

How is the Mainstream changed?

A Topic Model insight

Michele Ciruzzi*

6th February 2023

5 1. Why studying the mainstream?

To understand the reasons of this work and, eventually, its scientific relevance a premise is needed.

Economics as discipline presents very strong social norms and a very hierarchical structure. As result, it is possible, and quite easy, to identify a small group of journals
10 in which the scholars in the top US department often publish (Dusansky & Vernon, 1998; Ellison, 2003; Heck & Zaleski, 2006; Kim et al., 2006, 2009; Card & DellaVigna, 2013; Hamermesh, 2013). Those are the same journals that are able to assure a scholar the tenure by publishing one of their papers (Heckman & Moktan, 2020).

In other words there is a group of journals (the Top5 or the Blue Ribbon Eight) which
15 are liked (and managed) by those who are liked (and cited and funded) by the discipline.

This concentration of power and, as consequence, prestige have, inevitably, create a notion of prestige in every aspect of the discipline: there are more prestigious theories, more prestigious research topics, more prestigious departments and so on.

But the step is short in the underfunded academia to make the post prestigious altern-
20 atives the only alternatives, inducing conformism and, as a matter of necessity, excluding less prestigious theories from scientific debate.

The prestigious economics is, politely, knows as *mainstream economics* or, in a more colourful way, as *orthodox economics*. The idea of an orthodoxy, and so of some com-
peting heterodoxies, move the analysis to the semantic field of religion and faith. And
25 it is wanted.

As many, mostly heterodox, scholars have pointed out many of the hypotheses which support the mainstream economics have to be accepted by faith, since there are not real-world evidences of their realism¹. The consequence is that to pursue a career in

This version is intended to be submitted to the XV ESHET Conference

*mcciruzzi@uninsubria.it - <https://orcid.org/0000-0003-1485-1204>

¹Some orthodox scholars address this criticism by not considering realism as necessary in a scientific theory, but discussing this would be a too long digression.

economics and be able to get a prestigious (and so well-funded) position, a profession of
30 faith is required to every (young) scholar².

The last piece of the puzzle I require to justify this work is the role of economics in the society.

Economics is a social science which influences the kind of politics are realized by the governments. It provides theoretical justification, it carries out the forecast on the real-
35 world effects, it influences the public debate providing the words to convey different ideas of the economy (and so of the society), helping to move the Overton's window.

In other words, Economics is a scientific discipline with a strong impact on the society, and the social and political implications of any economic theory must not be overlooked³.

So, critically studying the most common economics is necessary to understand with
40 conceptual framework the discipline is providing to the society to understand its times, to unveil which cultural and political programme it is (more or less consciously) carried out by the discipline.

Nevertheless, this work only tries to describe the mainstream economics, without arguing on the social consequences of the features highlighted. Which is still a necessary
45 first step.

2. The interpretative hypothesis

I'm clearly not the first to try to describe mainstream economics and neither the first to do it quantitatively (I review some previous attempts in the next section).

Therefore, I can climb up on the shoulder of other scholars to be guided to what I can
50 look for and which interpretative frameworks I can use.

Two qualitative observation will be the core pieces to guide this work: the *empirical turn* hypothesis and the *mainstream pluralism* hypothesis.

The first lens with which observe the mainstream economics is the progressive shift toward empirical research (Backhouse & Biddle, 2000; Backhouse & Cherrier, 2016,
55 2017; Cherrier, 2022). Neoclassical economics, which is the theoretical framework on which mainstream economics has been based in the second half of the XX century, built

²An example of this kind of "religious" reasoning is proposed by Galbraith: "Accepted in reputable market orthodoxy is, as noted, the inherent perfection of the market. The market can reflect contrived or frivolous wants; it can be subject to monopoly, imperfect competition, or errors of information, but, apart from these, it is intrinsically perfect. Yet clearly the speculative episode, with increases provoking increases, is within the market itself. And so is the culminating crash. Such a thought being theologically unacceptable, it is necessary to search for external influences—in more recent times, the downturn in the summer of 1929, the budget deficit of the 1980s, and the "market mechanisms" that brought the crash of 1987. In the absence of these factors, the market presumably would have remained high and gone on up or declined gently without inflicting pain. In such fashion, the market can be held guiltless as regards inherently compelled error. There is nothing in economic life so willfully misunderstood as the great speculative episode." (Galbraith, 1994)

³The conceptual framework provided by Keynes' theories is necessary to historically and politically understand Roosevelt and the European social state of the 50s. Similarly, Margaret Thatcher's and Ronald Reagan's economic reforms would never be realized without the theoretical background provided by Friedman.

its fortune on the ability to provide clear models expressed using a very formal flavour of mathematics. In more recent times instead, in part as a reaction to the credibility crisis arose after the financial crises of the first decade of the XXI century, econometrics is becoming the standard mathematical tool of analysis and theory-driven models are being replaced by data-driven model. This empirical turn in the scientific practices of the discipline is moving researcher away from the comprehensive economic models of the past, towards a perilous land in which the only theory is the way in which data are analysed, and every economic insight can come only from data itself.

The second lens is the mainstream pluralism hypothesis (Davis, 2006; Cedrini & Fontana, 2018; Davis, 2019a, 2019b). It has been observed that the domain of research of economics is widening and the number of different computational and mathematical tools used is increasing⁴. These two tendencies have allowed the emergence of many new fields in the discipline, which are very focused on a particular topic (like environmental or health economics) or on a particular method (like economic complexity or behavioural economics). Researchers who are specialized in one of these fields rarely engage with the other niches, creating an archipelago of unconnected islands and an appearance of pluralism in economics. The mainstream pluralism hypothesis states that this process of specialization is not developing pluralism as historically intended in economics (i.e. as a plurality of competing ontological views of the economy and the economics), but rather it is creating a plurality of loosely connected fields which shares a common ancestral set of hypothesis (the neoclassical one) and relax some of them to be able to tackle specific problems, putting themselves in continuity rather than in contrast with neoclassical economics⁵.

In the quantitative exercise that constitute the core of this paper, I try to find some evidences in favour of these two hypotheses.

3. Distant Reading and quantitative methods in the History of Economic Thought

A common problem in the study of the history of ideas (of which the history of economic thought is surely part) is the impossibility to accurately read a big number of textual documents in a reasonable time. A first, and often adopted, solution is to select a sample of relevant works, like the books of a single author or a handful of representative masterpieces of a period, and analyse only them. An alternative approach is what Franco Moretti called Distant Reading (Moretti, 2013): instead of carefully (close) reading few works, a big number of documents are summarily analysed using statistical methods and the aid of computers. Different features of a text can be analysed, but generally the focus is on the textual content of the document (recovered by a digital version or

⁴This is also a consequence of the new data-driven methodologies, which are able to deal with a wider range of hypotheses than the neoclassical calculus-based approach (see Cherrier, 2022)

⁵In this sense they can view as subfields of a wider “post-neoclassical” economics, which aims to improve (or save) some parts of the neoclassical approach and legacy, rather than rebuild the discipline on completely different assumptions like the heterodoxies aim to do.

through OCRing) or some metadata (title, citations, authors' list).

95 The kind of data used and of research questions tackled, move Distant Reading close to bibliometrics and text mining studies. The fil rouge which ties together Distant Reading studies is their motivation: overcome with quantitative techniques the impossibility to (close) read thousands of document to reconstruct the history of some idea.

100 The history of neoclassical economics as mainstream economics covers at least thirty years and many journals. Moreover, its influence has been systemic, it has shaped the social norms of the discipline as a whole, and sampling some "representative master-pieces" (assuming the existence of a good enough criterion to select them) would not be able to return all the shades I'm looking for.

105 Since those are not groundbreaking truths, this is not the first attempt to apply quantitative methods to the history of economic thought in order to get a broad picture of the evolution of the discipline in the last half-century.

In 2018, an entire special issue of the *Journal of Economic Methodology* was dedicated to the topic (Cherrier & Svorenčík, 2018; Edwards et al., 2018), but the trail goes way back in the past (Backhouse et al., 1997). The following is a brief review of some interesting works which address similar questions to mine.

110 Kelly and Bruestle (2011), using JEL codes⁶, observe a reduction of microeconomics and macroeconomics papers and an increasing of finance and development economics papers. Moreover, the top journals over-represent microeconomics, mathematical and quantitative methods and labour economics.

115 Hamermesh (2013) shows that the quota of empirical articles in some top journals is increasing.

Card and DellaVigna (2013) observe a bias toward theoretical microeconomics papers (and to a lesser degree applied microeconomics and macroeconomics ones) in top journals, which is reducing in the last years as empirical papers become more frequent.

120 Claveau and Gingras (2016) clusterise the citation network to highlight seven subfields which are present during the entire studied period (1956-2014): econometrics, financial economics, development and trade economics, industrial economics and game theory (the largest subfield after 2000), history of economics and macroeconomic policies, macroeconomics and monetary policies, labour economics. In addition, they observe some clusters only at the beginning of the sequence, which are not preserved in subsequent years, and one cluster (environmental economics) which appears toward the end of the studied period. The increase in published papers makes the recognition of clusters more difficult at the end of the series, and it is possible that some clusters went undetected in the most recent years. Finally, they observe that it is typical for economics scholars to publish in more than one subfield.

130 Angrist et al. (2017) found that the share of microeconomics (mostly theoretical) papers is steadily increasing, as well as the share of development economics ones, while typical applied microeconomics fields (like labour economics or industrial organization)

⁶The *Journal of Economic Literature* (JEL) codes are a widely used taxonomy system to classify economics subfields, but their reliability is discussed, since their use appears to be highly subjective (Cherrier, 2017; Kosnik, 2018).

see their shares reducing over time. Moreover, they describe the increase of empirical papers as a methodological shift inside the subfields, more than a change in research topics.

Ambrosino et al. (2018) explore the whole JSTOR database on a decade-per-decade basis, using a Topic Model, going deep in describing the possibilities of this kind of approach.

Fontana et al. (2019) observe an increasing usage of treatment effects models at expense of time series, a growth of development economics, game theory and education economics and a decline of industrial economics and macroeconomics. This study uses LDA on papers' full-text.

Montesinos and Brice (2019) measure the most frequent words in the title of the most cited papers of the last seventy years, highlighting a shift from "theory" to "evidence" (and so empirical research).

Except for Ambrosino et al. (2018), all the studies listed focus on a small subset of the economic journals, usually the "Top5"⁷ or the "Blue Ribbon Eight"⁸.

4. The choice of the data

Going back to the aim of this paper, the two question to be answered pertain the evolution of economics at least since the seventies, but, ideally, since the post-war period to be able to observe both a pre- and a post-neoclassical period.

Moreover, the focus is on the mainstream economics, which allows a strong hypothesis in choosing the data: the journals perceived by the research community as the top journals are a representative sample of what is mainstream in the community, i.e. what is considered shared knowledge, high quality research and a signal of a path to follow.

To select the top journals one can rely on the "common knowledge" represented by the Blue Ribbon Eight, and the rankings published throughout the history of the discipline or looking at some bibliometric indicator. The choice is obviously restricted by the availability of different data type.

There are three potential data types to be used: citation network, articles' metadata (particularly JEL codes) and the articles' own text (or their abstract).

I exclude the JEL codes because they are highly subjective and can change multiple times during the editorial process (Kosnik, 2018), and they are used to signal the (desire of) belonging to a scientific community which can be interested in reading the paper. They are, indeed, more a way to observe where the scholars desire to be, and so the relations inside the discipline, rather than the evolution of the ideas.

Similar reasoning can be done with the citation network. Since, especially in more recent times, citations have become a currency with which buys funding and (tenured)

⁷The *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics* and *Review of Economic Studies* (Heckman & Moktan, 2020).

⁸The Top5 plus *International Economic Review*, *Journal of Economic Theory* and *Review of Economics and Statistics* (Dusansky & Vernon, 1998).

positions. According to the Goodhart's law, they have become a poor metrics to measure
170 the relevance of a paper, since the scholars have started to play the game.

This reasoning leads me toward the use of the text of the articles. This strategy has an
obvious downside (textual documents are more difficult to be analysed than a network
or a taxonomy) but reduces some disadvantages of the two previous ones. To be honest,
the text of a document is still influenced by the community in which a researcher put
175 himself (the same idea is expressed differently by scholars of different schools of thought)
but it is harder to be gamed (and, sincerely, maybe it is not a real downside). Finally,
the selection bias of the authors in writing the abstracts can be avoided by using the full
text of the articles, thanks to JSTOR.

This choice rises a problem: not every journal is indexed on JSTOR. Particularly, of
180 the Blue Ribbon Eight the *International Economic Review* and the *Journal of Economic
Theory* are missing. I believe this absence is not significant, since the Top5 are still
available, those two have slightly worse bibliometric indicators than the others and the
sample is still pretty large.

5. A strategy of analysis: the Topic Models

185 5.1. The choice of the algorithm

6. Preliminary steps

6.1. Corpus cleaning

6.2. Model selection, validation and description

7. Interpreting the model

190 7.1. The empirical turn

7.2. Mainstream pluralism?

8. Limitations

8.1. Choosing the corpus

8.2. The quest for a good Topic Modelling algorithm

195 8.3. Interpreting the topics

8.4. Assess the robustness

9. What's next?

9.1. Rephrasing the questions for a theory-driven Distant Reading

9.2. Delve into the hierarchy

200 9.3. Enhancing the database

Acknowledgments

References

- Ambrosino, A., Cedrini, M., Davis, J. B., Fiori, S., Guerzoni, M., & Nuccio, M. (2018). What topic modeling could reveal about the evolution of economics. *Journal of Economic Methodology*, 25(4), 329–348. <https://doi.org/10/gf359h>
- 205 Angrist, J., Azoulay, P., Ellison, G., Hill, R., & Lu, S. F. (2017). Economic Research Evolves: Fields and Styles. *American Economic Review*, 107(5), 293–297. <https://doi.org/10/gdgx8x>
- Backhouse, R., & Biddle, J. (2000). The Concept of Applied Economics: A History of Ambiguity and Multiple Meanings. *History of Political Economy*, 32, 1–24. https://doi.org/10.1215/00182702-32-Suppl_1-1
- 210 Backhouse, R., & Cherrier, B. (2016, October 1). 'It's Computerization, Stupid!' The Spread of Computers and the Changing Roles of Theoretical and Applied Economics (SSRN Scholarly Paper No. ID 2781253). Social Science Research Network. Rochester, NY. <https://doi.org/10.2139/ssrn.2781253>
- 215 Backhouse, R., & Cherrier, B. (2017). The Age of the Applied Economist: The Transformation of Economics since the 1970s. *History of Political Economy*, 49, 1–33. <https://doi.org/10.1215/00182702-4166239>
- Backhouse, R., Middleton, R., & Tribe, K. (1997). 'Economics Is What Economists Do', But What Do the Numbers Tell Us? <https://doi.org/10.2139/ssrn.2942733>
- 220 Card, D., & DellaVigna, S. (2013). Nine Facts about Top Journals in Economics. *Journal of Economic Literature*, 51(1), 144–161. <https://doi.org/10.1257/jel.51.1.144>
- Cedrini, M., & Fontana, M. (2018). Just another niche in the wall? How specialization is changing the face of mainstream economics. *Cambridge Journal of Economics*, 42(2), 427–451. <https://doi.org/10/gfxn7t>
- 225 Cherrier, B. (2017). Classifying Economics: A History of the JEL Codes. *Journal of Economic Literature*, 55(2), 545–579. <https://doi.org/10/gdh6pn>
- Cherrier, B. (2022, August 12). The Price of Virtue: Some Hypotheses on How Tractability Shaped Economics. Retrieved August 17, 2022, from <https://papers.ssrn.com/abstract=4189091>
- 230 Cherrier, B., & Svorenčik, A. (2018). The quantitative turn in the history of economics: Promises, perils and challenges. *Journal of Economic Methodology*, 25(4), 367–377. <https://doi.org/10.1080/1350178X.2018.1529217>
- Claveau, F., & Gingras, Y. (2016). Macrodynamics of Economics: A Bibliometric History. *History of Political Economy*, 48(4), 551–592. <https://doi.org/10/gf4g58>
- 235 Davis, J. B. (2006). The turn in economics: Neoclassical dominance to mainstream pluralism? *Journal of Institutional Economics*, 2(1), 1–20. <https://doi.org/10/bzn5sm>
- Davis, J. B. (2019a, January 28). Economics and Economic Methodology in a Core-Periphery Economic World. <https://doi.org/10.2139/ssrn.3324569>
- 240 Davis, J. B. (2019b). Specialization, fragmentation, and pluralism in economics. *The European Journal of the History of Economic Thought*, 26(2), 271–293. <https://doi.org/10.1080/09672567.2018.1555604>
- Dusansky, R., & Vernon, C. J. (1998). Rankings of U.S. Economics Departments. *Journal of Economic Perspectives*, 12(1), 157–170. <https://doi.org/10.1257/jep.12.1.157>

- 245 Edwards, J., Giraud, Y., & Schinckus, C. (2018). A quantitative turn in the historiography of economics? *Journal of Economic Methodology*, 25(4), 283–290. <https://doi.org/10.1080/1350178X.2018.1529133>
- Ellison, G. (2003, February 27). The Slowdown of the Economics Publishing Process. Retrieved February 6, 2023, from <https://papers.ssrn.com/abstract=327004>
- 250 Fontana, M., Montobbio, F., & Racca, P. (2019). TOPICS AND GEOGRAPHICAL DIFFUSION OF KNOWLEDGE IN TOP ECONOMIC JOURNALS. *Economic Inquiry*, 57(4), 1771–1797. <https://doi.org/10/ggz77h>
- Galbraith, J. K. (1994). *A short history of financial euphoria*. Whittle Books in association with Penguin Books.
- 255 Hamermesh, D. S. (2013). Six Decades of Top Economics Publishing: Who and How? *Journal of Economic Literature*, 51(1), 162–172. <https://doi.org/10.1257/jel.51.1.162>
- Heck, J. L., & Zaleski, P. A. (2006). The most frequent contributors to the elite economics journals: Half century of contributions to the “Blue ribbon eight”. *Journal of Economics and Finance*, 30(1), 1–37. <https://doi.org/10/drboxd5>
- 260 Heckman, J. J., & Moktan, S. (2020). Publishing and Promotion in Economics: The Tyranny of the Top Five. *Journal of Economic Literature*, 58(2), 419–470. <https://doi.org/10/ggz77h>
- Kelly, M. A., & Bruestle, S. (2011). TREND OF SUBJECTS PUBLISHED IN ECONOMICS JOURNALS 1969-2007. *Economic Inquiry*, 49(3), 658–673. <https://doi.org/10/cgw9zz>
- 265 Kim, E. H., Morse, A., & Zingales, L. (2006). What Has Mattered to Economics Since 1970. *Journal of Economic Perspectives*, 20(4), 189–202. <https://doi.org/10.1257/jep.20.4.189>
- 270 Kim, E. H., Morse, A., & Zingales, L. (2009). Are elite universities losing their competitive edge? *Journal of Financial Economics*, 93(3), 353–381. <https://doi.org/10.1016/j.jfineco.2008.09.007>
- Kosnik, L.-R. (2018). A SURVEY OF JEL CODES: WHAT DO THEY MEAN AND ARE THEY USED CONSISTENTLY?: A SURVEY OF JEL CODES. *Journal of Economic Surveys*, 32(1), 249–272. <https://doi.org/10/gf4g6n>
- 275 Montesinos, H., & Brice, B. (2019, November 25). *The Era of Evidence*.
- Moretti, F. (2013). *Distant reading*. Verso.

A. Results from different random seeds