The Korean Economic Review
Volume 27, Number 1, Summer 2011, 93-114.

The Economics of Citation*

Jeong-Yoo Kim** · Insik Min*** · Christian Zimmermann****

This paper studies the citation decision of a scientific author. When an author can make his argument more persuasive by citing a related work, this is called the correlation effect. On the other hand, when an author cites someone else's work, he gives the impression that he views the cited author as more competent than himself; this is called the signaling effect. These two effects are the main causes of citation bias. Using data from Research Papers in Economics or RePEc, a decentralized database of working papers, journal articles and professional books, we empirically show that a citation bias exists in this field. The empirical finding is obtained by controlling for many variables that affect citation patterns, such as network factors (co-authorship and an author's affiliation) and language.

JEL Classification: D81
Keywords: Citation Bias, Correlation Effect, Signaling Effect, Strategy

I. Introduction

Scientific progress is achieved through the cumulative efforts of scientists. High positive externalities are associated with scientific research, and as such, it should be socially encouraged through monetary or non-monetary rewards. Apart from the

---

Received: Feb. 23, 2011. Revised: April 6, 2011. Accepted: June 2, 2011.
* This research started when the first author was visiting ISER, Osaka University in the winter of 2005. We are grateful to the seminar audiences at the University of Connecticut, Keio University, Kookmin University, Kyoto University, Nagoya University, and Korea Foundation of Advanced Studies (KFAS). We also thank the participants of the Autumn Conference of the Japanese Applied Economics Association held at Chuo University and the 2008 Far Eastern Econometric Society Meeting held in Singapore for their helpful comments. We are also indebted to Kwanho Shin for the useful suggestions in the initial phase of this research.
** Corresponding author, Department of Economics, Kyung Hee University, 1 Hoegidong, Dongdaemunku, Seoul 130-701, Korea, E-mail: jyookim@khu.ac.kr
*** Department of Economics, Kyung Hee University, 1 Hoegidong, Dongdaemunku, Seoul 130-701, Korea, E-mail: imin@khu.ac.kr
**** Economics of Department, University of Connecticut, 341 Mansfield Road Unit 1063, Storrs, CT 06269-1063, U.S.A., E-mail: christian.zimmermann@uconn.edu

self-satisfaction achieved for one's hard work and effort, the main non-monetary reward for scientists is the recognition by their peers,[1] which is accorded to them when their work is cited. Thus, proper citations should be socially appreciated as an important impetus providing strong incentives for research.

As such, how often a scientist is cited can be a measure of his academic performance. Since the initiation of Shepard's Citations in 1873, the Institute for Scientific Information has introduced various citation indices with which to measure a scientist's contribution[2] to his discipline. Those indices have significantly influenced tenure, promotion and reappointment evaluations, as well as other decisions made within universities or research institutions, such as the granting of merit pay or endowed chairs. However, the reliability of these citation indices as measures of a scientist's contribution has been controversial. Do such indices properly reflect the actual contribution? Do scientific authors receive proper recognition for their research when they are recognized on the basis of citations? The answer to these questions depends on whether or not other authors tend to cite non-strategically, in the sense that they give precedence to more closely related precursors. If this is not the case, citation indices based on biased citation patterns do not represent a proper measure of the contribution made by authors.

In this regard, this study investigates whether or not there is a certain systematic bias in citation patterns and if so, where does it originate. To this end, we do not take for granted that citations are made honestly but consider the possibility of strategic citation as well. This approach follows directly from our fundamental underlying belief that "citing is a *strategy*."[3]

Authors may strategically take into account various factors in deciding whether or not to cite a given related work, not simply basing the decision on how closely the work is related to theirs. Many authors claim to have had the embarrassing experience of finding out that their work has not been cited by others in closely related works.[4] Why have the ensuing authors failed to cite a predating related work at the expense of embarrassing or even offending someone? This is because there

[1] Although the relative efficacy of recognition versus monetary rewards is ambiguous, recognition is widely considered as part of comprehensive performance improvement strategy (see Ryan 2005).

[2] The word "contribution" is rather ambiguous in this context. Note that quality and influence cannot be identified, although they may be correlated. Senn (2005) also reports a disparity between the two. In fact, he reports that the divergence of the two over time is innumerably observed in the history of economic thought.

[3] The view that the academic world has been driven, at least partly, by strategic motivations seems to be shared by many researchers. See, for example, Zamora Bonilla (2005).

[4] For example, Barry Palevitz (1997) introduces his own experience in this regard. The incidence in question occurred when he uncovered a paper, which omitted to cite his work, despite the fact that the paper was on a subject that was almost identical to that covered in his work; moreover, one of the authors was well acquainted with his work when they wrote the paper. The reader must surely have had similar experiences.

are some gains from neglecting to cite, that is, there are some costs of citing that exceed the benefits. The decision as to whether or not to cite a related work is made by comparing the costs and the benefits of doing so, not entirely citing with honesty or with a scholarly conscience.

The benefits that an author can obtain from citing a related work are apparent. Above all things,[5] it makes his argument more persuasive. Readers will believe that his argument is more likely to be correct or believable if it is supported by a closely related argument that has been made independently by someone else. As the effect is mainly due to the correlation between the truth of the two arguments, we call this the *correlation effect*.[6] The correlation effect of a clear citation becomes larger when the related argument has also been advanced by a more competent (reliable) author. For example, we say "Confucius said that…," but we seldom say "My friend Charles said that…" to try to convince others of our arguments.

This consideration may create some cost in citing others' work. If an author cites someone else's work, it may give the impression that he views the cited author as more competent than himself. This may make an author reluctant to cite works by others, especially those of authors who are less competent or perceived to be as such. We will call this the *signalling effect*. By neglecting to cite related works of less established authors, an author projects the idea that he believes himself more competent than the author he did not cite. Thus, failing to cite someone else's related work has a vaulting effect in the sense that an author intends to jump in reputation by using someone else as a vaulting tool.

Citing has a second obvious cost. It degrades the readers' evaluation of the originality of his work. One's concern for originality gives rise to a strong incentive not to cite other precursors. This may be a major reason for deliberately neglecting to cite. In order to pretend that his work is original, or at the very least, that he is not aware of the existence of the precursors, an author may purposely neglect to cite predated related works. We will call this the *originality effect*. Generally, a claim can be defined as original if there does not exist another claim from which it is induced; or even if it exists, the former is significantly developed from the latter.[7] If claim 1 is developed from claim 2 in a non-trivial way, citing claim 2 does not affect the originality of claim 1 very much, and claim 1 is likely to be cited. Otherwise, the

---

[5] Citing journal editors or potential referees has an obvious benefit, but the strategy will be neglected herein, because we cannot control for it in our empirical work.

[6] This effect can be interpreted in the paradigm of Latour (1987) as a political motivation to form alliances, with which to connect the current argument to those that appeared beforehand to create a doubt within the mind of those questioning the current work as to whether previous works must also be questioned.

[7] Claim 1 being significantly developed from claim 2 does not mean a low correlation between them, rather, a high degree (many steps) of induction. Given that explicitly modeling the steps of induction makes the analysis unnecessarily complicated, we omit the rigorous analysis associated with the originality effect in this paper.

originality of claim 1 becomes severely affected by citing claim 2, so the author will be reluctant to cite it. The best he could do might be to add the phrase, "This is an independent finding from claim 2," although such statement is hardly verifiable.

There are also minor costs of citing. An author cannot cite all the related works because this is burdensome to both the author and the readers. Moreover, it is costly to search for all the relevant works.[8] Unfortunately, this consideration is often used as an *ex post* excuse for failing to cite related precursors.[9]

When all the private benefits and costs generated from citing are taken into account, failing to cite may not necessarily have to be socially discouraged, at least from a static point of view. For example, citing a predated but only remotely related work may not help readers understand better at all; rather, they may only become confused. Nonetheless, viewed from a dynamic perspective, citing may very well remain socially desirable because authors might not have the sufficient incentives to produce scientific outputs without citations regarded as rewards.[10] Thus, citations may be undersupplied, and the undersupply can be measured by the gap between the *ought-to-be cited* and the *actually cited*. Although the undersupply of citations is, in general, socially undesirable, we are more concerned with a specific form of undersupply of citations. If all papers are equally less likely to be cited than they ought to be, it is less serious than if this happens to a subgroup of authors. Thus, there is a citation bias when, among the most closely related precursors that should be cited from the aspect of dynamic efficiency, some authors more frequently fail to be cited for some strategic reason.

The main purpose of this paper is not to identify all possible aspects of strategic citations, but to highlight citation biases due to strategic citations. Figure 1 illustrates our main point. It examines the correlation between an author's rank and the average rank of those he or she cites.[11] If there is no citation bias, the citation line in the figure would be horizontal. No matter who cites, the pool of cited works would also be similar. However, a positive slope of the citation line drawn in Figure 1 suggests that there is a bias in the citation pattern. In particular, the figure shows that authors tend to cite other authors, whose ranks are high relative to themselves. This means that the selection criterion for citation becomes stricter, as the author's rank is higher. The goal of this paper is to explain the phenomenon of such an

---

[8] Nevertheless, we believe that the searching cost is not as high as some authors argue thanks to the recent development in the Internet and in search engine technologies. An author can now easily obtain a rather comprehensive list of related literature with only a click after typing in some keywords.

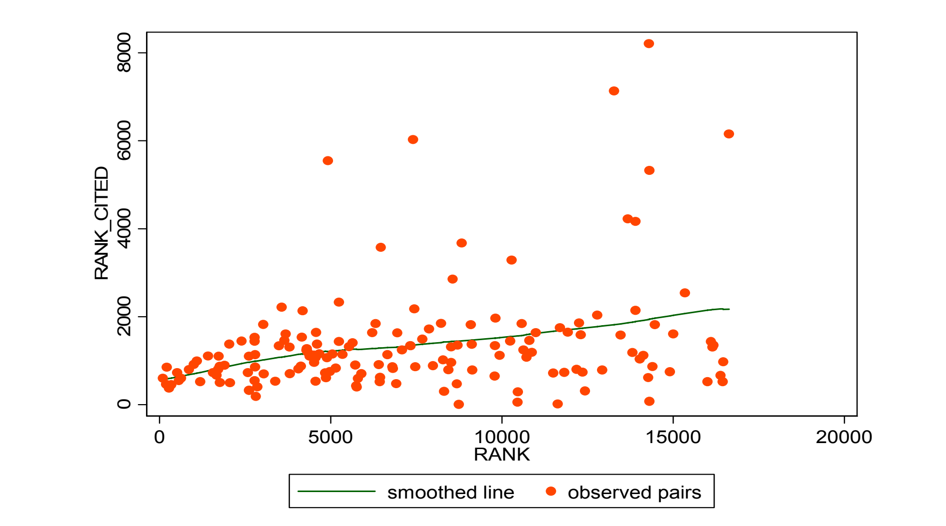
[9] Palevitz (1997) introduces an interesting anecdote describing this situation.

[10] This argument has a straightforward analogy with granting the property right to the early innovator in the case of sequential innovation. For detailed arguments, see Green and Scotchmer (1995).

[11] We use data from Research Papers in Economics (RePEc), which is a decentralized database of working papers, journal articles, and professional books. For more details of RePEc, see http://repec.org/, Krichel (2000) or Section V. Detailed variable descriptions are given in Table 1.

upward bias in citation.

[Figure 1] RANK_CITED vs. RANK



This paper consists of a theoretical and an empirical part. In the theory part, we build a simple model to explain an author's decision to cite. As argued above, we identify two main effects, the correlation effect and the signalling effect. With the correlation effect, an author tends to cite only competent authors whose claims are likely to be correct, because citing a related claim by less competent authors may make his own claim look less reliable. Meanwhile, the signalling effect makes an author—particularly one who is less visible (or reputed)—even more selective in citing. This is because citing a less competent author may give a bad signal about the ability of an author, whose academic ability has yet to become widely recognized. These two effects lead to a citation bias. We also briefly explain the originality effect, although the main focus of the paper is on the correlation effect and the signalling effect.

In the empirical part, we show that a citation bias does exist in Economics using data from Research Papers in Economics (RePEc). The most difficult part in this empirical research is choosing a proxy variable for the visibility of an author. We use an author's seniority as a proxy for his recognizability. We find a more severe citation bias amongst junior authors, that is, juniors tend to be more selective in citations, which is an evidence of the signalling effect. For the empirical task, we tried to control all other possible variables that could potentially affect citation patterns, such as network variables (co-authorship, affiliation) and language.

Our paper is organized as follows. We set up a model in Section II, and provide a theoretical analysis of an author's citation decision in Section III. To separate the correlation effect from the signalling effect, we consider two distinct cases when an author's ability is fully known to all other potential authors and when his ability is known only to a limited number of them. In Section IV, we discuss various factors that can affect citation patterns. In Section V, we present the empirical analysis supporting the results derived in Section III, by controlling the factors discussed in Section IV. Concluding remarks and caveats follow in Section VI.

II. Model

The following model is considered. A scientific author (author 1) makes a claim ω_1 in his writing. This claim can either be true (T) or false (F). Thus, the (average) prior probability (or belief) that his claim is true is $\mu_1 \in \Omega \equiv [0,1]$. We can interpret μ_1 as the ability of the author. He can also be informed of the existence of a related claim, ω_2, by another author (author 2), and it is his private information whether or not he is informed of ω_2. Other people (readers) believe that he is informed of ω_2 with probability γ. The average probability that author 2's claim is true is $\mu_2 \in \Omega$.

Now, author 1 decides whether or not to cite ω_2. We assume that the author is a risk-neutral Bayesian decision-maker, that is, he maximizes the posterior probability that his claim is true. Thus, he decides to cite ω_2 if it increases the posterior probability that ω_1 is true. Let $P(\omega_1 = T \mid \omega_2 = T) = \alpha_T$ and $P(\omega_1 = F \mid \omega_2 = F) = \alpha_F$. We assume that α_T and α_F are common knowledge. If $\alpha_T, \alpha_F > 1/2 (< 1/2)$, the two claims are positively (negatively) correlated, whereas they have no correlation if $\alpha_T, \alpha_F = 1/2$. We assume that $\alpha_T, \alpha_F > 1/2$, i.e., the two claims are positively correlated.[12] Note that a high value of α_T, α_F implies a high correlation.

Asymmetry is allowed in information regarding μ_1 and μ_2. We assume that a proportion λ of the population knows μ_1 for some known parameter $\lambda \in (0,1]$, while the rest does not know μ_1, only its distribution.[13] The distribution is given by $G(\mu_1)$ with mean $E(\mu_1) = \mu_m$, where $G(\mu_1)$ is defined over Ω. Note that we include the case of complete information about μ_1 ($\lambda = 1$). If $\lambda < 1$, then μ_1 is only known to a limited proportion of the public. When a higher λ is associated with author 1, it means that he is a more widely known author. On the other hand, we assume that μ_2 is common knowledge. One could imagine that author 2 is a well-established scholar in terms of name recognition, while author 1 is a junior

[12] This assumption implies that we do not consider negative citations that provide contradictory views or evidence. If $\alpha_T, \alpha_F < 1/2$, author 1 might make a negative citation. In fact, Wright and Armstrong (2007) documents evidence that authors have a tendency to be against negative citations. This empirical evidence justifies our assumption.

[13] For example, it is usual that the ability of a freshly minted Ph.D. is known only locally.

scholar who has just entered academics or a less known scholar affiliated with a small college. We will briefly discuss the case where μ_2 is also unknown in the proceeding section.

In addition to the effect on the posterior probability, citing ω_2 may incur the cost of degrading the evaluation for originality of an author by k. The size of k determines the originality effect. Except for brief discussions on the case that $k > 0$, the originality effect is ignored by assuming $k = 0$ throughout the analysis in order to focus on the correlation effect and the signaling effect. Thus, it is assumed that the cost of citing is only intrinsic (reputational).

III. Formal Analysis

We begin our analysis by defining the formal game. There are two players in this citation game, namely, author 1 and readers. Author 2, who may be cited, is not a valid player because he has no strategy to choose. The strategy set of author 1 is $S = \{cite, not\ cite\}$. We only allow pure strategies. Author 1 has two kinds of private information: μ_1 and awareness of ω_2. In other words, his type set is $T = \Omega \times \Upsilon$, where $\Upsilon = \{aware\ of\ \omega_2, not\ aware\}$. Then, his strategy s should be contingent on his type $s : T \to S$, which is called a strategy rule. The payoff of author 1 is the posterior probability (belief) that ω_1 is correct. After observing s, readers update the posterior belief based on s. Thus, the strategy of readers can be regarded as updating the posterior belief expressed as: $\hat{\mu}_1 : \Omega \times S \to \Omega$.[14,15] This strategy can be rationalized using the implicitly defined loss function, say $L = -E[(\sigma - \hat{\mu}_1)^2]$, where σ is their subsequent decision which is not defined in the model. In other words, readers update the posterior belief with their best decision in order to minimize the loss function.

The main part of our analysis aims to show the two main effects of making a formal citation. To separate the two effects, we start the analysis from the complete information case.

3.1. $\lambda = 1$ (Complete Information)

Given that author 1, who is uninformed of ω_2, has no choice but to not cite another author, we only consider the citation decision of informed author 1.

The intuition can be briefly described as follows. Given that $\lambda = 1$ so that μ_1 and μ_2 are both common knowledge, we then suppose that the prior probability

[14] Although the type of author 1 is originally 2-dimensional, only the belief on μ_1 matters.

[15] This setup is quite common, especially in the literature of reputational cheap talk. See, for example, Sharfstein and Stein (1990) and Trueman (1994).

that author 1's claim is correct is 0.7. If the prior probability that author 2's claim is correct is 0.9 and the correlation between the two claims is almost 1, citing author 2's claim increases the posterior probability that author 1's claim is correct to almost 0.9. Thus, author 1 benefits from the citation.

Formally, if author 1 cites ω_2, the posterior belief that claim 1 is true is given as:

$$P(\omega_1 = T \mid cite\ \omega_2) = P(\omega_1 = T \mid \omega_2 = T)P(\omega_2 = T) \\ + P(\omega_1 = T \mid \omega_2 = F)P(\omega_2 = F).$$

Then, the utility from citing ω_2 can be given by:

$$V_C \equiv E[P(\omega_1 = T \mid cite\ \omega_2)] = \alpha_T \mu_2 + (1-\alpha_F)(1-\mu_2).$$

Given that the expected probability that $\omega_1 = T$ with no citation is $V_N \equiv E[P(\omega_1 = T)] = \mu_1$,[16] the author then chooses to cite ω_2 if and only if

$$\mu_1 < \alpha_T \mu_2 + (1-\alpha_F)(1-\mu_2) \equiv \overline{\mu}_1. \tag{1}$$

Henceforth, we assume that $\overline{\mu}_1 \in (0,1)$. Since citing is socially desirable from the viewpoint of dynamic efficiency, we can say that there is a distortion in the sense that too few citations occur. Inequality (1) implies that a less capable author is more likely to cite another of given capability. The intuition is quite clear. A less capable author can increase the posterior belief that his claim is correct if he cites the claim by a reasonably competent author, whereas a more capable one simply decreases the posterior belief by citing the claim. We call this the *correlation effect* of citation. The distortion due to the correlation effect can be measured by $1-\overline{\mu}_1$.[17]

Rewriting inequality (1) leads to our result of selective citation in the case of complete information.

Proposition 1 *When* μ_1 *is publicly known, author 1 cites* ω_2 *if and only if* $\mu_2 > \overline{\mu}_2 \equiv \dfrac{\mu_1 + \alpha_F - 1}{\alpha_T + \alpha_F - 1}$.

Proof. Note that $\alpha_T + \alpha_F > 1$. Thus, it is clear that inequality (1) is equivalent to $\mu_2 > \overline{\mu}_2$.

[16] In this case, since μ_1 is known to all, the extra information that he failed to cite ω_2 does not deliberately affect the expected probability that $\omega_1 = T$ at all.

[17] If citing incurs a positive cost of k in terms of originality, author 1 cites ω_2 if and only if $\mu_1 < \alpha_T \mu_2 + (1-\alpha_F)(1-\mu_2) - k \equiv \overline{\mu}_1(k)$. In this case, we can say that $1-\overline{\mu}_1(0)$ is the pure correlation effect and $\overline{\mu}_1(0) - \overline{\mu}_1(k)$ is the originality effect, and that the sum of the two, which is $1-\overline{\mu}_1(k)$, is the (total) correlation effect.

Proposition 1 suggests that an author only cites claims made by competent authors and that he is reluctant to cite claims by unreliable author's ($\mu_2 < \overline{\mu}_2$). Given that only the works of less competent authors tend to be neglected in citation, we can say that the correlation effect leads to a citation bias. The intuition behind Proposition 1 is as follows. Given reasonably high α_T and α_F, ω_2 is likely to be correct if μ_2 is large; this, in turn, implies that ω_1 looks correct by citing ω_2 because of high α_T. Similarly, if μ_2 is small, ω_2 is likely to be false, implying that citing ω_2 makes ω_1 look false because of high α_F.

In addition, let us consider a specific case that $\alpha_T = \alpha_F \equiv \alpha$. If $\mu_2 > 1/2$, the citation benefit gets larger as α increases, so that author 1 becomes more willing to cite ω_2. In an extreme case that $\alpha \approx 1$, author 1 cites as long as the cited author's known ability is higher than his own. However, if $\mu_2 < 1/2$, the citation has a worse effect as α increases. The intuition is clear. Due to the fact that the two claims are more closely related, the truth of ω_2 is more likely to imply the truth of ω_1, while the falseness of ω_2 is more likely to imply the falseness of ω_1. When $\mu_1 = \mu_2 \equiv \mu$, inequality (1) holds if $\mu < 1/2$ but does not if $\mu > 1/2$, implying that an incompetent author ($\mu < 1/2$) always cites the claim by a comparable author, while a competent author does not.

3.2. $\lambda < 1$ (Incomplete Information)

If $\lambda < 1$, the citation decision of an author may convey some meaningful information about μ_1 to uninformed readers. Insofar as the citation decision depends on μ_1 under complete information, readers may be able to infer the author's unknown ability from his citation decision. Taking this into account, an author with unknown ability may cite more selectively to pretend to be more capable. We call this the *signalling effect* of citation.

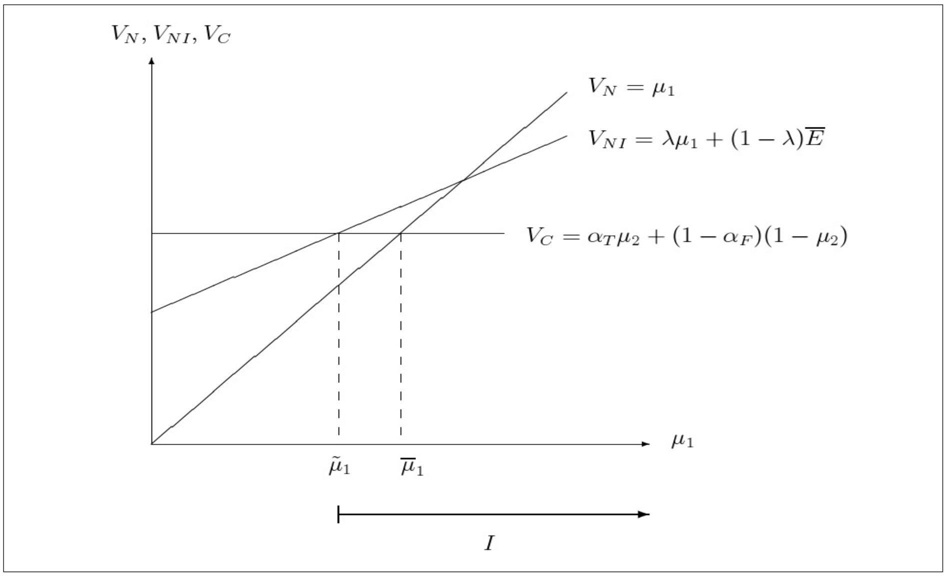
To demosntrate the signalling effect formally, we resort to the usual solution concept, the weak Perfect Bayesian Equilibrium (PBE), which requires the belief of the readers to be updated from the prior belief according to Bayes' law whenever possible. Our interest is confined to the equilibrium outcome that some types of author 1 cite while other types do not. In this semi-separating equilibrium, there must be a type who remains indifferent under incomplete information. The existence of a semi-separating equilibrium is ensured if μ_m is sufficiently high, such that $\overline{\mu}_1 < \mu_m$. Assuming this, let the cutoff type be $\tilde{\mu}_1(\lambda)$. Then, we arrive at Proprosition 2 below.

Proposition 2 *(i) Author 1 cites* ω_2 *if* $\mu_1 \le \tilde{\mu}_1(\lambda)$*, while he does not if* $\mu_1 > \tilde{\mu}_1(\lambda)$*, (ii)* $\tilde{\mu}_1(\lambda) < \overline{\mu}_1$*, and (iii)* $\tilde{\mu}_1(\lambda)$ *is strictly increasing in* λ.

Proof. See the appendix.

This proposition says that a more severe citation bias occurs due to the signalling effect. A less widely known author tends to be more reluctant to cite others. The intuition goes as follows. Apart from the originality effect, citation has two other effects. On one hand, citation directly increases the credibility of an author's claim (correlation effect); on the other hand, it has an indirect signalling effect, thereby adjusting the belief of the author's ability downwards. Thus, an author decides whether to cite or not by taking the two effects into account. Therefore, the citing decision of an author with a very high μ_1 (and a very low μ_1, respectively) will never (hardly respectively) be affected by the incomplete information. An agent with a medium range μ_1, especially close to $\bar{\mu}_1$, who would cite under complete information would rather opt to not cite under incomplete information if he takes into account the extra signalling effect. In Figure 2, the correlation effect is measured by $1-\bar{\mu}_1$ and the signalling effect is measured by $\bar{\mu}_1-\tilde{\mu}_1(\lambda)$.[18]

[Figure 2] Determination of $\bar{\mu}_1$ and $\tilde{\mu}_1$



In this model, an author's attempt to signal by deliberately neglecting to cite gives the same reputation benefit across author types, but is more costly to a type of lower μ_1, because he is giving up providing more convincing arguments to informed readers. Separation becomes possible due to a difference in this signaling cost.

Although our argument is focused on the the separation case, this does not mean that we deny the possibility of a pooling equilibrium. Depending on the prevailing

[18] If $k>0$, the pure correlation effect is $1-\bar{\mu}_1(0)$, and the pure signalling effect is $\bar{\mu}_1(0)-\tilde{\mu}_1(\lambda;0)$. The remaining part is the originality effect, which amounts to $\Delta=(\bar{\mu}_1(0)-\bar{\mu}_1(k))+[(\bar{\mu}_1(k)-\tilde{\mu}_1(\lambda;k))-(\bar{\mu}_1(0)-\tilde{\mu}_1(\lambda;0))]=\tilde{\mu}_1(0)-\tilde{\mu}_1(k)$.

parameter values, there may be two kinds of pooling equilibria: (1) *PP* equilibrium (no-citing equilibrium), in which no informed author 1 cites; and (2) *PS* equilibrium, in which informed author 1 cites for any μ_1.[19] However, the case turns out to be as stated below.

Proposition 3 *(i) There is no pooling equilibrium (all-citing equilibrium) in which all types cites* ω_2. *(ii) If* $(1-\lambda)\mu_m > \mu_1$, *there is a pooling equilibrium (no-citing equilibrium) in which no type of author 1 cites* ω_2.
Proof. See the appendix.

The intuition for the existence of a no-citing equilibrium is as follows. No-citing equilibrium is possible, especially if λ is small and μ_m is high. In this equilibrium, less competent authors can successfully pretend to be more competent by not citing the work of others. It is a large signalling effect by a large proportion of uninformed readers and a high expectation of author ability that enables an author to successfully mimic a more competent author.

An undersupply of citations occurs in the no-citing equilibrium. However, no citation bias occurs since all authors are equally unlikely to be cited. Thus, the citation bias occurs only in the semi-separating equilibrium.

IV. Other Factors Affecting Citations

In this section, we consider many other factors affecting citation decisions. Controlling the variables in the empirical analysis is essential in ensuring the reliable identification of citation bias based on the competence and the reputation of the potentially citing author.

First, an author's affiliation may matter in other authors' citation decisions. This tendency can be stark, especially when μ_2 in our model is not known. In this case, author 1 may determine whether to cite author 2, based on his/her affiliation or other observable public signals associated with him/her (for example, nationality or gender), insofar as the exact value of μ_2 is unknown. This consideration can lead to statistical discrimination in citations. To show this formally, let s_L and s_H be two imperfect signals for μ_2, respectively. If a higher μ_2 is more likely given the signal s_H, that is, the distribution of μ_2 given s_H stochastically dominates the distribution given s_L, it follows that $\mu_2^L < \mu_2^H$, where $E(\mu_2 \mid s_L) = \mu_2^L$ and

[19] Here, *PP* is short for pooling/pooling in the sense that author 1 uses the same pooling strategy of not citing whether or not he is informed of ω_2. Similarly, *PS* is short for pooling/separating. This pooling equilibrium is separating in the sense that any informed author 1 cites, while any uninformed author 1 neglects to cite a previous work.

$E(\mu_2 \mid s_H) = \mu_2^H$.[20] Then, if author 1 cites ω_2 given each signal, we have:

$$V_C(s_L) \equiv E[P(\omega_1 = T \mid cite\ \omega_2, s_L)] = \alpha_T \mu_2^L + (1-\alpha_F)(1-\mu_2^L),$$

$$V_C(s_H) \equiv E[P(\omega_1 = T \mid cite\ \omega_2, s_H)] = \alpha_T \mu_2^H + (1-\alpha_F)(1-\mu_2^H).$$

Given that $\alpha_T + \alpha_F > 1$, we have $V_C(s_L) < V_C(s_H)$, which implies that author 2 with signal s_L is less likely to be cited.

Second, an author's native language is another crucial factor. It is often alleged that authors from English-speaking countries are more frequently cited than those who come from non-English-speaking countries.[21] This citation pattern can be attributed either to the higher quality of publications produced by authors from English-speaking countries or to the parochial citation practices exhibited by these authors. In either case, there is merit in controlling for the author's language variable in order to obtain a more reliable empirical relation between citing authors and cited authors.

Another important feature in the citation decision, which is not modeled in this paper is the network effect in a broad sense. It is often reported that scholars that belong in a small group cite member authors more frequently than those outside the group. There are many possible reasons for this. A group of authors may be well aware of the outputs of one another if they are colleagues affiliated with the same institute, if they are in an advisor/student relation, or if they are coauthors. Since their research overlap significantly, there is a good chance that one author is assigned as a referee for the other. In such cases, citing the other's work even increases the chance of the other being assigned as the referee and the referee favoring the author. Such mutual favors are regarded as a mechanism, which facilitates collusion between authors belonging to the same group. To capture this network effect, one may define the distance between a pair of authors within the co-authorship network. For example, if $d = 1$, the pair of authors are the direct coauthors. We may then exclude citations from authors, whose distance is less than some number d. If $d = 0$, all self-cited articles are excluded, whereas if $d = 1$, the articles by all of his coauthors can be excluded as well.

Recently, some physicists[22] have identified a hub structure in scientific citation networks and explained it by using the notion of preferential attachment. Roughly stated, their argument states that a newcomer in a network (a newly written paper) is more likely to link to an article with more links. In other words, a newly written paper is more likely to cite an article, which is cited more often compared with the

[20] This is a standard result on the first-order stochastic dominance.

[21] See, for example, Seglen (1998), Kurmis (2003), and Leimu and Koricheva (2005).

[22] See, for example, Jeong et al. (2003).

others. In the citation network, they found that each node represents a paper, not an author.[23] Nevertheless, the network structure would be roughly preserved even if each node represented an author instead of an article. Therefore, our theory of citation bias based on the correlation effect and the signalling effect provides a rationale for the preferential attachment in this specific context of the citation network. If each node is identified with an author, the preferential attachment, which is very crucial to a hub structure, can also be interpreted as herding in economic terms,[24] or something like "an author tends to cite someone else simply because many people cite him." This may be another source of citation bias.

V. Empirical Evidence

We use citation data from RePEc. As of August 2008, the RePEc database holds close to 620,000 research items from 757 journals and 2,024 working paper series in Economics and related fields. In addition, 17,071 authors are registered through the RePEc Author Service,[25] with their contact information and list of published works catalogued in the database itself. Finally, the Citations in Economics (CitEc) project[26] performs citation analysis on items in RePEc, which then allows to constitute rankings of all registered authors.

An author's overall rank, *RANK,* is determined by taking a harmonic mean of his ranks in 31 different rankings based on citations, impact factors, paper counts, paper downloads and combinations thereof, removing the best and worst ranks. The impact factors are computed with the citation data collected in CitEc.[27] From 17,071 registered authors, we collect the information given in Table 1.[28] In Figure 1, we plot the *RANK_CITED* variable, which indicates the average rank of cited

[23] In Jeong et al. (2003), for instance, a node is associated with a paper published in 1988 in Physical Review Letters.

[24] See Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992) for informational explanations of herding.

[25] See http://authors.repec.org/ or Barrueco Cruz, Klink, and Krichel (2000).

[26] See http://citec.repec.org/ or Barrueco Cruz and Krichel (2005).

[27] Impact factors are computed using various methods. The exact procedures for the variables are too complex to provide in the text. Those who are interested in the details may refer to http://ideas.repec.org/top/ or Zimmermann (2007).

[28] Some suspect that there may exist alternative explanations to support the upward citation pattern, and suggest us to check whether stratification could be another possible explanation for it. They argue, for example, that big names tend to touch on major, general subjects (e.g., highly abstract theoretical economics), while relatively incompetent authors tend to work only on minor or special ones (e.g., agricultural economics). However, we do not agree that only high rankers tend to be associated with general issues. Moreover, we believe that even if it is the case, the explanation does not seem to be consistent with the identified pattern. If the explanation is correct, we would obtain a curve, which goes upward and then gets flat, because very low rankers also tend to cite only top rankers.

[Table 1] Variable Description

Variable	Description
RANK	Author's overall rank
RANK_NW	Author's rank determined by his number of works weighted by the simple impact factor of their series
RANK_CITED	Average rank of authors cited in this authors' works excluding self-citations: when a cited work is authored by several co-authors, the rank of the one with the highest ranking is used
RANK_CITED_CO	Average rank of authors cited in this authors' works excluding his coauthors
RANK_CITED_AFFI	Average rank of authors cited in this authors' works: excluding the authors affiliated in the same country (or the same state if the author is in North America) as the author
RANK_CITING	Average rank of authors citing this authors' works excluding self-citations
RANK_CITING_CO	Average rank of authors citing this authors' works, excluding his coauthors
RANK_CITING_AFFI	Average rank of authors citing this authors' works: excluding the authors affiliated in the same country (or the same state if the author is in North America) as the author
JUNIOR	Dummy variable whose value is 1 if an author belongs to the junior group who has the first published article in year 2005 or after
ENGLISH	Dummy variable whose value is 1 if this author's affiliation is located in the country whose first language is English
NW_CITING	The number of works citing this author
NW_WORKS	The number of this author's publications
AVE_CITNG	*NW_CITING / NW_WORKS*
AFFI	Author's affiliation: for multiple affiliations, the first affiliation is chosen

Note: Working papers are not included in computing all variables other than RANK already computed in RePEc.

authors, with respect to the author's rank (*RANK*). Here, the *lower* values of both variables indicate *better* researchers (*higher* rankers) in the usual sense. We exclude the authors whose *RANK_CITED* values are zero. It is possible that none of the cited authors are registered, or that references could not be found for any of the author's works, especially if he has only a few of them. Thus, 9,127 of 17,701 authors are considered in this simulation. We draw a random sample size of 200 out of 17,701 authors and investigate the citation pattern of observed pairs for *RANK_CITED* and *RANK* values.[29] Figure 1 reveals that the citation pattern line

[29] Given that a scatter plot rarely helps when the number of observations is 300 or more, we provide

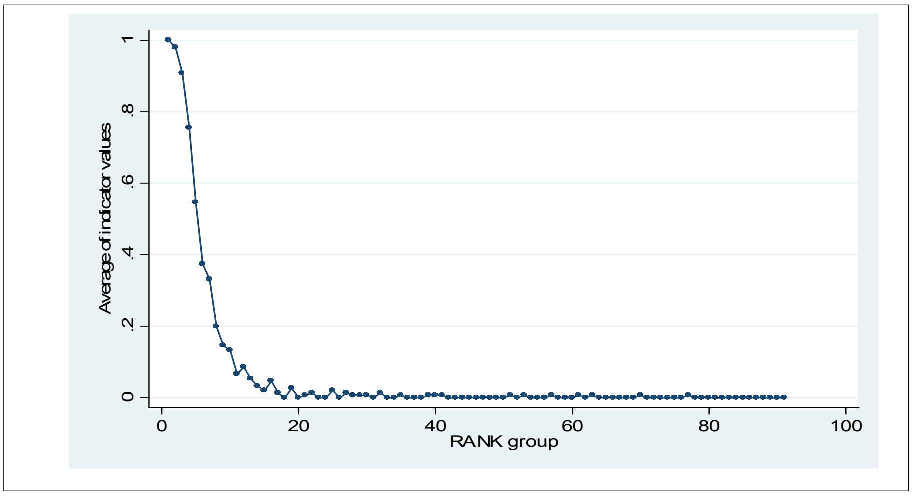
is not horizontal, i.e., the citation pattern is dependent on the author's rank (*RANK*), implying that citation bias does exist. To demonstrate that the slope of the citation pattern line is significantly different from zero, we estimate the following regression equation:

$$RANK\_CITED = \beta_0 + \beta_1 \times RANK + e.$$ [30]

Here, the estimate for β_1 is 0.065 with a standard error of 0.002; thus, we can reject the hypothesis that $\beta_1 = 0$. In addition, a positive slope of the citation pattern line is consistent with our theoretical result that authors tend to cite other authors with higher ranks than themselves.

To examine the citation pattern from another angle, we draw 91 rank groups by assigning about 150 authors to each group according to their overall ranking. For each author, 1 is given if the *RANK_CITED* value is larger than the *RANK* value[31] and 0 if otherwise. The average of the indicator values is then computed for each rank group. The graphical result is reported in Figure 3.

[Figure 3] Average of indicator values for each rank group



With no citation bias, the graph declines smoothly. In Figure 3, however, the graph falls rapidly and we clearly observe that the averaged indicator values are

the scatter plot with the smoothed line based on randomly drawn samples. For the limitation of the scatter plot (see Acock 2006). The citation pattern line is plotted using the Lowess smoothing method.

[30] The estimated coefficient for *RANK* variable is obtained from all 13,748 authors, and not from the randomly selected 200 authors.

[31] This means that the selected author's rank is higher than the average rank of his cited authors, since a low rank value implies a high rank.

recorded as zero from the 24th rank group,[32] indicating that the authors in the middle range are unlikely to cite the authors with lower ranks than their own. Accordingly, Figure 3 is consistent with Proposition 2, indicating that citation bias is more severe among less-established authors if such authors with intermediate ranks are interpreted as less established while top ranking authors are interpreted as established.

[Table 2] Estimation results for models (1.1), (1.2), and (1.3)

	Model 1.1	Model 1.2	Model 1.3
RANK	0.073 (0.002)***	0.065 (0.002)***	0.052 (0.002)***
JUNIOR	-149.49 (27.15)***	-91.04 (25.05)***	-44.13 (23.38)***

Note: *** indicates the estimate is significant at 1% significance level.

To see how robust our finding is, we use the following more refined regression models:[33]

$$RANK\_CITED = \beta_0 + \beta_1 RANK + \beta_2 JUNIOR + e, \tag{1.1}$$

$$RANK\_CITED\_CO = \beta_0 + \beta_1 RANK + \beta_2 JUNIOR + e, \tag{1.2}$$

$$RANK\_CITED\_AFFI = \beta_0 + \beta_1 RANK + \beta_2 JUNIOR + e, \tag{1.3}$$

where *JUNIOR* is a control variable indicating the seniority of an author. We consider three different dependent variables to control the network effect. Here, *RANK_CITED_CO* is a modification of *RANK_CITED*, which controls the tendency to cite coauthors more often, and *RANK_CITED_AFFI* controls the tendency to cite more often the authors affiliated in the same country. Table 2 gives detailed regression results for models (1.1), (1.2), and (1.3). In Table 2, all the coefficients of variable *RANK* are significantly positive, implying that our main result on the correlation effect is largely robust to controlling the variable *JUNIOR*. Note that variable *JUNIOR* is used as a proxy for the low reputation of an author. This can be justified on the grounds that an author's seniority can reflect his recognizability (high value of λ). Although a positive value of β_1 represents the bias due to the correlation effect, a negative value of β_2 implies that there is a bias due to the signalling effect, more specifically, that juniors are more selective in their citations, which is consistent with our intuition and theoretical result.

---

[32] Approximately 3607th–3768th ranked authors are allocated to the 24th rank group.

[33] In order to investigate the interaction effect of *RANK* and *JUNIOR*, we try the estimation models with interaction term *RANK*JUNIOR*. We find that all interaction terms become insignificant; thus, we report estimation results without the interaction term.

[Table 3] Estimation results for models (2.1), (2.2), and (2.3)

	Model 2.1 (IV)	Model 2.2 (IV)	Model 2.3 (OLS)
RANK	0.316 (0.009)***	0.267 (0.010)***	0.210 (0.010)***
RANK*ENGLISH	-0.122 (0.012)***	-0.069 (0.014)***	-0.015 (0.013)
ENGLISH	-152.60 (90.69)*	-167.92 (95.27)*	-271.05 (90.24)***

Notes: 1. *** , ** and * represent 1%, 5% and 10% significance levels, respectively.
2. The p-values for Wu-Hausman endogeneity test are 0.000 (Model 2.1), 0.000 (Model 2.2), and 0.697 (Model 2.3).

As discussed previously in Section IV, the first language of an author is another variable that must be controlled. It is important to note that what matters is the mother tongue of the cited author, not of the citing author. Thus, we reversely use an author's rank as an explanatory variable and the average rank of the *citing* authors as a dependent variable to models (1.1), (1.2) and (1.3). We consider the following regression models:

$$RANK\_CITING = \beta_0 + \beta_1 RANK + \beta_2 ENGLISH + \beta_3 RANK * ENGLISH + e, \quad (2.1)$$

$$RANK\_CITING\_CO = \beta_0 + \beta_1 RANK + \beta_2 ENGLISH + \beta_3 RANK * ENGLISH + e, \quad (2.2)$$

$$RANK\_CITING\_AFFI = \beta_0 + \beta_1 RANK + \beta_2 ENGLISH + \beta_3 RANK * ENGLISH + e. \quad (2.3)$$

Given that the data for the mother tongue of each individual author is not available, we use the variable *ENGLISH*, which is a dummy variable indicating whether or not the author's affiliation is in the country using English as the first official language. We do not include the variable *JUNIOR* in these regression models since the seniority of the *cited* authors is not relevant to the citation decision. One may suspect the endogeneity between *RANK* and *RANK_CITING* in the sense that the value of an author's rank is increased if more authors cite him. We select *RANK_NW*, the rank of an author according to the number of distinct works listed in RePEc, as an instrumental variable and apply two-stage least squares (2SLS) estimation in the model. According to the Wu-Hausman endogeneity test, the null hypothesis of exogeneity is rejected in models (2.1) and (2.2). Thus, Table 3 provides coefficient estimates from 2SLS for models (2.1) and (2.2). We can see that the estimates of β_1 are all significantly positive, implying that citation biases remain even after the language variable is controlled. The negative sign of the estimates for β_2 shows that the authors affiliated in countries using English officially are cited more often. Negative signs for β_3 in models (2.1) and (2.2) imply that the authors in English-speaking countries suffer less from the citation bias pertaining to the correlation effect.

VI. Conclusion and Caveats

In this paper, we provided a theoretical model of citation and tested the results empirically in the field of Economics using data from RePEc. Overall, the empirical results presented in this paper support the hypothesis that citation biases exist in Economics studies.

To confirm the correlation effect and the signalling effect empirically, we controlled many other variables that could possibly affect the citation decisions. However, there are still some aspects left uncontrolled mainly because relevant data are unavailable. For example, patterns citing working papers can be different from patterns citing published journal articles. Due to the fact that the qualities of working papers are yet to be proven, citing them may be attributed more to the signalling effect as compared to citing journal articles. It may be interesting to see whether or not the citation bias is stronger in top journals. To see this, we can divide the publications into two groups, top journals and other journals, estimate the regression models given in (1.1), (1.2), and (1.3) for each group, and then compare the values of coefficient β_1. One problem of this approach, however, is that the sets of authors publishing in top journals and in other journals overlap. Since the samples are not disjoint, it is difficult to test the disparity between the two coefficients. As Stigler and Friedland (1975) argue, the parochial effect exists in citation. Doctorates tend to cite the faculty of their *alma mater* more frequently, and this tendency often lasts even after completing the dissertation. Although this paper tries to control the general network effect by eliminating the citations by coauthors and authors from the same countries, we cannot directly control the parochial effect, because RePEc does not contain the information on the graduate school from which an author completed his doctorate degree. Overall, it is not easy to establish whether or not peculiar citation strategies that should be controlled are significant, especially that of adapting citations to the intended outlet (i.e., citing editors or potential referees, even being asked by referees to cite them). One may argue that better-established authors are less likely to give in to such games or that editors in better journals may not allow such behavior; however, this is only anecdotal evidence we cannot verify without data set. We look forward to future research that shall address these issues.

Appendix

Proof of Proposition 2:

(i) Let I be the set of μ_1 who does not cite in equilibrium. By the definition of $\tilde{\mu}_1$, we have:

$$V_C = \alpha_T \mu_2 + (1-\alpha_F)(1-\mu_2) = V_{NI}(\tilde{\mu}_1),$$

where $V_{NI}(\mu_1) = \lambda\mu_1 + (1-\lambda)\overline{E}$ and $\overline{E} = \gamma E(\mu_1 \mid I) + (1-\gamma)\mu_m$. Then, since $V_{NI}(\mu_1)$ is increasing in μ_1, it is clear that $V_C < V_{NI}(\mu_1)$ for all $\mu_1 > \tilde{\mu}_1$ and that $V_C > V_{NI}(\mu_1)$ for all $\underline{\mu}_1 < \tilde{\mu}_1$.

(ii) By the definition of μ_1, we have $V_C = \overline{\mu}_1$. This implies that:

$$\overline{\mu}_1 = \lambda\tilde{\mu}_1 + (1-\lambda)[\gamma E(\mu_1 \mid I) + (1-\gamma)\mu_m]. \tag{2}$$

We have $\mu_m > \overline{\mu}_1$ by assumption. Note that $E(\mu_1 \mid I) > \tilde{\mu}_1$, because $I = \{\mu_1 \mid \mu_1 > \tilde{\mu}_1\}$. Suppose $\tilde{\mu} \geq \overline{\mu}_1$, then the right hand side exceeds the left hand side in Equation (2), which is a contradiction. Therefore, it must be that $\overline{\mu}_1 > \tilde{\mu}_1$.

(iii) Total differentiation of (2) yields:

$$\left[\lambda + (1-\lambda)\gamma \frac{\partial E(\mu_1 \mid I)}{\partial \tilde{\mu}_1}\right] d\tilde{\mu}_1 + \left(\tilde{\mu}_1 - \overline{E}\right) d\lambda = 0.$$

Given that $\overline{E} > \tilde{\mu}_1$ and $\dfrac{\partial E(\mu_1 \mid I)}{\partial \tilde{\mu}_1} > 0$, the monotonicity of $\tilde{\mu}_1(\lambda)$ with respect to λ follows.

Proof of Proposition 3:

(i) All-citing equilibrium: Suppose an author does not cite in an all-citing equilibrium. Due to the fact that this is on the equilibrium path, (uninformed) readers must believe that it comes from an uninformed author 1. Therefore, we have:

$$V_{NI}(\mu_1) = \lambda\mu_1 + (1-\lambda)\mu_m.$$

The necessary and sufficient condition for the existence of the all-citing equilibrium is:

$$V_{NI}(\mu_1 = 1) = \lambda + (1-\lambda)\mu_m < V_C = \overline{\mu}_1.$$

Given that $\mu_m > \overline{\mu}_1$, this inequality cannot hold, thus completing the proof.

(ii) No-citing equilibrium: In this equilibrium, $V_{NI}(\mu_1 = 0) = (1-\lambda)\mu_m$ since $I = \Omega$; thus, no belief updating follows. If $(1-\lambda)\mu_m > V_C$, clearly $V_{NI}(\mu_1) > V_C$ for all $\mu_1 \in \Omega$, since V_{NI} is strictly increasing in μ_1. This implies that all $\mu_1 \in \Omega$ prefer no citing to citing.

References

Acock, A. C. (2006), A Gentle Introduction to Stata, Stata Press Publication, College Station, TX.

Banerjee, A. (1992), "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, pp. 797-817.

Barrueco Cruz, J., M. Klink and T. Krichel (2000), "Personal Data in a Large Digital Library," In: Proceedings 4th European Conference on Research and Advanced Technology for Digital Libraries, Lisboa.

Barrueco Cruz, J. and T. Krichel (2005), "Building an Autonomous Citation Index for Grey Literature: RePEc, The Economics Working Papers Case," *The Grey Journal: An International Journal on Grey Literature*, 1 (2) pp. 91-97.

Bikhchandani, S., D. Hirshleifer and I. Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, pp. 992-1026.

Green, J. and S. Scotchmer (1995), "On the Division of Profit in Sequential Innovation," *RAND Journal of Economics*, 26, pp. 20-33.

Jeong, H., Z. Neda and A-L. Barabasi (2003), "Measuring Preferential Attachment in Evolving Networks," *Europhysics Letters*, 61, pp. 567-572.

Kim, J.-Y. and J. Park (2006), "On Prejudice," *Scottish Journal of Political Economy*, 53, pp. 505-522.

Krichel, T. (2000), "Working Towards an Open Library for Economics: The RePEc project,"presented at the 'PEAK 2000 Conference: The Economics and Use of Digital Library Collections,' http://openlib.org/home/krichel/papers/myers.html

Kurmis, A. (2003), "Understanding the Limitations of the Journal Impact Factor," *Journal of Bone and Joint Surgery*, 85-A, pp. 2449-2454.

Latour, B. (1987), "Science in Action: How to Follow Scientists and Engineers through Society," Harvard University Press, Cambridge.

Leimu, R. and J. Koricheva (2005), "What Determines the Citation Frequency of Ecological Papers?," *Trends in Ecology and Evolution*, 20, pp. 28-32.

Palevitz, B. (1997), "The Ethics of Citation: A Matter of Science's Family Values," *The Scientist*, 11, pp. 8.

Ryan, S. (2005), "Rewards and Recognition," available at http://edweb.sdsu.edu/people/ARossett/pie/Interventions/incentivesrewards_2.htm

Seglen, P. O. (1998), "Citation Rates and Journal Impact Factors are not Suitable for Evaluation of Research," *Acta Orthopaedica Scandinavica*, 69, pp. 224-229.

Senn, P. (2005), "Influence and the Referee Process," *European Journal of Law and Economics*, 19, pp. 199-206.

Sharfstein, D. and J. Stein (1990), "Herd Behavior and Investment," *American Economic Review*, 80, pp. 465-479.

Stigler, G. and C. Friedland (1975), "The Citation Practices of Doctorates in Economics," *Journal of Political Economy*, 83, pp. 477-507.

Trueman, B. (1994), "Analyst Forecasts and Herding Behavior," *Review of Financial Studies*, 7, pp. 97-124.

Wright, M. and J. S. Armstrong (2007), "The Ombudsman: Verification of Citations: Fawlty Towers of Knowledge?," MPRA Paper No. 4149, available at http://mpra.ub.uni-muenchen.de/4149/

Zamora Bonilla, J. P. (2005), "Scientific Studies and the Theory of Games," forthcoming in Perspectives on Science.

Zimmermann, C. (2007), "Academic Rankings with RePEc," Working Paper 2007-36, University of Connecticut.