Journal of Economic Issues

ISSN: 0021-3624 (Print) 1946-326X (Online) Journal homepage: http://www.tandfonline.com/loi/mjei20

The Power of Scientometrics and the Development of Economics

Matthias Aistleitner, Jakob Kapeller & Stefan Steinerberger

To cite this article: Matthias Aistleitner, Jakob Kapeller & Stefan Steinerberger (2018) The Power of Scientometrics and the Development of Economics, Journal of Economic Issues, 52:3, 816-834, DOI: 10.1080/00213624.2018.1498721

To link to this article: https://doi.org/10.1080/00213624.2018.1498721

Published online: 28 Aug 2018.

JOURNAL OF ECONOMIC ISSUES
Volume LII No. 3 September 2018
DOI 10.1080/00213624.2018.1498721

The Power of Scientometrics and the Development of Economics

Matthias Aistleitner, Jakob Kapeller, and Stefan Steinerberger

Abstract: Citation metrics and its related indices and rankings have become increasingly important in the evaluation of research. Such indices are part of a more general tendency aiming for the simplification of complex and interconnected phenomena through quantification. The purpose of our contribution is to analyze the impact of such quantitative indices on the further development of science, with a special emphasis on economics. In this case, we observe a multitude of interesting effects on both the level of individual scientists and the global development of the discipline.

Keywords: academic reproduction, citation metrics, paradigms, pluralism, quantification, reactivity

JEL Classification Codes: A14, B50

Understanding the impact of scientometrics on the development of academic disciplines is a complex problem of great current significance. Its relevance stems from a recent trend to introduce numerical measurements of scientific performance in order to evaluate research activities and facilitate comparisons on various levels, e.g., across different researchers, institutions, or publication outlets. These comparisons usually take the form of rankings that are intended to "measure" the "quality" of universities, scientists, scientific articles, and journals on a single scale.[1] This development also points to a methodological shift inside scientometrics that has its conceptual origins in an interpretative analysis of scientific

Matthias Aistleitner is a researcher in the Institute for Comprehensive Analysis of the Economy (ICAE) at Johannes Kepler University in Linz, Austria. Jakob Kapeller is a post-doctoral researcher in ICAE and the Department of Economics at the same university. Stefan Steinerberger is an assistant professor in the Department of Mathematics at Yale University, USA. *This work was funded by the Institute for New Economic Thinking (INET) under grant numbers INO1500038 and INO1500039.*

[1] Examples for internationally known rankings in academia are the *Times Higher Education World Reputation Ranking* and the *QS World University Ranking*. In the field of journal evaluation rankings, the *Journal Citation Reports* from *Thomson Scientific* stand out.

communication aimed at understanding the discursive properties of academic publishing (e.g., de Solla Price 1965; Rip and Courtial 1984). This approach, sometimes dubbed "cognitive scientometrics," is increasingly making way for new forms of "evaluative scientometrics" that try to define indicators of research quality based on an analysis of citation frequencies inside the scientific literature or through peer review and survey instruments (Adler, Ewing and Peter Taylor 2008; Mingers and Leydesdorff 2015).

The increasingly prominent role of quantitative evaluation in academia can be interpreted as part of a more general social trend toward the numerical assessment of social issues. One major historical impetus of this process of an increased "quantification of social phenomena" (Espeland and Stevens 2008, 401) is to make social conditions politically predictable and controllable. An archetypical example in this context is the development of the system of national accounts (SNA) and its corresponding parameters like the gross domestic product (GDP), which serves the purpose of assessing a nation's economic activities and provides a yardstick for measuring economic development. Aside from administrative interests, indicators like GDP were also developed to serve scientific purposes since they facilitate the quantification and measurement of theoretically postulated aggregate concepts like GDP, inflation, or capital.

From a theoretical perspective, one can understand the "quantification of social phenomena" as a "general sociological phenomenon" that covers "the production and communication of numbers" and deals with "regimes of measurement" (cf. Espeland and Stevens 2008, 401). From this point of view, numbers fulfill two functions. For one, they serve as symbolic placeholders representing single entities or events (e.g., "9/11"). Such placeholders do indeed contribute to the "quantification of social phenomena" as a general tendency, but do not yet establish any specific "regimes of measurement" by themselves. However, numerical information also directs attention toward the relative properties of various entities. For example, when assigning numerical values to different entities of the same class, numerical information allows for creating an ordering that has the double function of *unifying* different objects across a uniform scale, which simultaneously makes it easier to *distinguish and differentiate* between these objects (cf. Espeland and Stevens 2008).

Recently, this process of quantification has gained additional momentum within science due to the introduction of regimes of measurement in the sphere of research evaluation. The interplay between supply and demand for evaluation of academic performance has the potential to create ever new tools for and facets of numerical evaluation procedures in academia. While the rise of evaluative scientometrics within academia is clearly part of a broader phenomenon, we focus more particularly on its role in and impact on the field of economics. Our main contribution is in providing a theoretical embedding of the notion of quantification in a more general account of scientific development and in fostering a better

understanding of its impact on the development of specific research fields and the behavior of individual researchers. Our resulting argument not only provides an integrated assessment of the "power of evaluative scientometrics," but also supports this assessment with small case studies focused on economics that are suitable for empirically illustrating the underlying argument.

We focus on three specific aspects. First, scientometric routines in research assessment are always embedded in conventional patterns of academic reproduction. Thus, the intensified use of these routines has to be understood against this backdrop, which is our main focus in the second section of the article. Second, scientometric evaluation is based on mechanical procedures that are easily reproducible and, therefore, cause incentives for strategic behavior among researchers. In the third section of the article, we discuss some aspects of reactivity with a special focus on the impact of reactive evaluation routines on the citation behavior of economists. Third, we consider the special case of economics as a "contested discipline" (Lee and Elsner 2011). Unlike other social sciences, economics has only one dominant paradigm: the axiomatic core of the so-called mainstream economics is largely based on neoclassical economic principles. These principles guide the majority of the economic scientific community. Their conceptual predominance remains largely unchallenged except for the protest of a small minority of heterodox economists (Dobusch and Kapeller 2012a) and, more recently, also students of economics (ISIPE 2014). In this context, our fourth section selectively summarizes and extends past approaches analyzing interparadigmatic engagement in economics, and discusses the impact of quantitative evaluation regimes against the backdrop of paradigmatic divisions in economics. The final section summarizes our main arguments.

Patterns of Academic Reproduction: From Matthew Effects to Path-Dependency

> *The competition between paradigms is not the sort of battle that can be resolved by proofs.*
>
> *– Kuhn (1962, 148)*

One classical finding of scientometric research is that academic attention—mostly measured on the basis of citation frequencies—is highly skewed. The distribution of attention, influence, and prestige among a given quantity of single researchers or research articles follows a simple underlying structure: most researchers or research articles receive very little or no attention, whereas a few researchers or research articles receive a great deal of attention. From a formal point of view, this specific distribution of attention, influence, and prestige in science has similar characteristics to, for instance, the distribution of wealth or the attractiveness of websites and follows a "power law" Pareto/Zipf distribution at the top. Such types of distributions are common within social contexts. For example, the population of cities, the

number of received phone calls, the number of words used in a text, or the sales figures of book titles tend to follow such a power law (Newman 2006).

One main feature of this kind of distribution is that just a few elements at the top of the respective distribution collect a disproportionately large share of the variable of interest. One example, derived from the population structure of Germany, is that the inhabitants of only four different cities comprise about 10 percent of the German population. In the scientific discourse, a majority of inner-academic attention focuses only on a small fraction of the respective scientific literature.[2] Figure 1 illustrates the characteristics of such distributions and shows three potency-distributed measures in descending order: the population of the 82 biggest cities in Germany, the wealth distribution in Austria, and the distribution of citations to articles published in the *American Economic Review* between 1981 and 1985.[3]

To better illustrate the properties of these samples, we compare these three distribution to that of the birth weight of newborn babies, which is a normally distributed random variable. Figure 1 reminds us of an important property of power-law distributions: namely, that the differences between median, mean, and maximum values are significantly larger than they would be for a quantity that follows a normal distribution. Indeed, for some Pareto distributions with heavy tail the notion of mean is not even defined. The simple interpretation of this pattern is that, in the case of power-law distribution, more extreme values occur and hence these extreme values are quantitatively more important than in the case of normally distributed properties. For distributions following a power law, we observe a remarkable difference between median and mean as a large amount of the total quantity is concentrated in the edge of the distribution.

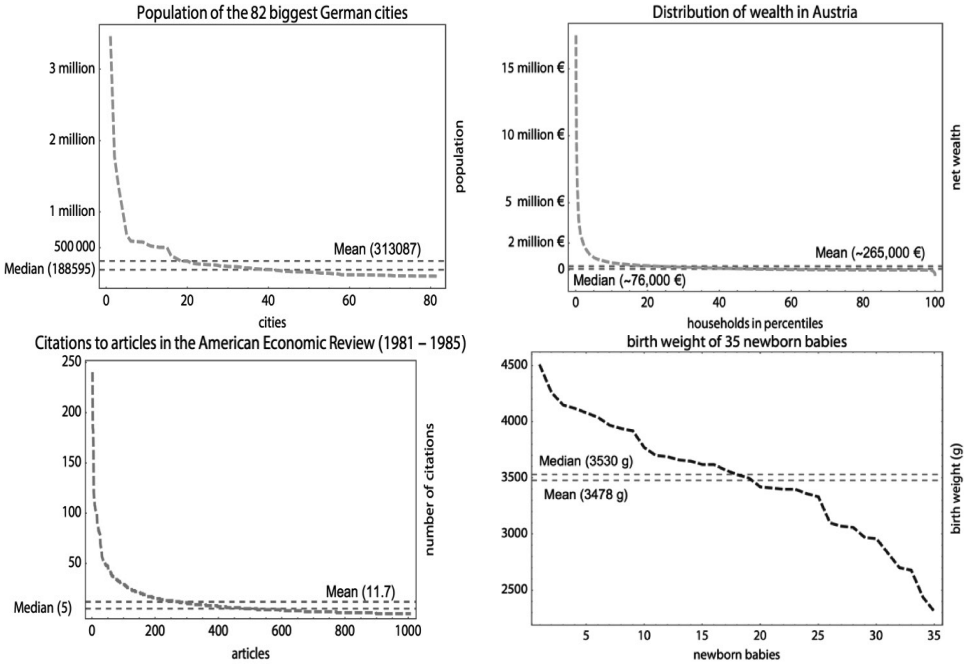
In 1965, Derek J. de Solla Price undertook the first systematic study on the distribution of attention in scientific discourse. Based on an analysis of citation data, he (1965) postulated that current generations of scientific results only refer to a small number of past contributions. However, the mechanism from which this uneven distribution emerges was not apparent at first and was only later clarified by Robert K. Merton's classical work on the "Matthew Effect in Science" (1968). Merton (1968) postulated that the acquisition of prestige and attention given to scientific work is closely correlated with the amount of attention acquired in the past. He based his argument on citation data as well as interviews conducted with Nobel Prize laureates. The implicit logic of this mechanism—"whosoever has, will be given more"—today is aptly called the "Matthew Effect" in reference to the passage in the gospel of Matthew.[4] "[T]he Matthew effect consists in the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable

[2] For economists, such distributions have been known for a long time, especially in the context of the analysis of income and wealth distributions, where they are discussed on the basis of the Pareto formula.

[3] The citations assigned to these articles cover the period from 1981 to 1990.

[4] "For to everyone who has will more be given, and he will have abundance; but from him who has not, even what he has will be taken away" (Matthew 25:29).

Figure1. Comparing Four Distributions



Source: Authors' own work.[5]

repute and the withholding of such recognition from scientists who have not yet made their mark" (Merton 1968, 58).

Matthew effects can be characterized more generally as self-reinforcing effects that, in the special case of Matthew effects in science, have the following structure: If an author/article *x* is cited, then *x* will become a more attractive point of reference in the future. Self-reinforcing effects appear in similar form in a series of social contexts that, for instance, have been discussed in detail in economic literature about path-dependency (Dobusch and Schüßler 2013; Sydow, Schreyögg and Koch 2009). The main thesis of this theoretical strand is that the establishment of technical, organizational, or social standards can lead to self-reinforcing effects which enable the relative or absolute dominance of these standards. The former kind of dominance (relative dominance) can be observed in most cases in the form of a distribution of power-law type, we discussed above.

Classical applications of path-dependency theory can be found in examples, such as the persistence of specific technological standards (e.g., the QWERTY keyboard design; see

[5] Data sources: The population of the German cities is taken from the Mathematica10 database. The estimate of property assets is based on data from the Household Finance and Consumption Survey (HFCS) of the Austrian national bank. The number of citations of the American Economic Review is retrieved from the Web of Science database and the birth weight data of the newborns represents the collected number of births in the first two weeks of February 2015, which took place in an Austrian hospital. All data are available from the authors upon request.

David 1985), the evolution of monopolies in software markets (Shapiro and Varian 1999), the relative attractiveness of file-sharing networks and social networking sites, or the cluster structure of the high-tech industry (Arthur 1994). Moreover, the development of the Pareto distribution as a "natural" form of wealth distribution can be simulated in a theoretical model by integrating self-reinforcing effects (Levy and Levy 2003). Self-reinforcing mechanisms can be empirically identified and range from direct network effects (the more a standard is used, the more attractive it becomes) over learning effects (the better a standard has been understood, the more attractive it becomes) to indirect network effects (if a certain standard is a requirement for the use of other products, this standard will become more attractive).

Based on the examples we presented above, the Mertonian Matthew effect can be interpreted as a central mechanism of academic reproduction and posits—in terms of path-dependency theory—a so-called "direct network effect": the attractiveness of a point of reference correlates with the number of past references. The consequence of such self-reinforcing effects is unevenly distributed attention inside the scientific discourse, whereby a few contributions receive disproportionately large attention, while many others remain largely unnoticed. It is indeed observed that most scientometric indicators follow a power-law distribution.

These patterns of academic reproduction do not only hold for the level of individual researchers and articles, but also apply to scientific institutions (e.g., academic journals) as well as specific schools of thought and paradigmatic traditions (in the sense of Kuhn 1962). Such a perspective, which applies the basic idea of Matthew effects to specific fields and approaches instead of single authors and contributions, is exemplified by the simulation-based study of John D. Sterman and Jason Wittenberg (1999) who provide a detailed account of the path-dependent properties of scientific evolution. For example, Sterman and Wittenberg (1999) found that the intrinsic quality of the core ideas of a single paradigm is only of minor importance for its success, which provides a theoretical rationale for non-linear developments in the field of scientific knowledge and stands in line with Thomas Kuhn's historical observations. A prime example in this context is the advent of Copernican theory that, in its beginning, lagged behind the geocentric view in terms of its empirical explanatory power and precision. The main reason for the initial superiority of the Ptolemaic theory was that it could employ a sophisticated theoretical apparatus with numerous correction terms (the so-called "epi-cycles"), introduced to improve the empirical performance of the underlying models.

For the purpose of this article, it seems promising for us to take a closer look at the structure of Sterman and Wittenberg's model, in which three essential positive feedback loops emerge. These feedback loops stabilize the persistence of paradigms and thereby provide a theoretical justification of the Kuhnian observation by means of path-dependence

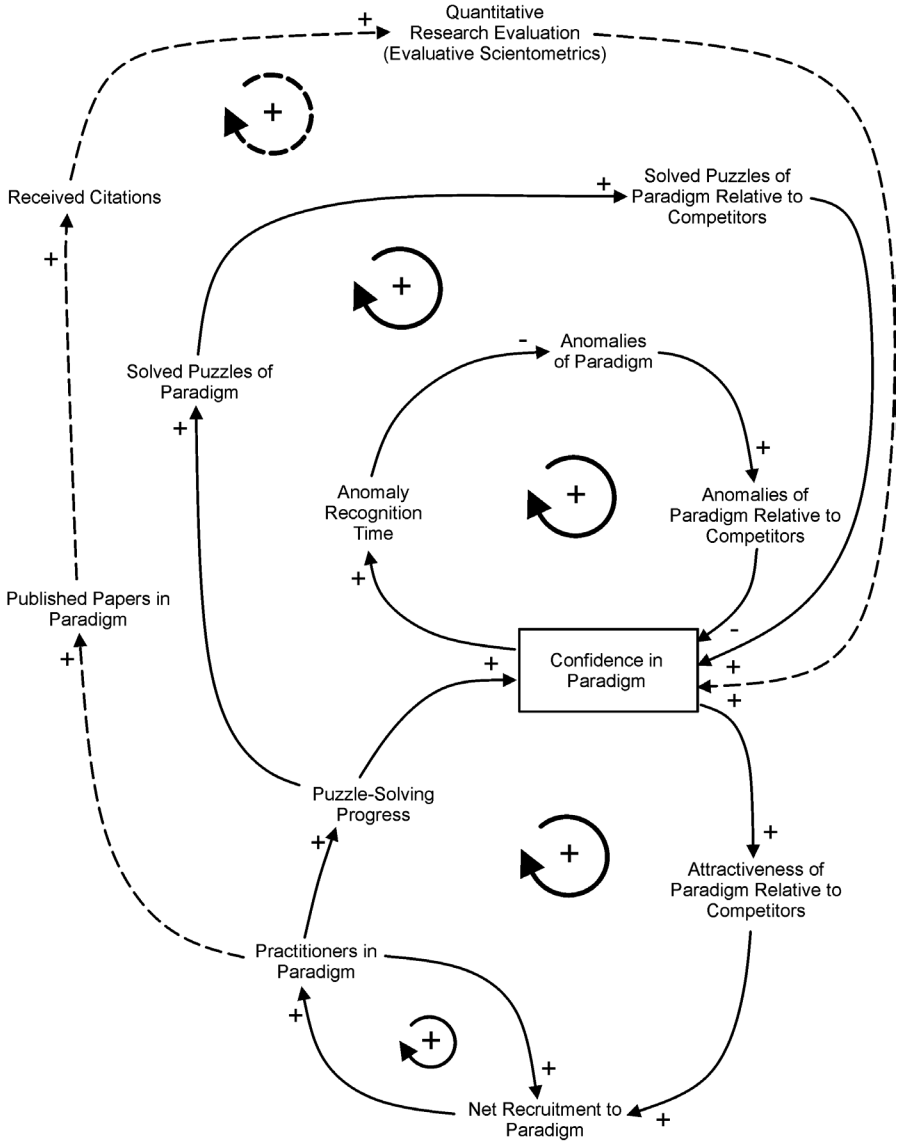
theory. The feedback loops take the form of direct network effects and refer to the academic labor market, the perceived relative explanatory power of paradigms, and the role of obvious anomalies. In all three cases, a greater number of practitioners within a paradigm leads to self-reinforcing effects because the number of recruitments, solved problems, and rationalized anomalies is proportional to a paradigm's size. In turn, these size-dependent factors contribute in a positive way to the attractiveness of the paradigm and, therefore, stabilize the dominance of established patterns of thought in the sense of a path-dependency of the paradigmatic development. "The prevalence of positive feedback processes in paradigm development means that the evolution of the system as a whole is strongly path-dependent" (Sterman and Wittenberg 1999, 333).

Building on this argument (also summarized in Figure 2), we now sketch a theoretical argument contextualizing the advent of evaluative scientometrics within the past twenty years on academic reproduction.[6] Evaluative scientometrics tries to measure influence inside the sciences by determining standardized citation frequencies (usually the number of citations received by an author, an article, or a journal), which in turn is interpreted as a measure of quality. The self-referential logic of this process is striking since quality is essentially equated with influence and impact. Measuring impact then serves as a basis for institutional evaluation that further redistributes influence inside the scientific community. In this context, the implementation of ranking systems in order to evaluate research performance leads to a further concentration of academic attention by attributing a high level of quality to those authors, articles, and research fields on which the initial level of attention has already been high (Dobusch and Kapeller 2009). Accordingly, a fourth feedback loop can be added to the model of Sterman and Wittenberg that we formulate in the following way: the bigger a paradigm or research field is, the higher the amount of received citations in this field will be—as the saying goes, "big is beautiful." Finally, this number of received citations is used as a rarely questioned hallmark of scientific quality and, therefore, further improves the attractiveness of the respective paradigm (see Figure 2).

This additional feedback loop appears in the form of an indirect network effect. The establishment of scientometric indicators as a standard in terms of scientific quality evaluation favors "bigger" fields of research and makes them more attractive to those researchers, who have internalized the ruling quality standards of scientific evaluation. While these rather general arguments on additional feedback loops in academic reproduction are rather difficult to empirically illustrate, it is possible to make a clear-cut argument

[6] Although citation data has already been widely used in the twentieth century to assess the relative impact of academic literature (in economics, already the traditional "Diamond list' of distinguished economics journals, published in Diamond [1989], was based on citation data), its direct institutional impact has strongly increased in the last 15 years due to the introduction of various means of quantitative research evaluation (e.g., Lee 2007) and the associated emergence of popular journal rankings, most notably ISI's/*Thomson Scientific*'s *Journal Citation Reports* (starting in 1999).

Figure 2. The Role of Evaluative Scientometrics as Additional Feedback Loop (Dashed Line) in Stabilizing a Scientific Paradigm



Source: Authors' own work based on Sterman and Wittenberg (1999, 333).

on the expected effects of this additional feedback loop of rankings—and the associated visibility—on academic journals. The inclusion of a journal in a ranking should boost

Table 1. Aggregated Submission Data from Three Economic Journals Joining the JCR (2010–2011)

Average number of annual submissions before inclusion into the JCR	Average number of annual submissions after inclusion into the JCR	Average growth in annual submissions after inclusion into the JCR
87 submissions/year	156 submissions/year	78.8 percent

its visibility and lead to an increase in attractiveness for potential authors, i.e., result in an increase in received submissions. Regrettably, submission data is hardly shared and often treated confidentially by editors and/or publishers. Nonetheless, we managed to acquire submission data for three anonymous economic journals of comparable size and character, which joined the most important ranking of journals—Thomson Scientific's *Journal Citation Reports* (JCR)—in the years 2010/2011.[7] Table 1 shows their aggregate average submission numbers in the years before and after inclusion in the ranking.

While Table 1 gives a first indication as to the impact of the indirect network effect introduced by evaluative scientometrics, it focuses on special cases: newcomers, so to say, for whom the inclusion in the JCR obviously implies a boost in attractiveness. For established journals, institutions, and paradigms, the very same effect leads to a reinforcement of inherent path-dependencies by rewarding the already rewarded and thereby shedding light on those who are already visible (see also Demange 2012). Thus, it is possible to state our first finding regarding the power of (evaluative) scientometrics as it reinforces existing patterns of academic reproduction by increasing the number of self-reinforcing feedback loops operating in the distribution of academic attention and interests.

Evaluation Routines and the Role of Reactive Measurement Procedures

There are many quantitative evaluation methods. Most of them are based on conventional empirical techniques, such as surveys (e.g., the peer-review process for journal evaluation of the German Academic Association for Business Research[8]) or counting event frequencies (e.g., citation frequencies). In the context of empirical social research, such techniques are required to conform to certain minimal methodological requirements. Among these minimal requirements are the validity (which means really measuring what is meant to be measured) and reliability (repeated measurements should lead to similar results) of measurement techniques. Both aspects require that empirical measurement procedures are non-reactive, which means that their application must not influence the observed behavior.

This idealized methodological viewpoint is widely contradicted by the fact that the

[7] We obtained this data through personal correspondence with a series of editorial offices.
[8] See also http://vhbonline.org/service/jourqual/vhb-jourqual-3/.

reactivity phenomenon occupies a central role in inner scientific evaluation routines. Reactivity means that the application of an investigation instrument can lead to a change in behavior of the observed subjects. In the social sciences, reactivity is primarily seen as a methodological problem that may cause biases if data is collected repeatedly. In the case of methodologically guided evaluations, this bias translates into a problem of individuals anticipating the specific evaluation criteria, which leads to a change in behavior in accordance to the criteria imposed for purposes of evaluation. A significant contribution to understanding the consequences of reactive measurement procedures in scientific evaluation comes from the work of Wendy N. Espeland and Michael Sauder (2007). In a comprehensive study about the effects of a law school ranking introduced by the magazine *U.S. News*, they find that evaluative measures, especially rankings, can have a strong influence on the social environment as well as on single actors in the scientific community.

Specifically, the authors show that behavior and perception of relevant agents in the U.S. law school context (deans and faculty, students, public institutions, donors, etc.) is influenced by two central mechanisms of reactivity: *self-fulfilling prophecies* and *commensuration*. The first mechanism, *self-fulfilling prophecies*, means that the law school ranking leads to a behavioral change that further reinforces and polarizes the ranking position: the ranking assumes a self-affirming character. The ordering of the law schools according to a ranking score signalizes decisive differences, which induces a behavioral change of the social environment (students, public, donors, etc.) and, furthermore, leads to a reinforcement of the respective ranking position or trend (downward or upward) in the ranking process. For example, former ranking scores are not only used to determine the allocation of financial resources, but also impacts the peer review process itself when former ranking positions are interpreted as indicators of quality by reviewers. Toward the lower end of a ranking list, the power of such a self-fulfilling prophecy can lead to a downward spiral: the lower the ranking position, the lower the equipment with financial resources through external financiers (which play a central role in the U.S. higher education system), the harder it becomes to move up or even maintain one's position in the ranking. For law schools, the internalization of such factors leads to a stronger focus on aspects that influence the position in the ranking list. At the same time, aspects that are irrelevant for the ranking are neglected (Espeland and Sauder 2007, 11-14). "Rankings create self-fulfilling prophecies by encouraging schools to become more like what rankings measure, which reinforces the validity of the measure" (Espeland and Sauder 2007, 15).

The second mechanism, *commensuration*, is characterized by a transformation of human cognition, caused by a numerical-competitive framing of a complex social object. Here, the conceptual clarity of rankings plays a central role. It suggests that the comparative analysis of the relative performance of educational institutions is achieved by means of a down-to-earth

and trustworthy yardstick. As Espeland and Sauder (2007) put it:

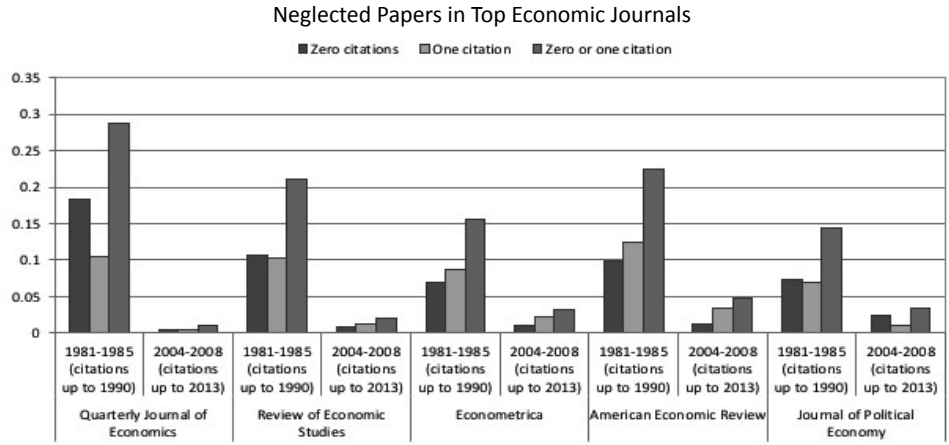
> [Commensuration] changes the locus and form of attention, both creating and obscuring relations among entities. Commensuration is characterized by the transformation of qualities into quantities that share a metric, a process that is fundamental to measurement. (Espeland and Sauder 2007, 16)

In the context of the law school ranking, the *commensuration* mechanism is triggered by the hierarchical relationship that emerges from an explicit comparison of scores. In the context of economics, journal rankings serve as main reference points for research evaluation (Bloch 2010; Corsi, D'Ippoliti and Lucidi 2010; Lee 2007) and thus as the main medium for commensuration making the journals the prime level for comparisons of all sorts. "By simultaneously unifying and distinguishing objects … rankings classify, reward and punish, and organize interventions" (Espeland and Stevens 2008, 416). In Espeland and Stevens's case, the imposition of a ranking leads to a series of specific patterns of action on the side of the law schools, like increasing expenses for marketing to raise the chance of successful future peer review processes,[9] creating specific administrative departments to obtain information about the employment status of graduates or lowering "acceptance rates" to signal selectivity to external assessors. As students' test results are relevant for the ranking, many law schools increased the number of merit-based scholarships in order to attract better students and decreased the importance of other evaluation criteria when selecting students, and so on. In short, the *U.S. News Ranking proved to be highly reactive.*

Following this logic, we would expect ranking-based evaluation procedures in economics to induce similar systematic incentives. Since journal rankings play a core role in economics and provide an arena for the mechanisms of commensuration and self-fulfilling prophecies, such effects should eventually materialize on the level of economic journals. We would expect reactive effects induced by the increasing visibility and importance of journal rankings to affect journals more strongly than single authors or articles, as journals reside in the focal point of the rankings in question. This argument implies a transmission of prestige toward scientific journals. If correct, *all* articles of a highly ranked journal should become attractive, simultaneously leading to a more balanced distribution of citations within a highly ranked journal. Figure 3 shows that this expectation indeed holds for the case of economics by means of a period-based comparison of citations to articles in five high-quality economic journals. We compare the number of articles receiving at most one citation within a fixed time period (articles published in 1981–1985 and 2004–2008, respectively, with citations counted up to 1990 and 2013, respectively.)

[9] Therefore, the brochures are sent primarily to other law schools (cf. Espeland und Sauder 2007, 26).

Figure 3. Comparison of Citations to Published Articles in Five Prestigious Economic Journals

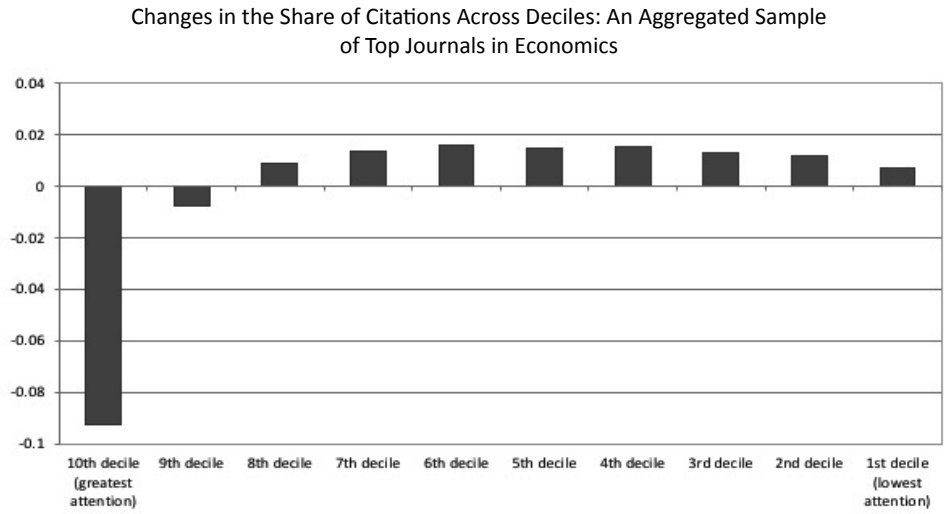


Source: Authors' own work based on data from Thomson Scientific.
Note: Published document types, such as editorials, book reviews or biographic items, were excluded in this analysis.

We already mentioned de Solla Price's (1965) insight that the majority of the scientific literature—in our case, the majority of articles in a journal—receives little to no attention from future articles, and this insight has been confirmed empirically by several works (Garfield 2006, 91; *Nature* 2005; Seglen 1997). In light of this, the results shown in Figure 3 are quite surprising as *they indicate a substantial reduction in the number of articles that are neglected in terms of citations.* However, at this point, one could argue that the change of this citation pattern may be due to an overall increase of the citation frequency. Put differently, the decrease in neglect of articles could point to an intensification and diversification of economic research. Thus, we provide a more nuanced analysis in Figure 4 by inspecting the changes in shares of received citations across the whole distribution of articles published in each period. The resulting pattern again supports our theoretical expectations that a decrease in the share of citations (by roughly 9 percent), dedicated to the upper limit of the distribution, is complemented by a corresponding increase in the lowest eight deciles.

Overall, these findings provide some support of our hypothesis that academic attention is undergoing a shift away from attention focused on certain authors and contributions to attention being focused on highly ranked journals. Further work to better disentangle the sources of the observed patterns would be highly relevant. In a first step, additional evidence on this issue can be obtained from a more large-scale empirical analysis, indicating that the distribution of attention, and hence the discursive properties of economics, are increasingly concentrated on a limited number of authors, departments, and journals (Glötzl and Aigner 2015). This latter trend thereby stands in contrast to the general development of scientific discourse that shows patterns of an increasing diversification of attention for most disciplines

Figure 4. Relative Change of Citation Patterns in Top Journals



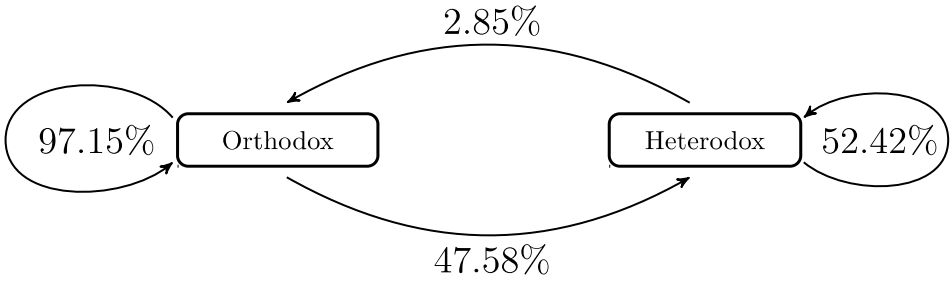
Source: Authors' own work based on data from Thomson Scientific.
Note: Journal sample from Figure 3.

(Larivière, Gingras and Archambault 2009). Regarding the power of scientometrics, we find that the currently practiced form of scientometrics has the power to influence the behavior of researchers and scientific institutions. In the case of economics, this identified power crystallizes especially (as shown in this section) in the institutional field of economic journals.

Economic Discourses and the Influence of Scientometrics on the Future of Economics

The scientific discourse in economics is different from other social sciences because it is dominated by a single paradigm—neoclassical economic theory—that significantly shapes academic teaching and economic research. We refer to the works of Roger E. Backhouse (2005) and Leonhard Dobusch and Jakob Kapeller (2009) for a discussion of the historical roots of this paradigmatic dominance. In the domain of a more narrowly defined research discourse, the 1970s and 1980s stand out. In this period, not only Keynesian macroeconomic approaches were driven out by neoclassical and monetaristic theories, but also the journal culture experienced a significant theoretical narrowing. This led to a virtually complete exclusion of critical and alternative scientific contributions from the economic discourse. With some exceptions, such articles were generally rejected during the review procedures. The confrontation resulting from this exclusion from an "official" economic discourse led to the founding of now eminent heterodox economic journals like the *Cambridge Journal of Economics,* the *Journal of Post Keynesian Economics*, or the *Journal of Economic Issues* (King 2003, 134-136). In this sense, economics is still a "contested discipline" (Lee and Elsner 2011). For

Figure 5. Interaction Pattern Between Mainstream Economics and Heterodoxy (1989-2008)



Source: Dobusch and Kapeller (2012b, 474).

the same reason, it is not surprising that alternative or heterodox schools of thought only constitute a small fraction inside the economic discipline, since they are confronted with such "exclusion routines" on several levels.

Of course, the description of this specific constellation per se does not suffice to draw a conclusion regarding the representation of alternative economic approaches within the mainstream economic discourse. As a consequence, the question of how heterodox ideas are regarded within the mainstream discourse has to be answered primarily from an empirical perspective. Past works have analyzed the interaction between heterodoxy and mainstream by comparing citation patterns associated with mainstream and heterodox economic journals. A representative example of this literature is reproduced in Figure 5, which compares the relative citation flows among 26 economics journals (13 highly ranked mainstream and 13 heterodox) over a period of twenty years, from 1989 to 2008.[10] It shows that heterodox journals exhibit a quite balanced citation pattern (heterodox and mainstream journals are cited equally), while the citation behavior of mainstream journals is drastically in favor of other mainstream journals. In this view, heterodoxy is more open or pluralist, whereas orthodoxy is relatively closed or monistic. This assertion is reinforced by an inspection of absolute citation flows, i.e., net transfer of ideas, as we conduct below.

A more detailed analysis of the data from Figure 5 further shows that the percentage of citations from the top 13 heterodox journals exported into mainstream journal literature considered here (2.85 percent of total references) is driven heavily by statistical outliers. Measured in absolute figures, 2.85 percent represent 753 citations. Of these, the majority (613 citations) is caused by only three journals that hold a special position in the economic

[10] The basis for our sample was the *Web of Science* database from *Thomson Reuters*. For the selection of relevant journals, we used the *Journal Citation Report* 2007, whereby the 13 best ranked journals were interpreted as the "top 13 orthodox" journals. We identified the "top 13 heterodox" journals by using the *Heterodox Economics Directory* (www.heterodoxnews.com/hed).

discourse.[11] Within the remaining 23 journals, only a minimal transfer of ideas in the form of 140 citations over a period of 20 years can be found.

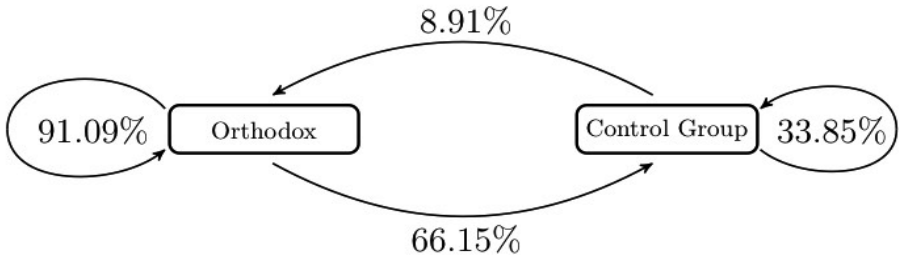
Even if Figure 5 serves as sufficient evidence for the underrepresentation of alternative economic approaches in mainstream economic literature, it is still not clear whether the observed pattern can be explained by a paradigmatic divide or by a strong focus of the economic discourse on articles published in highly ranked journals. After all, the orthodox sample represents the top 13 journals from the *Journal Citation Report* 2007, whereas the heterodox sample is scattered between ranking position 17 and 130. Considering the high level of self-reference as well as a strong elitist orientation inherent in the scientific discourse in economics, this is a strong argument. These aspects were recently documented by Marion Fourcade, Etienne Ollion, and Yann Algan (2015), who showed in detail that the economic literature is less inclined to refer to other disciplines (self-reference) as compared to other scientific disciplines, and has a stronger focus on a small group of top journals whose authors primarily stem from a small and homogeneous quantity of universities (concentration) (see also Hodgson and Rothman 1999). This does not only refer to a more focused attentiveness within the economic discipline, but also to a relatively tightly structured and hierarchical internal organization (elitist orientation). Therefore, it is not surprising that an analysis of the development of economic "top journals," recently published in the *Journal of Economic Literature*, was limited to the observation of only five journals (Card and DellaVigna 2013).

We will now replicate the procedure shown in Figure 5, with a corresponding control group, in order to find out to what extent the neglect of alternative theoretical approaches is due to an elitist approach to journal rankings. Figure 6 shows a replication of Figure 5, where the heterodox sample is replaced by a control group consisting of those 13 mainstream journals in the JCR 2007 ranking that are—in each case—one position below the heterodox journal sample. The result in Figure 6 shows that the observed discursive pattern cannot exclusively be explained by the relative ranking position of the heterodox journals. On the contrary, the citation frequency of the control group is more than three times higher than in the heterodox sample. Indeed, paradigmatic factors seem to play a central role and suggest a systematic discursive exclusion of alternative theoretical approaches.

Of course, this aspect is not the only weakness in our examination of the inter-paradigmatic discourse in economics. The time period we consider here—namely, before the crisis—could be another possible point of criticism. The financial and economic crisis could be understood as a central anomaly in sharp contrast to basic postulates of neoclassical theory (such as the efficient market hypothesis and the associated arbitrage-based thinking).

[11] On the heterodox side of the sample, the *Journal of Economic Behavior and Organization*–which is attractive both for heterodox and mainstream articles—accounts for 340 exported citations. On the mainstream side of the sample, we included two journals from the field of economic geography (*Journal of Economic Geography* und *Economic Geography*) that import further 273 citations. In this connection, the relative openness of economic geography for heterodox approaches is disproportionately higher than in the strict mainstream economics.

Figure 6. Interaction Pattern Between Top Mainstream Journals and a Control Group (1989–2008)



Source: Authors' own calculation based on data from Thomson Scientific.[12]

It is certainly conceivable that the crisis has changed the perception of basic facts of the economy and has led to a more inclusive pluralist approach.

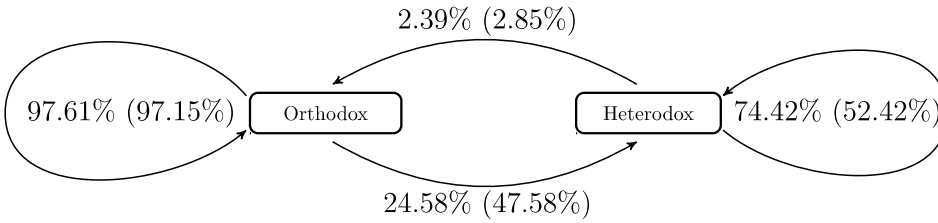
A theoretical answer is provided by Kuhn (1962), who argued that a dominant paradigm facing a significant anomaly will try to resolve this dissonance by an adaption of already established theses, models, and methods and simultaneously aim to avoid fundamental debates. While it is beyond question that economic research has changed in some ways in response to the recent crisis (e.g., Young 2014), a more nuanced analysis seems necessary to assess whether these changes have also led to an increased reception of alternative economic approaches in mainstream outlets, or whether this reaction follows a Kuhnian pattern of an "internal" adaption of existing models. Figure 7 provides a further replication of Figure 5 which focuses on the current journal literature in the period from 2009 to 2013, instead of the pre-crisis period.

Here again, supporting Kuhn's prediction, no substantial change in the citation behavior of the dominant paradigm can be observed. Indeed, the behavior of the economic mainstream remains widely constant. So far, the anomaly of the financial and economic crisis has not intensified the reception of alternative theoretical approaches within the mainstream economic discourse.

We can now establish the following thesis: inner-scientific criteria, with citation metrics and associated institutionalized evaluation routines among them, exert a stronger influence on the distribution of attention in economics than the actual economic development (the most drastic example being the financial and economic crisis). One major reason for this lies in the "size bias" of citation metrics. By definition, citation metrics certify quantitatively meaningful research disciplines like mainstream economics as high-quality disciplines. Especially in the case of economics as a "contested discipline," characterized by an extremely unequal distribution of initial resources, it is obvious that the consideration of scientometric

[12] A comparison of the relative change in citation behavior of the orthodox journals with data from Figure 5 shows a statistically significant difference on the 1-percent level.

Figure 7. Interaction Pattern Between Mainstream Economics and Heterodoxy (2009–2013)



Source: Authors' own calculation based on data from Thomson Scientific.
Note: Values from Figure 5 in brackets.

criteria will further stabilize the dominant economic paradigm.

First indications for such a development, for example, can be found in France, where citation metrics have started to play a significant role in the centralist appointment policy of professors in economics. Citation metrics in France were included in a formal scoring system in 2005. Assuming the same level of productivity, heterodox economists achieve—due to this "size bias" of citation metrics—a significantly lower score, making the appointment of heterodox economists appear less attractive. This circumstance is indeed reflected in the French appointment policy. Parallel to the introduction of the scoring system, the relative amount of heterodox economists newly appointed in France fell from almost 18 percent in the period from 2000 to 2004 to 5 percent in the period from 2005 to 2011 (FAPE 2014). Similar trends can also be identified in other countries, where evaluation routines based on citation metrics are used. For example, in Great Britain, the so-called *Research Assessment Exercise*, which evaluates thematically related institutes, also influences the appointment policy of these institutes. Therefore, ranking criteria are anticipated in favor of mainstream economics, leading to a positive discrimination of mainstream economic applicants (see also Lee [2007] who analyzes the situation at the beginning of the 2000s). Summarizing, we see that another effect of evaluative scientometrics lies in the stabilization of the dominant role of a prevailing scientific paradigm and, therefore, promotes an increasing homogenization of scientific disciplines.

Conclusion

The purpose of our examination was to provide a theoretical overview of the consequences of scientometric evaluation routines, with a special emphasis on economics. We identified three main findings. First, evaluative scientometrics contributes to a reinforcement and

stabilization of existing patterns of academic reproduction and further increases any existing bias in the distribution of attention, prestige, and resources (second section). Second, there are various instances of reactive effects that are a natural consequence of employing evaluative scientometrics and actively influence the behavior of scientific agents (third section). Third, evaluative scientometrics significantly contributes to the stabilization of a dominant economic paradigm and limits the influence of alternative or critical approaches in the scientific discourse (fourth section). Our evidence supported the indication that the mechanisms we discussed can be assessed empirically and can provide inspiration for further research on the role of evaluative citation metrics in academic reproduction and scientific development.

References

Adler, Robert, John Ewing and Peter Taylor. "Citation Statistics." Report from the International Mathematical Union, 2008. Available at http://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf. Accessed March 10, 2015.

Arthur, Brian W. *Increasing Returns and Path-Dependence in the Economy.* Ann Arbor, Michigan: University of Michigan Press, 1994.

Backhouse, Roger E. "The Rise of Free Market Economics: Economists and the Role of the State Since 1970." *History of Political Economy* 37, no. 5 (2005): 355-392.

Bloch, Harry. "Research Evaluation Down Under: An Outsider's View from the Inside of the Australian Approach." *American Journal of Economics and Sociology* 69, no. 5 (2010): 1530-1552.

Card, David and Stefano DellaVigna. "Nine Facts About Top Journals in Economics." *Journal of Economic Literature* 51, no. 1 (2013): 144-161.

Corsi, Marcella, Carlo D'Ippoliti and Federico Lucidi. "Pluralism at Risk? Heterodox Economic Approaches and the Evaluation of Economic Research in Italy." *American Journal of Economics and Sociology* 69, no. 5 (2010): 1495-1529.

David, Paul A. "Clio and the Economics of QWERTY." *American Economic Review* 75, no. 2 (1985): 332-337.

Demange, Gabrielle. "On the Influence of a Ranking System." *Social Choice and Welfare* 39, no. 2 (2012): 431-455.

Diamond, Arthur M. "The Core Journals of Economics." *Current Contents* 1 (1989): 4-11.

Dobusch, Leonhard and Jakob Kapeller. "Why Is Economics Not an Evolutionary Science? New Answers to Veblen's Old Question." *Journal of Economic Issues* 43, no. 4 (2009): 867-898.

——. "Heterodox United vs. Mainstream City? Sketching a Framework for Interested Pluralism in Economics." *Journal of Economic Issues* 46, no. 4 (2012a): 1035-1057.

——. "A Guide to Paradigmatic Self-Marginalization—Lessons for Post-Keynesian Economists." *Review of Political Economy* 24, no. 3 (2012b): 469-487.

Dobusch, Leonhard and Elke Schüßler. "Theorizing Path Dependence: A Review of Positive Feedback Mechanisms in Technology Markets, Regional Clusters and Organizations." *Industrial and Corporate Change* 22, 3 (2013): 617-647.

Espeland, Wendy N. and Michael Sauder. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." *American Journal of Sociology* 113, no. 1 (2007): 1-40.

Espeland, Wendy N. and Mitchell L. Stevens. "A Sociology of Quantification." *European Journal of Sociology* 49, no. 3 (2008): 401-436.

FAPE. "Evolution of Economics Professors' Recruitment Since 2000 in France: The End of Pluralism." Working Paper No. 2014. FAPE, 2014. Available at http://www.heterodoxnews.com/HEN/attach/hen167/FAPE_pluralism_france.pdf. Accessed February 30, 2014.

Fourcade, Marion, Etienne Ollion and Yann Algan. "The Superiority of Economists." *Journal of Economic Perspectives* 29, no. 1 (2015): 89-114.

Garfield, Eugene. "The History and Meaning of the Journal Impact Factor." *Journal of the American Medical Association* 295, no. 1 (2006): 90-93.

Glötzl, Florentin and Ernest Aigner. "Pluralism in the Market of Science? A Citation Network Analysis of Economic Research at Universities in Vienna." Ecological Economic Papers No. 5. Vienna: Austria: WU Vienna University of Economics and Business, 2015. Available at http://epub.wu.ac.at/4730/1/EcolEcon_WorkingPaper_2015_5.pdf. Accessed August 09, 2017.

Hodgson, Geoffrey M. and Harry Rothman. "The Editors and Authors of Economics Journals: A Case of Institutional Oligopoly?" *Economic Journal* 109, no. 453 (1999): F165-F186.

ISIPE. "An International Student Call for Pluralism in Economics." ISIPE, 2014. Available at http://www.isipe.net/open-letter. Accessed July 26 , 2016.

King, John E. *A History of Post-Keynesian Economics Since 1936*. Cheltenham: Edward Elgar, 2003.

Larivière, Vincent, Yves Gingras and Éric Archambault. "The Decline in the Concentration of Citations, 1900–2007." *Journal of the Association for Information Science and Technology* 60, no. 4 (2009): 858-862.

Lee, Frederic S. "The Research Assessment Exercise, the State and the Dominance of Mainstream Economics in British universities." *Cambridge Journal of Economics* 31, no. 2 (2007): 309-325.

Lee, Frederic S. and Wolfram Elsner. *Evaluating Economic Research in a Contested Discipline: Ranking, Pluralism, and the Future of Heterodox Economics*. Hoboken: John Wiley & Sons, 2011.

Levy, Moshe and Haim Levy. "Investment Talent and the Pareto Wealth Distribution: Theoretical and Experimental Analysis." *Review of Economics and Statistics* 85, no. 3 (2003): 709-725.

Kuhn, Thomas. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

Merton, Robert K. "The Matthew Effect in Science." *Science* 159, no. 3810 (1968): 56-63.

Mingers, John and Loet Leydesdorff. "A Review of Theory and Practice in Scientometrics." Working Paper. Archiv.org, 2015. Available at http://arxiv.org/abs/1501.05462. Accessed 23 February 2015.

Nature. "Not-So-Deep Impact." *Nature* 435, no. 7045 (2005): 1003-1004.

Newman, Mark E.J. "Power Laws, Pareto Distributions and Zipf's Law." Cornell University Library, 2006. Available at http://arxiv.org/abs/cond-mat/0412004. Accessed December 10, 2014.

Rip, Arie and Jean-Pierre Courtial. "Co-Words Maps in Biotechnology: An Example of Cognitive Scientometrics." *Scientometrics* 6, no. 6 (1984): 381-400.

Seglen, Per O. "Why the Impact Factor of Journals Should Not Be Used for Evaluating Research." *British Medical Journal* 314, 7079 (1997): 498-502.

Shapiro, Carl and Hal R. Varian. *Information Rules: A Strategic Guide to the Network Economy*. Boston: Harvard Business School Press, 1999.

De Solla Price, Derek J. "Networks of Scientific Papers." *Science* 149, no. 3683 (1965): 510-515.

Sterman, John D. and Jason Wittenberg. "Path Dependence, Competition, and Success in the Dynamics of Scientific Revolution." *Organization Science* 10, no. 3 (1999): 322-341.

Sydow, Jörg, Georg Schreyögg and Jochen Koch. "Organizational Path Dependence: Opening the Black Box." *Academy of Management Review* 34, no. 4 (2009): 689-709.

Young, Brigitte. "Financial Crisis: Causes, Policy Responses, Future Challenges." Brussels, Belgium: Directorate-General for Research and Innovation, European Commission, 2014. Available at http://ec.europa.eu/research/social-sciences/pdf/policy_reviews/KI-NA-26554_EN-C.pdf. Accessed February 24, 2016.