# WHAT DO WE THINK AN ECONOMIST SHOULD KNOW?

## A MACHINE LEARNING INVESTIGATION OF RESEARCH AND INTERMEDIATE-LEVEL TEXTBOOKS

Samuel Bowles, Wendy Carlin, Simon D. Halliday, Sahana Subramanyam

23 August, 2019

The content of intermediate-level undergraduate textbooks represents a consensus in the discipline about what a student trained in economics should know. We use topic modeling to explore both this conceptual benchmark in leading textbooks and the content of economic research in journal publications over 115 years. Our mapping of content to 3-dimensional meta-topic spaces in microeconomics and macroeconomics reveals that the conceptual frameworks used in research have diverged over the last four decades from the benchmarks conveyed to majors through textbooks. We suggest that the origins of the divergences and the implications for economics education differ between microeconomics and macroeconomics.

JEL: C11 Bayesian analysis; A22 Undergraduate economic education; B20 History of economic thought since 1925.

.

Affiliations: Bowles, Behavioral Sciences Program, Santa Fe Institute samuel.bowles@gmail.com; Carlin, Department of Economics, University College London and CEPR w.carlin@ucl.ac.uk, Halliday, Department of Economics, Smith College simon.d.halliday@gmail.com, Subramanyam, Department of Economics, Bocconi University sahana.subramanyam@gmail.com

Acknowledgements: We would like to thank the Behavioral Sciences Program of the Santa Fe Institute for support of this research, and Ingvild Almås, Antonio Cabrales, Louis Carlin, Simon DeDeo, Doyne Farmer, Jessica Flack, Joshua Garland, Albert Kao, Christopher Lee, Alan Kirman, Manfred Laubichler, Suresh Naidu, Michael Price, Paul Seabright, Rajiv Sethi, Joel Sobel, Margaret Stevens and Sylvie Thoron for helpful comments and criticisms, and Scott Cohn for excellent research assistance.

We use topic modeling to explore Thomas Kuhn's concept of a scientific paradigm in the field of economics – roughly, "what an economist should know" – and how this paradigm has evolved over the past century along with changes in economic research.

For Kuhn, a discipline's paradigm (or benchmark, to use a term more familiar to economists) provides "[a]nswers to questions like these … What are the fundamental entities [under study]? How do these interact with each other…? What questions may legitimately be asked about such entities and what techniques employed in seeking solutions?" (Kuhn 1970) :5 "The [scientific] community's paradigms," he wrote "[are] revealed in its textbooks, lectures, and laboratory exercises. By studying them and practicing with them, the members of the corresponding community learn their trade." (p.43)

In many fields of study – history, for example, or political science – there would be little consensus on answers to the questions Kuhn asked above, so it would be difficult to identify a paradigm. In others, a relatively well-defined and widely agreed upon set of questions and methods of answering them is evident in the second- and third-year undergraduate courses required for majors. In physics, for example, second- or third-year courses in classical mechanics, statistical physics (including thermodynamics), electromagnetism and quantum physics are standardly required. And the content of such courses differs little from professor to professor or from institution to institution.

As is the case in physics, in economics, mastery of the content of intermediate microeconomics and macroeconomics courses, as Kuhn put it, is what constitutes "the educational initiation that prepares and licenses the student for professional practice" (p.5). The vast majority of people who have studied economics "learn their trade" as undergraduate economics majors (not as doctoral students). The data for our investigation of changes in the paradigm in economics will therefore be from the intermediate courses rather than from PhD or other graduate level instruction. And we do not draw upon the content of the introductory courses, because such courses are mainly taken by students who will not take another economics module. In the words of one of the most successful introductory textbook writers: "I am guided by the fact that, in introductory economics, the typical student is not a future economist but is a future voter." (Mankiw 2016):170.

One way to investigate and calibrate changes in the benchmark that intermediate-level students are expected to learn would be a deep reading of the relevant textbooks, parallel to Roger Backhouse's study of Samuelson's 1948 introductory textbook and the environment in which it was written (Backhouse 2017). Alternatively, one could collect syllabi and exam papers from the relevant courses. But it would be virtually impossible to find a representative set of these documents extending back more than a few decades. Here we adopt an alternative approach, where we use textbooks as the source data on what students learn in intermediate courses and a methodology – machine-learning – that reduces the role of the researchers' judgments about content in favor of a more data-centered approach.

To do this we draw upon a second set of data – research published in the top economics journals since 1900 – which we refer to as the research corpus. By applying a machine-learning technique -- topic modeling -- to this research corpus, we are able to produce a lens in the form of a set of economically meaningful topics that can be used for measuring themes and their relative importance in any work in economics. This lens can be focused on the object of interest, which in our case is the content of the intermediate-level textbooks, allowing comparisons with minimal subjective judgement by the researcher. In a second step, to map the distance between the research frontier and the benchmark represented by the textbooks, we then focus the same lens (the set of topics) on the research corpus itself, allowing us to track the evolution of the themes that have been prominent among leading research economists over the past century or so.

We find that immediately following the Second World War, the topic distribution in both the research corpus and undergraduate instruction were remarkably similar. This was the case in both microeconomics (where Marshallian topics predominated) and macroeconomics (where Keynesian topics were prominent). Since then the research corpus has diverged from benchmark models taught in the intermediate level courses.

In macroeconomics this reflects the broad acceptance of aggregate demand and other Keynesian concepts as both essential to what an economist should know but no longer a focal point of economic research. In microeconomics, by contrast, the eclipse of Marshallian themes in the research corpus by models of strategic interaction and incomplete information represents a widely-accepted view of market structure and competition among research economists that has not been accorded the status – in undergraduate instruction – of what an economist should know.

## I. Topic modeling of the research corpus

Topic modeling is a Bayesian machine-learning technique that treats a corpus of observed data (in our study, the research papers that are the documents making up the research corpus) as arising from a hidden data-generating process, the structure of which is to be estimated (Blei, Ng, and Jordan 2003a, b, Blei 2012, Gentzkow, Kelly, and Taddy 2019). The method has proven insightful in the subjects that we address, that is, both economics and intellectual history (Ash, Chen, and Naidu 2018, Klingenstein, Hitchcock, and DeDeo 2014, Barron et al. 2018).[1] To study the benchmark taught to economics majors we develop a measure of economic content based on the themes present in research papers in top economics journals since 1900. These themes, formally a set of topics, can be used to explore the content of any work in economics.

The observed data is a set of N unique tokens, that is, words or bigrams (two-word couplets like 'minimum wage') located in a set of D documents. Thus, our observed data is shown by the matrix **W**:

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1D} \\ w_{21} & w_{22} & \cdots & w_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & w_{ND} \end{bmatrix}$$

where $w_{id}$ is the number of occurrences of word $i$ in document $d$.

Based on these data the estimated topic model then delivers two matrices. The first comprises the K topic vectors whose elements – the N token weights in each topic vector – are the probability that the token will contribute to the document's "bag of words" conditional on the topic being drawn to contribute to the document. So, denoting $\beta_{ik}$ as the weight of token $i$ in topic $k$, we have the matrix of K topic vectors in matrix **B**, where each vector of weighted tokens, $\beta_k$ , is called a topic:

[1] The simplest and most widely used topic model is the latent Dirichlet allocation or LDA model based on the discrete distribution due to the 19th century German mathematician Gustav Lejeune Dirichlet. The LDA model is best understood as a type of principal components analysis applied to discrete data (the presence of a particular word in a topic or topic in a document, in this case).

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1K} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N1} & \beta_{N2} & \cdots & \beta_{NK} \end{bmatrix} \qquad \beta_k \ (k \in [1, K]) \qquad \beta_{ik} \text{ with } \sum_{i=1}^{N} \beta_{ik} = 1$$

The second matrix, $\mathbf{\Theta}$, is the set of D document vectors, the elements of which ($\theta_{kd}$) are the probability that topic *k* will be drawn to contribute tokens to document *d.* Each document in the corpus is mapped on to a document vector of weighted topics, $\theta_d$.

$$\mathbf{\Theta} = \begin{bmatrix} \theta_{11} & \theta_{12} & \cdots & \theta_{1D} \\ \theta_{21} & \theta_{22} & \cdots & \theta_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{K1} & \theta_{K2} & \cdots & \theta_{KD} \end{bmatrix} \qquad \theta_d \ (d \in [1, D]) \qquad \theta_{kd} \text{ with } \sum_{k=1}^{K} \theta_{kd} = 1$$

Topic modeling uses an expectation maximization algorithm to provide maximum likelihood estimates of the process (the matrices **B** and $\mathbf{\Theta}$) that could hypothetically have generated the observed data, matrix **W**.

Other than the choice of the number of topics K and the corpus of documents making up the observed data (see below), the topics $\beta_k$ and the document vectors $\theta_d$ are generated by the expectation maximization algorithm without input from the researcher.[2] Neither the meaning, the order in a document nor the temporal order of documents in the corpus is used in generating the topics. Thus, each document is treated as a "bag of words" where the only observed structure is the presence and frequency of words in documents.

Any particular **B** and $\mathbf{\Theta}$ will generate a predicted distribution of words across the set of documents. The data-generating process by which words are deposited into the bag of words making up a particular document, *d*, occurs as follows. A topic is selected to contribute a word (meaning, a token) to the bag, topic *k* being drawn with probability $\theta_{kd}$; then a word is drawn from topic *k*'s vector of words, word *i* being drawn with probability $\beta_{ik}$. The process is then repeated until the document has its complement of words. This is done for all documents in the corpus. If the iterative process of the expectation maximization algorithm converges, the

[2] To run the LDA model we also selected the two hyperparameters (at the default values in the gensim package) governing the variance of token weights in the topics and topic weights in the document vectors (See Appendix A).

resulting **B** and **Θ** are then said to describe the generative process most likely to have produced the observed data, that is, the actual distribution of words in the research corpus.

Once the topics (the vectors making up **B**) have been generated, we confirm by inspection that they are interpretable in economic terms and give them names. These hundred named topics constitute the instrument or lens that we then use to measure the content of any document, textbook or time slice of the research corpus.

To implement this method, we proceed in three steps. First, we select a corpus of documents from which to generate topics. This corpus is economics research comprising all articles published in the major economics journals in the UK and USA between 1900 and 2014, a total of 27,436 articles as shown in the top panel of Figure 1.

| | 1900 - 1924 | 1925 - 1949 | 1950 - 1974 | 1975 - 1999 | 2000 - 2014 | Proportion of articles in journal |
|---|---|---|---|---|---|---|
| *Journal of Political Economy (1900 - 2014)* | 0.028 | 0.031 | 0.051 | 0.049 | 0.021 | 0.179 |
| *The American Economic Review (1911 - 2012 )* | 0.024 | 0.038 | 0.035 | 0.097 | 0.039 | 0.233 |
| *The Economic Journal (1900 - 2011)* | 0.013 | 0.012 | 0.021 | 0.040 | 0.018 | 0.104 |
| *The Quarterly Journal of Economics (1900 - 2011)* | 0.019 | 0.029 | 0.043 | 0.046 | 0.019 | 0.155 |
| *The Review of Economics and Statistics (1919 - 2003)* | 0.006 | 0.019 | 0.038 | 0.050 | 0.026 | 0.139 |
| *Econometrica (1933 - 2014)* | | 0.010 | 0.025 | 0.039 | 0.016 | 0.090 |
| *The Review of Economic Studies (1933 - 2012)* | | 0.010 | 0.025 | 0.040 | 0.023 | 0.099 |
| **Proportion of articles in time frame** | 0.091 | 0.151 | 0.238 | 0.359 | 0.161 | 1.000 |

| Title | Authors | Latest Edition | First Published | Edition(s) Used |
|---|---|---|---|---|
| Microeconomics: Competition, Conflict and Coordination | Bowles and Halliday | 1st ed, 2020 | 2020 | 1st ed, 2020 |
| Microeconomics and Behavior | Frank | 9th ed, 2015 | 1990 | 7th ed, 2008 |
| Microeconomics | Perloff | 8th ed, 2018 | 1998 | 7th ed, 2014 |
| Microeconomics | Pindyck and Rubinfeld | 9th ed, 2017 | 1988 | 9th ed, 2017 |
| The Theory of Price | Stigler | 4th ed, 1966 | 1942 | 4th ed, 1966 |
| Intermediate Microeconomics | Varian | 9th ed, 2014 | 1987 | 9th ed, 2014 |
| Macroeconomics | Blanchard | 7th ed, 2017 | 1991 | 4th ed, 2006 and 7th ed, 2017 |
| Macroeconomics: Imperfections, Institutions and Policies | Carlin and Soskice | 1st ed, 2006 | 2006 | 1st ed, 2006 |
| Macroeconomics: Institutions, Instability and the Financial System | Carlin and Soskice | 1st ed, 2015 | 2015 | 1st ed, 2006 |
| Macroeconomics | Jones | 4th ed, 2017 | 2008 | 1st ed, 2008 and 3rd ed, 2014 |
| Macroeconomics | Mankiw | 10th ed, 2019 | 1996 | 6th ed, 2006 and 9th ed, 2016 |
| Outlines of Economics | Ely | 6th ed, 1937 | 1893 | 5th ed, 1930 |
| Economics | Samuelson | 19th ed, 2009 | 1948 | 1st ed, 1948 |

**Figure 1: The corpus of documents (27,436 research papers since 1900, top panel) and the intermediate-level textbooks (in microeconomics and macroeconomics, bottom panel).**

The corpus is processed by 'stemming' to collect as a single token the set of words that are present in different forms such as a noun, a verb or an adjective ("competition"; "compete"; "competitive") and by using dictionaries to remove so-called stop words that are without

informative content for our purposes (conjunctions, pronouns, prepositions). This processing results in a vocabulary of 10,849 unique tokens.

Second, we set the number of topics K = 100 and successfully generated the topics, matrix **B**.[3] We stuck with one hundred topics because the topics generated are interpretable and the level of granularity is similar to the 111 2-digit JEL codes (excluding 'general' and 'non-substantive' ones). To proceed, we used our judgement based on the token weights making up each vector to name the topics.[4]

Figure 2 illustrates Topic 4 and Topic 10, two of the hundred topics using word clouds (the numbering of topics is not meaningful). In the word clouds, the size of the font is proportional to $\beta_{ik}$., the probability that the word or bigram contributed to a document's bag of words, conditional on that topic being drawn to contribute. In the word cloud in the left panel, the most heavily weighted tokens, are "quality" with a token weight of 0.296 and "car" with a weight of 0.069, meaning that if Topic 4 is selected to contribute to a document, these two tokens will be contributed to the document's bag of words with probabilities 29.6% and 6.9% respectively. We named this topic "Adverse selection; 'lemons'".

The five articles for which Topic 4 is most heavily weighted and their topic weights ($\theta_{4d}$, the estimated probability that Topic 4 will be drawn to contribute tokens to the article in question) are shown in bottom left of Figure 2. The first left-hand column entry in the left panel, for example, means that for any particular draw in generating the bag of words represented by Hendel, et al, 1999, Topic 4 would be selected to contribute with probability 0.36, and similarly for the other papers. The panel on the right in Figure 2 presents similar information for the topic we named: "Bargaining and incomplete information."

[3] We use the same 100 topics based on the research corpus since 1900 to analyze introductory textbooks in another paper (Bowles and Carlin 2019).

[4] In the online appendix of a related paper (Bowles and Carlin 2020), we present word clouds of the 100 named topics also used here, showing the top hundred tokens along with the names we assigned to each topic.

Topic 4 : Adverse selection; "lemons"


| | |
|---|---|
| 0.3619 | Igal Hendel, Alessandro Lizzeri (1999). Adverse Selection in Durable Goods Markets. The American Economic Review |
| 0.3364 | Alessandro Gavazza, Alessandro Lizzeri, Nikita Roketskiy (2014). A Quantitative Analysis of the Used-Car Market. The American Economic Review |
| 0.3361 | Jae-Cheol Kim (1985). The Market for "Lemons" Reconsidered: A Model of the Used Car Market with Asymmetric Information. The American Economic Review |
| 0.3355 | Christopher L. House, John V. Leahy (2004). AnsSModel with Adverse Selection. Journal of Political Economy |
| 0.3196 | Igal Hendel, Alessandro Lizzeri, Marciano Siniscalchi (2005). Efficient Sorting in a Dynamic Adverse-Selection Model. The Review of Economic Studies |

Topic 10 : Bargaining and incomplete information


| | |
|---|---|
| 0.5892 | Drew Fudenberg, David K. Levine, Jean Tirole (1987). Incomplete Information Bargaining with Outside Opportunities. The Quarterly Journal of Economics |
| 0.5677 | Thomas H. Noe, Jun Wang (2004). Fooling All of the People Some of the Time: A Theory of Endogenous Sequencing in Confidential Negotiations. The Review of Economic Studies |
| 0.5369 | Peter C. Cramton (1984). Bargaining with Incomplete Information: An Infinite-Horizon Model with Two- Sided Uncertainty. The Review of Economic Studies |
| 0.5325 | Janet Neelin, Hugo Sonnenschein, Matthew Spiegel (1988). A Further Test of Noncooperative Bargaining Theory: Comment. The American Economic Review |
| 0.4976 | Peter C. Cramton (1992). Strategic Delay in Bargaining with Two-Sided Uncertainty. The Review of Economic Studies |

**Figure 2. Left panel, top: word weights ($\beta_{i4}$); bottom: top papers and their topic weights ($\theta_{4d}$) for Topic 4: "Adverse selection: lemons. Right panel similar information for Topic 10: "Bargaining and incomplete information".** The size of the font in the top of the panels is proportional to the word weight. The word clouds include the 100 top-weighted tokens.

The third step: we then use the document vectors making up matrix $\boldsymbol{\Theta}$ to visualize the content of documents. Figure 3 presents the topic weights ($\theta_{kLucas75}$) for a document (Lucas 1975) in our corpus. The large weights on the topics, "equilibrium stability…" and "growth models" show that the topic model's hypothetical generation of that paper would have drawn on these topics.

Comparing the topic weights in Lucas, 1975, with another article allows us to identify similarities or differences in their contents. Similarly, we can use topic weight comparisons to track the contents of the research corpus over time (treating all documents published in a time-slice as a single document) and compare these with the content of intermediate microeconomic and macroeconomic textbooks.

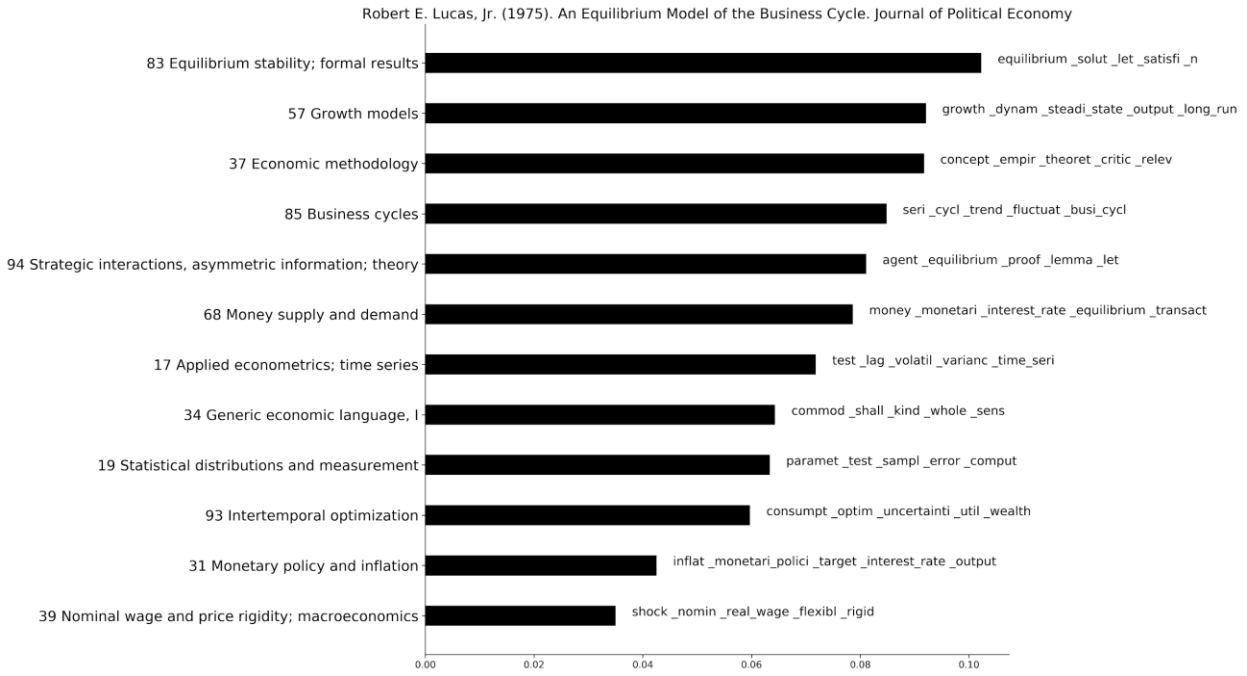


**Figure 3. Topic weights $\theta_{kLucas75}$ for the document Lucas (1975).** The words to the right of each bar are the tokens with the largest weight ($\beta_{ik}$) in the topic indicated.

Looking ahead, Figure 4 provides a road map for the remainder of the paper. In the bottom left outlined in red is the observed data – the corpus of research papers. The blue elements are the interventions by the researcher and the green ones are the computational steps. In this section, we have shown the path from the research corpus to the 100 named topics, involving both reasoning and computation. In the next sections we introduce respectively the intermediate-level microeconomics and macroeconomics textbooks and compare their content with that of the research corpus using topic weights.

But with 100 topics, it is a challenge to map the overall trajectory of the research corpus in a way that can be compared with the content of textbooks. For this reason, we used our reading of the evolution of economics over the past century along with the hundred topics, to identify what we call meta-topics. The meta-topics, three each in microeconomics and macroeconomics, are constructed by us from subsets of topics. This then allows a mapping of time-slices of the research corpus and textbooks into a simplex (one for microeconomics and one for macroeconomics), the vertices of which are labelled by the meta-topics. In the penultimate

section of the paper, we perturb the topic-composition of the meta-topics to test the sensitivity of our results to plausible minor redefinitions of the meta-topics. We begin by illustrating our method and providing some background.

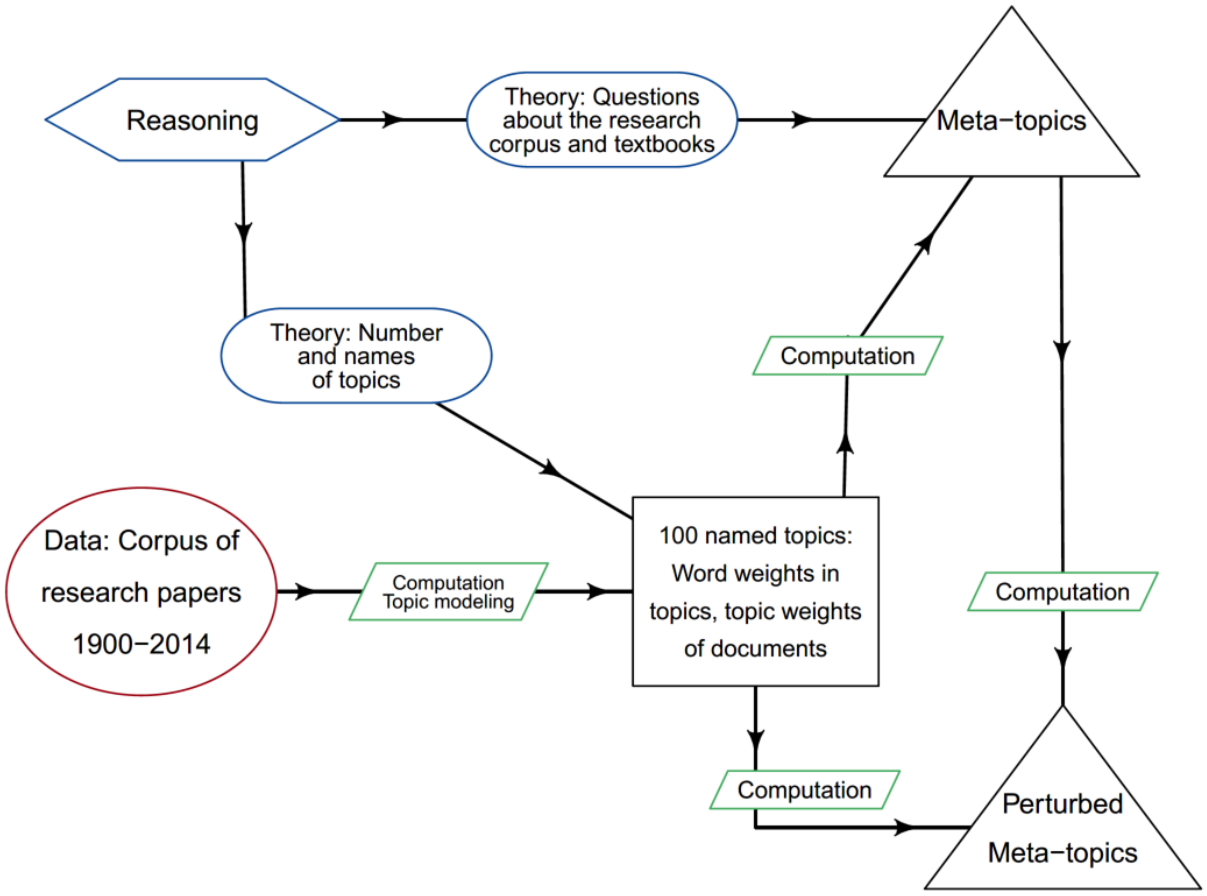


**Figure 4. Roadmap: from conceptual questions and research corpus to meta-topics and a robustness check.**

## II. The impact of the Great Depression on the research corpus and the textbooks

In Figure 5 we compare the topic weights for the 2489 research papers before 1925 (treated as a single document) and during the decade and a half following the end of the Second World War. The length of each outline bar measures the importance of that topic (strictly: the probability that it contributed to the corpus) for the early years (in the bars to the right of the vertical axis) and for the later years (in the bars to the left.) The solid bars show the difference in the weight on the

topic between the two corpora. Large solid bars to the right show a greater weight in the early corpus than in the later corpus, and vice versa.

The black bars to the right show that the earlier corpus was far more empirical and institutional than the later one with larger topic weights for example, on "named organizations…", "public regulation," "banking; institutions," "business entrepreneurship and organization" and "economic history." The distinctive characteristics of the later corpus (the grey bars to the left of 0) include what later came to be called "macroeconomics" ("fluctuations in aggregate demand", "aggregate demand, consumption" and "business cycles") and microeconomic theory ("elasticity of demand and supply", "utility functions", "production functions" and "competition and market structure.") Figure 5 documents a substantial change in the content of the research corpus: only a single topic "empirical studies of industry" has substantial weight in both corpora.

Among economists, the dominant work in the early 20th century was Alfred Marshall's *Principles of Economics*, first published in 1890 but not widely used as an undergraduate textbook. We therefore explore how the shift in focus in the research corpus was reflected in what undergraduates learned by looking at general economics textbooks for the early period. Throughout most of the 20th century prior to the Second World War, the US market was dominated by a work by Richard T. Ely (and a series of coauthors), *Outlines of Economics*, written at the same time as Marshall's work and published first in 1893 (Ely et al. 1930, Backhouse, Bateman, and Medema 2010).

Not surprisingly the Great Depression and the publication of Keynes' *General Theory* stimulated a number of new introductory textbooks, including Lorie Tarshis' initially very successful *Elements of Economics* (Tarshis 1947). But these, along with Ely et. al., were quickly eclipsed by Paul Samuelson's *Economics, an Introductory Analysis* (Samuelson 1948).

Figure 6 shows the topic weights for the two textbooks. The topics "banking: institutions" "empirical studies of industry" and "income tax: institutional" are weighted heavily in both textbooks. But Ely's book places more weight on the topics characteristic of the research corpus earlier in the century, for example "public regulation" and "business entrepreneurship and organization" that have lesser weights mid-century.

Samuelson narrowed the gap between the text and the contemporaneous research corpus, as can be seen in the greater emphasis on the macroeconomic topics "fluctuations in aggregate demand" and "aggregate demand: consumption" as well as "elasticity of demand and supply" and "market structure and competition", all of which were prominent in the mid-century research corpus, but not earlier.

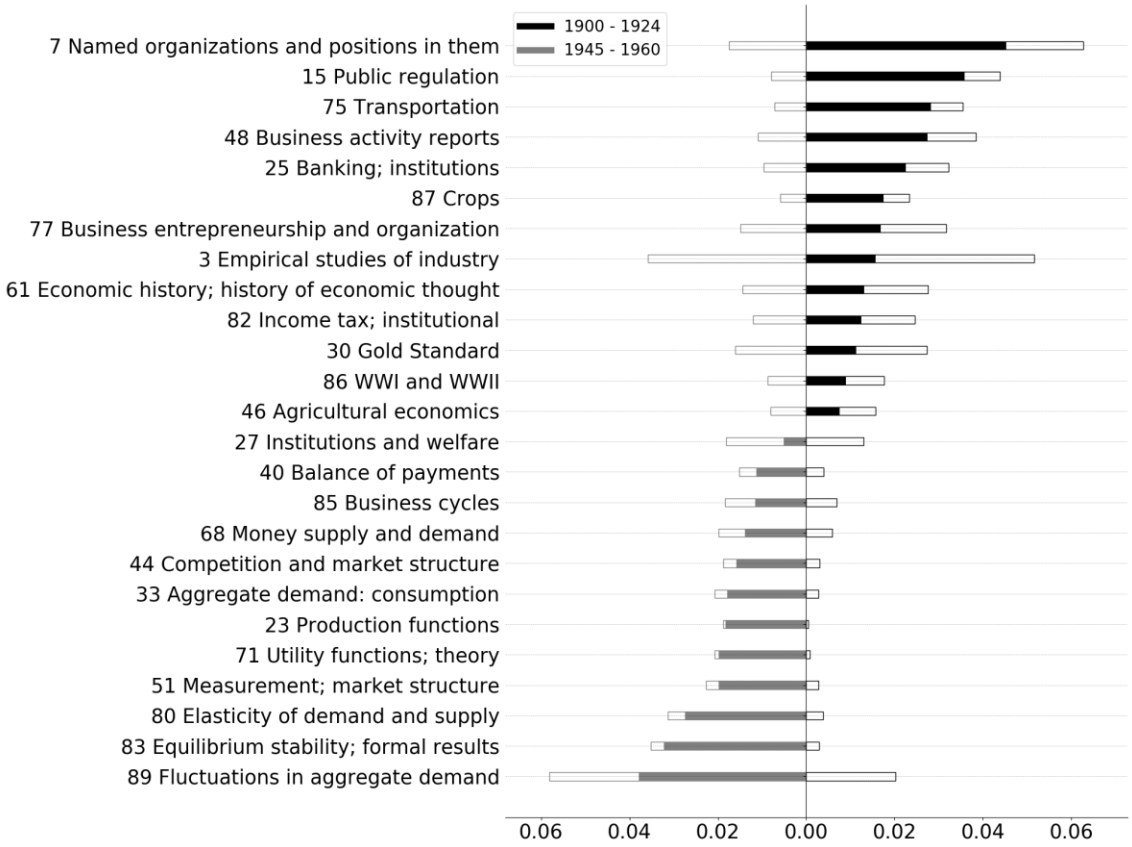


**Figure 5. A shift in the research corpus: Topic weights $\theta_{kd}$ of the research corpora of the early in the 20th century and mid-century.** In this and subsequent similar figures, a topic is excluded if it has a weight less than 0.015 in both of the textbooks or the token with the greatest weight is less than 0.01. Each outline bar is the topic weight $\theta_{kd}$ for the earlier research corpus (in the bars to the right of the vertical axis) and the later corpus (in the bars to the left.) The solid bars show the difference in the topic weight between the two corpora. The topics are ordered by the difference in topic weight between the two periods.

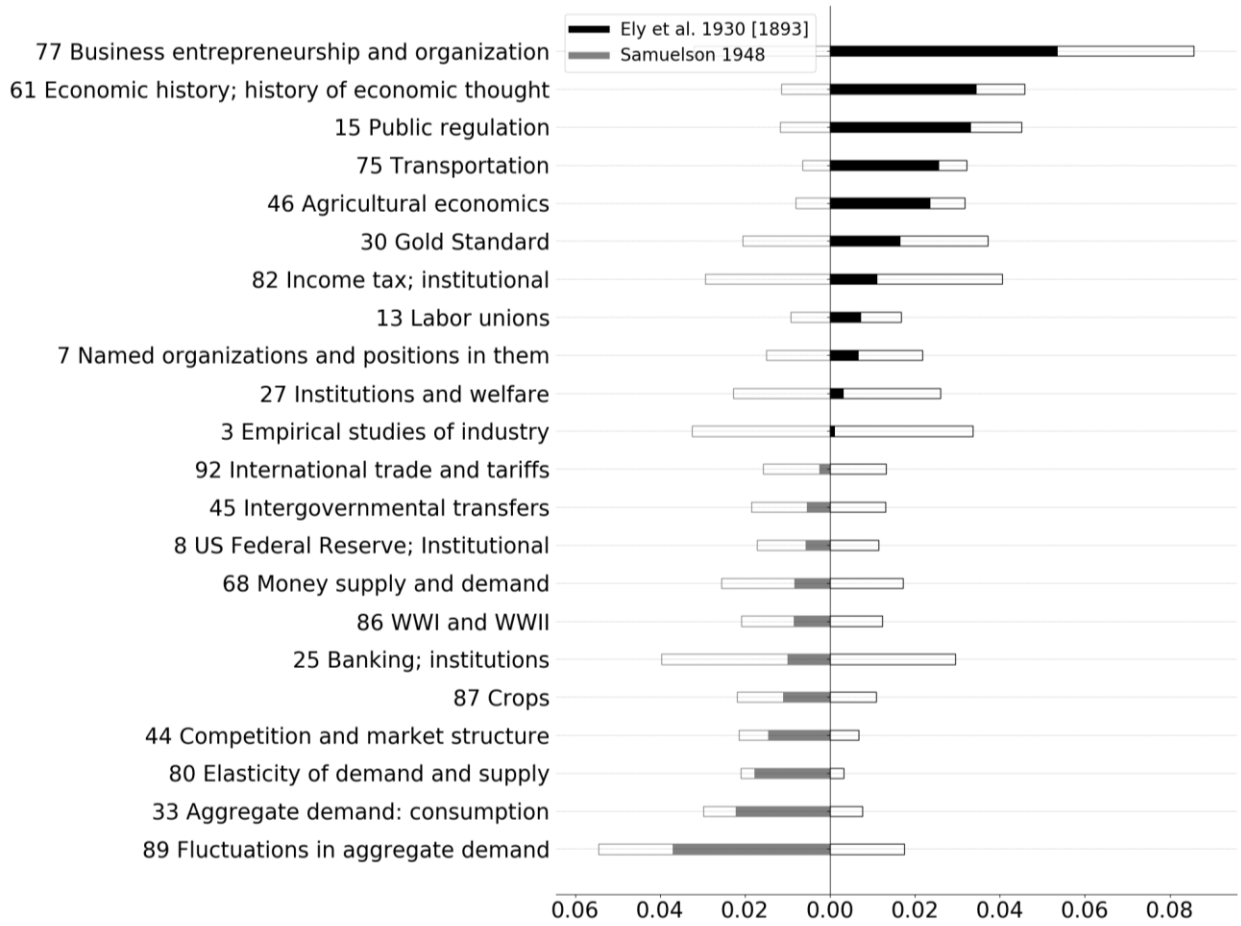


**Figure 6. Samuelson 1948 replaces Ely et al. 1930.** As before, the length of each outline bar is the topic weight $\theta_{kd}$ for Ely et al. (in the bars to the right of the vertical axis) and for Samuelson, 1948 (in the bars to the left.) The solid bars show the difference in the topic weight between the two corpora. The topics are ordered by the difference in topic weight between the two textbooks.

George Stigler's *The Theory of Price* (1942) was the first intermediate-level textbook to make prominent these 'Marshallian' microeconomic topics, as well as "utility functions" and "production functions" (Stigler 1942). In this, like Samuelson, Stigler dramatically reduced the difference between the conceptual frameworks research economists were using and what economics majors were learning.[5]

[5] We performed a similar analysis of Abba Lerner's *The Economics of Control: Principles of Welfare Economics,* a text suitable for second-year economics published two years after Stigler's *The Theory of Price* (Lerner 1944). With Lerner's few macroeconomic chapters removed, the topic weights in the Lerner and Stigler textbooks is virtually identical.

## III. Research and intermediate microeconomics textbooks since World War II

The research corpus, of course, continued changing, due to more extensive use of mathematics, the moderation of the business cycle and other influences. Figure 7 charts the shifting topic weights from the early post-war period to the present.

Notably less heavily weighted in the research corpus of the recent period are the macroeconomic topics that rose to prominence in the aftermath of the Great Depression, along with the Marshallian microeconomics topics ("competition and market structure", "elasticity of demand and supply"). Also reflecting a shift away from Walras-Marshall style microeconomics, the recent research corpus is heavily weighted on strategic interactions under incomplete information ("strategic interactions, asymmetric information," "game theory and behavioral economics," and "equilibrium signaling..."). The topic weights in the recent corpus also reflect a revival of interest in empirical studies (Angrist et al. 2017), but now using econometric and experimental rather than the largely descriptive methods of the early 20th century ("applied econometrics..." "distributions and measurement," and "experimental design").

Even more than was the case in Figure 5 (comparing research in the early and middle parts of the 20th century), what is striking about the comparison in Figure 7 is how little of the mid-century corpus is retained in 21st century research. The only topic with substantial weight in in both periods shown in Figure 8 is "equilibrium stability…". The topics with greater weights at mid-century ("fluctuations in aggregate demand," "empirical studies of industry," and "elasticity of demand and supply") are almost entirely absent in the recent corpus. And correspondingly the newly prominent topics ("strategic interactions…" and "applied econometrics…") have very small weights in the mid-century corpus.

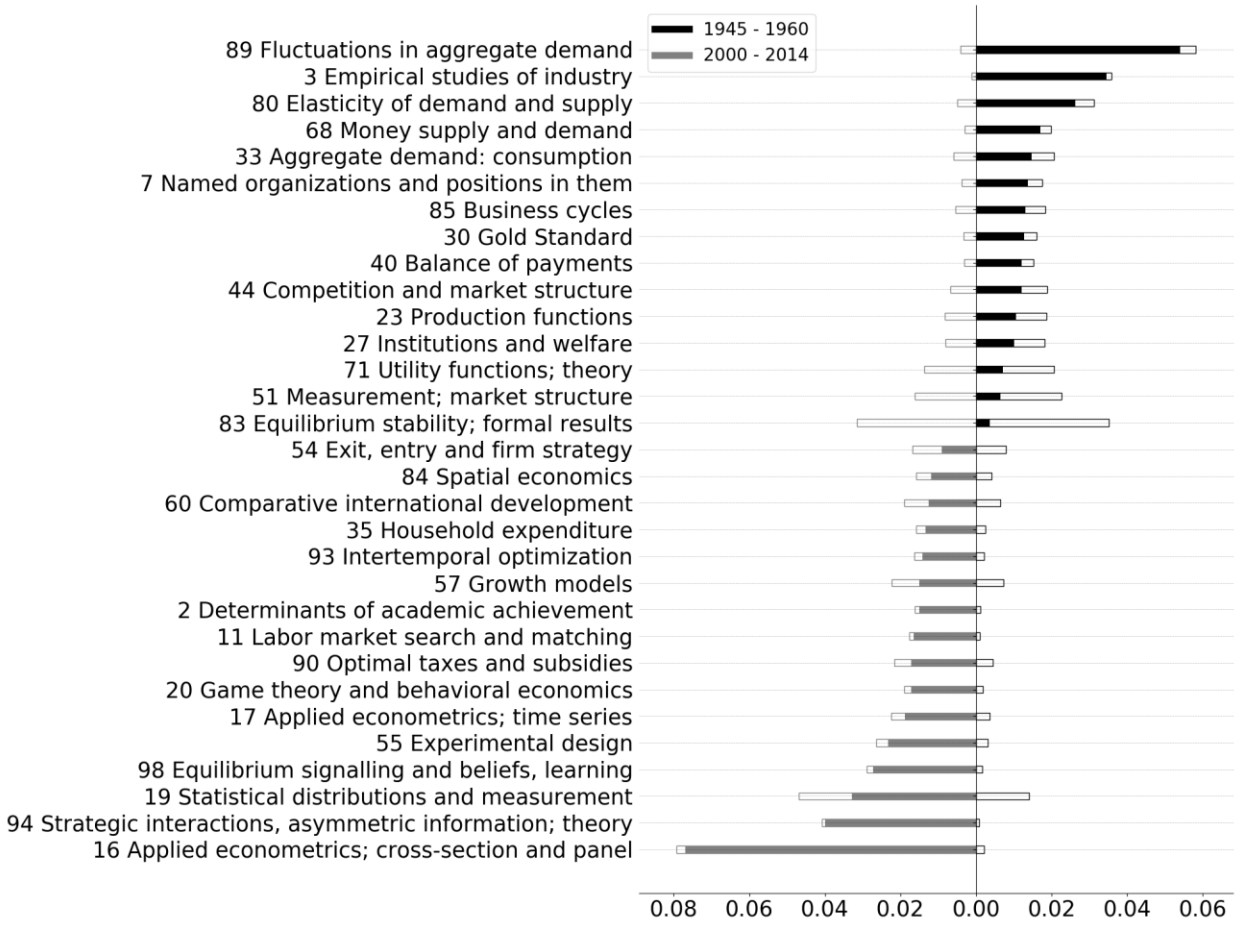


**Figure 7. A shift in the research corpus: from the mid-20th century to the early 21st century**
As before, the length of each outline bar measures the topic weight $\theta_{kd}$ for the early corpus (in the bars to the right of the vertical axis) and for the later corpus (in the bars to the left.) The topics are ordered by the difference in topic weight between the two periods.

A measure of conservation of content is the sum of topic weights in the newer corpus where these are less than the topic weights for the older corpus (the length of the entirely open bars to the left of the vertical axis at the top of the figure). By this measure conservation of older content is 43 per cent greater in mid-century than in the early 21st century corpus. A measure of discovery (or novelty) is the sum of the (absolute value of the) difference in the topic weights for the topics more highly weighted in the recent period (the lengths of the grey bars to the left of the axis at the bottom of the figure). By this measure the early 21st century corpus is 56 per cent more novel than is the mid-twentieth century corpus.

By contrast, what second-year students are learning appears to have changed remarkably little. In Figures 8 and 9 we compare the topic weights of the mid-century Stigler with the leading contemporary textbooks by Pindyck and Rubinfeld, and Varian. The weights on "elasticity of

demand …" and "competition and market structure," by far the largest in Stigler, remain dominant in the recent textbooks. Nevertheless, the two recent textbooks themselves differ in important ways. From the substantial topic weight on "game theory and behavioral economics," and the lesser but nonetheless important weight on "intertemporal optimization", Varian appears to be closer to the research corpus of the early 21st century.

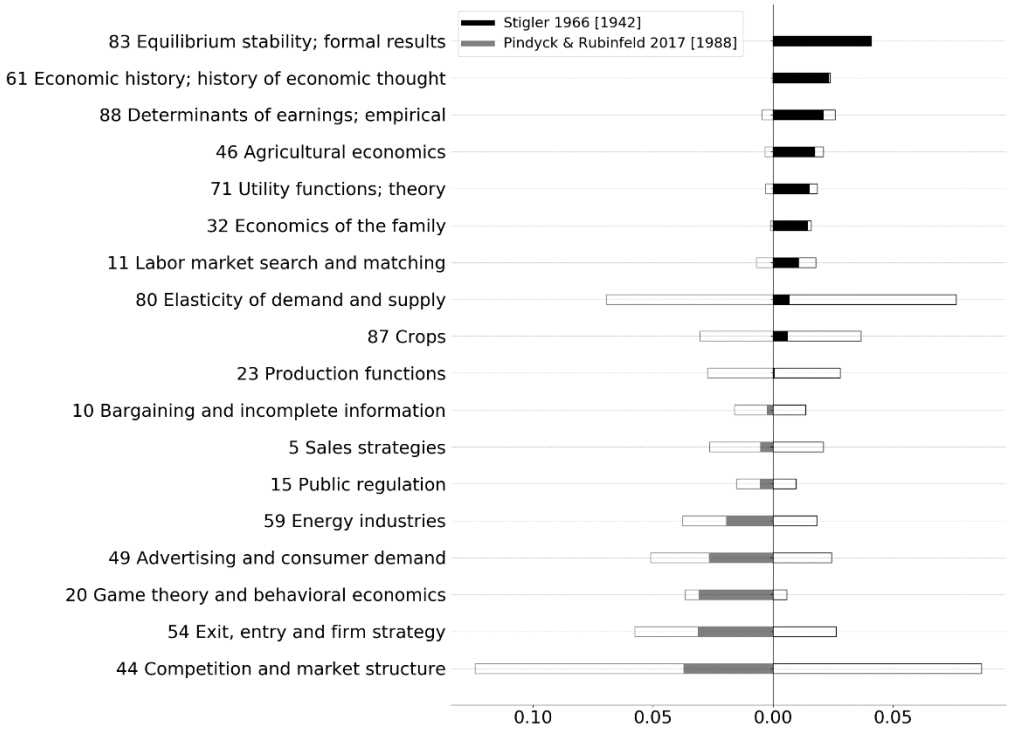


**Figure 8. Conservation and novelty in microeconomics textbooks since mid-20th century: Topic weights Stigler and Pindyck & Rubinfeld.** The topics are ordered by the difference in topic weight between the two textbooks.

## IV. Meta-topics in microeconomics: Co-evolution of research and textbooks

These applications show that topic modeling provides a lens with which to analyze content. Yet with the 100 topics extracted from the research corpus, it is difficult to summarize the changes over time in research and in the content of textbooks. To provide a more synthetic picture of the whole and a method of visualizing trajectories, we identified what we call meta-topics based on the questions we were seeking to answer in light of our own reading of the evolution of economics over the past century.

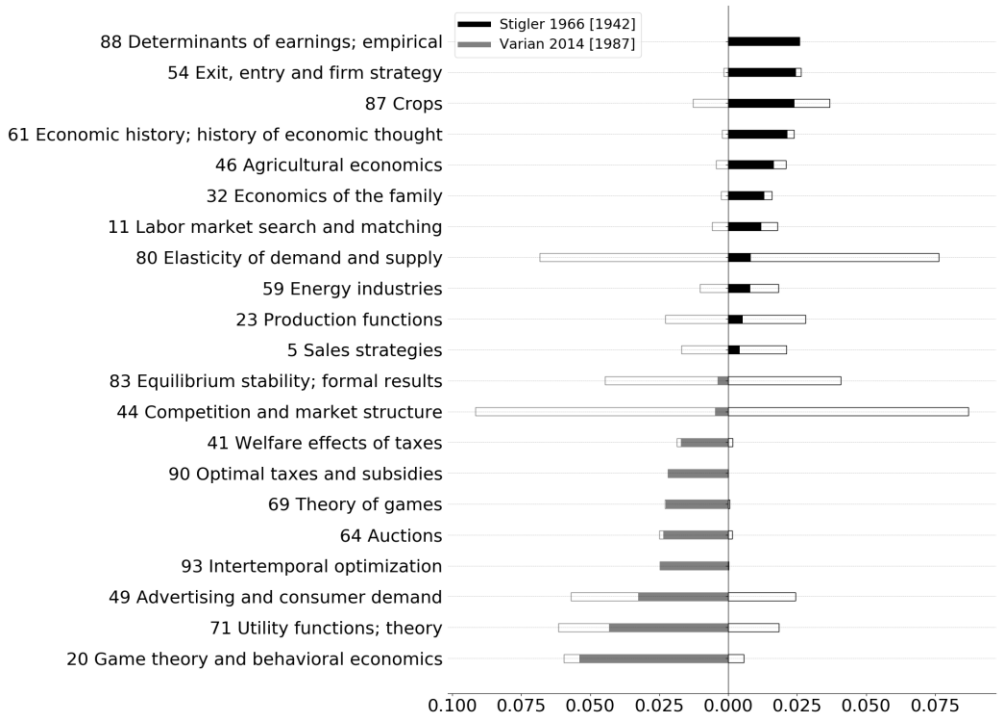


**Figure 9. Conservation and novelty in microeconomics textbooks since mid-20th century: Stigler and Varian**

The topics used in the separate analysis of micro and macro textbooks, and in comparing the movement of the research corpus over the past century or so are the single set of 100 topics extracted from the entire research corpus. A meta-topic is the average of a subset of the K =100 topics selected by us to reflect the themes making up the meta-topic. A microeconomics meta-topic is defined as a vector $\bar{\beta}_k (k \in [1,3])$, the word weights of which are the sum of the word weights of the topics drawn (from our full set of K = 100 topics) in the set of topics constituting the meta-topic, divided by the number of topics in that set. So, we have the three microeconomic meta-topic word weights:

$$\bar{\beta}_{ik} \text{ with } \sum_{i=1}^{N} \beta_{ik} = 1 \text{ and } k \in [1,3]$$

The flowchart in Figure 4 shows where researcher and computation intervene in the construction of the meta-topics. Our interventions, circled in blue in the figure, were based on informed reasoning of what topics should constitute a meta-topic. We check our conception of the meta-topics for robustness later by simulating alternative definitions of each meta-topic.

In Figure 10 we show the topics making up each of the three microeconomics meta-topics, and by size of font in the word clouds, the 100 top token weights in each. We chose these three meta-topics because they capture differences in theorizing and in the dimensions of an appropriate educational benchmark. Meta-topic 0 covers Marshallian aspects of market structure and competition; Meta-topic 1 is about individual optimization and expected utility; and Meta-topic 2 encompasses strategic interaction, incomplete information and behavioral economics.

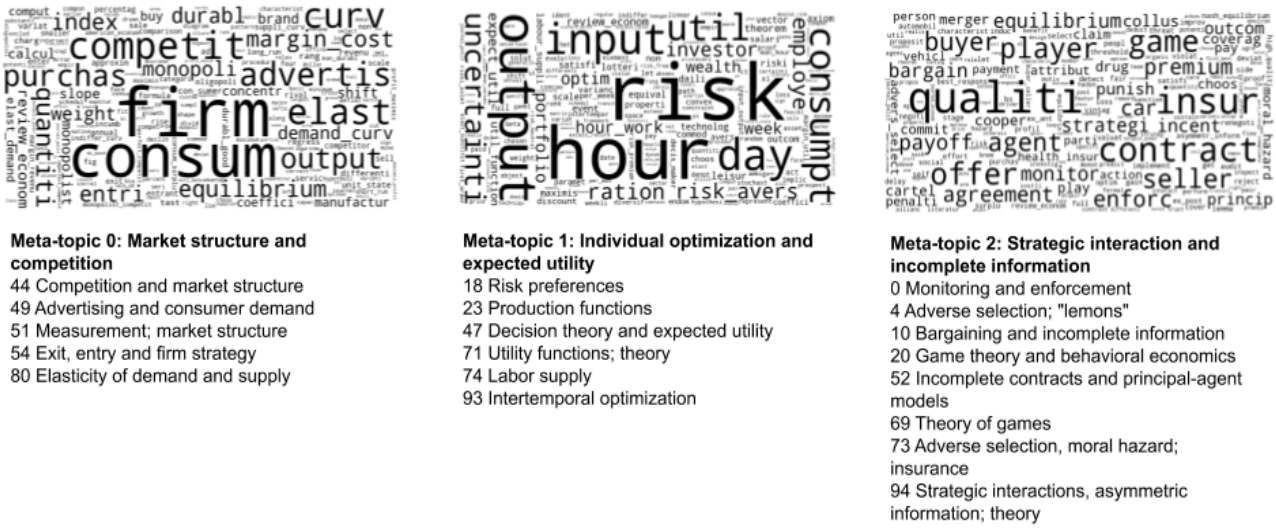


**Figure 10. Microeconomic meta-topics and token weights $\bar{\beta}_{ik}$.** The topics making up the meta-topic are listed below the name we have given to the meta-topic. The size of the font in the word clouds is proportional to the token weights in the meta-topic.

In Figure 11 the colored dots locate the meta-topic weights (normalized to sum to one) for the time slices of the research corpus from 1900 to the present (shown in the legend on the left of the figure), connected by arrows from an earlier to a later period. The early 20th century ascendency of Marshallian economics is evident in the increased weight on the meta-topic "market structure and competition." But by mid-century, research had moved in the direction of "individual optimization …", a process that continued until the mid-1980s. The final three time slices starting in 1985 have seen a substantial increase in the weight on "strategic interaction and incomplete information," which rose from just over 10 percent to over 40 percent in the most recent time period, making it the most heavily weighted microeconomics meta-topic.

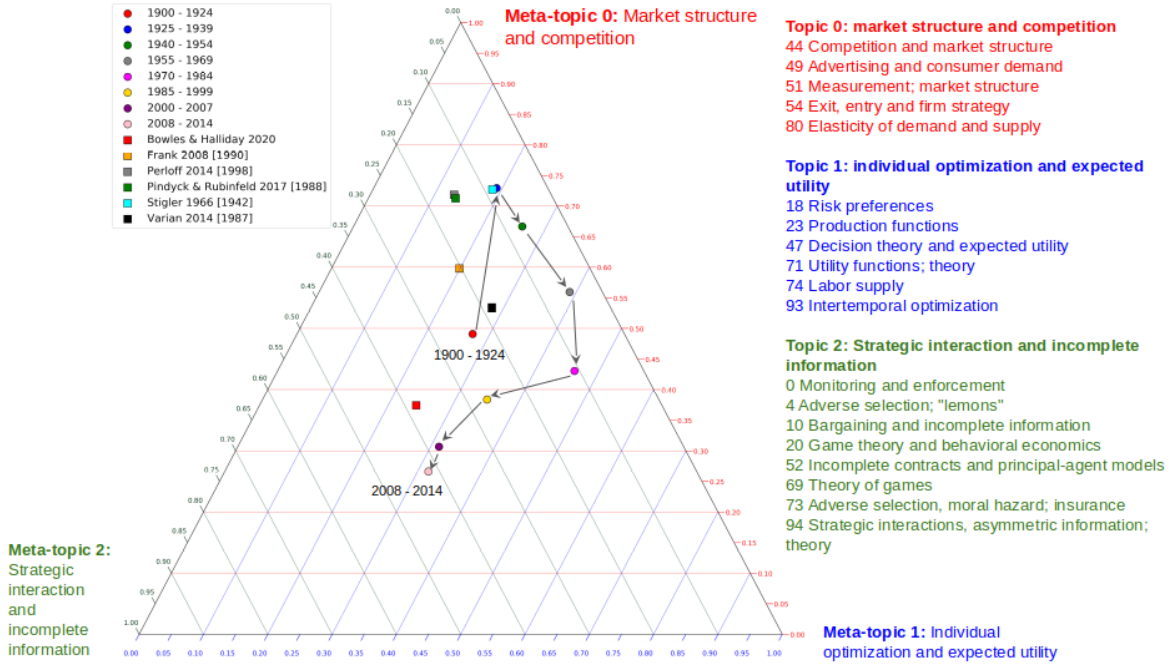


**Figure 11. Microeconomics: The evolution of the research corpus and the textbooks.** Coordinates of a point in the simplex sum to one and give the distribution of meta-topic weights for that point. The initial red dot (the research corpus prior to 1925) indicates a 0.49 weight on Meta-topic 0, 0.27 on Meta-topic 1 and 0.24 on Meta-topic 2

In the same simplex we also locate the microeconomics textbooks we have mentioned, and include two additional ones – Perloff and Frank. Three of them (Stigler, Pindyck and Rubinfeld, and Perloff) are close to the Meta-topic 0 vertex, all with a weight of more than seventy percent on "market structure and competition" more or less exactly where the research corpus was between 1925 and 1939. Varian, however, and to a lesser extent Frank, more closely map the initial movement of the research trajectory after 1940 away from Meta-topic 0 (toward Meta-topic 1 – individual optimization) though Meta-topic 0's weight in Varian is still greater than one half.[6]

We have also included in the simplex the meta-topic distribution of a new intermediate microeconomics textbook written by two of the authors of this paper (Bowles and Halliday 2020). As with the other textbooks, there are substantial weights on "elasticity of demand and

[6] Frequently used in PhD microeconomics courses, the textbook by Mas-Colell, Whinston and Green – like any other text – can be located in the simplex (Mas-Colell, Whinston, and Green 1995). It is closer to Meta-topic 1 (Individual optimization and expected utility) with a weight of 0.55 than either the closest textbook (Varian with a weight of 0.28) or the research corpus in any time-slice. And it is much more distant from Topic 0 with a weight of 0.20 than any of the intermediate textbooks or the contemporary research corpus.

supply", "utility functions; theory" and "competition and market structure". However, like Varian, this new work follows the research corpus away from "market structure and competition" as the preeminent meta-topic for intermediate microeconomics teaching. But in contrast to Varian, Meta-topic 2 "strategic interaction…" is more heavily weighted than "individual optimization and expected utility". We can see in Figure 12 that the difference from Varian is especially pronounced in the topics making up this meta-topic, especially "game theory and behavioral economics" but also "theory of games," "incomplete contracts …" and "bargaining and incomplete information."

## V. Meta-topics in macroeconomics: co-evolution of research and textbooks

We constructed meta-topics in macroeconomics in the same way as the microeconomics ones; the topics making up the meta-topics and the meta-topic word weights appear in Figure 13. We chose three themes in macroeconomics that span research over the century: Meta-topic 0 is aggregate demand including business cycles and fiscal and monetary policy; Meta-topic 1 is economic growth; and Meta-topic 2 comprises the supply-side components of the aggregate economy and of macroeconomic models – labor markets, credit markets and financial markets.

The trajectory of the research corpus for the macroeconomics meta-topics shown in Figure 14 is remarkably similar to that for microeconomics. Starting from the time-slice before the Great Depression (1900-1924) there is an initial move in the research corpus towards Meta-topic 0, which we interpret as conventional Keynesian aggregate demand-based themes, including associated fiscal and monetary policies for cyclical stabilization. The peak weights at just over 60 percent on Meta-topic 0 for the research corpus are for the period 1925-54. This coincides with the first use of the term macroeconomics in 1933 attributed to Ragnar Frisch (Hoover and DeVroey 2005) when working with Jan Tinbergen to build statistically estimated models of the aggregate economy.

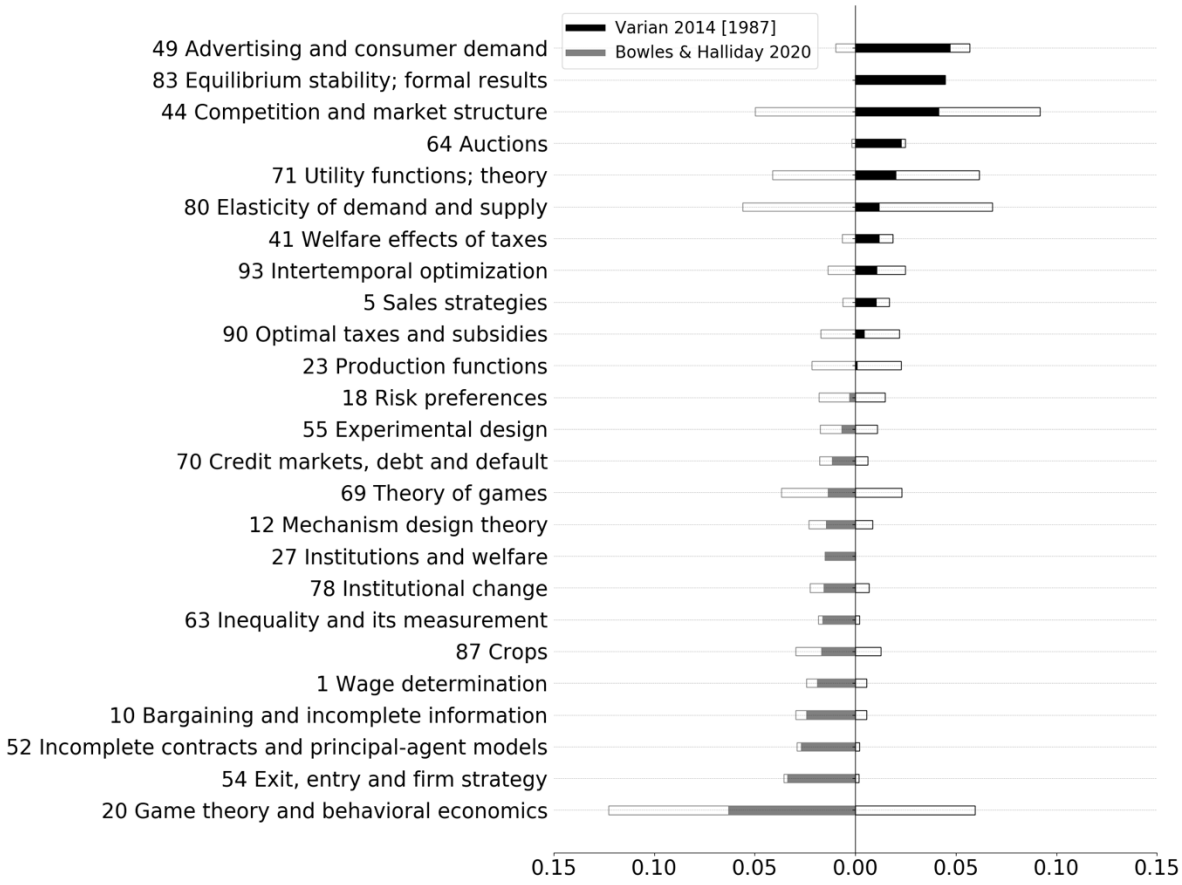


**Figure 12. Conservation and novelty in microeconomics textbooks since mid-century: Varian and Bowles & Halliday**

But as the US, Western Europe and Japan experienced historically unprecedented growth rates of per capita GDP during the decades following the Second World War rather than another depression, this emphasis on what came to be called Keynesian economics (Meta-topic 0) was replaced by a focus on models of growth (Meta-topic 1). A role may also have been played by the concern that successful growth of living standards in market economies of the Third World would be important in winning the Cold War.

From the mid-1980s in the aftermath of the Great Stagflation, the drift away from Meta-topic 0 (aggregate demand) continued and we see a somewhat increased emphasis in macroeconomics on the supply-side components of the aggregate economy – Meta-topic 2: labor, credit and financial markets, which is now more heavily weighted (over 30 percent) than Meta-topic 0 in the research corpus (less than 20 percent). The sustained weight of around one half on the growth meta-topic from 1970 to the present reflects new empirical techniques, endogenous and

Schumpeterian growth models, and microeconomic datasets deployed to investigate the role in growth of market structure and other institutions.

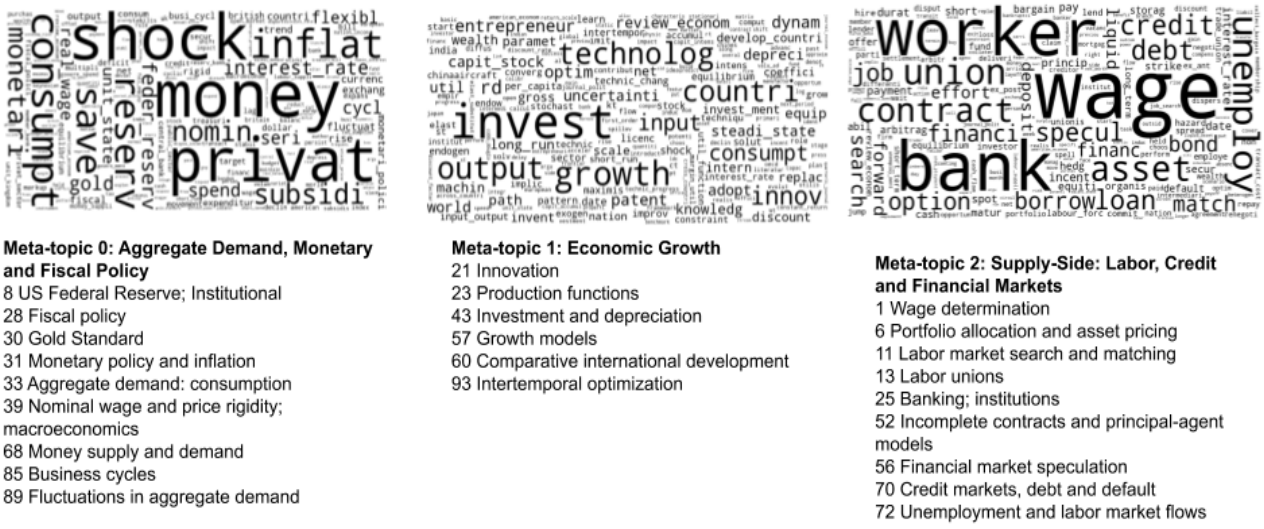


**Figure 13. Macroeconomic meta-topics and their token weights.** The topics making up the meta-topic are listed below the name we have given to the meta-topic. The size of the font in the word cloud is proportional to the token weights in the meta-topic.

In economics education through the 1950s, macroeconomic topics were subsumed in courses on money, and monetary and fiscal policy, as well as in public economics courses. Hansen's 1953 textbook brought Keynes' *General Theory* and, in particular, Hicks' IS-LM interpretation of Keynes' model, into the new intermediate-level macroeconomics classroom. Like Stigler's *The Theory of Price* in microeconomics, Hansen's *A Guide to Keynes* (1953) is located close to the research corpus and to Meta-topic 0 whose weight for Hansen is over 75 percent.

The modern macroeconomics textbooks reflect the movement of the research corpus away from aggregate demand and toward growth. Three of the most popular contemporary intermediate macroeconomics textbooks are placed in the simplex: Mankiw (with weights on Meta-topics 1-3 of 0.52,0.33,0.16), Blanchard (0.41,0.39,0.20), and Jones (0.32,0.56,0.12). All three retain a greater emphasis on the aggregate demand topic than was the case for the research corpus from 1970 onwards. Because of its emphasis on growth models, Jones is closest to the 1970-84 research corpus.

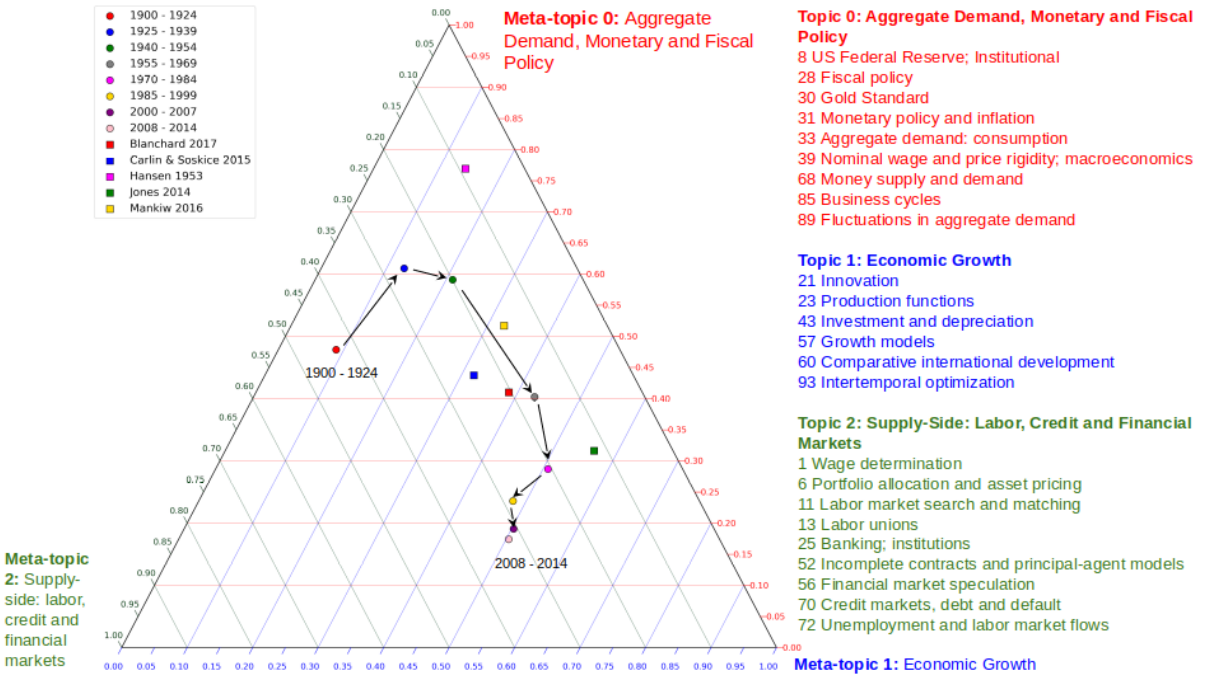


**Figure 14. Macroeconomics: The evolution of the research corpus and the textbooks.** Coordinates of a point in the simplex sum to one and give the distribution of meta-topic weights for that point. The initial red dot (the research corpus prior to 1925) indicates a weight of 0.48 in topic 0, 0.09 in topic 1 and 0.43 in topic 2.

A fourth intermediate textbook co-authored by one of the authors of this paper is also shown in the simplex. Carlin and Soskice published macroeconomics textbooks in 1990, 2006 and 2015, choosing to publish new books rather than new editions. The most recent is closest to Blanchard in the simplex, and a bit closer than the other three to Meta-topic 2, towards which the research corpus has moved somewhat since 1985.

Because we have focused on three dimensions of both microeconomic and macroeconomic theory in constructing the simplexes, there are major themes in the literature that we have chosen not to investigate and that as a result are not captured by our meta-topics. Of the total of a hundred topics, forty-three are used in constructing our six meta-topics (19 in microeconomics, 24 in macroeconomics). If there are major changes in the weights of the 57 topics that do not appear in our six meta-topics, then substantial changes in the research corpus could occur without generating movements in our two simplexes (Figures 11 and 14).

To check that our two sets of meta-topics are measuring a substantial part of the research corpus, we show in Figure 15 the total weights of the set of topics making up our microeconomic and

macroeconomics meta-topics. We also show the total weights on a new meta-topic that we have constructed: econometrics, the constituent topics of which are shown in the figure caption. Panels A and B of the figure show that the sets of topics comprising our meta-topics constitute a significant fraction of the total topic weights, excepting the microeconomic topics in the first period, when the research corpus was highly descriptive (see Figure 5). Panel C shows the rising importance of econometrics, as expected from previous research (Angrist et al. 2017, Hamermesh 2013). Panel D shows the weight in the research corpus of a single topic "equilibrium stability: formal results", an indication of the importance (albeit declining over the past four decades) of a theoretical topic that is not part of our set of meta-topics.

## VI. Macroeconomics since the financial crisis

The global financial crisis of 2008 highlighted the neglect by macroeconomists of the potential for bank-lending-based financial instability originating in the United States. In its aftermath, the lessons of the Great Depression also returned to prominence, which motivated the rapid policy intervention to bail out banks and apply both monetary and fiscal policy to stabilize aggregate demand. In consequence, after the crisis one might have expected a movement in the research trajectory to greater weights on both Meta-topic 0 (Aggregate demand) and Meta-topic 1 (Supply side: labor, credit, and financial markets). It is striking that the distribution of the research corpus up to 2014 within the macroeconomics simplex was unmoved either by the Great Moderation from the early 1990s to 2007 or by the 2008 financial crisis (as can be seen from the virtually identical coordinates of the three last colored dots in the research corpus trajectory in Figure 14).[7]

Although the macroeconomics textbooks we consider added some coverage of the financial crisis, their location in our conceptual meta-topic space changed little (Figure 16). Moreover, what did change appears to have little connection with the causes and policy responses to the financial crisis. Appendix Figures A.3.1 and A.3.2 compare topic weights in textbooks by the

[7] Given the publication lags in economics, we split the last period's data into two to see if a shift in the location of research had occurred by the years 2013-14. Contrary to the expectation that the trajectory might have moved toward Meta-topics 0 and 2, the movement was in the opposite direction, further toward Economic growth (Meta-topic 1). A closer look at the 100 topic weights in the most recent research corpus underscores this finding.

same author before and after 2008. The weights of topics that would seemingly be most likely to be given more attention in the aftermath of the crisis – "credit markets, debt and default" and "banking institutions" – remain for the most part unchanged, suggesting that unlike the Great Depression, the financial crisis did not lead to a rethinking of "what an economist should know".

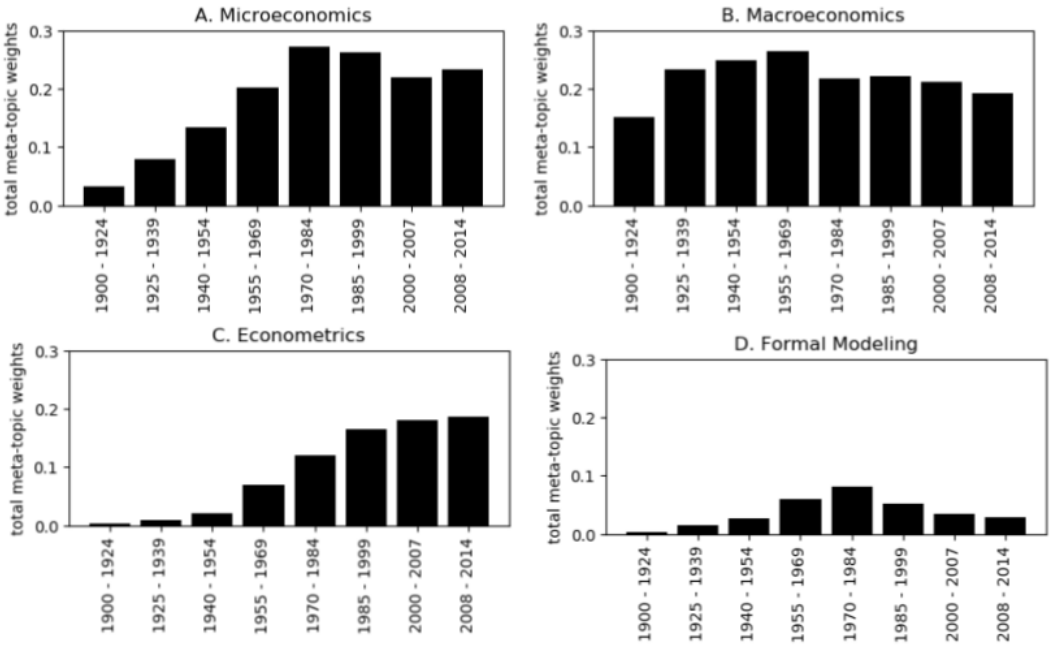


**Figure 15. Composition of the research corpus by topic** A. Microeconomics B. Macroeconomics, C. Econometrics (comprising 16 applied econometrics: cross-section and panel; 17 applied econometrics time series; 19 statistical distributions and measurement; 22 forecasting; 55 experimental design) and D. Formal modeling (83 equilibrium stability; formal results).

However, our topic modeling does not pick up major substantive changes in the macroeconomics textbooks that did occur at this time. Most notably, although Olivier Blanchard's textbook registers a small change in the topic-based measure related to the financial crisis, he

implemented a major pedagogical shift by dropping both the money-supply based LM curve, and the entire AD/AS apparatus (Blanchard 2016).[8]

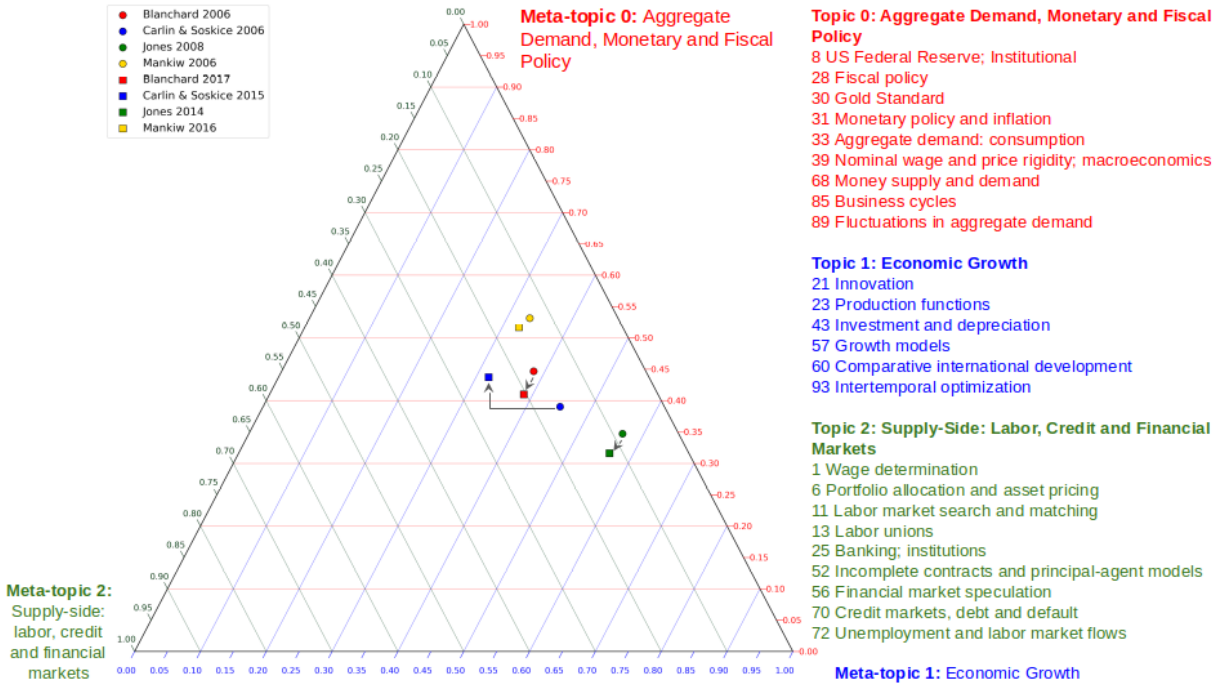


**Figure 16. Macroeconomics textbooks: What difference did the financial crisis make?**

## VII. Visualizing the uncertainty of our estimates

Uncertainty in our estimates could arise not only from such standard sources as sampling error, but also from variations in how a researcher would choose to categorize the subject matter of one or more of the meta-topics. The meta-topics we have used and the topics selected as constituting them are not generated by the data; they are determined by the set of questions we wish to ask, and they reflect our understanding of the approaches and fields of study represented by our meta-topic names (see Figure 4). By "economic growth" (macroeconomic Meta-topic 1), for example, we mean the body of work that draws upon the topics "production functions," "innovation,"

[8] The macroeconomic modeling in Carlin and Soskice (2006 and 2015) does not use either a money-supply based LM curve, or the AD/AS apparatus.

"intertemporal optimization," and the others listed as constituting that meta-topic. The list of subtopics is what we mean by the meta-topic.

But would a somewhat different conception of the meta-topics – omitting the topic "innovation" from the meta-topic "economic growth," and replacing it by the topic "exit, entry, and firm strategy," for example – yield a substantially different set of results? To answer this question, we have explored how small perturbations of the composition of the meta-topics would affect the topic weights and hence locations in the simplexes for each of the textbooks and for the historical periods of the research corpus.

For each set of meta-topics in the microeconomics and macroeconomics applications, we generate an alternative set of meta-topic weights in two steps. First where $n$ is the number of topics constituting the set of microeconomic or macroeconomic meta-topics ($n$ = 19 for micro, $n$ = 24 for macro), we eliminate each topic it the set of micro or macro meta-topics with probability $1/n$. Thus, on average one of the topics making up the set of meta-topics will be eliminated, but with a small probability a substantial number will be excluded.

Second, we then replace each excluded topic with the topic, from among those not already included in the meta-topic of which the excluded topic was a member, whose topic weight ($\theta_{kd}$) is most highly correlated with the excluded topic across the corpus of 27,436 research papers since 1900. For example, if "innovation" were randomly selected to be excluded, we would then select from the 95 topics that are not members of the "economic growth" meta-topic the one most correlated with "innovation." By this procedure, the topic "exit, entry, and firm strategy" replaces "innovation" in the "economic growth" meta-topic. Appendix figure [to come] to see the replacement topics selected by this procedure. We replace any additional randomly eliminated topics by the same procedure, giving us a new perturbed set of meta topics. We implement this process 1,000 times, giving us a large set of alternative perturbed meta-topic weights that might have produced the textbooks and documents from the research corpus in each period.

The uncertainty that we explore using this method, arises not from statistical errors but from hypothetical conceptual differences. How demanding an assessment of the uncertainty

surrounding our estimates one thinks our procedure delivers depends on the extent of variations in the substantive conceptual content of the meta-topics that one is prepared to entertain.

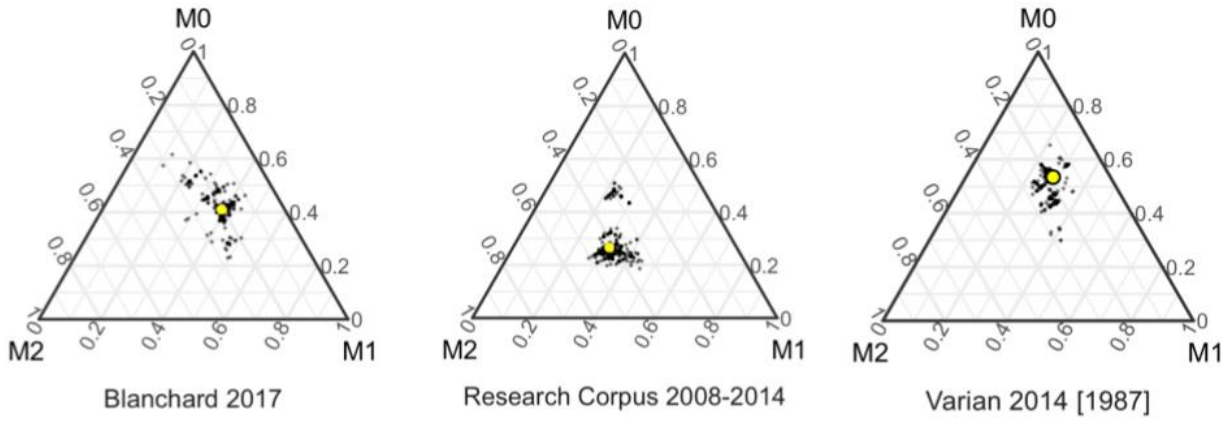


**Figure 17. Perturbed estimates of the location in meta-topic space of two textbooks and a time slice of the research corpus**. Each of the one thousand black dots in each simplex represents the location of the text or time slice under the perturbation process described in the text: Blanchard in the macroeconomics meta-topic space (left panel), Varian in the microeconomics meta-topic space (right panel) or research corpus time slice in the micro-economic meta-topic space (middle panel). The yellow dots are the unperturbed case. In the left panel: M0 - Aggregate demand, monetary and fiscal policy; M1 – Economic growth; M2 - Supply-Side: labor, credit, and financial markets. In the right panel: M0 - Market structure and competition; M1 - Individual optimization and expected utility; M2 - Strategic interaction and incomplete information

The three panels of Figure 17 illustrate our perturbation procedure by showing the distribution in a simplex of the resulting 1,000 sets of topic weights for Varian, Blanchard, and the research corpus in the most recent period. Each of the data points in one of these simplexes represents an alternative set of meta-topic weights for the textbook or time slice of the research corpus that would have resulted had the investigator entertained a conception of the topics constituting the meta-topics that is somewhat different from the ones we proposed.

From these data we then use Mahalanobis distances to construct 95 percent confidence regions based on these perturbed meta-topics.[9] For a point in the simplex with coordinates summing to 1,

[9] The Mahalanobis distance is the Euclidian distance normalized using the variance-covariance matrix of the perturbed data (Mahalanobis 1936, Maesschalck, Jouan-Rimbaud, and Massart 2000). We used the ggtern package in R to create the simplex plots and the confidence regions, based on a log ratio transformation (Hamilton and Ferry 2018).

$\{\theta_0^M, \theta_1^M, \theta_2^M\}$, the confidence level associated with some textbook (e.g. Varian) gives the degree of certainty that we can assign to the hypothesis that the generative process given by $\{\theta_0^M, \theta_1^M, \theta_2^M\}$ did not write Varian.

The results for the movement of the research corpus and for the leading textbooks in macroeconomics are shown in Figure 18, and the corresponding results for microeconomics in Figure 19.

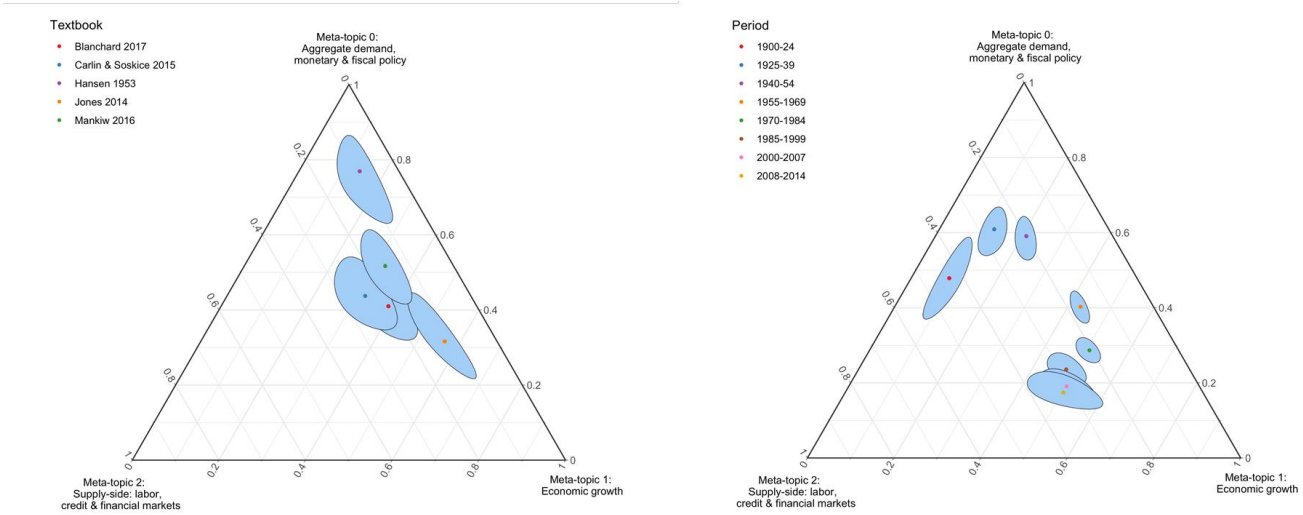


**Figure 18. Uncertainty in the macroeconomics estimates**: Ninety-five percent confidence regions for textbooks (left panel) and time slices of the research corpus (right panel). The dots in the confidence regions are the estimates for the unperturbed case.

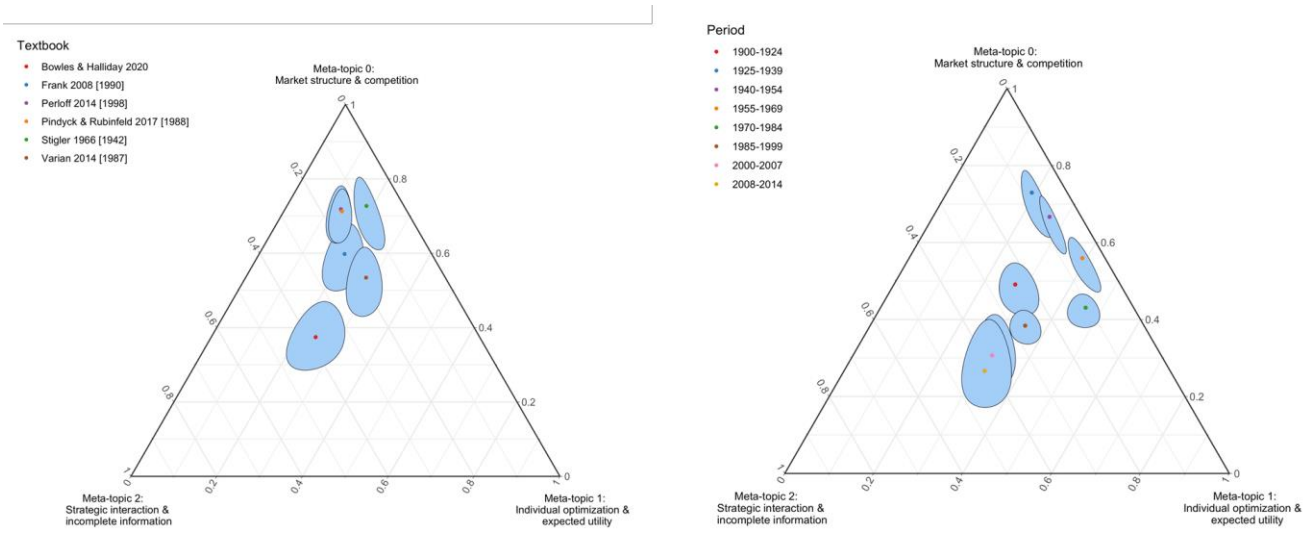


**Figure 19. Uncertainty in the microeconomics estimates**: Ninety-five percent confidence regions for textbooks (left panel) and time slices of the research corpus (right panel). The dots in the confidence regions are the estimates for the unperturbed case.

The confidence regions in the right panel of Figure 18, for example, show that with the exception of the three recent time slices, it is very unlikely that the generative processes accounting for the distribution of words in the research corpus remained unchanged from period to period. The left panel of Figure 19 shows, for example, that the generative process accounting for the distribution of words in Varian is very unlikely to account for the distribution of words in either Bowles and Halliday or in Perloff or Pindyck and Rubinfeld.

Three aspects of this robustness check should be noted. First, the method implemented is illustrative but far from exhaustive. The perturbation of the topics constituting the meta-topics, for example, could be varied so that the exclusion of each topic occurs not with probability 1/n but some other probability (2/n, 1/2n or another).

Second, we could explore the effects of excluding but not replacing topics, or the opposite, adding topics to the meta-topic without excluding topics. Our "eliminate and replace" procedure is a more stringent test. The combination of elimination and replacement adds a new dimension

to the meta-topic while eliminating one of its constituent elements, thereby introducing more noise than would occur by either elimination or addition alone.[10]

Third, this aspect of the procedure is illustrated in the middle panel in Figure 17 by the small "island" of alternative generative processes giving much greater weight to microeconomics Meta-topic 0 in the research corpus. The points making up the island are without exception generated by the random elimination of the topic "market structure: measurement" from the Meta-topic "Market structure and competition". This topic has a small weight in the recent research corpus and its replacement by the topic "applied econometrics: cross-section and panel", which is much more heavily weighted in the research corpus, displaced the Meta-topic weights toward Meta-topic 0.

For Jones' macroeconomic text, similar "islands" of alternative generative processes appear when the topic "growth models" is randomly eliminated from the meta-topic, "growth" (Appendix A). Our data-driven consideration of alternative compositions of the meta-topics thus includes a number of perturbations that do not represent plausible modifications. Neither the inclusion of "applied econometrics: cross-section and panel" as a component of market structure and competition, nor the exclusion of "growth models" from the meta-topic "growth" could be convincingly motivated on conceptual grounds. A result is that the confidence regions shown may be larger than would be appropriate for the thought experiment motivating this robustness test.

## VIII. Discussion

For economics, the content of our second-year micro and macroeconomics courses is an indicator of what Kuhn termed the reigning paradigm; it defines the subject matter of the field, the types of questions to be asked, and appropriate methods for answering them. Topic modeling applied to a corpus of 27 thousand papers in top research journals in the past century provides a tool for measuring the content of the intermediate textbooks as a representation the discipline's consensus on "what an economist should know" and allows a mapping between this and the

[10] We experimented extensively with the "elimination only" procedure, and as expected, the resulting confidence regions were generally smaller.

evolution of the conceptual frameworks used in the research papers published in highly regarded journals.

Is there any reason why economists should be concerned if the consensus about what any economist should know (the benchmark taught in the intermediate textbook) is substantially different from the content of contemporary research? It depends on the nature of the difference.

We have shown that in varying degree prominent microeconomics textbooks have not reflected the movement since the early 1960s of the research corpus away from Marshallian market analysis, initially to a greater emphasis on individual optimization and expected utility, and subsequently towards strategic interactions under incomplete information. While macroeconomics textbooks followed the research trajectory's movement away from a near-exclusive focus on Keynesian themes to a greater emphasis on growth, aggregate demand and its management remain an important component of the benchmark taught to economics majors.

By contrast, in the dominant macro textbooks, vibrant research topics – about search and matching, quantity constraints, and dynamics in labor, credit and financial markets – receive little attention. There has been remarkably little change in either the research corpus (over the last 40 years) or in the content of macroeconomics textbooks (since the financial crisis), especially by comparison to the changes following the Great Depression.

Even taking account of the perturbations to the definitions of meta-topics in the previous section, our topic modeling of macroeconomics textbooks and the evolution of the research corpus has shown a substantial divergence between the two. The Keynesian content that was introduced rapidly to textbooks following the shift in the research corpus after 1929 (our macroeconomic Meta-topic 0) has remained important in the textbooks.

A likely explanation we think is that fluctuations in aggregate demand and the possibility of their attenuation by public policy (via automatic stabilizers, fiscal and monetary policy), while enduring as a centrally important aspect of economic knowledge for each generation of economics majors, has, since the 1970s been regarded by researchers as largely understood, and by policy makers, as a substantially solved problem. The movement of the research trajectory away from Keynesian economics, if our interpretation is correct, does not reflect a conviction

that it is mistaken, but rather that in its modern form – including, for example, inflation-targeting (in Meta-topic 0) – it is settled.

By contrast, the fact that in microeconomics most textbooks have not tracked the evolution of the research corpus, we think, represents a matter of greater concern for the education of economists. When Stigler wrote his *The Theory of Price* in 1942, its Marshallian economics (captured by our Meta-topic 0) was regarded as up to date, and Stigler's content closely matched the conceptual frameworks that top research economists used. The coordinates of his textbook in the simplex are very close to the research corpus 1925-54. Unlike the case of macroeconomics, where little of essential importance to the theory of aggregate demand management was added to the research corpus in the past half century, major theoretical developments have taken place in microeconomic theory over that time, including in competition and functioning of markets, which have been associated with the theoretical developments in Meta-topic 2.

The dramatic reduction since 1970 in the weight on the "Market structure and competition" meta-topic in the research corpus does not arise from a consensus that Marshallian economics has settled the important questions about how markets and competition work and as a result provided adequate tools for appropriate policy responses. Of course, Marshallian concepts such as elasticities of supply and demand remain essential in formulating competition policy and should surely be part of the benchmark. But in contrast to macroeconomics, the benchmark for how markets and competition work has moved considerably beyond Marshall and Stigler to include the tools provided by contemporary models of strategic interaction under asymmetric information addressed to problems of intellectual property and near zero marginal costs.

## References


Angrist, Joshua, Pierre Azoulay, Glenn Ellison, Ryan Hill, and S. F. Lu. 2017. "Economic Research Evolves: Fields and Styles." *American Economic Review* 107 (5):293-297.

Ash, Elliott, Daniel Chen, and Suresh Naidu. 2018. "Ideas have consequences: The impact of law and economics on American Justice." *Unpublished manuscript, Columbia University*.

Backhouse, Roger. 2017. *Founder of Modern Economics: Paul A. Samuelson*. Vol. I. Becoming Samuelson, 1015-1948. Oxford: Oxford University Press.

Backhouse, Roger, Bradley Bateman, and Steven Medema. 2010. "The reception of Marshall in the United States." In *The impact of Alfred Marshall's ideas*, edited by T Raffaelli, 59-80. Chettenham: Edward Elgar.

Barron, Alexander, Jenny Huang, Rebecca Spang, and Simon DeDeo. 2018. "Individuals, institutions and innovation in the debates of the French Revolution." *Proceedings of the National Academy of Science* 115 (18):4607-4612.

Blanchard, Olivier. 2016. "How to teach intermediate macroeconomics after the crisis?" *Peterson Institute for International Economics Realtime Economics Issues Watch, https://piie.com/blogs/realtime-economic-issues-watch/how-teach-intermediate-macroeconomics-after-crisis* (2 June).

Blei, David. 2012. "Probabilistic Topic Models." *Communication of the Association for Computing Machinery* 55 (4):77-84.

Blei, David, Andrew Ng, and Michael Jordan. 2003a. "Hierarchical Bayesian Models for Applicaitons in Information Retrieval." *Bayesian Statistics* 7:25-43.

Blei, David, Andrew Ng, and Michael Jordan. 2003b. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3:993-1022.

Bowles, Samuel, and Wendy Carlin. 2020. "What students learn in economics 101: Time for a change." *Journal of Economic Literature* forthcoming [https://tinyco.re/1812530].

Ely, Richard T., Thomas S. Adams, Max O. Lorenz, and Allyn A. Young. 1930. *Outlines of Economics* New York: McMillan Company [first published 1893].

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as Data." *Journal of Economic Literature* forthcoming.

Hamermesh, Daniel. 2013. "Six decades of top economics publishing: Who and how? ." *Journal of Economic Literature* 51 (1):162-172.

Hamilton, N.E., and M Ferry. 2018. "ggtern: Ternary diagrams using ggplot2." *Journal of Statistical Software* 87 (3):1-17.

Hoover, K. D, and Michel DeVroey. 2005. "The IS-LM model: Its rise, fall, and strange persistence: Introduction to the Special Issue." *History of Political Economy* 36:1-11.

Klingenstein, Sarah, Tim Hitchcock, and Simon DeDeo. 2014. "The civilizing process of London's Old Bailey." *Proceedings of the National Academy of Science*. doi: 10.1073/pnas.1405984111.

Kuhn, Thomas. 1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lerner, Abba. 1944. *The economics of control: Principles of welfare economics*. New York: Macmillan.

Lucas, Robert. 1975. "An equilibrium model of the business cycle." *Journal of Political Economy* 83 (6):1113-44.

Maesschalck, R. De, D Jouan-Rimbaud, and D.L. Massart. 2000. "The Mahalanobis distance." *Chemometrics and Intelligent Laboratory Systems* 50:1018.

Mahalanobis, P.C. 1936. "On the generalized distance in statistics." *Proceedings of the National Institute of Sciences of India* 2 (1):49-55.

Mankiw, N. Gregory. 2016. "Response: The tradeoff between nuance and clarity." *Eastern Economic Journal* 42:169-170.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1995. *Microeconomic Theory*. New York: Oxford University Press.

Samuelson, Paul. 1948. *Economics, an Introductory Analysis*. New York: McGraw Hill.

Stigler, George. 1942. *The Theory of Price*. New York City: Irwin.

Tarshis, Lorie. 1947. *Elements of Economics*. New York: Houghton Mifflin.

# What Do We Think An Economist Should Know? Appendices for online publication

Samuel Bowles, Wendy Carlin, Simon D. Halliday,
and Sahana Subramanyam

23 August 2019

# Contents

# Appendix A

# Further Results

## A.1 Introduction

This appendix can be read in conjunction with Appendix B: Methods and additional figures; Appendix C: The R Markdown document analyzing the perturbed meta-topic data (to be provided upon request); and, Appendix D: The Python code for the topic models and robustness simulations (to be provided upon request).

## A.2 Top-weighted topics over time

The following bar charts for each time slice show all topics that fit the following criteria: first, they have weights greater than 0.015 and, second, they do not belong to any of the following topics 79: Publication details, 34: Generic economic language, or 26: French language.

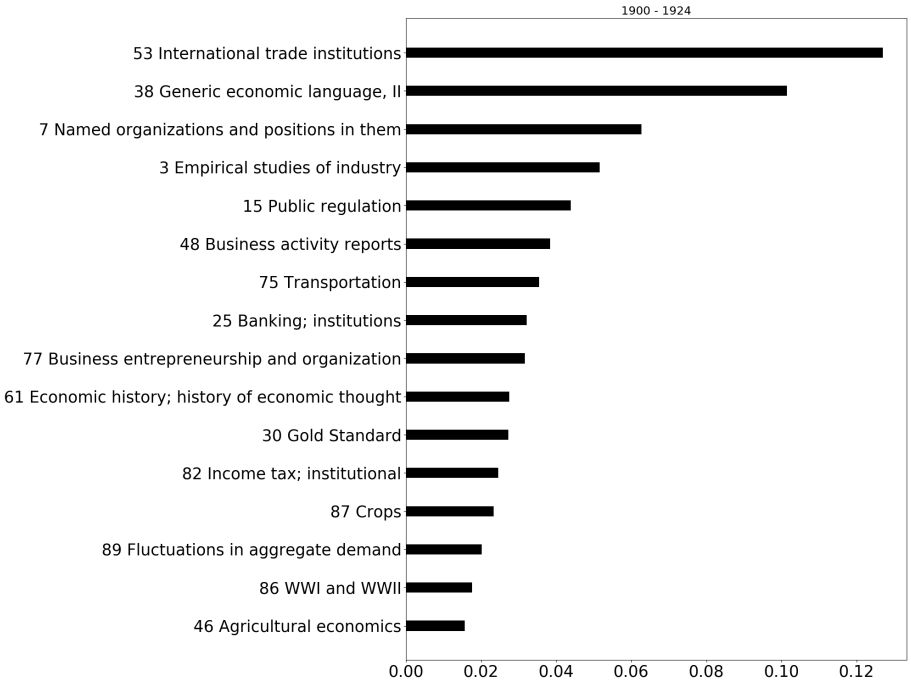


Figure A.1: Top topics for the period 1900 - 1924. Each solid bar is the topic weight $\theta_{kd}$, where the document $d$ is the sum of documents published in this time slice treated as one document. The bars are ranked from highest weight to lowest weight with a cut-off of $\theta_{kd} = 0.015$.

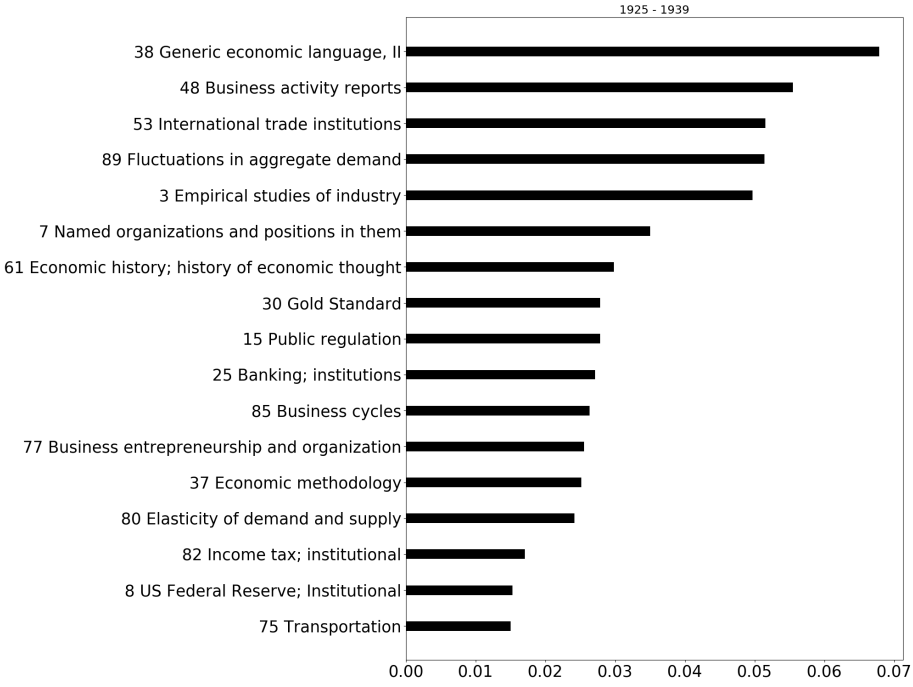


Figure A.2: Top topics for the period 1925 - 1939. Each solid bar is the topic weight $\theta_{kd}$, where the document $d$ is the sum of documents published in this time slice treated as one document. The bars are ranked from highest weight to lowest weight with a cut-off of $\theta_{kd} = 0.015$.

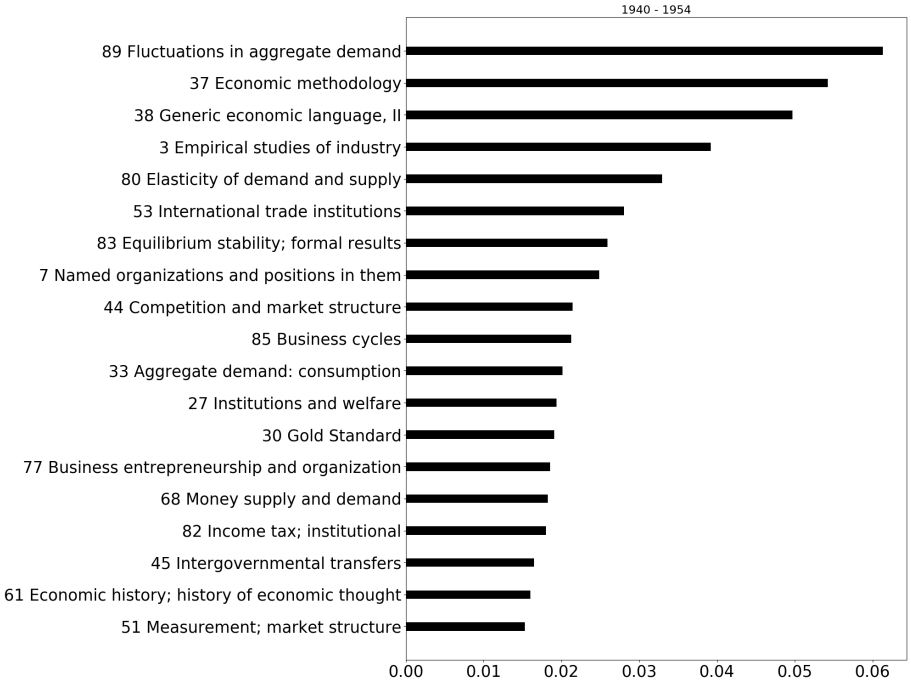


Figure A.3: Top topics for the period 1940 - 1954. Each solid bar is the topic weight $\theta_{kd}$, where the document $d$ is the sum of documents published in this time slice treated as one document. The bars are ranked from highest weight to lowest weight with a cut-off of $\theta_{kd} = 0.015$.

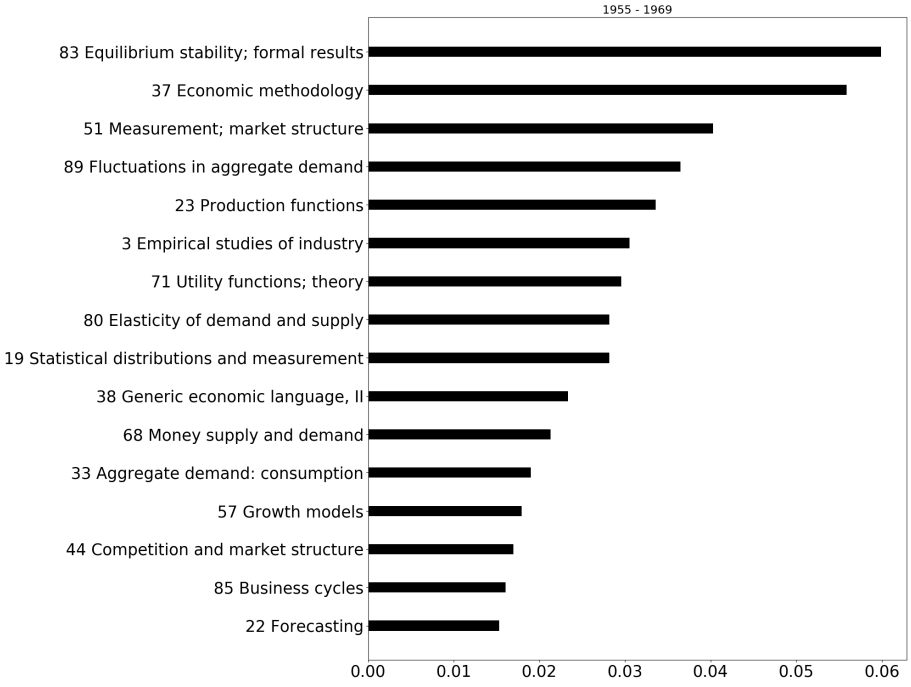


Figure A.4: Top topics for the period 1955 - 1969. Each solid bar is the topic weight $\theta_{kd}$, where the document $d$ is the sum of documents published in this time slice treated as one document. The bars are ranked from highest weight to lowest weight with a cut-off of $\theta_{kd} = 0.015$.

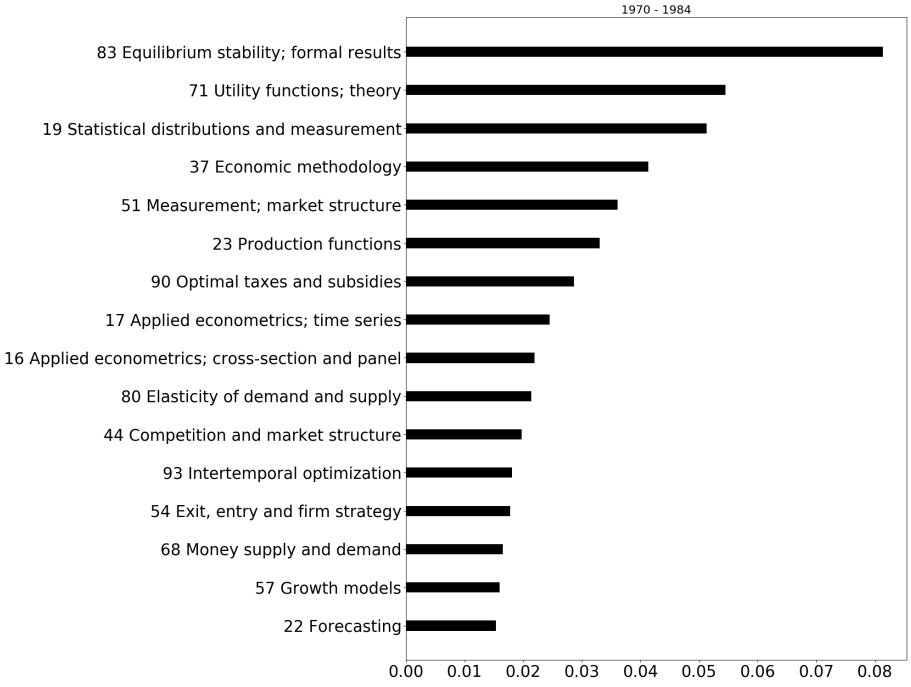


Figure A.5: Top topics for the period 1970 - 1984. Each solid bar is the topic weight $\theta_{kd}$, where the document $d$ is the sum of documents published in this time slice treated as one document. The bars are ranked from highest weight to lowest weight with a cut-off of $\theta_{kd} = 0.015$.

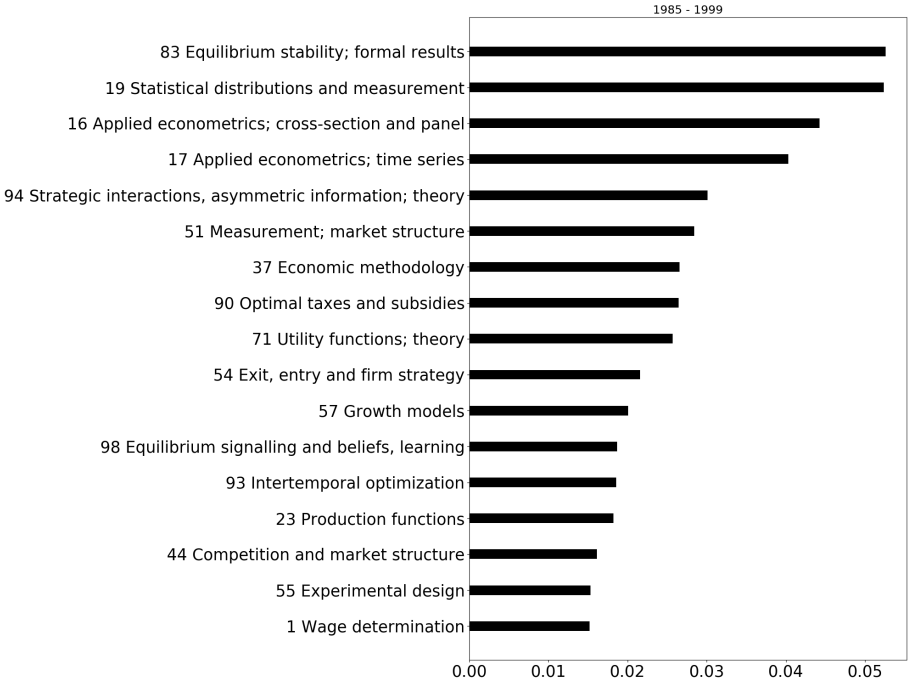


Figure A.6: Top topics for the period 1985 - 1999. Each solid bar is the topic weight $\theta_{kd}$, where the document $d$ is the sum of documents published in this time slice treated as one document. The bars are ranked from highest weight to lowest weight with a cut-off of $\theta_{kd} = 0.015$.

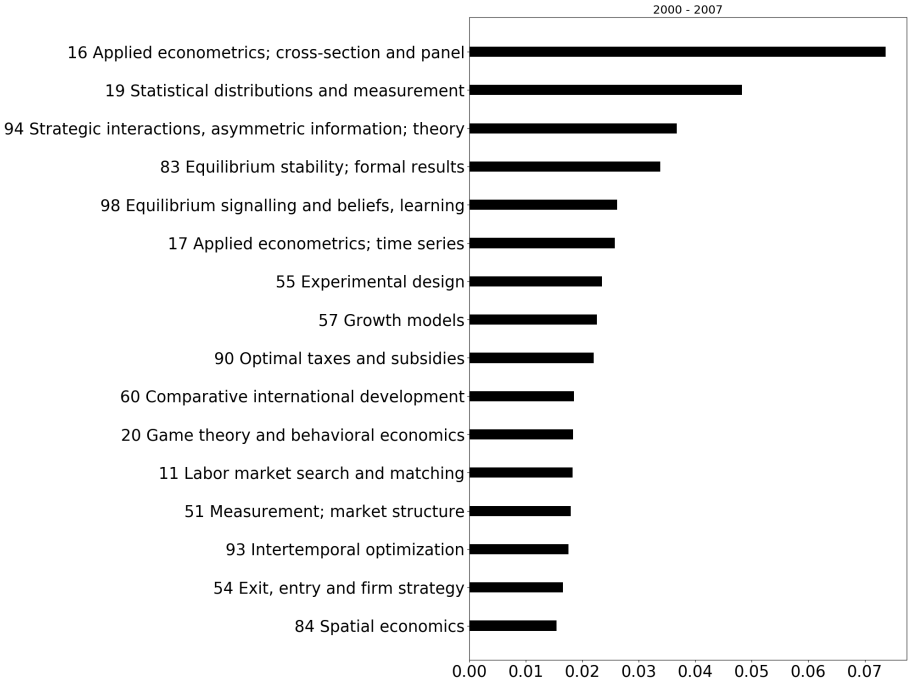


Figure A.7: Top topics for the period 2000 - 2007. Each solid bar is the topic weight $\theta_{kd}$, where the document $d$ is the sum of documents published in this time slice treated as one document. The bars are ranked from highest weight to lowest weight with a cut-off of $\theta_{kd} = 0.015$.

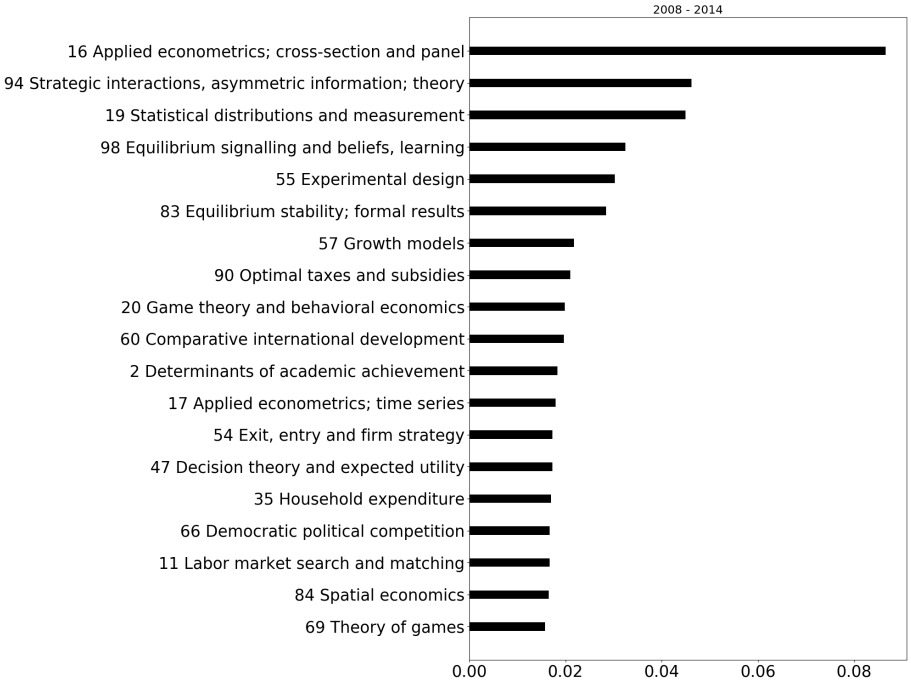


Figure A.8: Top topics for the period 2008 - 2014. Each solid bar is the topic weight $\theta_{kd}$, where the document $d$ is the sum of documents published in this time slice treated as one document. The bars are ranked from highest weight to lowest weight with a cut-off of $\theta_{kd} = 0.015$.

## A.3 Before and after the global financial crisis: within-author textbook comparisons

### A.3.1 Topic weights before and after the financial crisis

The following figures compare four intermediate macroeconomic textbooks before and after the 2008 crisis.

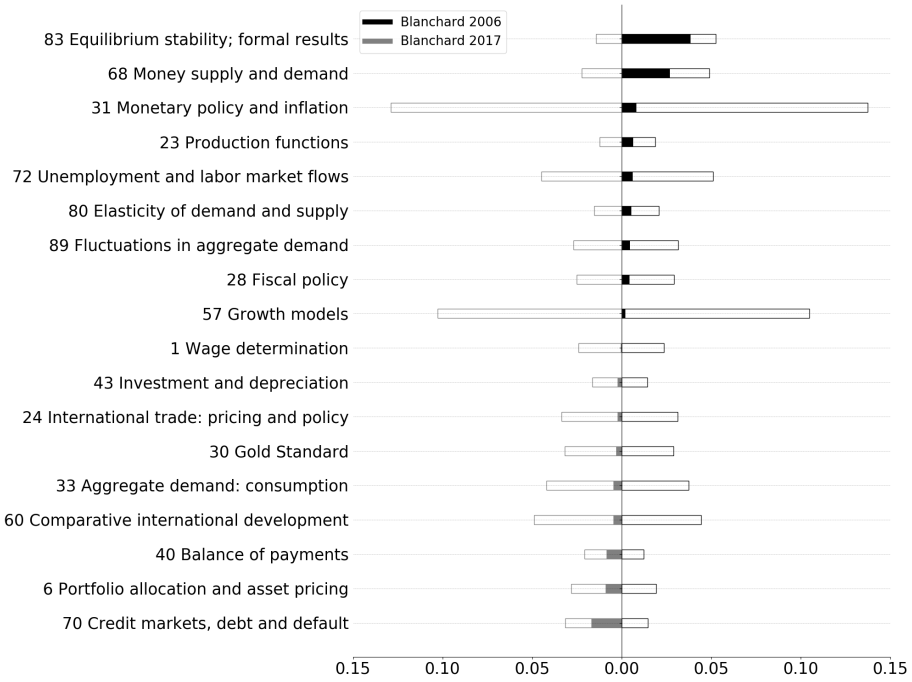


Figure A.9: Blanchard (2006) vs. Blanchard (2017). Each outline bar is the topic weight $\theta_{kd}$ for the earlier (bars to the right of the vertical axis) and the later (bars to the left). The solid bars show the difference in the topic weight between the two.

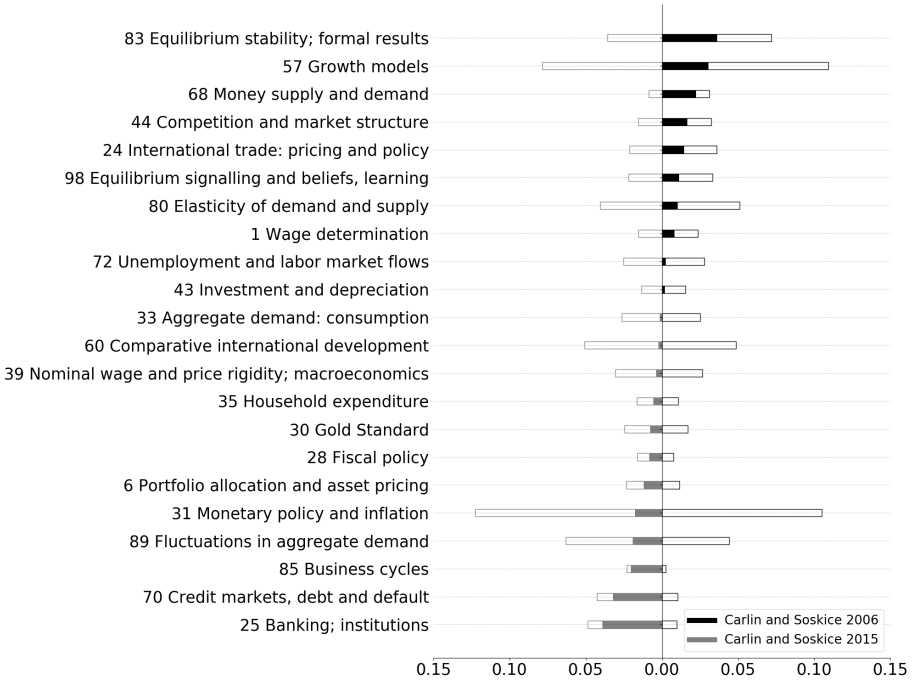


Figure A.10: Carlin & Soskice (2006) vs. Carlin & Soskice (2015). Each outline bar is the topic weight $\theta_{kd}$ for the earlier (bars to the right of the vertical axis) and the later (bars to the left). The solid bars show the difference in the topic weight between the two

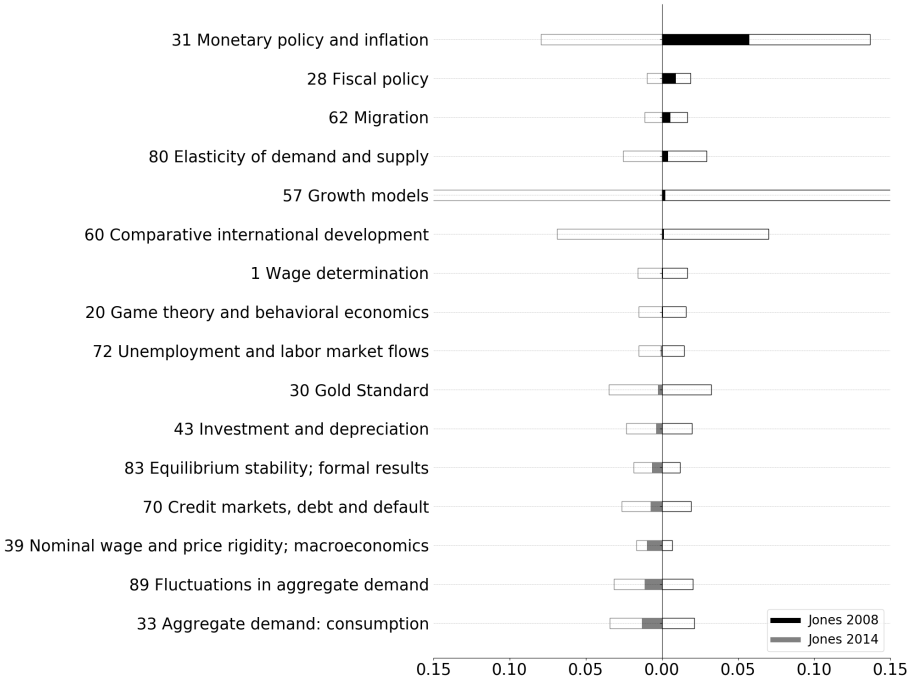


Figure A.11: Jones (2008) vs. Jones (2014). Each outline bar is the topic weight $\theta_{kd}$ for the earlier (bars to the right of the vertical axis) and the later (bars to the left). The solid bars show the difference in the topic weight between the two.

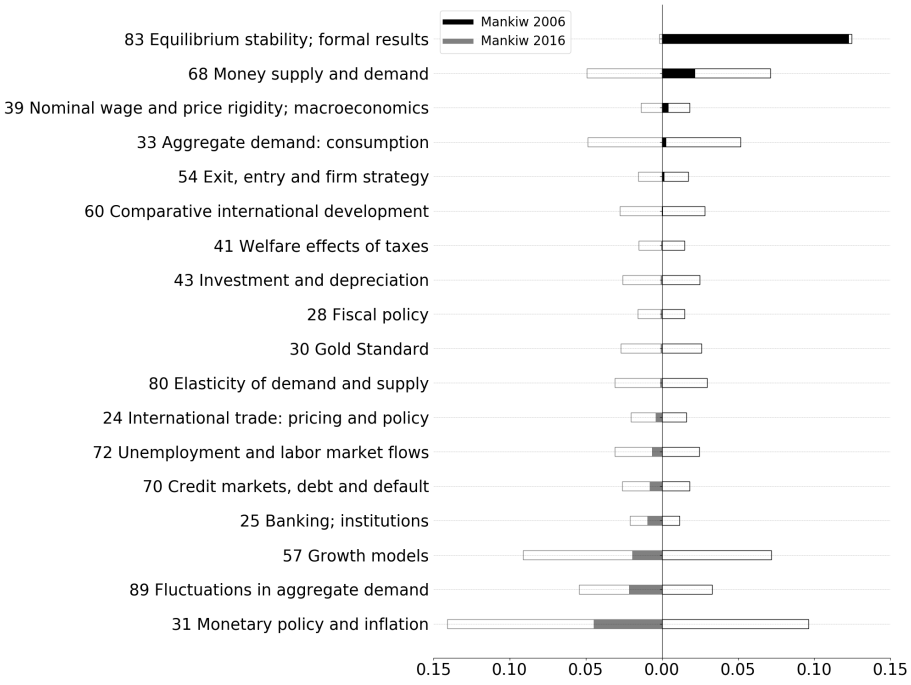


Figure A.12: Mankiw (2006) vs. Mankiw (2016). Each outline bar is the topic weight $\theta_{kd}$ for the earlier (bars to the right of the vertical axis) and the later (bars to the left). The solid bars show the difference in the topic weight between the two.

### A.3.2 Confidence regions before and after the financial crisis

We use the following figures for the comparisons of the macroeconomics textbooks before and after the financial crisis and to understand the confidence regions of their topic perturbations.

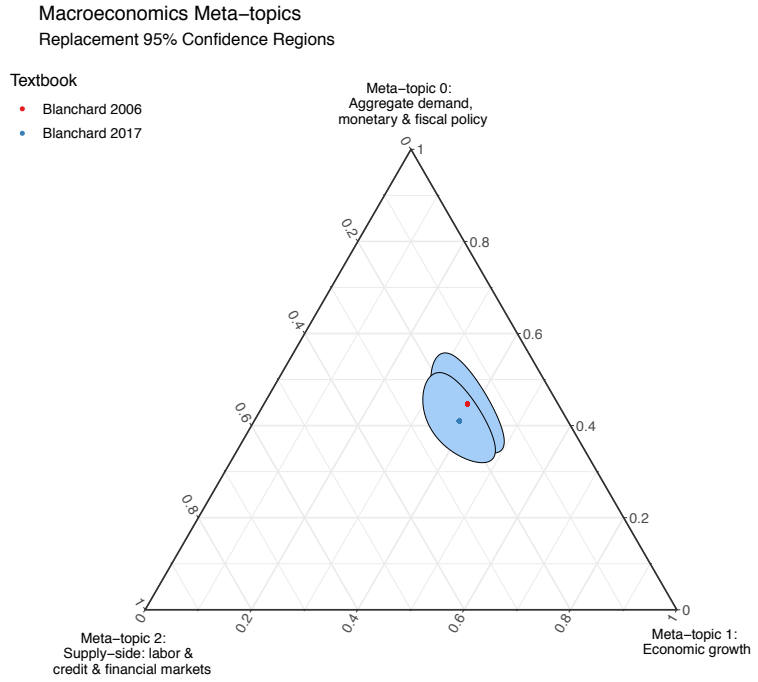


Figure A.13: Blanchard (2006) vs. Blanchard (2017), before and after the financial crisis.

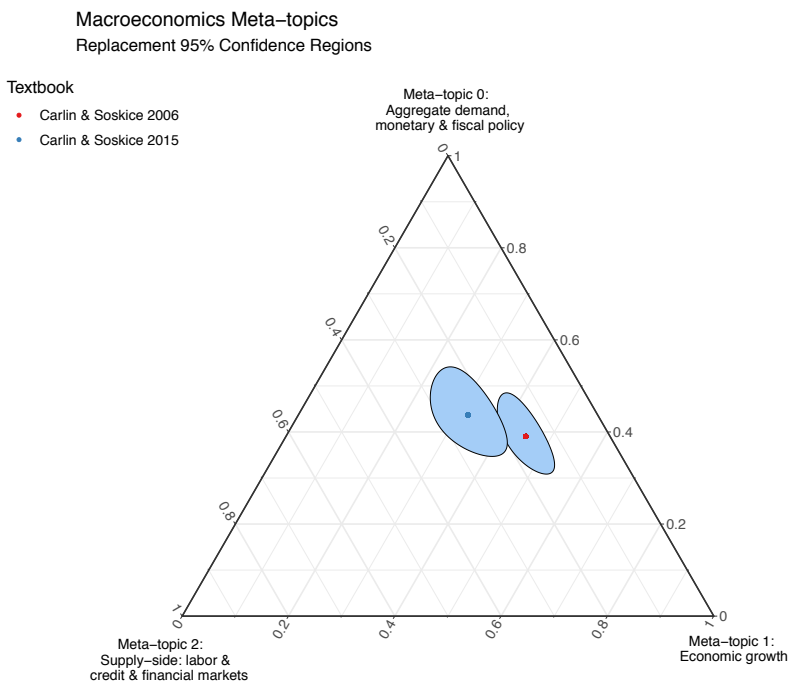


Figure A.14: Carlin & Soskice (2006) vs. Carlin & Soskice (2015) before and after the financial crisis.

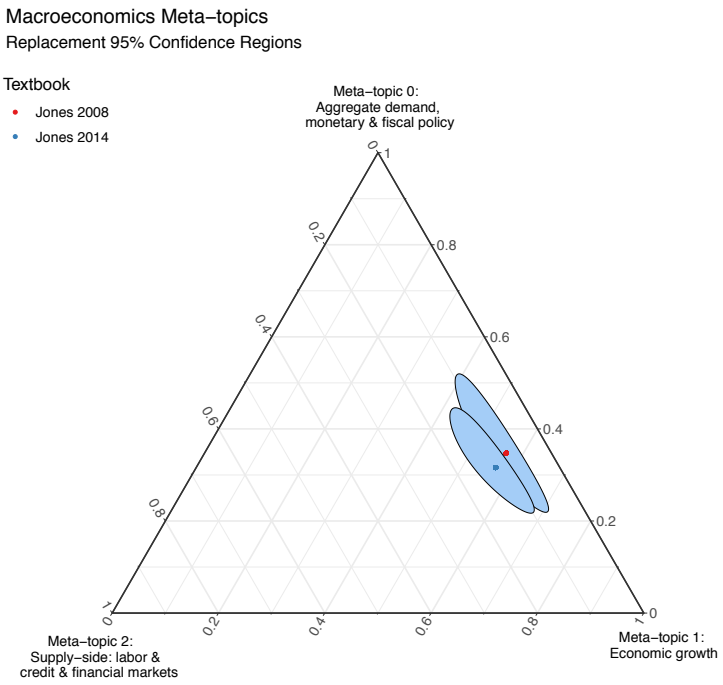


Figure A.15: Jones (2008) vs. Jones (2014), before and after the financial crisis.

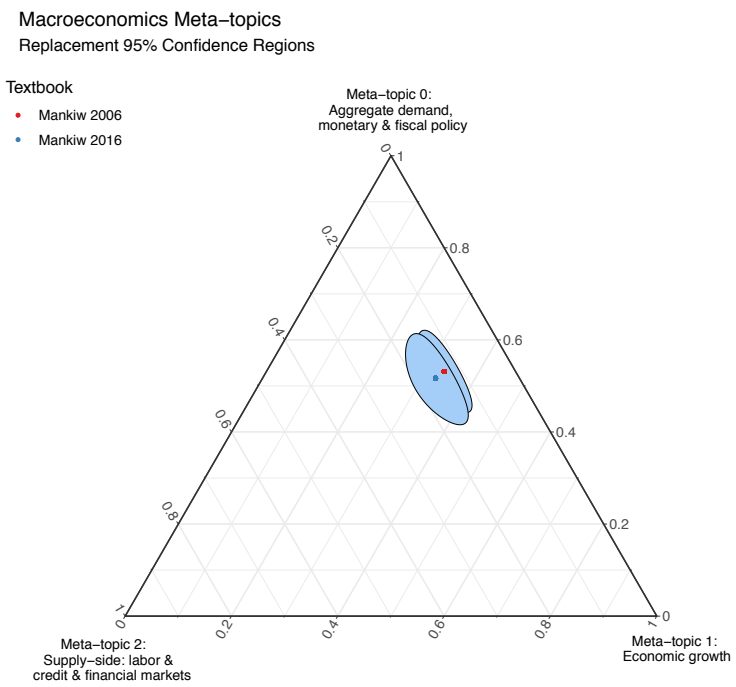


Figure A.16: Mankiw (2006) vs. Mankiw (2016), before and after the financial crisis.

## A.4 Replacement topics for meta-topics

| micro mtopic | topic | correlation | topic | correlation | topic | correlation |
|---|---|---|---|---|---|---|
| **Meta-topic 0: Market structure and competition** | | | | | | |
| 44 Competition and market structure | 54 Exit, entry and firm strategy | 0.308464 | **90 Optimal taxes and subsidies** | 0.178587 | 80 Elasticity of demand and supply | 0.167568 |
| 49 Advertising and consumer demand | **4 Adverse selection; "lemons"** | 0.184286 | 5 Sales strategies | 0.144479 | 44 Competition and market structure | 0.141862 |
| 51 Measurement; market structure | **16 Applied econometrics; cross-section and panel** | 0.191311 | 23 Production functions | 0.179041 | 60 Comparative international development | 0.128642 |
| 54 Exit, entry and firm strategy | 44 Competition and market structure | 0.308464 | **21 Innovation** | 0.167688 | 11 Labor market search and matching | 0.119436 |
| 80 Elasticity of demand and supply | 44 Competition and market structure | 0.167568 | **92 International trade and tariffs** | 0.097024 | 83 Equilibrium stability; formal results | 0.088416 |
| **Meta-topic 1: Individual optimization and expected utility** | | | | | | |
| 18 Risk preferences | **6 Portfolio allocation and asset pricing** | 0.235105 | 93 Intertemporal optimization | 0.151869 | 47 Decision theory and expected utility | 0.125812 |
| 23 Production functions | **51 Measurement; market structure** | 0.179041 | 14 Skills, technology, trade and training | 0.135796 | 57 Growth models | 0.127923 |
| 47 Decision theory and expected utility | **55 Experimental design** | 0.183164 | 71 Utility functions; theory | 0.144425 | 18 Risk preferences | 0.125812 |
| 71 Utility functions; theory | **83 Equilibrium stability; formal results** | 0.300463 | 94 Strategic interactions, asymmetric information; theory | 0.21681 | 12 Mechanism design theory | 0.147176 |
| 74 Labor supply | **1 Wage determination** | 0.198096 | 88 Determinants of earnings; empirical | 0.114076 | 16 Applied econometrics; cross-section and panel | 0.093463 |
| 93 Intertemporal optimization | **94 Strategic interactions, asymmetric information; the** | 0.179567 | 83 Equilibrium stability; formal results | 0.159933 | 57 Growth models | 0.159353 |
| **Meta-topic 2: Strategic interaction and incomplete information** | | | | | | |
| 0 Monitoring and enforcement | 52 Incomplete contracts and principal-agent models | 0.11187 | **90 Optimal taxes and subsidies** | 0.103651 | 20 Game theory and behavioral economics | 0.092619 |
| 4 Adverse selection; "lemons" | **49 Advertising and consumer demand** | 0.184286 | 5 Sales strategies | 0.086945 | 16 Applied econometrics; cross-section and panel | 0.080577 |
| 10 Bargaining and incomplete information | **98 Equilibrium signalling and beliefs, learning** | 0.159723 | 94 Strategic interactions, asymmetric information; theory | 0.157553 | 69 Theory of games | 0.157081 |
| 20 Game theory and behavioral economics | **55 Experimental design** | 0.238165 | 69 Theory of games | 0.204723 | 98 Equilibrium signalling and beliefs, learning | 0.147298 |
| 52 Incomplete contracts and principal-agent models | 94 Strategic interactions, asymmetric information; theory | 0.230244 | **90 Optimal taxes and subsidies** | 0.15991 | 10 Bargaining and incomplete information | 0.126322 |
| 69 Theory of games | **98 Equilibrium signalling and beliefs, learning** | 0.34025 | 94 Strategic interactions, asymmetric information; theory | 0.242783 | 55 Experimental design | 0.2308 |
| 73 Adverse selection, moral hazard; insurance | **99 Social security** | 0.164577 | 81 Healthcare | 0.141228 | 18 Risk preferences | 0.092189 |
| 94 Strategic interactions, asymmetric information; theory | **98 Equilibrium signalling and beliefs, learning** | 0.363898 | 12 Mechanism design theory | 0.301176 | 69 Theory of games | 0.242783 |

Figure A.17: Replacement topics for microeconomic meta-topics. The replacement topic for the topic in the left column is shown in bold, along with its correlation with the replaced topic.

| macro mtopic | topic | correlation | topic | correlation | topic | correlation |
|---|---|---|---|---|---|---|
| **Meta-topic 0: Aggregate demand, monetary and fiscal policy** | | | | | | |
| 8 US Federal Reserve; Institutional | **25 Banking; institutions** | 0.388265 | 89 Fluctuations in aggregate demand | 0.158028 | 48 Business activity reports | 0.140408 |
| 28 Fiscal policy | **45 Intergovernmental transfers** | 0.195144 | 41 Welfare effects of taxes | 0.166587 | 65 Local public goods | 0.120227 |
| 30 Gold Standard | **40 Balance of payments** | 0.25166 | 89 Fluctuations in aggregate demand | 0.192933 | 24 International trade: pricing and policy | 0.182935 |
| 31 Monetary policy and inflation | 39 Nominal wage and price rigidity; macroeconomics | 0.206933 | **17 Applied econometrics; time series** | 0.190163 | 68 Money supply and demand | 0.173409 |
| 33 Aggregate demand: consumption | **35 Household expenditure** | 0.194324 | 43 Investment and depreciation | 0.128939 | 93 Intertemporal optimization | 0.113557 |
| 39 Nominal wage and price rigidity; macroeconomics | 31 Monetary policy and inflation | 0.206933 | **57 Growth models** | 0.141542 | 17 Applied econometrics; time series | 0.130844 |
| 68 Money supply and demand | 31 Monetary policy and inflation | 0.173409 | **6 Portfolio allocation and asset pricing** | 0.15881 | 34 Generic economic language, I | 0.105035 |
| 85 Business cycles | **48 Business activity reports** | 0.171913 | 51 Measurement; market structure | 0.100315 | 3 Empirical studies of industry | 0.089473 |
| 89 Fluctuations in aggregate demand | 30 Gold Standard | 0.192933 | 8 US Federal Reserve; Institutional | 0.158028 | **40 Balance of payments** | 0.146553 |
| **Meta-topic 1: Economic growth** | | | | | | |
| 21 Innovation | **54 Exit, entry and firm strategy** | 0.167688 | 57 Growth models | 0.115309 | 14 Skills, technology, trade and training | 0.080269 |
| 23 Production functions | **51 Measurement; market structure** | 0.179041 | 14 Skills, technology, trade and training | 0.135796 | 57 Growth models | 0.127923 |
| 43 Investment and depreciation | 57 Growth models | 0.145157 | **67 Project evaluation** | 0.130324 | 33 Aggregate demand: consumption | 0.128939 |
| 57 Growth models | **14 Skills, technology, trade and training** | 0.164866 | 93 Intertemporal optimization | 0.159353 | 43 Investment and depreciation | 0.145157 |
| 60 Comparative international development | **16 Applied econometrics; cross-section and panel** | 0.174525 | 24 International trade: pricing and policy | 0.167684 | 62 Migration | 0.149161 |
| 93 Intertemporal optimization | **94 Strategic interactions, asymmetric information; the** | 0.179567 | 83 Equilibrium stability; formal results | 0.159933 | 57 Growth models | 0.159353 |
| **Meta-topic 2: Supply-side: labor, credit and financial markets** | | | | | | |
| 1 Wage determination | 11 Labor market search and matching | 0.25975 | **88 Determinants of earnings; empirical** | 0.244691 | 72 Unemployment and labor market flows | 0.209417 |
| 6 Portfolio allocation and asset pricing | **18 Risk preferences** | 0.235105 | 36 Stock prices and trading | 0.201345 | 70 Credit markets, debt and default | 0.199512 |
| 11 Labor market search and matching | 1 Wage determination | 0.25975 | 72 Unemployment and labor market flows | 0.167994 | **54 Exit, entry and firm strategy** | 0.119436 |
| 13 Labor unions | **7 Named organizations and positions in them** | 0.240397 | 1 Wage determination | 0.187662 | 53 International trade institutions | 0.115522 |
| 25 Banking; institutions | **8 US Federal Reserve; Institutional** | 0.388265 | 70 Credit markets, debt and default | 0.155947 | 48 Business activity reports | 0.1324 |
| 52 Incomplete contracts and principal-agent models | **94 Strategic interactions, asymmetric information; the** | 0.230244 | 90 Optimal taxes and subsidies | 0.15991 | 10 Bargaining and incomplete information | 0.126322 |
| 56 Financial market speculation | **36 Stock prices and trading** | 0.113182 | 18 Risk preferences | 0.105306 | 6 Portfolio allocation and asset pricing | 0.090254 |
| 70 Credit markets, debt and default | 6 Portfolio allocation and asset pricing | 0.199512 | 25 Banking; institutions | 0.155947 | **8 US Federal Reserve; Institutional** | 0.091443 |
| 72 Unemployment and labor market flows | 1 Wage determination | 0.209417 | 11 Labor market search and matching | 0.167994 | **96 Welfare and social insurance programs** | 0.1595 |

Figure A.18: Replacement topics for macroeconomic meta-topics. The replacement topic for the topic in the left column is shown in bold, along with its correlation with the replaced topic.

# Appendix B

# Methods and Data

## B.1 Introduction

In this appendix we outline the methods for creating the topic models and we provide examples and explanations of the relevant methods and packages we have used.

## B.2 Methods

### B.2.1 Topic modeling

#### Pre-processing

The corpus of 27,436 research articles were collected in the form of plain text files with corresponding xml metadata containing publication details such as author names, publication date, journal, issue etc. The text files were pre-processed using the following steps:

- Cleaning: This first step splits sentences into words, strips away all punctuation and numerals retaining only words with alphabets with length greater than one. Finally all spelling is standardized to British English and made lower case.
- Eliminating stop words: Commonly occurring English words such as articles and prepositions that are not useful in distinguishing topics in texts are discarded. Additional words specific to the corpus such as "exercise", "table" etc. were also removed. A complete list of stop words removed can be found on the GitHub page (which will be made public upon publication).
- Stemming: In this step we use the NLTK PorterStemmer from Bird and Klein (2009) to reduce inflected words to their corresponding word stems.

For example, "utility" and "utils" stem to "util" but "utilitarianism" stems to "utilitarian".

- Bigramming: To ensure that commonly co-occuring words are combined, we use the Gensim Phraser by Řehůřek and Sojka (2010) to generate bigrams. An example of this is "risk_neutral".

- Filtering: To further trim the corpus we retain only tokens that occur in at least one percent of the documents in the corpus. This helps get rid of highly specialized vocabulary and rare text errors.

The process results in a vocabulary of 10,849 unique tokens that form the dictionary used in estimating the topic model.

**Corpus reduction**

The initial raw corpus contains 42,152 articles. However some of these articles are not important for our purposes since they are special non-research articles such as book reviews and notes. To filter out these articles we used the 'subject' field of the metadata. Normally, a "none" value in the subject field indicates a standard research article. These articles were retained. Any article with the subject column populated in the metadata file was excluded. Table B.1 shows the top 10 subjects that are excluded. The highest exclusions occur due to "no data found" which implies that the article had no input meta data. Going down this table does reveal certain erroneous exclusions due to incorrect source metadata. For instance, some articles had themes of the work stored under subject incorrectly in the source .xml files. Since these were negligible and most of the exclusions as shown in the table were valid, this was ignored. A complete list of all subjects that were excluded can be found on GitHub. Reducing the corpus in this way gave us a total of 27,436 articles that enter the topic model.

| Subject | Number of articles |
|---|---|
| no data found | 8650 |
| Reviews and New Books | 1970 |
| Communications | 1485 |
| Notes and Memoranda | 1053 |
| Shorter Papers | 538 |
| Notes and Comments | 517 |
| Notes | 328 |
| Comments | 84 |
| Errata | 77 |
| Confirmations and Contradictions | 61 |

Table B.1: Excluded article types: top 10

**LDA model**

The type of topic model used is a Latent Dirichlet Allocation model, from the highly cited Blei, Ng, and Jordan (2003) and implemented in python using Gensim by Řehůřek and Sojka (2010). The $\alpha$ and $\eta$ hyperparameters that inform the Dirichlet distribution were both set to their Gensim defaults of a symmetric 1.0/number of topics (in our case 0.01) prior. The $\alpha$ hyperparameter controls document topic distributions and $\eta$ the word distribution in topics. The same model, hyperparameter values and 100 topics have been used in Bowles and Carlin (2020). The underlying programming is done in python using Gensim by Řehůřek and Sojka (2010) for the analysis and Harper et al. (2015) for the visualization.

### B.2.2 Topic exclusion

When producing the bar charts with the topic weights, we do not include all of the topics (Figures A.1 through A.8). A topic will be excluded if it meets any of the following criteria.

1. It has a weight less than or equal to 0.015 in at least one of the texts compared (this applies to all figures); or
2. If the highest weighted token in the topic is less than 0.01 (this applies to comparison figures only); or
3. If it belongs to any of the following topics: 79: Publication details, 34: Generic economic language, I, or 26: French language (see discussion below).

The figures illustrate the top-weighted topics that are not excluded by these criteria.

Considering the three excluded topics, for clarity, we present the top 10 words in each topic:

- 26: French language: 0.136*"de" + 0.061*"et" + 0.053*"franc" + 0.048*"le" + 0.044*"la" + 0.042*"french" + 0.036*"en" + 0.031*"de_la" + 0.020*"du" + 0.016*"un"
- 34: Generic economic language: 0.010*"commod" + 0.008*"shall" + 0.007*"kind" + 0.006*"whole" + 0.006*"sens" + 0.006*"quantiti" + 0.006*"concept" + 0.005*"let_us" + 0.005*"principl" + 0.005*"get"
- 79: Publication details: 0.047*"book" + 0.039*"author" + 0.023*"articl" + 0.019*"field" + 0.019*"institut" + 0.019*"univers" + 0.017*"econom_societi" + 0.016*"economist" + 0.013*"publish" + 0.010*"depart"

The words in these topics do not form substantive economic ideas that can be included in the relevant meta-topic or they are sufficiently broad that they could be included in all of the meta-topics (thus making their inclusion redundant,

such as in the case of generic economic language, e.g. "commod"(commodity). Word clouds showing the 100 top-weighted words in each of the 100 topics (including these excluded topics) are presented in Bowles and Carlin (2020).

### B.2.3 Replacement and perturbation

To explore the robustness of our estimated parameters, we simulate topic exclusion and replacement in the meta-topics and identify how the exclusion and replacement perturbs the location of the corpus in a given period (for different time slices) or for a given textbook.

We identify a replacement topic for each topic that makes up the microeconomics and macroeconomics meta-topics by finding the most correlated topic (using the highest Pearson correlation). A replacement topic cannot be selected from the same meta-topic group as the topic being replaced. For instance in Figure A.17, the highest correlated topic to topic 44 is topic 54. However, since topic 54 is a topic in meta-topic 0, it is excluded. The next-highest correlated permissible topic is topic 90.

The tables in Figures A.17 and A.18 show the top 3 candidate replacement topics ordered from highest to lowest Pearson correlation. The topics in bold are the ones selected, and the coloring scheme indicates the meta-topic to which the subtopic belongs. In terms of this coloring scheme, as in the previous example one would simply exclude a candidate replacement topic if it shares the same color as the topic to be replaced. The table by design has excluded instances of topics 79, 34 and 26 to make picking candidates easier.

### B.2.4 Confidence regions

To understand the dispersion of the simulated texts around the original estimate provided in the paper, we produce 95% confidence regions around the original estimate.[1]

To construct and analyze the simulated data statistically, we use Wickham (2017) for initial data work. We use `ggtern` to produce and analyze the confidence regions of the perturbed topic models (Hamilton and Ferry 2018). The confidence regions are generated based on log ratio transformations of the Mahalanobis distances generated for the points within the simplexes (see Katz and King (1999), Egozcue et al. (2003) and Filzmoser and Hron (2008)).

[1] One can easily re-specify the code in the accompanying R Markdown Document in Appendix C such that `confidencebreaks ← c(0.95)` instead become `confidencebreaks ← c(0.90)` to produce the 90% confidence regions.

# References


Bird Steven, Edward Loper and Ewan Klein. *Natural Language Processing with Python*. O'Reilly Media Inc., 2009.

Blei, David M, Andrew Y Ng, and Michael I Jordan. "Latent dirichlet allocation". *Journal of machine Learning research* 3.Jan (2003), pp. 993–1022.

Bowles, Samuel and Wendy Carlin. "What Students Learn in Economics 101: Time for a Change". *Journal of Economic Literature* (2020).

Egozcue, J. J. et al. "Isometric Logratio Transformations for Compositional Data Analysis". *Mathematical Geology* 35.3 (2003), pp. 279–300.

Filzmoser, Peter and Karel Hron. "Outlier Detection for Compositional Data Using Robust Methods". *Mathematical Geosciences* 40.3 (2008), pp. 233–248.

Hamilton, Nicholas E and Michael Ferry. "ggtern: Ternary diagrams using ggplot2". *Journal of Statistical Software* 87.1 (2018), pp. 1–17.

Harper et al., Marc. *python-ternary: Ternary Plots in Python*. Zenodo. 2015.

Katz, Jonathan N. and Gary King. "A Statistical Model for Multiparty Electoral Data". *American Political Science Review* 93.1 (1999), 15–32.

Řehůřek, Radim and Petr Sojka. "Software Framework for Topic Modelling with Large Corpora". English. In: *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. `http://is.muni.cz/publication/884893/en`. Valletta, Malta: ELRA, May 2010, pp. 45–50.

Wickham, Hadley. *tidyverse: Easily Install and Load the 'Tidyverse'*. R package version 1.2.1. 2017.