest that the ontology of the 1991-2007 period is

thus much more framed by policy issues. Although in 1975-1990 period there was a distinct class that spoke of inflation and its policies, this class was placed in the domain of cause and effect. Both the budget debates and the national insurance controversy in the later period, elicit a distinct vocabulary, polemical and argumentative, and they sit in what we call the domain of political economy. Our other finding is that the worlds of business and international politics that had been gripped by the ideological divides of the Cold War moved to the domain of the causal economy, and become remarkably beyond the political.

5. Professors vs journalists (1975-1990)

A second element in our analysis compares the authors that penned these opinion columns. The prominent academic credentials of some and the journalistic achievements of others, makes this is a formidable group. Insights into the style of their writing and their favored topics allows us to peer into the cultural worlds of economic opinion leaders. We examine herein the 1975-1990 period where several authors wrote alongside each other.

We began by asserting that authorship alone could not explain the emergence of the classes discussed above. This claim is partly corroborated by examining how an author's output split in terms of the classes we have above identified (see illustration 4).

All authors contributed to the observed classes. Somewhat surprisingly given their status as opposites in the spectrum of American politics, Paul Samuelson and Milton Friedman looked very similar: while they derived different policy recommendations from their analyses, their vocabulary was alike. Both had the classes we have labelled the domain of the causal economy taking well above 50% of their writings, as much as 65% for Friedman, and 58% for P. Samuelson, Friedman's texts being more over-represented in the "oil crisis" class and Samuelson's in the "exchange rates" class. Robert Samuelson wrote on these classes no more than 15% of the time. Even more impressive is that class 6 is so small for Robert Samuelson while it takes up 35%-40% of Friedman and P. Samuelson's writing. The class denotes a concern for the policies of the state but in terms that Robert Samuelson might have found too academic. For instance, from Friedman, an analysis of the Federal reserve policy, elicits the following text,

"Presumably the recent monetary explosion will also be followed by a corresponding swing in income, but it is too soon to know whether it has yet be begun.

(1) To smooth out erratic movements, all rages of growth are averaged over six-month periods.

Had the Fed avoided the swings in money, there might still have been swings in income, but they surely would have been much milder." -

Milton Friedman, June 1976

The same non-narrative approach is taken by Paul Samuelson looking ahead to the economy of 1979, as in

“The private economy still looks pretty strong. Consumers, despite more debt, still buy cars, still crowd the stores. Housing holds up. Inventories seem balanced. It is argued that, if the Fed does not bring the real growth rate down below our long-run trend rate of 3 to 4 per cent per year, the inflation that accelerated in the last half of 1978 will further accelerate in 1979-1980.” - Paul Samuelson, January 1979

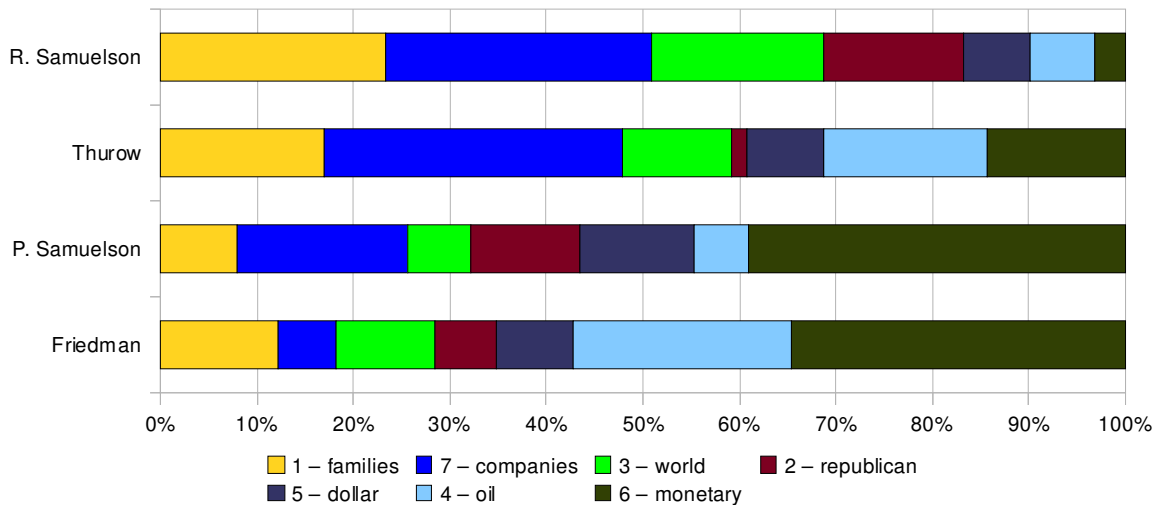


Illustration 4: Share of each word class on the writings of each author 1975-1990.

In contrast, for both Robert Samuelson and Lester Thurow classes 1 and 7, families and companies, dominate, and with class 3, addressing world politics, they group to

encompass over half of their output. All four “political economy” classes are, highly over-represented in Robert Samuelson's columns, the “politics” and “companies” classes being under-represented in texts by Friedman and the “families” and “world” classes in texts by Paul Samuelson. One might argue that these patterns reflect the authors' tenure at the magazine, which came after 1982 and faced a different set of world and economic events. However, Friedman fully overlapped with Thurow's two years at the magazine. What is more likely is an editorial re-orientation of the magazine. Thurow joined the magazine for his strengths at discussing business affairs, and Robert Samuelson had an even broader range.²⁵ Friedman and Samuelson were policy commentators, former policy advisers, and we see markers of this identity in their language.

The most striking result of the author analysis is a confirmation and clarification of Robert Samuelson's colloquial style (see illustration 5). In the list of the twenty words over-represented in his writing we find: *it, there, huge, they, that*, and the contractions: *'re, isn't, didn't, doesn't, can/can't* (note that in this analysis we take “function words” as well as “content words” into account). Indeed the best discriminant between the one journalist (R. Samuelson) and the academics (all others) is one word “*it*”. The word is the most under-represented in the academics' writing and the most over-represented in

²⁵ It is worth recalling that we read here distinctions in topical vocabulary. It is not that R. Samuelson did not write on the “causal economy”, as we have seen, even with some changes in its formulation, the same cluster exists in the magazine's script during Samuelson's solitary tenure post 1990. The argument is that in this earlier period his writing was more distinctive and focused on this first sphere of vocabulary.

Robert Samuelson's text. The ambiguity and prosaicness of "it" is seemingly embraced by the journalist but abhorred by the scholar. Among the "function words" most over-represented in the class of sentences about politics (and diapers), we find R. Samuelson using variants of the first person: I, my, me, as well as "you" and "yes": his style is both direct, close to orality, and highly personal, establishing closeness with the reader (in the 1991-2007 period, when he is the only writer, this style is best expressed in class 7, that of book reviews and politics). Even in the more impersonal, geopolitical class 3 of 1975-90, R. Samuelson often uses "our" and "we", as well as "you" and "your" in the "companies" class. The use of personal pronouns, abundant in R. Samuelson's writing, is however not limited to him alone. Class 6, the most characteristic of the economist's economics, with its statements about money and inflation, also over-represents "I" and "am" but their use is substantively different. Friedman and P. Samuelson as they talked about "M1" and "M2" express "their" views on the controversial matters, so "I" refers to the voice of the expert reflecting on a controversial subject, not of an "I" actor in the economy as in the writing of R. Samuelson.

Présence	khi2	Présence	khi2	Présence	khi2	Présence	khi2
monetar	258	carter	167	it	353	america	170
govern	151	stock	118	percent	331	firm	159
produce	128	recover	112	compan	185	american	137
effect	113	bum	100	there	151	saving	132
amendment	111	gold	92	huge	137	competitor	123
feed	109	out	87	re	121	mexican	113
cent	94	kennedy	80	they	108	industrial	113
control	86	cent	72	global	106	coll	111
inflation	77	ford	72	isn	104	negative	86
africa	73	recession	72	japan	92	net	82
propose	72	forecaster	71	farm	90	necessary	77
hence	71	tum	67	bush	85	elderly	73
produc	69	dr	62	technolog	79	fire	68
enterprise	67	load	62	didn	73	reagan	58
taxpayer	64	float	61	can	68	blue	52
taxes	63	concern	61	idea	68	work	48
volunt	63	index	60	trade	67	current	48
impose	59	down	59	that	66	consumptio	48
addition	59	arthur	58	doesn	59	eastern	46
draft	58	real	56	american	59	security	44
rhodesia	58	team	55	children	59	personal	43

Absence	khi2	Absence	khi2	Absence	khi2	Absence	khi2
it	-128	percent	-156	cent	-223	it	-17
percent	-115	govern	-88	monetar	-211	stock	-9
american	-114	it	-75	carter	-137	write	-9
econom	-90	spend	-69	president	-97	congress	-9
think	-66	high	-63	inflation	-64	re	-8
compan	-66	compan	-52	down	-55	feed	-8

Illustration 5: Over-represented and under-represented words for (left to right) Friedman, P. Samuelson, L. Thurow and R. Samuelson.

For the other authors the analysis also reveals interesting patterns (see illustration 5). Friedman's focus on monetary subjects hardly warrants commentary. Among his top words are: monetar-, govern-, and inflation. Not surprisingly also the words: control and voluntary, building on his advocacy for economic freedom. Also over-represented is the term "Mr.", which is consistent with Friedman's observed use of "politeness" strategies in addressing critics.²⁶ In contrast, P. Samuelson was more fond of quoting academic authority and using the term "Dr." We also observe that Samuelson was fond

²⁶ Dudley-Evans, Tony (1993) "The Debate over Milton Friedman's Theoretical Framework: An Applied Linguist's View." in *Economics and Language*, edited by Willie Hender, Tony Dudley-Evans and Roger Backhouse. Routledge: 132-152.

of naming Presidents: Carter, Ford, Kennedy are highest on his most salient words.

Thurow in light of his *Zero Sum Society* was most preoccupied to locating: America, and American in matters of competitiveness.

6. Summing up

The statistical analysis of word collocation in *Newsweek's* economic opinion offers a rich fabric of language patterns. We argue that this record affords invaluable insights into the dynamics of economic culture in the past four decades, and into the language of professionals and journalists. Consistent throughout the two periods under study, 1975-1990 and 1991-2007, two domains can be distinguished. One encompasses discussions of subjects where choice and decision making lie at the core and which we have labelled the domain of “political economy”. In contrast, there is a grouping of topics where cause and effect, the laws of economics, seem to apply without requiring individual action, but never dispensing elucidation, reflection, and opinion. Although the two domains can be observed pre and post 1990, what is included in each domain changes. Crucially, with the unwinding of the Cold War, international affairs have come to fall on the sphere of the “causal economy”. Also notable, is the refocusing of subjects along the lines of policy controversies, the budget process and debates about health insurance reform.

We were interested in examining how authors with different backgrounds differed in their choice of topics, in the way they put them together, and in their style of writing.

Our findings plausibly establish differences in terms of subject matter: professional economists tended to concentrate on what we called the domain of cause and effect; the journalist and one of the more publicly oriented academics, on topics of political economy. Yet, the most salient difference lay in the colloquial and personal style of the journalist's writing, as opposed to the more deferential preferences of the scholars. The former elicited closeness, the latter elicited distance.