

Research Governance in Academia: Are there Alternatives to Academic Rankings?

Margit Osterloh
Bruno S. Frey

CESIFO WORKING PAPER NO. 2797
CATEGORY 1: PUBLIC FINANCE
SEPTEMBER 2009

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

Research Governance in Academia: Are there Alternatives to Academic Rankings?

Abstract

Peer reviews and rankings today are the backbone of research governance, but recently came under scrutiny. They take explicitly or implicitly agency theory as a theoretical basis. The emerging psychological economics opens a new perspective. As scholarly research is a mainly curiosity driven endeavor, we include intrinsic motivation and supportive feedback by the peers as important determinants of scholarly behavior. We discuss whether a stronger emphasis on selection and socialization offers an alternative to the present regime of academic rankings.

JEL Code: O00.

Keywords: peer reviews, rankings, research governance, agency theory, psychological economics, new public management, economics of science, control theory.

Margit Osterloh
Institute of Organization and Administrative
Science
Universitätsstrasse 84
8006 Zurich
Switzerland
osterloh@iou.uzh.ch

Bruno S. Frey
University of Zurich
Institute for Empirical Research in
Economics
Winterthurerstrasse 30
8006 Zurich
Switzerland
bsfrey@iew.uzh.ch

Peer reviews and academic rankings are generally considered the backbone of research governance in academia. The recent, lively discussion about the quality of peer reviews (e.g., Abramo, Angelo, & Caprasecca, 2009; Frey, 2003; Starbuck, 2005, 2006) and academic rankings (e.g., Adler & Harzing, 2009; Lawrence, 2002, 2003; Weingart, 2005) focused mainly on the issues of method and how to improve it. However, the question was not raised on which theoretical background peer reviews and rankings are based and whether this background is adequate for research governance in academia. Until today, there is no discussion about whether there are alternatives to the dominant principles of academic research governance.

We argue that these principles explicitly or implicitly follow mainly from the principal agent view, which in the literature on corporate governance has come under fire due to corporate failures and scandals (e.g., Benz & Frey, 2007; Daily, Dalton, & Canella, 2003). A corresponding, critical discussion in the field of research governance in academia is lacking. In our paper, we contribute to this discussion by confronting two different perspectives, agency theory and the newly emerging psychological economics.¹

The latter approach builds on psychologically informed economics. In line with an understanding of scholarly research as a mainly curiosity driven endeavor, it includes intrinsic motivation as a major determinant of scholarly behavior. We combine this approach with managerial control theory based on the work of Ouchi (1977, 1979) and discuss the implications derived from the two different perspectives. While agency theory

counts on the refinements of indicators and the measurement process, psychological economics builds mainly on the careful selection and socialization of scholars, on supporting feedback, and on symbolic benefits like awards.

We begin by analyzing the theoretical basis of the current dominant governance system for academic research, namely agency theory combined with the economics of science. The second section presents empirically based findings on present research governance, discussing the advantages and disadvantages of its backbones, namely peer reviews and academic rankings. Psychological economics then is suggested as a theoretical basis for academic governance, appreciating the unique features of research in academia. The fourth section considers the implications of the two perspectives for research governance in academia. The last section concludes by arguing that psychological economics presents a promising, novel avenue for research in academic governance.

THEORETICAL BASIS OF THE PRESENT GOVERNANCE SYSTEM FOR ACADEMIC RESEARCH

Over the past years, universities have increasingly adopted the idea that the governance of academic research should be subjected to the same governance as for-profit enterprises. This is reflected in procedures transferred from private companies. The most prominent examples are pay-for-performance for scholars according to output measures like rankings, ratings, and competitive fundraising. Overall, the reforms are aimed at the

establishment of an “enterprise university” (e.g., Bok, 2003; Clark, 1998; Donoghue, 2008; Marginson & Considine, 2000; for business schools, see Khurana, 2007).

This concept is based on new public management and economics of science. The proponents of *new public management* draw on the principal agent view (Kaboolian, 1998) as proposed by Jensen and Murphy (1990). This view dominates new public management (Burgess & Ratto, 2003) in the same way it dominates corporate governance (Daily et al., 2003). Economics that “has won the battle for theoretical hegemony in academia and society as a whole” (Ferraro, Pfeffer, & Sutton, 2005: 10) has come to dominate the analysis of all spheres of life, for instance, the family, art, sport, and religion (Becker, 1976). Today this approach is also applied to academia, either implicitly (e.g., Worell, 2009) or explicitly (e.g., Deem, 2004; Schimank, 2005). According to agency theory, scholars have to be monitored and sanctioned in the same way as managers. The underlying assumption is that control and correctly administered pay-for-performance schemes contain the potential for opportunistic behavior, boost productivity, and lead to an efficient allocation of resources (Lavy, 2007; Swiss, 2005).

According to the *economics of science* in academia the evaluation by the market has to be substituted for the evaluation of peers in the self-governed “republic of science” (Polanyi, 1962). This is the case because of two fundamental characteristics of science, its high uncertainty and its public nature (Dasgupta & David, 1994; Nelson, 1959, 2004; Stephan, 1996).

The fundamental uncertainty of scientific endeavors is due to the fact that success in academia is reflected by success in the market often only after a long delay or sometimes not at all (Bush, 1945; Nelson, 2006). In addition, research often produces serendipity effects; that is, it provides answers to unposed questions (Stephan, 1996). As it is often not predictable how useful a particular research endeavor produces is and whether it ever will be marketable, peers instead of the market have to evaluate whether a piece of research represents an advance.

The public nature of scientific discoveries has been intensively discussed by Arrow (1962) and Nelson (1959, 2006). A discovery must be communicated by scholars as quickly as possible to the community of peers in order to be recognized as the discoverer of a new scientific idea (Dasgupta & David, 1994). In contrast, in profit-oriented enterprises, incentives to transform scientific results into a public good are normally absent.²

As a consequence of these characteristics of research in academia, the “priority rule” has been established as the main success criterion (Dasgupta & David, 1994; Merton, 1957; Stephan, 1996; Gittelman & Kogut, 2003). Only peers can establish scientific priority. Consequently, the peer review system is taken to be the founding stone of academic research evaluation. Instances are awards, honorary doctorates, or membership in prestigious academies (Stephan, 1996; Frey & Neckermann, 2008). Its main form for the majority of scholars consists of publications and citations in

professional journals with high impact factors. Such indicators are provided by academic rankings, based on peer-reviewed publications, citations, and the impact factor of journals like Thomson Scientific's Impact Factor (JIF) (see Garfield, 2006, for a historical review) and the recent h-index (Hirsch, 2005).

Indeed, a well-designed governance system based on peer reviews and academic rankings seems to combine perfectly an output-oriented evaluation of researchers, as postulated by new public management, on the one side, with the requirements of a peer-based evaluation system, as postulated by the economics of science on the other side. Therefore, today these measures are adopted almost universally in academia for most things that matter: tenure, salary, postdoctoral grants, and budget decisions.

However, in recent times a broad discussion arose about the quality of peer reviews (e.g., Starbuck, 2005, 2006) and academic rankings (e.g., Adler, Ewing, & Taylor, 2008; Adler & Harzing, 2009; Lawrence 2002, 2003). It focused mainly on the issues of method, while the theoretical background on which these measures are based was not questioned.

EMPIRICALLY BASED FINDINGS ON THE PRESENT RESEARCH

GOVERNANCE IN ACADEMIA

Findings on Qualitative Peer Reviews

Peer reviews are the backbone of the research governance and evaluation system in academia. However, in recent times, it has been argued that the present peer review

system has major problems (e.g., Abramo et al., 2009; Bedeian, 2004; Campanario, 1996; Frey, 2003; Gillies, 2005, 2008; Starbuck, 2005, 2006; Tsang & Frey, 2007; Wenneras & Wold, 1999).³

Low inter-rater reliability

There is an extensive literature on the low extent to which reviews conform to each other (Cole, 1992; Miner & MacDonald, 1981; Weller, 2001). The correlation between the judgments of two peers falls between 0.09 and 0.5 (Starbuck, 2005).⁴ The correlation is higher for papers rejected than for papers accepted (Cichetti, 1991). This means that peer reviewers are better able to identify academic low performers; that is, it is easier to identify papers that do not meet minimum quality standards than those of high performers and those that are a result of excellent research (Moed, 2007). The reliability thus is particularly low with regard to the opinion of peers among published papers in top journals (Lindsey, 1991).

Low prognostic quality

The reviewers' rating of manuscript quality is found to correlate only 0.24 with later citations (Gottfredson, 1978). According to Starbuck (2006: 83–84), the correlation of a particular reviewer's evaluation with the actual quality as measured by later citations of the manuscript reviewed is between 0.25 and 0.3; this correlation rarely rises above 0.37. Although there is evidence that higher prestige journals publish more high-value articles (Judge, Cable, Colbert, & Rynes, 2007), there is much randomness in editorial selections

(Starbuck 2005). As a result, one editor even advises rejected authors to “Just Try, Try Again” (Durso 1997). However, this strategy overburdens reviewers and tends to lower the quality of reviews. For example, reviewers have neither enough time nor the incentive to check the quality of the data and of the statistical methods employed, as some striking examples in economics demonstrate (Hamermesh 2007).

Reviewers’ biases

Many rejections in highly ranked journals are documented, even regarding papers that later were awarded high prizes, including the Nobel Prize (Campanario, 1996; Gans & Shepherd, 1994; Horrobin, 1996; Lawrence, 2003). Reviewers find methodological shortcomings in 71 percent of papers contradicting the mainstream, compared to only 25 percent of papers supporting the mainstream (Mahoney, 1977).

Findings on Bibliometrics

Advantages of bibliometrics

As a reaction to the criticism of qualitative peer reviewing, bibliometric methods, that is, rankings and ratings based on the number of publications, citations, and impact factors have become more prominent.⁵ This procedure is expected to produce several advantages over qualitative peer reviews (e.g., Abramo et al., 2009).

First, it is more objective because it is based on more than the three or four evaluations typical for qualitative approaches. Although it is based on qualitative peer

reviews because the articles counted must have passed peer evaluation, there may occur a balance of reviewers' biases by the aggregation of many reviewers' evaluations by scientific statistical methods (Weingart, 2005).

Second, the influence of the old boys' network may be avoided. An instrument is provided to dismantle unfounded claims to fame. Rankings can serve as fruitful, exogenous shocks to some schools and make them care more about the reactions of the public (Khurana, 2007: 337).

Third, it is cheaper than qualitative reviews, at least in terms of time. It admits updates and rapid intertemporal comparisons.

Fourth, outsiders to the scientific community, for example, politicians, administrators, journalists, and students, may get a transparent and easy to comprehend picture of scholarly activity. The evaluation process is externalized and has been said to have unlocked the "secrets of the world of research" (Weingart, 2005: 119). In particular, politicians and deans consider rankings an objective measure to allocate resources and to provide compensation packages (e.g., Worrell, 2009). Scholars themselves use them to assess the research quality of their peers.

However, in recent times, the disadvantages of bibliometric methods have been hotly discussed (Adler & Harzing, 2009; Adler et al., 2008; Butler, 2007; Donovan, 2007b; Weingart, 2005). There are three groups of problems. Until now, mainly technical

and methodological problems were highlighted (van Raan, 2005). The third group, the dysfunctional reaction of scholars and institutions has been discussed less.

Technical problems

Technical problems consist of errors in the citing-cited matching process, leading to a loss of citations to a specific publication. First, it is estimated that this loss amounts on average to 7 percent of the citations. In specific situations, this percentage may even be as high as 30 percent (Moed, 2002). Second, there are many errors made in attributing publications and citations to the source, for example, institutes, departments, or universities. In the popular ranking of the Shanghai Jiao Tong University, these errors led to differences of possibly 5 to 10 positions in the European list and about 25 to 50 positions in the world list (Moed, 2002). The most important impact factor, Thomson's ISI Web of Science, is accused of having many faults (Monastersky, 2005; Taylor, Perakakis, & Trachana, 2008). It is unlikely that the errors are distributed equally. Kotiaho, Tomkin, & Simmons (1999) find that names from unfamiliar languages lead to a geographical bias against non-English speaking countries. Third, it has been shown that small changes in measurement techniques and classifications can have large effects on the position in rankings (Ursprung & Zimmer, 2006).

Methodological problems

Methodological problems of constructing meaningful and consistent indices to measure scientific output recently have been widely discussed (Adler & Harzing, 2009; Adler et al., 2008; Lawrence, 2002, 2003; Frey, 2003; Frey, forthcoming). Therefore, we briefly mention the main problems discussed in the literature.

First, there are selection problems. Only journal articles are selected for incorporation in the rankings, although books, proceedings or blogs contribute considerably to scholarly work. Other difficulties include the low representation of small research fields, non-English papers, regional journals, and journals from other disciplines even if they are highly ranked in their respective disciplines. Hence, collaboration across disciplinary boundaries is not furthered.

Second, citations can have a supportive or rejective meaning or merely a halo or herding effect. The probability of being cited is a function of previous citations according to the “Matthew effect” in science (Merton 1968). Simkin and Rowchowdhury (2005) estimate that, according to an analysis of misprints turning up repeatedly in citations, about 70–90 percent of scientific citations are copied from the list of references used in other papers; that is, 70–90 percent of the papers cited have not been read. Consequently, incorrect citations are endemic. They are promoted by the increasing use of meta-analysis, which generally does not distinguish between high and low quality analyses (Todd & Ladle, 2008). In addition, citations may reflect fleeting references to fashionable “hot topics.”

Third, using the impact factor of a journal as a proxy for the quality of a single article leads to substantial misclassification. Singh, Haddad, & Chow (2007) and Starbuck (2005) found that in management research many top articles are published in non-top journals, and many articles in top journals generate very few citations (see for economics Laband & Tollison, 2003; Oswald, 2007; for the journal *Nature* Campbell, 2008). A study of the “International Mathematical Union” even concludes that the use of impact factors can be “breathtakingly naïve” (Adler et al., 2008: 14) because it leads to large error probabilities.

Fourth, there are difficulties comparing citations and impact factors between disciplines and even between subdisciplines (Bornman, Mutz, Neuhaus, & Daniel, 2008).

However, even if these technical and methodological problems could be resolved, there are problems caused by the dysfunctional reactions of scholars and institutions. Even more and better indicators could not overcome these kinds of problems (Osterloh & Frey, 2009). They will occur even if the bibliometric system were to work perfectly.

Individual dysfunctional reactions.

The dysfunctional reactions of individual scholars consist of goal displacement and counterstrategies to “beat the system.” Goal displacement (Perrin, 1998) means that people maximize indicators that are easy to measure and disregard features that are hard to measure. This problem is also discussed as the multiple-tasking effect (Holmstrom &

Milgrom, 1991; Ethiraj & Levinthal, 2009). There is much evidence of this effect in laboratory experiments (Gilliland & Landis, 1992; Ordonez, Schweitzer, Galinsky, & Bazerman, 2009⁶; Schweizer, Ordonez, & Douma, 2004; Staw & Boettger, 1990). For example, Fehr and Schmidt (2004) show that output-dependent financial incentives lead to the neglect of non-contractible tasks. This problem is avoided when principals are offered a fixed wage.

Empirical field evidence of goal displacement in academia is the “slicing strategy” (Weingart, 2005: 125) whereby scholars divide their research results into a “least publishable unit” by breaking them into as many papers as possible. This reaction is amplified when funding is dependent on the quantity of published papers. This was demonstrated in a study for Australia (Butler, 2003). The mid-1990s saw a linking of the number of peer-reviewed publications to the funding of universities and individual scholars. The number of publications increased dramatically, but the quality as measured by citations decreased. It could be argued that a remedy to this problem consists of resorting to citation counts. While this remedy overcomes some of the shortcomings of publication counts, it is subject to the technical and methodological problems mentioned.

Counterstrategies are more difficult to observe than goal displacement (Butler, 2007). They consist of altering research behavior itself in order to “beat the system” (Moed, 2007). Numerous examples can be found in educational evaluation (e.g., Haney,

2002; Heilig & Darling-Hammond, 2008; Nichols, Glass, & Berliner, 2006). The following behaviors are of special relevance in academia.

Scholars distort their results to please, or at least not to oppose, prospective referees. Bedeian (2003) finds evidence that no less than 25 percent of authors revised their manuscripts according to the suggestions of the referee although they knew that the change was incorrect. Frey (2003) calls this behavior “academic prostitution”.

Authors cite possible reviewers because the latter are prone to judge papers more favorably that approvingly cite their work, and these same reviewers tend to reject papers that threaten their previous work (Lawrence, 2003: 260).⁷ Authors willingly adapt to editors who pressure them to cite their respective journals in order to raise their impact rankings (Garfield, 1997; Smith, 1997; Monastersky, 2005).

To meet the expectations of their peers—many of whom consist of mainstream scholars—authors may be discouraged from conducting and submitting creative and unorthodox research. (Armstrong, 1997; Gillies, 2008; Horrobin 1996; Prichard & Wilmott 1997).

Institutional dysfunctional reactions

Dysfunctional reactions on the institutional level are manifold. Most importantly, the ranking system based on self-organized peer evaluation paradoxically results in an intensified control from outside the “republic of science,” in particular, by administrators

and politicians. “Managers are stealing power from scientists” (Lawrence, 2003: 259). Quantitative output indicators give politicians and administrators a handy instrument to manage academia from outside without knowing the process and content of research. However, managerial control theory suggests that such output control is inefficient when outputs are ambiguous and subject to change, and there are information asymmetries between the controller and the controlled. In these cases, only clan control is appropriate (Ouchi, 1977, 1980; Eisenhardt, 1985; Turner & Makhija, 2006), which is exerted by peers. Therefore, the intervention of administrators and politicians into academic self-governance produces unintended effects on the academic system.

First, in academia, output control systems handling research from outside the system create a *lock-in effect*. Even those scholars and academic institutions that are aware of the deficiencies of bibliometrics and the administrative interventions based on them do well not to oppose them. If they do, they are not only accused of being afraid of competition, but also of not wanting to increase the prestige and resources of their department or university. Therefore, it is a better strategy to follow the rules and to play the game. A self-enforcing cycle sets in. For example, in several countries, highly cited scientists are hired immediately before the evaluation of departments and programs are scheduled to take place in order to raise publication and citation records. Such “stars” are highly paid although they often have little involvement with the respective university. (Brook, 2003; Stephan, 2008).

Second, a negative *walling-off effect* sets in. Scholars themselves are inclined to apply output criteria to evaluate each other in order to gain more resources for their research group or department. In addition, it is easier to count the publications and citations of colleagues than to evaluate the content of their scholarly contributions. By doing this, scholars delegate their own judgment to the counting exercise behind bibliometrics, although, by using such metrics, they admit their incompetence in that subject (Browman & Stergiou 2008). This practice is defended by arguing that specialization in science has increased so much that even within disciplines it is impossible to evaluate the research in neighboring fields (Swanson, 2004; van Fleet, McWilliams, & Siegel, 2000). However, this practice in turn reinforces specialization and furthers a walling-off effect between disciplines and subdisciplines. By using output indicators instead of communicating on the contents, the knowledge in the various fields becomes increasingly disconnected. This widens the gap between theory and practice and hampers the ability to create radical innovations that often cross disciplinary borders (Amabile, Conti, Coon, Lazenby, & Herron, 1996; Dogan, 1999).

Third, *research is increasingly homogenized*. Research endeavors tend to lose the diversity that is necessary for a creative research environment. This consequence was pointed out for business schools by Gioia & Corley (2002). For economics, Great Britain provides an example: the share of heterodox, not strictly neoclassical economics sank drastically since the ranking of departments became based mainly on citation counts.

Heterodox journals have become less attractive for researchers due to their smaller impact factor when compared to mainstream journals (Lee, 2007; see also Holcombe, 2004).

Fourth, it is argued that a *positional competition* or a rent-seeking game takes place instead of an enhancement of research quality by the increased investment by universities and journals in evaluating research (Ehrenberg, 2000). It has been shown that the percentage of “dry holes” (i.e., articles in refereed journal which have never been cited) in economic research during 1974 to 1996 has remained constant (Laband & Tollison, 2003), though the resources to improve the screening of papers have risen substantially.

Despite the various criticisms, there seems to be a consensus that there is no alternative to bibliometrics and rankings as the main basis for research governance in academia. The theoretical basis for this system is not discussed. In particular, it is not questioned whether new public management and agency theory really do match the conditions of scientific work. Therefore, the dominant view emphasizes that a strong effort must be made to improve the present system (Adler & Harzing, 2009; Albers, 2009; Butler, 2007; Moed, 2007; Starbuck, 2009).

PSYCHOLOGICAL ECONOMICS AS THE THEORETICAL BASIS FOR ACADEMIC GOVERNANCE

In research on corporate governance in recent times, agency theory as the dominating theoretical frame of corporate governance has been questioned (Daily, Dalton, & Canella, 2003; Hillman & Daziel, 2003; Sundaramurthy & Lewis, 2003). The same is true for new public management (Bogh-Andersen, 2007; Giauque, 2003; Osterloh, Frey, & Homberg, 2008). Yet, in the present governance system of academic research, agency theory and its implications are still in place,⁸ and its theoretical foundations are not discussed.

To do this, we draw on the newly emerging field of psychological economics. It has emerged because of criticism about the assumptions of homo economicus, the standard economic model of human behavior (for reviews, see Fehr & Falk, 2002; Rabin, 1998). It seeks to reintroduce psychology into economics after standard economics had driven it out (Bruni & Sugden, 2007). Psychological economists investigate deviations from homo economicus in three main directions (Frey & Benz, 2004). First, individuals are boundedly rational. They are not able to maximize their expected utility. Second, individuals are boundedly self-interested. Depending on the circumstances, persons are not driven only by external rewards, but also by intrinsic motivation and prosocial preferences. Third, the utility concept of homo economicus is bounded. Psychological economists investigate subjective well-being or happiness as a measure for utility that

goes beyond financial income (Frey & Stutzer, 2002a, 2002b). In our analysis, we focus on the aspect of bounded self-interest, since this is arguably the most contested aspect of the agency paradigm.

Agency theory is based on the idea that individuals respond systematically in a self-interested way on extrinsic incentives from outside. This view disregards intrinsic preferences. People do not only react in an instrumental way, but they also act for their own sake or because of reasons lying within their own person. Examples are compliance with civic virtues, social or professional norms, self image (Ajzen, 1988), a flow experience in a fascinating activity (Cikszentmihalyi, 1975), or curiosity.

It is generally acknowledged that for academic research intrinsically motivated curiosity is of decisive importance (Amabile, 1996, 1998; Stephan, 1996). In standard economics and agency theory, these kinds of preferences are assumed to be a given and can be treated as constant. However, there exists considerable empirical evidence in psychology and psychological economics that this is not the case. Rather, there is a crowding-out effect of intrinsic motivation by externally imposed goals and incentives as well as by perceived unfair treatment provided that intrinsic motivation exists in the first place (Deci, Koestner, & Ryan, 1999; Falk & Kosfeld, 2006; Frey, 1992, 1997; Ordonez et al., 2009; for a survey of the empirical evidence, see Frey & Jegen, 2001).⁹ According to self-determination theory (Gagne & Deci, 2005), intrinsic motivation relies on two preconditions, autonomy and a supportive feedback helping to enhance ones

competence.¹⁰ Crowding-out intrinsic motivation is explained by a reduction of autonomy and a controlling instead of a supportive feedback.

From the point of view of psychological economics, output oriented rankings have four disadvantages that cannot be managed by improving the present governance system in academic research based on agency theory. First, intrinsically motivated curiosity to do research tends to be crowded out and is in danger of being substituted by extrinsic motivation to score high in rankings. Content loses importance (Kruglansky, 1975). Autonomy can be reduced by quantitative output measurements, in particular, if they are linked to incentives. A supportive feedback is not provided by quantitative output measurements because in contrast to qualitative peer reviews they do not tell scholars how to improve their research.

Second, if intrinsic motivation is crowded out and extrinsic motivation prevails, the dysfunctional reactions of scholars like goal displacement and counterstrategies are enforced because they are not constrained by intrinsic preferences. The inducement to “game the system” in an instrumental way may get the upper hand.

Third, a negative self-selection effect takes place, in particular, when monetary rewards are linked to the position in rankings. According to Merton (1973), in academia, there exists a special incentive system called “taste for science”. It is characterized by a relatively low importance of monetary incentives and a high importance of peer recognition and autonomy. People are attracted to research for which, at the margin, the

autonomy to satisfy their curiosity and to gain peer recognition is more important than money. They value the possibility of following their own scientific goals more than financial rewards. These scholars are prepared to trade-off autonomy against money, as empirically documented by Stern (2004): scientists pay to be scientists. The preference for autonomy to choose their own goals is important for innovative research in two ways. It leads to a useful self-selection effect, and autonomy is the most important precondition for intrinsic motivation, which in turn is required for creative research (Amabile, 1998; Amabile et al., 1996; Mudambi, R., Mudambi, S., & Navarra, 2007).

Fourth, a negative self-fulfilling prophecy of agency theory sets in by institutional designs (incentive system, measurement practice, selection process), social norms (obeying the norm of self-interest not to appear as foolish or naïve), and language (evoking a gain frame instead of a community frame) (Ferraro et al., 2005). If intrinsic motivation is crowded out, only extrinsic rewards work—the assumption of agency theory has become true (Ghoshal, 2005; Ghoshal & Moran, 1996; Gibbons, 1998).

IMPLICATIONS FOR THE IMPROVEMENT OF ACADEMIC RESEARCH GOVERNANCE

The two approaches lead to different implications of how improvements of academic research governance can be achieved. Agency theory builds on ever more refined measurements to monitor and control academic researchers. In contrast, psychological

economics builds on careful selection and socialization to academic research, as well as on supporting the intrinsic motivation to undertake meaningful and creative research. These implications are discussed in turn.

Implications from Agency Theory

The proponents of the principal-agency theory for academia are well aware of some of its shortcomings. Three proposals are made to improve the present governance system in academia within the conventional paradigm, in particular, to improve rankings as the backbone of this system.

First, a temporary moratorium of rankings is suggested “until more valid and reliable ways to assess scholarly contributions can be developed” (Adler & Harzing, 2009: 72). As is the case for most authors, they believe that the identification of particular shortcomings should serve as a stepping stone to develop a more reliable research evaluation system (see also Abramo et al., 2009; Starbuck, 2009).

Second, it is suggested that bibliometric indicators should not be used as ready-to-go indicators lacking the competence to understand what is being measured (van Raan, 2005). Therefore, standards of good practice for the analysis, interpretation, and presentation of bibliometric data should be developed and adhered to when assessing research performance. This needs a lot of expertise (Bornmann et al., 2008), which

constrains considerably the responsible use of rankings as a handy instrument for politicians and journalists to assess academic performance.

Third, a combination of qualitative peer reviews and bibliometrics, so-called informed peer reviews, should be applied. It is argued that they can balance the advantages and disadvantages of these two methods (Butler, 2007; Moed, 2007; Weingart, 2005).

While the three proposals may help to avoid some of the disadvantages of bibliometrics, they cannot avoid or balance strategic reactions in the form of goal displacement and counterstrategies of scholars and institutions. This applies even if qualitative and quantitative measures worked perfectly (Osterloh & Frey, 2009).

Implications from Psychological Economics

The application of psychological economics to the governance of academic research is in its infancy; it is therefore only possible to outline some implications in need of more theoretical and empirical analysis.

From the point of view of psychological economics, the following aspects are to be considered. Intrinsic motivation is necessary for academic research but it is undermined by rankings because they curtail autonomy and give no supportive feedback. According to the “taste of science” (Merton, 1973) extrinsic motivation mainly in the

form of peer recognition is important. Monetary compensation plays a role, though a secondary one. Two implications follow, which should be further analyzed.

First, instead of treating scholars as agents who have to be monitored permanently, it should be considered whether it is more appropriate to carefully socialize and select aspiring scholars in order to downplay the controlling role of peer reviews and rankings. The main idea is to find out whether scholars master the state of the art and are creative and intrinsically motivated for research—and then trust that he or she will indeed perform well. This approach to research governance was emphasized by the famous President of Harvard University James Bryan Conant (Renn, 2002):¹¹ “There is only one proved method of assisting the advancement of pure science—that is picking men of genius, backing them heavily, and leaving them to direct themselves.” This view is still part of the “Principles Governing Research at Harvard,” which states:¹² “The primary means for controlling the quality of scholarly activities of this Faculty is through the rigorous academic standards applied in selection of its members.” Such a system may lead a limited number of researchers, after having received tenure, to misuse their autonomy. However, it may be the price that has to be paid for creative research to flourish.

Though autonomy is taken to be essential in this approach, it still requires to some extent informed peer reviews in spite of their deficiencies. This applies during restricted periods, for example, the selection and socialization process and whenever scholars apply

to a new position or for a grant, or submit a paper. However, there is a great difference between being under pressure to publish permanently and being submitted to control during a certain phase, knowing that once this phase is over one will enjoy a wide range of autonomy. If the pressure to publish is low, peer reviews change their role. They can be perceived as supportive instead of controlling and thus will further intrinsic motivation instead of undermining it.

Such governance principles also are employed in other professions characterized by a low degree of observable outputs, such as in the life-tenured American judiciary (Posner, forthcoming). These ideas are in accordance with empirical findings in psychological economics. They show that intrinsically motivated people do not shirk when they are given autonomy (Frey, 1992; Gneezy & Rustichini, 2000; Fong & Tosi, 2007). Instead, they raise their efforts when they perceive that they are trusted (Falk & Kosfeld 2006). This is of decisive importance for knowledge work (Kogut & Zander, 1996; Osterloh & Frey, 2000).

Second, since researchers also are motivated extrinsically, awards may serve as an externally mediated recognition (Frey, 2007; Frey & Neckermann, 2008). In contrast to variable pay for performance, awards are not perceived as controlling. Instead, they are of a symbolic nature that gives supportive feedback. Empirical research suggests that symbolic rewards do not crowd out intrinsic motivation (Heckhausen, 1991). In addition,

criteria for awards are usually not clearly specified ex-ante and thus provide considerably lower incentives for goal displacement.

As already mentioned, psychological economics is in its infancy and needs to be further developed theoretically and empirically. Systematic applications to issues of research governance in academia to our knowledge have not been undertaken. For example, little is known to what relative extent the “taste for science” (Merton, 1973) in different stages of a scholar’s career contains intrinsic elements, the desire for peer recognition, and monetary interests. Another open issue is the implications for the allocation of resources for research. Gillies (2008) suggests that each research unit that has passed the rigorous selection processes should be allocated basic funds sufficient to do meaningful research. Horrobin (1996) argues that the present concentration of resources to huge “centers of excellence” or “research empires” only rarely achieves more than would be possible had the same funds been distributed to small research units. This is in accordance with the considerations that giving more and more resources to a few “research empires” may hinder outsiders from participating in the resource allocation (Burris, 2004; Viner, Powell, & Green, 2004) and cause a decreasing marginal effect of additional research resources. While there exists some empirical work in this regard (Etzkowitz & Leydesdorff, 2000; Jansen, Wald, Frenke, Schmoch, & Schubert, 2007), this issue must be further elaborated.

CONCLUSION

Our paper contributes to governance research, in particular, to the neglected field of governance of academic research. The theoretical foundations of the present dominant view of research governance have rarely been seriously analyzed. Implicitly or explicitly, agency theory and its application in the form of new public management are taken for granted as a theoretical basis, as has long been the case in corporate governance. This approach relies on monitoring and sanctioning to govern agents' behavior. We identify the major shortcomings of the present research governance in academia, in particular, peer reviews and rankings, and confront this view with an alternative approach. We suggest that the new, emerging psychological economics presents a fruitful avenue for research governance. In contrast to agency theory, psychological economics extends the motivational foundations beyond extrinsic preferences. In line with an understanding of scholarly research as a mainly curiosity driven endeavor, we include intrinsic motivation as a major determinant of scholarly behavior. In addition, following Merton (1973), the recognition by peers and supportive feedback is an important part of the motivational bundle of researchers that he aptly calls "taste for science."

We also confront the different implications of the two approaches. Agency theory counts on the refinements of indicators and the measurement process. While these refinements may help to improve research governance, they cannot avoid the strategic

reactions of scholars and institutions. They also lead to a negative lock-in effect and a self-fulfilling prophecy. In contrast, psychological economics counts on a broader theoretical foundation of behavior including scholars' intrinsically motivated curiosity, as well as the desire for peer recognition. In accordance with managerial control theory based on the work of Ouchi (1977, 1979), this alternative approach emphasizes selection and socialization of scholars and symbolic benefits in order to downplay the impact of rankings.

We believe the theoretical ideas presented here provide a useful foundation for future research in a number of areas. In particular research governance should be extended to a general academic governance including teaching which we have not dealt with in this paper. Future research in academic governance could also link up to the discourse on professionalization recently directed to managers (Khurana 2007). It may be that agency theory has contributed to a de-professionalization of the scholarly community, possibly leading to an erosion of professional codes of ethics. In view of the experiences with the recent behavior of some managers, the possibility of such a development is an issue of concern.

REFERENCES

- Abramo, G. D, Angelo, C. A., & Caprasecca. A. 2009. Allocative efficiency in public research funding: Can bibliometrics help? *Research Policy* 38: 206–215.
- Adler, R., Ewing, J., & Taylor, P. 2008. *Citation statistics, A report from the joint committee on quantitative assessment of research* (IMU, ICIAM, IMS). A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS).
- Adler, N. J., & Harzing, A.-W. 2009. When knowledge wins: Transcending the sense and nonsense of academic rankings. *Academy of Management Learning*, 8: 72–95.
- Ajzen, I. 1988. *Attitudes, personality, and behaviour*, Milton Keynes: Open University Press.
- Albers, S. 2009. Misleading rankings of research in business. *German Economic Review*, 3: 352-363.
- Amabile, T. 1996. *Creativity in context: Update to the social psychology of creativity*. Boulder.
- . 1998. How to kill creativity. *Harvard Business Review*, 76: 76–87.
- Amabile, T. M., Conti, R., Coon, H., Lazenby, J. & Herron, M. 1996. Assessing the work environment for creativity. *Academy of Management Journal*, 39: 1154–1184.

- Armstrong, J. S. 1997. Peer review for journals: Evidence on quality control, fairness, and innovation. *Science and Engineering Ethics*, 3: 63–84.
- Arrow, K. 1962. Economic welfare and the allocation of resources for invention. In R. Nelson (Ed.), *The rate and direction of inventive activity: Economic and social factors*: 609–626. Princeton, NJ: Princeton University Press.
- Becker, G. S. 1976. *The economic approach to human behavior*. Chicago: University of Chicago Press.
- Bedeian, A. G. 2003. The manuscript review process: The proper roles of authors, referees and editors. *Journal of Management Inquiry*, 12: 331–338.
- . 2004. Peer review and the social construction of knowledge in the management discipline. *Academy of Management Learning and Education*, 3: 198–216.
- Benz, M., & Frey, B. S. 2007. Corporate governance: What can we learn from public governance? *Academy of Management Review*, 32 (1): 92–104.
- Bogh-Andersen, L. 2009. What determines the behaviour and performance of health professionals? Public service motivation, professional norms and/or economic incentives. *International Review of Administrative Sciences*, 75(1): 79-97.

- Bok, D. 2003. *Universities in the marketplace. The commercialization of higher education*. Princeton and Oxford.
- Bornmann, L., Mutz, R. Neuhaus, C., & Daniel H. D. 2008. Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*, 8 (June): 93–102.
- Brook, R. 2003. Research survival in the age of evaluation. In *Science between evaluation and innovation: A conference on peer review*: 61–66. München: Max-Planck-Gesellschaft.
- Browman, H. I., & Stergiou, K. I. 2008. Factors and indices are one thing, deciding who is scholarly, why they are scholarly, and the relative value of their scholarship is something else entirely. *Ethics in Science and Environmental Politics*, 8: 1–3.
- Bruni, L., & Sugden, R. 2007. The road not taken: How psychology was removed from economics, and how it might be brought back. *The Economic Journal*, 117: 146–173.
- Burgess, S., & Ratto, M. 2003. The role of incentives in the public sector: Issues and evidence. *Oxford Review of Economic Policy*, 19 (2): 285–300.
- Burris, V. 2004. The academic caste system: Prestige hierarchies in PHD exchange networks. *American Sociological Review*, 69: 239–264.

- Bush, V. 1960 [1945]. *Science: The endless frontier*. Report to the president. Washington DC: National Science Foundation.
- Butler, L. 2003. Explaining Australia's increased share of ISI publications—the effects of a funding formula based on publication counts. *Research Policy*, 32: 143–155.
- . 2007. Assessing university research: a plea for a balanced approach. *Science and Public Policy*, 34: 565–574.
- Campanario, J. M. 1996. Using citation classics to study the incidence of serendipity in scientific discovery. *Scientometrics*, 37: 3–24.
- Campbell, P. 2008. Escape from the impact factor. *Ethics in Science and Environmental Politics*, 8 (June): 5–7.
- Cicchetti, D. V. 1991. The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation. *Behavioral and Brain Sciences*, 14: 119–135.
- Cikszentmihalyi, M. 1975. *Beyond boredom and anxiety*. San Francisco: Jossey-Bess.
- Clark, B. R. 1998. *Creating entrepreneurial universities. Organizational pathways of transformation*. Surrey: Pergamon Press.
- Cole, S. 1992. *Making science. Between nature and society*. Cambridge, MA: Harvard University Press.

- Daily, C. M., Dalton, D. R., & Cannella, A. A. 2003. Corporate governance: Decades of dialogue and data. *Academy of Management Review*, 28: 371–382.
- Dasgupta, P., & David, P. A. 1994. Toward a new economics of science. *Research Policy*, 23: 487–521.
- Deci, E. L., Koestner, R., & Ryan, R. M. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125: 627–668.
- Deem, R. 2004. The knowledge worker, the manager-academic and the contemporary UK university: New and old forms of public management? *Financial Accountability and Management*, 20 (2): 107–128.
- Dogan, M. 1999. Marginality. *Encyclopedia of Creativity*, 2: 179–184.
- Donoghue, F. 2008. *The last professors. The corporate university and the fate of the humanities*. New York: Fordham University Press.
- Donovan, C. 2007a. Introduction: Future pathways for science policy and research assessment: metrics vs. peer review, quality vs. impact. *Science and Public Policy*, 34: 538–542.
- . 2007b. The qualitative future of research evaluation. *Science and Public Policy*, 34: 585–597.

- Dosi, G., Marengo, L., & Pasquali, C. 2006. How much should society fuel the greed of innovators? On the relations between appropriability, opportunities and rates of innovation. *Research Policy*, 35: 1110–1121.
- Durso, T. W. 1997. Editor’s advice to reject authors: Just try, try again. *The Scientist*, 11: 13.
- Earley, P. C.; Connolly, T., & Ekegren, G. 1989. Goals, strategy development, and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology*, 74: 24–33.
- Ehrenberg, R. G. 2000. *Tuition rising: Why college costs so much*. Cambridge, MA: Harvard University Press.
- Eisenberger, R., & Cameron, J. 1996. Detrimental effects of reward: Reality or myth? *American Psychologist* 51 (11): 1153–1166.
- Eisenhardt, K. M. 1985. Control: Organizational and economic approaches. *Management Science*, 31 (2): 134–149.
- Ethiraj, S. K., & Levinthal, D. 2009. Hoping for A to Z while rewarding only A: Complex organizations and multiple goals. *Organization Science*, 20: 4–21.
- Etzkowitz, H., & Leydesdorff, L., 2000. The dynamics of innovation: From national systems and “mode 2” to a triple helix of university–industry–government relations. *Research Policy* 29: 109–123.

- Falk, A., & Kosfeld, M. 2006. The hidden cost of control. *American Economic Review*, 96: 1611–1630.
- Fehr, E., & Falk, A. 2002. Psychological foundations of incentives. *European Economic Review* 46: 687–724.
- Fehr, E., & Schmidt, K. M. 2004. Fairness and incentives in a multi-task principal-agent model. *Scandinavian Journal of Economics*, 106: 453–474.
- Ferraro, F., Pfeffer, J., & Sutton, R. I. 2005. Economics language and assumptions: How theories can become self-fulfilling. *Academy of Management Review*, 30: 8–24.
- Fong, E. A., & Tosi, H. L., Jr. 2007. Effort, performance, and conscientiousness: An agency theory perspective. *Journal of Management*, 33: 161–179.
- Frey, B. S. 1992. Tertium datur: Pricing, regulating and intrinsic motivation. *Kyklos*, 45: 161–185.
- . 1997. *Not just for the money: An economic theory of personal motivation*. Cheltenham, UK.
- . 2003. Publishing as prostitution? – Choosing between one’s own ideas and academic success. *Public Choice*, 116: 205–223.
- . 2007. Awards as compensation. *European Management Review*, 4 (1): 6–14.
- . Forthcoming. Economists in the PITS. *International Review of Economics*.

- Frey, B. S., & Benz, M. 2004. From imperialism to inspiration: A survey of economics and psychology. In J. Davis, A. Maricano, & J. Runde (Eds.), *The Elgar companion to economics and philosophy*: 61–83. Cheltenham: Edward Elgar Publishing.
- Frey, B. S., & Jegen, R. 2001. Motivation crowding theory. *Journal of Economic Surveys*, 15 (5): 589–611.
- Frey, B. S., & Neckermann, S. 2008. Awards – A view from psychological economics. *Journal of Psychology*, 216: 198–208.
- Frey, B. S., & Stutzer, A. 2002a. *Happiness and economics: How the economy and institutions affect human well-being*. Princeton, NJ: Princeton University Press.
- . 2002b. What can economists learn from happiness research? *Journal of Economic Literature*, 40 (2): 402–435.
- Gagné, M., & Deci, E. L. 2005. Self-determination theory and work motivation. *Journal of Organizational Behavior*, 26: 331–362.
- Gans, J. S., & Shepherd, G. B. 1994. How are the mighty fallen: Rejected classic articles by leading economists. *Journal of Economic Perspectives*, 8: 165–179.
- Garfield, E. 2006. The history and meaning of the journal impact factor. *The Journal of the American Medical Association (JAMA)*, 295 (1): 90–93.

- Garfield, E. 1997. Editors are justified in asking authors to cite equivalent references from same Journal. *British Medical Journal*, 314: 1765.
- Ghoshal, S. 2005. Bad management theories are destroying good management practices. *Academy of Management Learning and Education*, 4: 75–92.
- Ghoshal, S., & Moran, P. 1996. Bad for practice: A critique of the transaction cost theory. *Academy of Management Review*, 21: 13–47.
- Giauque, D. 2003. New public management and organizational regulation: The liberal bureaucracy. *International Review of Administrative Sciences*, 2003 (69): 567–592.
- Gibbons, R. 1998. Incentives in organizations. *The Journal of Economic Perspectives*, 12 (4): 115–132.
- Gillies, D. 2005. Hempelian and Kuhnian approaches in the philosophy of medicine: The Semmelweis case. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36: 159–181.
- . 2008. *How should research be organised?* College Publication King's College: London, UK.
- Gilliland, S. W., & Landis, R. S. 1992. Quality and quantity goals in a complex decision task: Strategies and outcomes. *Journal of Applied Psychology*, 77 (5): 672–681.

- Gioia, D. A., & Corley, K. G. 2002. Being good versus looking good: Business school rankings and the Circean transformation from substance to image. *Academy of Management Learning and Education*, 1: 107–120.
- Gittelman, M., & Kogut, B. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science* 49 (4): 366–382.
- Gneezy, U., & Rustichini, A. 2000. Pay enough or don't pay at all. *Quarterly Journal of Economics*, 115: 791–810.
- Gottfredson, S. D. 1978. Evaluating psychological research reports: Dimensions, reliability, and correlates of quality judgments. *American Psychologist*, 33: 920–934.
- Haney, W. M. 2002. Ensuring failure: How a state's achievement test may be designed to do just that. *Education Week*, (10 July): 56–58.
- Hamermesh, D. 2007. **Replication in economics**. IZA Discussion Paper No. 2760. Available at SSRN: <http://ssrn.com/abstract=984427>.
- Heckhausen, H. 1991. *Motivation and action*. New York: Springer.
- Heilig, J. V., & Darling–Hammond, L. 2008. Accountability Texas-style: The progress and learning of urban minority students in a high–stakes testing context. *Educational Evaluation and Policy Analysis*, 30 (2): 75–110.

- Hillman, A. J., & Dalziel, T. 2003. Boards of directors and firm performance: Integrating agency and resource dependence perspectives. *The Academy of Management Review*, 28: 383–396.
- Hirsch, J. E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* 102: 16569–16572. <http://dx.doi.org/10.1073/pnas.0507655102>.
- Holcombe, R. G. 2004. The national Research Council ranking of research universities: Its impact on research in economics. *Econ Journal Watch*, 1: 498–514.
- Holmstrom, B. P., & Milgrom, P. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, 7: 24–52.
- Horrobin, D. F. 1996. Peer review of grant applications: a harbinger for mediocrity in clinical Research? *Lancet*, 348: 1293–1295.
- Jansen, D., Wald, A., Frenke, K., Schmoch, U., Schubert, T. 2007. Drittmittel als Performanzindikator der wissenschaftlichen Forschung. Zum Einfluss von Rahmenbedingungen auf Forschungsleistung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 59: 125–149.
- Jensen, M. C., & Murphy, K. J. 1990. CEO-incentives – It's not how much you pay, but how. *Harvard Business Review*, 68 (3): 138–153.

- Judge, T. A., Cable, D. M., Colbert, A. E., & Rynes, S. L. 2007. What causes a management article to be cited – article, author, or journal? *Academy of Management Journal*, 50: 491–506.
- Kaboolian, L. 1998. The new public management: Challenging the boundaries of the management vs. administration debate. *Public Administration Review*, 58 (3): 189–193.
- Khurana, R. 2007. *From higher aims to hired hands: the social transformation of business schools and the unfulfilled promise of management as a profession*. Princeton, NJ: Princeton University Press.
- Kogut, B., & Zander, U. 1996. What firms do?: Coordination, identity, and learning. *Organization Science*, 7: 502–518.
- Kotiaho, J. S., Tomkin, J. L., & Simmons, L. W. 1999. Unfamiliar citations breed mistakes. *Nature*, 400 (July): 307.
- Kruglanski, A. W. 1975. The endogenous-exogenous partition in attribution theory. *Psychological Review*, 82: 387–406.
- Laband, D. N., & Tollison, R. D. 2003. Dry holes in economic research. *Kyklos*, 56: 161–174.
- Lavy, V. 2007. Using performance-based pay to improve the quality of teachers. *Future of Children*, 17 (1): 87–109.

- Lawrence, P. A. 2002. Rank injustice: the misallocation of credit is endemic in science. *Nature*, 415 (21 February): 835–836.
- . 2003. The politics of publication – authors, reviewers, and editors must act to protect the quality of research. *Nature* 422 (20 March): 259–261.
- Lazear, E. P., & Shaw, K. L. 2007. Personnel economics: The economist’s view of human resources. *Journal of Economic Perspectives*, 21: 91–114.
- Lee, F. S. 2007. The research assessment exercise, the state and the dominance of mainstream economics in British universities. *Cambridge Journal of Economics*, 31: 309–325.
- Lindsey, D. 1991. Precision in the manuscript review process: Hargens and Herting revisited. *Scientometrics*, 22: 313–325.
- Locke, E. A., & Latham, G. P. 2009. Has goal setting gone wild, or have its attackers abandoned good scholarship? *Academy of Management Perspectives*, 21, 17–23.
- Mahoney, M. J. 1977. Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy Research*, 1: 161–175.
- Marginson, S. & Considine, M. 2000. *The enterprise university: Power, governance and reinvention in Australia*. Cambridge: Cambridge University Press.
- Merton, R. K. 1957. Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, 22 (6): 635–659.

- . 1968. The Matthew effect in science. *Science*, 159: 56–63.
- . 1973. *The sociology of science: Theoretical and empirical investigation*.
Chicago, IL: University of Chicago Press.
- Miner, L., & McDonald, S. 1981. Reliability of peer review. *Journal of the Society of Research Administrators*, 12: 21–25.
- Moed, H. F. 2002. The impact factors debate; the ISI's uses and limits. *Nature*, 415: 731–732.
- Moed, H. F. 2007. The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy*, 34: 575–583.
- Monastersky, R. 2005. The number that's devouring science. *Chronicle of Higher Education*, 52 (8): A12.
- Mudambi, R., Mudambi, S., & Navarra, P. 2007. Global innovation in MNCs: The effects of subsidiary self-determination and teamwork. *Journal of Product Innovation Management*, 24: 442–455.
- Nelson, R. R. 1959. The simple economics of basic scientific research. *Journal of Political Economy*, 67: 297–306.
- . 2004. The market economy, and the scientific commons. *Research Policy*, 33: 455–471.

- . 2006. Reflections on “The simple economics of basic scientific research”:
Looking back and looking forward. *Industrial and Corporate Change*, 15: 903–
917.
- Nichols, S. L., Glass, G. V., & Berliner, D. C. 2006. High–stakes testing and student
achievement: Does accountability pressure increase student learning? *Education
Policy Analysis Archives*, 14 (1): 1–175.
- Ordonez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. 2009. Goals
gone wild: the systematic side effects of overprescribing goal setting. *Academy of
Management Perspectives*, 23: 6–16.
- Osterloh, M. & Frey, B. S. 2000. Motivation, knowledge transfer, and organizational
forms. *Organization Science*, 11: 538–550.
- . 2004. Corporate governance for crooks. The case for corporate virtue. In A.
Grandori (Ed.), *Corporate governance and firm organization*: 191–211. Oxford:
Oxford University Press.
- . 2009. Are more and better indicators the solution? Comment to William
Starbuck. *Scandinavian Journal of Management*, 25 :225-227.
- Osterloh, M., Frey, B. S., & Homberg, F. 2008. *Pay for performance: Does it really
motivate public officials?* Working Paper University of Zurich.

- Oswald, A. J. 2007. An examination of the reliability of prestigious scholarly journals: Evidence and implications for decision-makers. *Economica*, 74: 21–31.
- Ouchi, W. G. 1977. The relationship between organizational structure and organizational control. *Administrative Science Quarterly*, 22: 95–113.
- . 1979. A conceptual framework for the design of organizational control mechanisms. *Management Science*, 25: 833–848.
- . 1980. Markets, bureaucracies and clans. *Administrative Science Quarterly*, 25: 129–141.
- Perrin, B. 1998. Effective use and misuse of performance measurement. *American Journal of Evaluation*, 19: 367–379.
- Perry, J. L., Mesch, D., & Paarlberg, L. 2006. Motivating employees in a new governance era: The performance paradigm revisited. *Public Administration Review*, 66 (4): 505–514.
- Peters, D., & Ceci, S. J. 1982. Peer review practices of psychological journals: The fate of published articles, submitted again. *The Behavioral and Brain Sciences*, 5: 187–195.
- Polanyi, M. 1962. The republic of science: Its political and economic theory. *Minerva*, 1: 54–73.

- Posner, R.A. Forthcoming. From the new institutional economics to organization economics: Four applications to corporate governance, government agencies, and legal institutions. *Journal of Institutional Economics*.
- Prichard, C., & Willmott, H. 1997. Just how managed is the McUniversity? *Organization Studies*, 18: 287–316.
- Rabin, M. 1998. Psychology and economics. *Journal of Economic Literature*, 36: 11–46.
- Renn, J. 2002. *Challenges from the past. Innovative structures for science and the contribution of the history of science*. Paper presented at the Max Planck Forum 5, Innovative structures in basic decision research, Ringberg Symposium, München.
- Schimank, U. 2005. New public management and the academic profession: Reflections on the German situation. *Minerva*, 43: 361–376.
- Schweitzer, M. E., Ordonez, L., & Douma, B. 2004. Goal setting as a motivator of unethical behavior. *Academy of Management Journal*, 47: 422–432.
- Simkin, M. V., & Roychowdhury, V. P. 2005. Copied citations create renowned papers? *Annals Improb. Res.*, 11 (1): 24–27.
- Simon H. A. 1985. Human nature in politics – the dialogue of psychology with political science. *American Political Science Review*, 79 (2): 293–304

- Singh, G., Haddad, K. M., & Chow, S. 2007. Are articles in “top” management journals necessarily of higher quality? *Journal of Management Inquiry*, 16: 319–331.
- Smith, R. 1997. Journal accused of manipulating impact factor. *British Medical Journal*, 314: 463.
- Starbuck, W. H. 2005. How much better are the most prestigious journals? The statistics of academic publication. *Organization Science*, 16: 180–200.
- . 2006. The production of knowledge. The challenge of social science research. Oxford University Press.
- . 2009. The constant causes of never-ending faddishness in the behavioural and social sciences. *Scandinavian Journal of Management*, 25: 225-227.
- Staw, B. M., & Boettger, R. D. 1990. Task revision: A neglected form of work performance. *Academy of Management Journal*, 33: 534–559.
- Stephan, P. E. 1996. The economics of science. *Journal of Economic Literature*, 34: 1199–1235.
- . 2008. Science and the university: Challenges for future research. *CESifo Economic Studies*, 54: 313–324.
- Stern, S. 2004. Do scientists pay to be scientists? *Management Science*, 50: 835–853.
- Sundaramurthy, C., & Lewis, M. 2003. Control and collaboration: Paradoxes of governance. *Academy of Management Review*, 28: 397–415.

- Swanson, E. 2004. Publishing in the majors: A comparison of accounting, finance, management, and marketing. *Contemporary Accounting Research*, 21 (1): 223–225.
- Swiss, J. E. 2005. A framework for assessing incentives in results-based management. *Public Administration Review*, 65 (5): 592–602.
- Taylor, M, Perakakis, P., & Trachana, V. 2008. The siege of science. *Ethics in Science and Environmental Politics*, 8 (June): 17–40.
- Todd, P. A., & Ladle, R. J. 2008. Hidden dangers of a “citation culture.” *Ethics in Science and Environmental Politics*, 8 (June): 13–16.
- Tsang, E. W. K., & Frey, B. S. 2007. The as-is journal review process: Let authors own their ideas. *Academy of Management Learning & Education*, 6: 128–136.
- Turner, K. L., & Makhija, M. V. 2006. The role of organizational controls in managing knowledge, *Academy of Management Review*, 31: 197–217.
- Ursprung, H. W., & Zimmer, M. 2006. Who is the “Platz-Hirsch” of the German economics profession? A citation analysis. *Jahrbücher für Nationalökonomie und Statistik*, 227: 187–202.
- van Raan, A. F. J. 2005. Fatal attraction: Conceptual and methodological problems in the rankings of universities by bibliometric methods. *Scientometrics*, 62: 133–143.

- van Fleet, D., McWilliams, A., & Siegel 2000. A theoretical and empirical analysis of journal rankings: the case of formal lists. *Journal of Management*, 26 (5): 839–861.
- Viner, N., Powell, P., & Green, R. 2004. Institutionalized biases in the award of research grants: A preliminary analysis revisiting the principle of accumulative advantage. *Research Policy*, 33: 443–454.
- Weibel, A., Rost, K., & Osterloh, M. Forthcoming. Pay for performance for the public sector – benefits and (hidden) costs. *Journal of Public Administration Research and Theory*.
- Weingart, P. 2005. Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, 62: 117–131.
- Weller, A. C. 2001. *Editorial peer review. Its strength and weaknesses*. Medford, NJ: American Society for Information Science and Technology.
- Wenneras, C., & Wold, A. 1999. Bias in peer review of research proposals in peer reviews in health sciences. In F. Godlee & T. Jefferson (Eds.), *Peer review in health sciences*: 79–89. London: BMJ Books.
- Worrell, D. 2009. Assessing business scholarship: The difficulties in moving beyond the rigor-relevance paradigm trap. *Academy of Management Learning*, 8: 127–130.

ENDNOTES

¹ We prefer the expression “psychological economics” rather than “behavioral economics” for two reasons.

First, economists had already examined human behavior before this new field emerged. Second, Simon (1985) points out that the term „behavioral” is misleading since it may be confounded with the „behaviorist” approach in psychology.

² Patents should fulfil the task of transforming the public good “discovery” into a private good and at the same time to communicate the discovery. Patents on the one hand provide an incentive to invest in innovations by a legally enforced temporal monopoly, and on the other hand force to disclose the patent specification. However, it is questionable whether they really fulfil this task. Many discoveries are not patentable or the cost of disclosing are greater than the gains attainable from patenting. Moreover, there is an extensive discussion that today patenting in some fields, in particular university patenting, might have negative impact on the rates of innovation, see Nelson (2004, 2006); Dosi, Marengo & Pasquali (2006).

³ See also the special issue of *Science and Public Policy* (2007) and the Special Theme Section on “The use and misuse of bibliometric indices in evaluation scholarly performance” of *Ethics in Science and Environmental Politics*, 8, June 2008.

⁴ The most discussed study of peer reviewing was conducted by Peters and Ceci (1982). They resubmitted 12 articles to the top-tier journals that had published them only 18 to 32 months earlier, giving the articles fictitious authors at obscure institutions. Only three out of 38 editors and reviewers recognized that the articles had already been published. From the remaining nine articles, eight were rejected.

⁵ For example the British government decided to replace its Research Assessment Exercise based on qualitative evaluations with a system based on bibliometrics. Interestingly, the Australian

Government, which has used mostly bibliometrics in the past, plans in the future to introduce qualitative peer review methods (Donovan, 2007a).

⁶ Locke and Latham (2009) in a rejoinder provide counterevidence to Ordonez et al. (2009). However, they disregard that goal setting may well work for simple but not for complex tasks within an organization. For the latter case, see Earley, Connolly, & Ekegren (1989) and Ethiraj & Levinthal (2009).

⁷ Such problems of sabotage in tournaments have been extensively discussed in personnel economics, see Lazear & Shaw (2007).

⁸ “Pay for performance” which has become scrutinized in new public management e.g. by Perry 2006; Osterloh, Frey & Homberg (2009) as well as in the for-profit management field, e.g., Osterloh & Frey (2004) in academia in some countries like UK and Australia prevail or has recently been introduced like in Germany or Austria.

⁹ The crowding-out effect does not always takes place, e.g. Gerhard & Rynes (2003); Locke & Latham (2009), or is contested, e.g. Eisenberger & Cameron (1996). However the empirical evidence for complex tasks and actors intrinsically motivated in the first place is strong, see Deci, Koestner and Ryan (1999); Weibel, Rost & Osterloh (2009).

¹⁰ A third precondition is social relatedness, see Gagne & Deci (2005).

¹¹ Letter to the *New York Times*, 13 August 1945.

¹² See <http://www.fas.harvard.edu/research/greybook/principles.html>.

CESifo Working Paper Series

for full list see www.cesifo-group.org/wp

(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

- 2733 Francesco Cinnirella and Joachim Winter, Size Matters! Body Height and Labor Market Discrimination: A Cross-European Analysis, July 2009
- 2734 Samuel Bowles and Sandra Polanía Reyes, Economic Incentives and Social Preferences: A Preference-based Lucas Critique of Public Policy, July 2009
- 2735 Gary Burtless, Lessons of the Financial Crisis for the Design of National Pension Systems, July 2009
- 2736 Helmuth Cremer, Firouz Gahvari and Pierre Pestieau, Fertility, Human Capital Accumulation, and the Pension System, July 2009
- 2737 Hans Jarle Kind and Frank Stähler, Market Shares in Two-Sided Media Industries, July 2009
- 2738 Pamela Campa, Alessandra Casarico and Paola Profeta, Gender Culture and Gender Gap in Employment, August 2009
- 2739 Sebastian Gechert, Supplementary Private Health Insurance in Selected Countries: Lessons for EU Governments?, August 2009
- 2740 Leif Danziger, Endogenous Monopsony and the Perverse Effect of the Minimum Wage in Small Firms, August 2009
- 2741 Yan Dong and John Whalley, A Third Benefit of Joint Non-OPEC Carbon Taxes: Transferring OPEC Monopoly Rent, August 2009
- 2742 Valentina Bosetti, Carlo Carraro and Massimo Tavoni, Climate Change Mitigation Strategies in Fast-Growing Countries: The Benefits of Early Action, August 2009
- 2743 Christina Felfe, The Willingness to Pay for Job Amenities: Evidence from Mothers' Return to Work, August 2009
- 2744 Jörg Franke, Christian Kanzow, Wolfgang Leininger and Alexandra Väh, Effort Maximization in Asymmetric N-Person Contest Games, August 2009
- 2745 Bruno S. Frey and Paolo Pamini, Making World Heritage Truly Global: The Culture Certificate Scheme, August 2009
- 2746 Frank N. Caliendo, Is Social Security behind the Collapse of Personal Saving?, August 2009
- 2747 Caterina Liesegang and Marco Runkel, Corporate Income Taxation of Multinationals and Fiscal Equalization, August 2009

- 2748 Chrysovalantou Milliou and Apostolis Pavlou, Upstream Horizontal Mergers and Efficiency Gains, August 2009
- 2749 Rüdiger Pethig and Christian Wittlich, Interaction of Carbon Reduction and Green Energy Promotion in a Small Fossil-Fuel Importing Economy, August 2009
- 2750 Kai Carstensen, Oliver Hülsewig and Timo Wollmershäuser, Monetary Policy Transmission and House Prices: European Cross-country Evidence, August 2009
- 2751 Olaf Posch, Explaining Output Volatility: The Case of Taxation, August 2009
- 2752 Beatrice Scheubel, Daniel Schunk and Joachim Winter, Don't Raise the Retirement Age! An Experiment on Opposition to Pension Reforms and East-West Differences in Germany, August 2009
- 2753 Daniel G. Arce, Dan Kovenock and Brian Roberson, Suicide Terrorism and the Weakest Link, August 2009
- 2754 Mario Larch and Wolfgang Lechthaler, Comparative Advantage and Skill-Specific Unemployment, August 2009
- 2755 Horst Raff and Nicolas Schmitt, Buyer Power in International Markets, August 2009
- 2756 Seppo Kari, Hanna Karikallio and Jukka Pirttilä, The Impact of Dividend Taxation on Dividends and Investment: New Evidence Based on a Natural Experiment, August 2009
- 2757 Mirco Tonin and Michael Vlassopoulos, Disentangling the Sources of Pro-social Behavior in the Workplace: A Field Experiment, August 2009
- 2758 Nicole Grunewald and Inmaculada Martínez-Zarzoso, Driving Factors of Carbon Dioxide Emissions and the Impact from Kyoto Protocol, August 2009
- 2759 Yu-Fu Chen and Michael Funke, Booms, Recessions and Financial Turmoil: A Fresh Look at Investment Decisions under Cyclical Uncertainty, August 2009
- 2760 Jan-Egbert Sturm and Jakob de Haan, Does Central Bank Communication really Lead to better Forecasts of Policy Decisions? New Evidence Based on a Taylor Rule Model for the ECB, August 2009
- 2761 Larry Karp, Sacrifice, Discounting and Climate Policy: Five Questions, August 2009
- 2762 Marianna Belloc and Samuel Bowles, International Trade, Factor Mobility and the Persistence of Cultural-Institutional Diversity, August 2009
- 2763 Charles Noussair and Fangfang Tan, Voting on Punishment Systems within a Heterogeneous Group, August 2009
- 2764 Birgit Bednar-Friedl and Karl Farmer, Internationally Coordinated Emission Permit Policies: An Option for Withdrawers from the Kyoto Protocol?, August 2009

- 2765 Pierre M. Picard and David E. Wildasin, Labor Market Pooling, Outsourcing and Labor Contracts, August 2009
- 2766 Stefan Voigt and Lorenz Blume, The Economic Effects of Federalism and Decentralization – A Cross-Country Assessment, August 2009
- 2767 David S. Jacks, Christopher M. Meissner and Dennis Novy, Trade Booms, Trade Busts, and Trade Costs, August 2009
- 2768 Mario Jametti and Thomas von Ungern-Sternberg, Hurricane Insurance in Florida, August 2009
- 2769 Alessandro Balestrino, Kind of Black: The Musicians' Labour Market in Italy, August 2009
- 2770 Yosr Abid Fourati and Cathal O'Donoghue, Eliciting Individual Preferences for Pension Reform, August 2009
- 2771 Christian Breuer and Chang Woon Nam, VAT on Intra-Community Trade and Bilateral Micro Revenue Clearing in the EU, August 2009
- 2772 Choudhry Tanveer Shehzad, Jakob De Haan and Bert Scholtens, Growth and Earnings Persistence in Banking Firms: A Dynamic Panel Investigation, August 2009
- 2773 Erdal Yalcin, Uncertain Productivity Growth and the Choice between FDI and Export, August 2009
- 2774 Klaus Abberger, Wolfgang Nierhaus and Shynar Shaikh, Findings of the Signal Approach for Financial Monitoring in Kazakhstan, September 2009
- 2775 Sascha O. Becker, Francesco Cinnirella and Ludger Woessmann, The Trade-off between Fertility and Education: Evidence from before the Demographic Transition, September 2009
- 2776 Thomas Aronsson and Erkki Koskela, Optimal Income Taxation, Outsourcing and Policy Cooperation in a Dynamic Economy, September 2009
- 2777 Joel Slemrod, Old George Orwell Got it Backward: Some Thoughts on Behavioral Tax Economics, September 2009
- 2778 Cagri Seda Kumru and Athanasios C. Thanopoulos, Social Security Reform and Temptation, September 2009
- 2779 Alessandro Buccioli and Roel M. W. J. Beetsma, Inter- and Intra-generational Consequences of Pension Buffer Policy under Demographic, Financial and Economic Shocks, September 2009
- 2780 Eduardo Strube and Marcelo Resende, Complementarity of Innovation Policies in the Brazilian Industry: An Econometric Study, September 2009

- 2781 Henry Tulkens and Vincent van Steenberghe, "Mitigation, Adaptation, Suffering": In Search of the Right Mix in the Face of Climate Change, September 2009
- 2782 Maria L. Loureiro, Anna Sanz-de-Galdeano and Daniela Vuri, Smoking Habits: Like Father, Like Son, Like Mother, Like Daughter, September 2009
- 2783 Momi Dahan, Tehila Kogut and Moshe Shalem, Do Economic Policymakers Practice what they Preach? The Case of Pension Decisions, September 2009
- 2784 Eytan Sheshinski, Uncertain Longevity and Investment in Education, September 2009
- 2785 Nannette Lindenberg and Frank Westermann, How Strong is the Case for Dollarization in Costa Rica? A Note on the Business Cycle Comovements with the United States, September 2009
- 2786 Leif Danziger, Noncompliance and the Effects of the Minimum Wage on Hours and Welfare in Competitive Labor Markets, September 2009
- 2787 Gerlinde Fellner, Rupert Sausgruber and Christian Traxler, Testing Enforcement Strategies in the Field: Legal Threat, Moral Appeal and Social Information, September 2009
- 2788 Gabriel J. Felbermayr, Mario Larch and Wolfgang Lechthaler, Unemployment in an Interdependent World, September 2009
- 2789 Sebastian G. Kessing, Federalism and Accountability with Distorted Election Choices, September 2009
- 2790 Daniel Gros, Global Welfare Implications of Carbon Border Taxes, September 2009
- 2791 Louis N. Christofides, Michael Hoy and Ling Yang, The Gender Imbalance in Participation in Canadian Universities (1977-2005), September 2009
- 2792 Jan K. Brueckner and Robert W. Helsley, Sprawl and Blight, September 2009
- 2793 Vidar Christiansen and Stephen Smith, Externality-correcting Taxes and Regulation, September 2009
- 2794 John Beirne, Guglielmo Maria Caporale, Marianne Schulze-Ghattas and Nicola Spagnolo, Global and Regional Spillovers in Emerging Stock Markets: A Multivariate GARCH-in-mean Analysis, September 2009
- 2795 Rüdiger Pethig and Frieder Kollé, Asymmetric Capital-Tax Competition, Unemployment and Losses from Capital Market Integration, September 2009
- 2796 Ngo Van Long, Horst Raff and Frank Stähler, Innovation and Trade with Heterogeneous Firms, September 2009
- 2797 Margit Osterloh and Bruno S. Frey, Research Governance in Academia: Are there Alternatives to Academic Rankings?, September 2009