ICAE Working Paper Series - No. 46 - March 2016 (Updated: August 2017)

The Power of Scientometrics and the Development of Economics

Matthias Aistleitner
Jakob Kapeller
Stefan Steinerberger

Institute for Comprehensive Analysis of the Economy
Johannes Kepler University Linz
Altenbergerstraße 69, 4040 Linz
icae@jku.at
www.jku.at/icae

JOHANNES KEPLER UNIVERSITY LINZ

INSTITUTE FOR COMPREHENSIVE ANALYSIS OF THE ECONOMY

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 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 as well as the global development of the discipline.

Keywords

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

JEL classification

A14, B50

⋆ This work was funded by the Institute for New Economic Thinking (INET) under grant numbers INO1500038 and INO1500039. The manuscript has been accepted for publication in the *Journal of Economic Issues*.

[*] Johannes Kepler University Linz, Institute for Comprehensive Analysis oft he Economy (ICAE), matthias.aistleitner@jku.at

[**] Johannes Kepler University Linz, Institute for Comprehensive Analysis oft he Economy (ICAE) & Department of Economics, jakob.kapeller@jku.at

[***] Yale University, Department of Mathematics, stefan.steinerberger@yale.edu

1. Introduction

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, which aim 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, which has its conceptual origins in an interpretative analysis of scientific communication aiming to understand the discursive properties of academic publishing (e.g. De Solla Price 1965, Rip and Courtial 1984). This approach, which has sometimes been dubbed as 'cognitive scientometrics', is increasingly making way for new forms of 'evaluative scientometrics', which 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 (Mingers and Leydesdorff 2015, Adler et al. 2008).

The increasingly prominent role of quantitative evaluation in academia can be interpreted as a part of a more general social trend oriented towards the numerical assessment of social issues. One major historical impetus of this process of an increased 'quantification of social phenomena' (Espeland and Stevens 2008, p. 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 thereby provides a yardstick for

measuring economic development. Aside from administrative interests, indicators like the 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', which covers 'the production and communication of numbers' and deals with 'regimes of measurement' (cf. Espeland and Stevens 2008, p. 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 they by themselves do not yet establish any specific 'regimes of measurement'. However, numerical information also directs attention towards the relative properties of various entities: when assigning numerical values to different entities of the same class, numerical information allows to create an ordering that has the double function of *unifying* different objects across a uniform scale, which, at the same time, 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 have 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, this paper focuses more particularly on its role and impact within the field of economics. Its main contribution is to provide a theoretical embedding of the notion of quantification in a more general account of scientific development and to better understand its impact on the development of specific research fields and the behavior of individual researchers. The 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, which are suitable for empirically illustrating the underlying argument.

We focus on three specific aspects: firstly, scientometric routines in research assessment are always embedded in conventional patterns of academic reproduction; hence, the intensified use of these routines has to be understood against this backdrop, which is the main focus of section 2. Secondly, scientometric evaluation is based on mechanical procedures that are easily reproducible and therefore cause incentives for strategic behavior among researchers. In section 3 we discuss some aspects of reactivity with a special focus on the impact of reactive evaluation routines on the citation behavior of economists. Thirdly, 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 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, section 4 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. Section 5 concludes and summarizes the main arguments of this paper.

2. 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, p. 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 no or just very little 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. To name a few examples, 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% 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 1981-1985[3]. To better illustrate the properties of said 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.

<Figure 1 near here>

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 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 only later clarified by Robert K. Merton's classical work on the 'Matthew Effect in Science' (Merton 1968): he postulated that the acquisition of

prestige and attention given to scientific work is closely correlated with the amount of attention acquired in the past. Merton based his argument on citation data as well as interviews conducted with Nobel Prize laureates. The implicit logic of this mechanism, namely 'whosoever has, will be given more' is nowadays 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 repute and the withholding of such recognition from scientists who have not yet made their mark.' (Merton 1968, p. 58)

Matthew effects can be characterized more generally as self-reinforcing effects, which, 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, which for instance has been discussed in detail in economic literature about path dependency (Sydow et al. 2009; Dobusch and Schüßler 2013). The main thesis of this theoretical strand is that the establishment of technical, organizational or social standards can lead to self-reinforcing effects, which lead to a 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 as 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 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 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 a '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, where 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 a Matthew-effects to specific fields and approaches instead of single authors and contributions, is exemplified by

simulation-based study of John D. Sterman and Jason Wittenberg (1999), which provides a detailed account on the path-dependent properties of scientific evolution. For example, Sterman and Wittenberg 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 S. Kuhn's historical observations. A prime example in this context is given by the advent of Copernican theory, which in its beginning – in terms of its empirical explanatory power and precision – lagged behind the geocentric view. The main reason for the initial superiority of the Ptolemaic theory was that it could employ a sophisticated theoretical apparatus with numerous correction terms (so-called 'epi-cycles') introduced to improve the empirical performance of the underlying models.

For the purpose of this paper, it seems promising to take a closer look at the structure of the Sterman and Wittenberg 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, p. 333)

Building on this argument, which is also summarized in Figure 2 we now sketch a theoretical argument contextualizing the advent of evaluative scientometrics within the past one-and-a-half decades[5] on academic reproduction. 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, which 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 attesting a high level of quality to those authors, articles and research fields whose initial level of paid attention is already high (Dobusch and Kapeller 2009). Accordingly, a fourth feedback loop can be added to the model of Sterman and Wittenberg, which can be formulated in the following way: The bigger a paradigm or research field is, the higher the amount of received citations in this field will be – ‘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).

<Figure 2 near here >

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 for 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 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 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 2010/11.[6] Their aggregate average submission numbers in the years before and after inclusion in the ranking look as follows.

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%

Table 1: Aggregated submission data from three economic journals joining the JCR in 2010/11.

While Table 1 gives a first intuition on 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 obvious implies a boost in attractiveness. For the established journals, institutions and paradigms the very same effect obviously leads to a reinforcement of inherent path-dependencies by rewarding the already rewarded once more and thereby shedding light on those, who are already visible (see also Demange 2012). Hence, it is possible to state a 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.

3. Evaluation Routines and the Role of Reactive Measurement Procedures

There are a large number of 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[7]) 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 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, which 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 the criteria imposed for purposes of evaluation.

A significant contribution for the understanding of the consequences of reactive measurement procedures in scientific evaluation comes from 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 US-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 which 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 peer-review processes itself when former ranking positions are interpreted as indicators of

quality by reviewers. Towards 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 US-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, p. 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, p. 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.

> ‘[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, p. 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 (Lee 2007; Bloch 2010; Corsi et al. 2010) and, hence, 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, p. 416)

In Espeland and Stevens' 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[8], 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 said rankings. This argument implies a transmission of prestige towards 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 (papers published in 1981-1985 and 2004-2008, respectively, with citations counted up to 1990 and 2013, respectively.)

<Figure 3 near here>

The insight of de Solla Price (1965) 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 was already mentioned in Chapter 2 and has been confirmed empirically by several works (Garfield 2006, p. 91; Nature 2005, Seglen 1996). 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 articles could point to an intensification and diversification of economic research. Hence, we provide a more nuanced analysis in Figure 4 by inspecting the changes in shares of received citations across the whole distribution of papers published in each period. The resulting pattern again supports our theoretical expectations – a decrease in the share of citations (by roughly 9%) dedicated to the upper limit of the distribution is complemented by a

corresponding increase in lowest eight deciles.

<Figure 4 near here>

Overall, these findings weakly support our hypothesis that academic attention is undergoing a shift away from attention focused on single 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 appreciated. In a first step, additional evidence on this issue can be obtained from more large-scale empirical analysis, which indicates that the distribution of attention and, hence, the discursive properties of economics are increasingly concentrated upon 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, which shows patterns of an increasing diversification of attention for most disciplines (Larivière et al. 2009).

Regarding the power of scientometrics, we find in this section 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 at the end of this chapter – in the institutional field of economic journals.

4. 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 – which significantly shapes academic teaching and economic research. We refer to Backhouse (2005) or Dobusch and Kapeller (2009) for a discussion of the historical roots of this paradigmatic dominance. In the domain of a more narrowly defined research discourse especially 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; such articles were – with some exceptions – rejected in review procedures. The resulting confrontation caused by this exclusion from an 'official' economic discourse led to the foundation of – nowadays eminent – heterodox-economic journals, like the *Cambridge Journal of Economics,* the *Journal of Post Keynesian Economics* or the *Journal of Economic Issues* (King 2003, p. 134-136). In this sense, economics is still a 'contested discipline' (Lee and Elsner 2011). For 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 which 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 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 between 26 economics journals (13 highly ranked mainstream and 13 heterodox) over a period of twenty years (1989-2008).[9] 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 – an assertion, that is reinforced by an inspection of absolute citation flows, i.e. net transfer of ideas, as conducted below.

<Figure 5 near here>

A more detailed analysis of the data from Figure 5 furthermore shows that the percentage of citations from the top thirteen heterodox journals exported into mainstream journal literature considered here (2,85% of total references) is driven heavily by statistical outliers. Measured in absolute figures, 2,85% represent 753 citations. Of these, the majority (613 citations) is caused by only three journals that hold a special position within the economic discourse.[10] Within the remaining 23 journals only a minimal transfer of ideas in the form of 140 citations within 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 within economic discourse on articles published in highly ranked journals. After all, the orthodox sample represents the top 13 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 to the scientific discourse in economics, this is a strong argument. These aspects were recently documented by Fourcade et al. (2015), who showed in detail that the economic literature is – compared to other scientific disciplines – less inclined to refer to other disciplines (self-reference) 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 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 understand to which extent the neglect of alternative theoretical approaches is due to an elitist approach towards journal rankings. Figure 6 shows a replication of Figure 5, where the heterodox sample is replaced by a control group consisting of those thirteen mainstream journals in the JCR 2007 ranking which 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.

<Figure 6 near here>

Of course, this aspect is not the only weakness in our examination about interparadigmatic discourse in economics. The time period considered here – namely before the crisis – could be another possible point of criticism. The financial and economical 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). It is certainly conceivable that the crisis has changed the perception of basic facts of economy and has lead to a more inclusive pluralist approach.

A theoretical answer is provided by Kuhn (1962), who argues 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, at the same time, aim to avoid fundamental debates. While it is beyond question that economic research has changed in some way in response to the recent crisis (e.g. Young 2014), a more nuanced analysis seems necessary to assess whether these changes 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 – instead of the pre-crisis period – on current journal literature from the period 2009 to 2013.

<Figure 7 near here>

Here again – supporting the prediction of Kuhn – 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.

The following thesis can now be established: 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 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 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 fell from almost 18% in the period of 2000-2004 to 5% in the period of 2005-2011 (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. Hereby, 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.

5. Conclusion

The purpose of this examination was to provide a theoretical overview on the consequences of scientometric evaluation routines with a special emphasis on economics. Three main findings could be identified: firstly, evaluative scientometrics contribute to a reinforcement and stabilization of existing patterns of academic reproduction and further increase any existing bias in the distribution of attention, prestige and resources (Chapter 2). Secondly, there exist various instances of reactive effects, which are a natural consequence of employing evaluative scientometrics and actively influence the behavior of scientific agents (Chapter 3). Thirdly evaluative scientometrics significantly contribute to the stabilization of a dominant economic paradigm and limit the influence of alternative or critical approaches within the scientific discourse (Chapter 4). Illustrative evidence was supported to indicate how the discussed mechanisms can be assessed empirically and to provide inspiration for further

research on the role of evaluative citation metrics in academic reproduction and scientific development.

References

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

Arthur, Brian W. *Increasing returns and path-dependence in the economy*. 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, 5 (2005): 355-392.

Bloch, Harry. "Research Evaluation Down Under: An Outsider's View from the Inside of the Australian Approach." *The American Journal of Economics and Sociology* 69, 5 (2010): 1530-1552.

Card, David and Stefano DellaVigna. "Nine facts about top journals in economics." *Journal of Economic Literature* 51, 1 (2013): 144-161.

Corsi, Marcella, D'Ippoliti, Carlo and Federico Lucidi. "Pluralism at Risk? Heterodox Economic Approaches and the Evaluation of Economic Research in Italy." *The American Journal of Economics and Sociology* 69, 5 (2010): 1495-1529.

David, Paul A. "Clio and the economics of QWERTY." *The American Economic Review* 75, 2 (1985): 332–337.

Demange, Gabrielle. "On the influence of a ranking system." *Social Choice and Welfare* 39, 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, 4 (2009): 867-898.

Dobusch, Leonhard and Jakob Kapeller. "Heterodox United vs. Mainstream City? Sketching a framework for interested pluralism in economics." *Journal of Economic Issues* 46, 4 (2012a): 1035-1057.

Dobusch, Leonhard and Jakob Kapeller. "A guide to paradigmatic Self-marginalization - Lessons for Post-Keynesian Economists." *Review of Political Economy* 24, 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, 1 (2007): 1-40.

Espeland, Wendy N. and Mitchell L. Stevens. "A Sociology of Quantification." *European Journal of Sociology* 49, 3 (2008): 401-436.

FAPE. “Evolution of Economics Professors’ Recruitment since 2000 in France: The End of Pluralism.” Working Paper, 2014. Accessed February 30, 2014. http://www.heterodoxnews.com/HEN/attach/hen167/FAPE_pluralism_france.pdf

Fourcade, Marion, Ollion, Etienne and Yann Algan. “The superiority of economists.” *Journal of Economic Perspectives* 29, 1 (2015): 89-114.

Garfield, Eugene. “The History and Meaning of the Journal Impact Factor.” *Journal of the American Medical Association* 295, 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 5, 2015. WU Vienna University of Economics and Business, Vienna. Accessed August 09, 2017. http://epub.wu.ac.at/4730/1/EcolEcon_WorkingPaper_2015_5.pdf

Hodgson, Geoffrey M. and Harry Rothman. “The editors and authors of economics journals: a case of institutional oligopoly?” *The Economic Journal* 109, 453 (1999): F165-F186.

ISIPE. “An international student call for pluralism in economics.” 2014. Accessed July 26 , 2016. http://www.isipe.net/open-letter/

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

Larivière, Vincent, Gingras, Yves and Éric Archambault. “The decline in the concentration of citations, 1900–2007.” *Journal of the Association for Information Science and Technology* 60, 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, 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*. John Wiley & Sons, 2011.

Levy, Moshe and Haim Levy. “Investment talent and the pareto wealth distribution: Theoretical and experimental analysis.” *The Review of Economics and Statistics* 85, 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, 3810 (1968): 56-63.

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

Nature. “Not-so-deep impact.” *Nature* 435, 7045 (2005): 1003-1004.

Newman, Mark E. J. “Power laws, Pareto distributions and Zipf's law.” 2006. Accessed December 10, 2014. http://arxiv.org/abs/cond-mat/0412004

Rip, Arie and Jean-Pierre Courtial. “Co-Words Maps in Biotechnology: An Example of Cognitive Scientometrics.” *Scientometrics* 6, 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, 3683 (1965): 510-515.

Sterman, John D. and Jason Wittenberg. “Path Dependence, Comeptition, and Success in the Dynamics of Scientific Revolution.” *Organization Science* 10, 3 (1999): 322-341.

Sydow, Jörg, Schreyögg, Georg and Jochen Koch. “Organizational Path Dependence: Opening the Black Box.” *Academy of Management Review* 34, 4 (2009): 689–709.

Young, Brigitte. “Financial Crisis: Causes, policy responses, future challenges.” Directorate-General for Research and Innovation, European Commission, Brussels, 2014. Accessed February 24, 2016. http://ec.europa.eu/research/social-sciences/pdf/policy_reviews/KI-NA-26554_EN-C.pdf

Figures

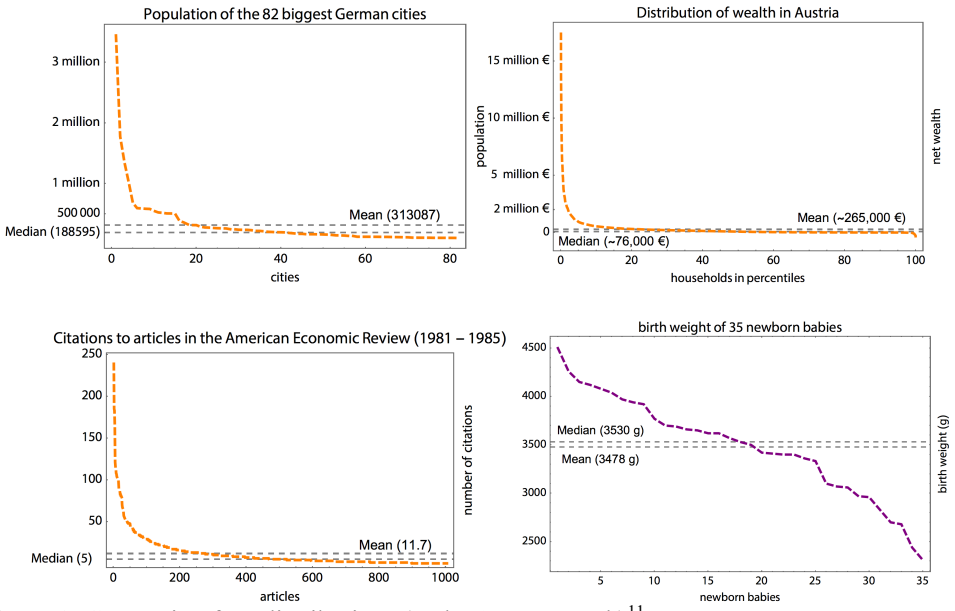

Figure 1: Comparing four distributions (author's own graph)[11]

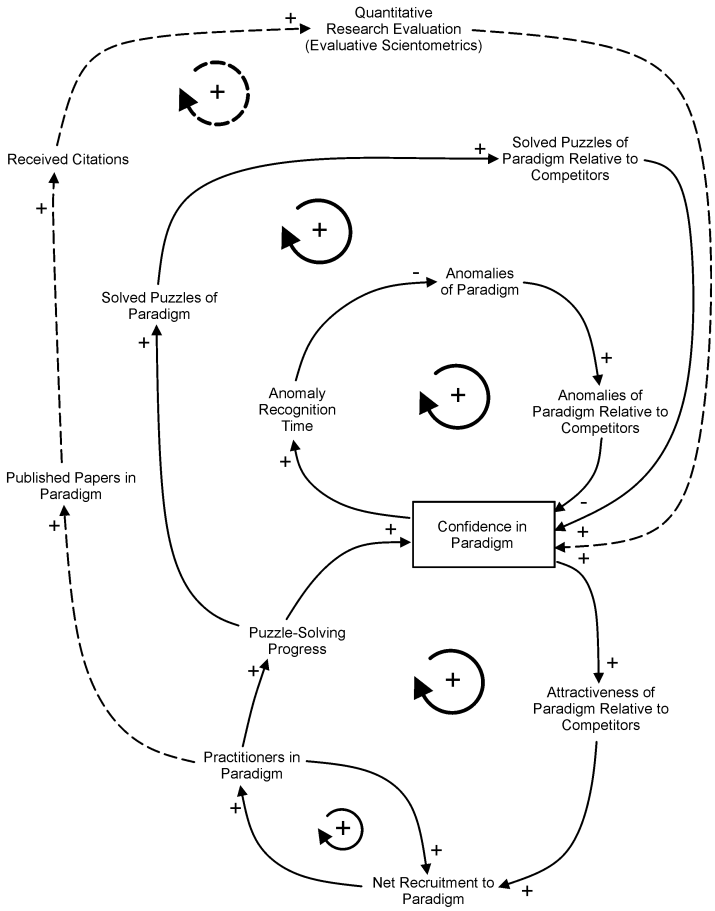

Figure 2: The role of evaluative scientometrics as additional feedback loop (dashed line) in stabilizing a scientific paradigm (author's own graph based on Sterman and Wittenberg 1999, 333).

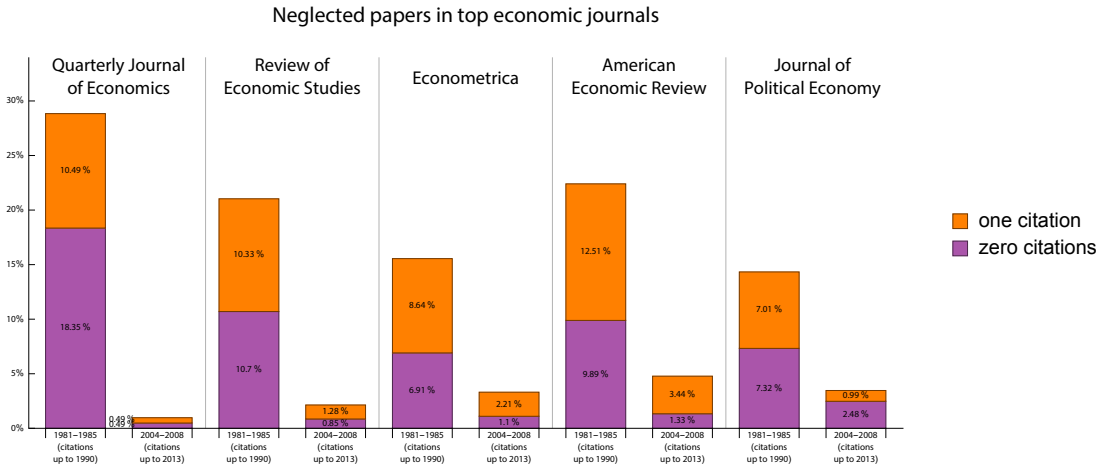

Figure 3: Comparison of citations to published articles in five prestigious economic journals (authors' own graph based on data from *Thomson Scientific*). Note: Published document types such as editorials, book reviews or biographic items were excluded in this analysis.

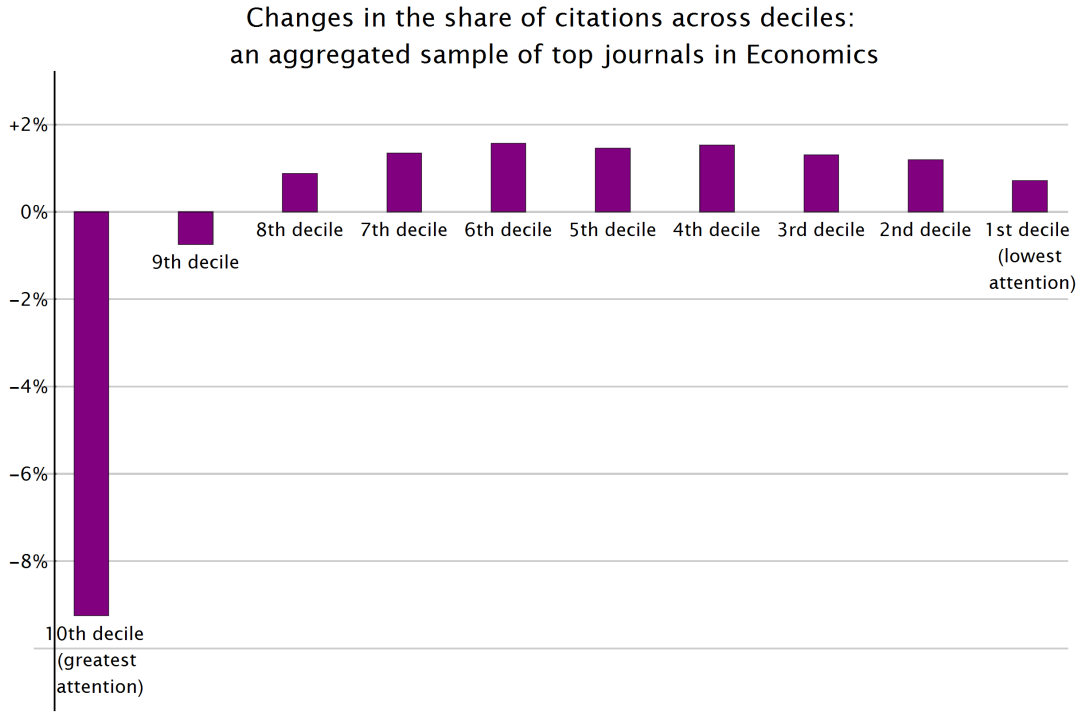

Figure 4: Relative change of citation patterns in top journals (journal sample from Figure 3). (authors' own graph based on data from Thomson Scientific).

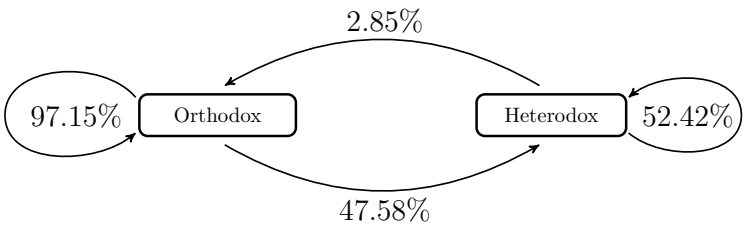

Figure 5: Interaction pattern between mainstream economics and heterodoxy (1989-2008; taken from Dobusch and Kapeller 2012b, p. 474).

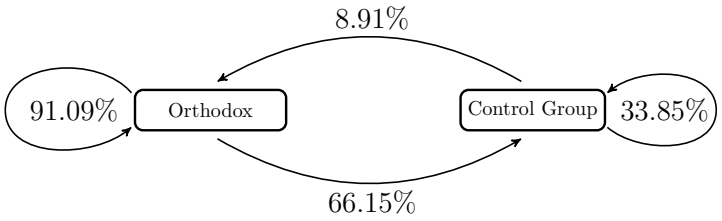

Figure 6: Interaction pattern between top mainstream journals and a control group (1989-2008; author's own calculation based on data from *Thomson Scientific*). [12]

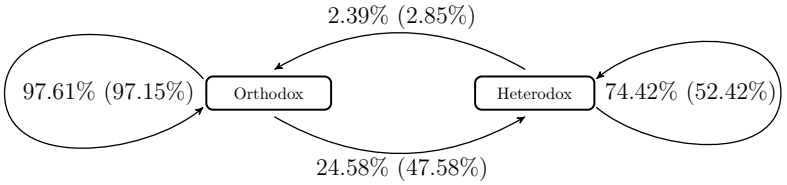

Figure 7: Interaction pattern between mainstream economics and heterodoxy (2009-2013; author's own calculation based on data from *Thomson Scientific;* values from Figure 5 in brackets).

Notes

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

[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 stem from 1981-1990.

[4] 'For to every one 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.' (Mt 25,29)

[5] Although citation data has already been widely used in the 20th 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 fifteen 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).

[6] We obtained this data through personal correspondence with a series of editorial offices.

[7] See also: http://vhbonline.org/service/jourqual/vhb-jourqual-3/

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

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

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

[11] 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 to 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 on request.

[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%-level.