TI 2004-113/1
Tinbergen Institute Discussion Paper

Signals in Science

Hendrik P. van Dalen[1]
Kène Henkens[2]

[1] *Department of Economics, Erasmus Universiteit Rotterdam, SEOR-ECRI, and Tinbergen Institute,*
[2] *Netherlands Interdisciplinary Demographic Institute (NIDI), The Hague.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Amsterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to driessen@tinbergen.nl.
Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

Signals in Science - On the Importance of Signaling in Gaining Attention in Science

Hendrik P. van Dalen[1,2] and Kène Henkens[2]

October 15, 2004

[1] *Erasmus University Rotterdam, Department of Economics, SEOR-ECRI and Tinbergen Institute, Rotterdam (The Netherlands), P.O. Box 1738, NL-3000 DR, Rotterdam, The Netherlands, email: vandalen@few.eur.nl*

[2] *Netherlands Interdisciplinary Demographic Institute (NIDI), The Hague (the Netherlands), P.O. Box 11650, NL-2502 AR The Hague, The Netherlands, email: dalen@nidi.nl or henkens@nidi.nl*

Keywords: journals, signaling, first citations, duration dependence

JEL-codes: C41, D83, O31, O33

Abstract

Which signals are important in gaining attention in science? For a group of 1,371 scientific articles published in 17 demography journals in the years 1990-1992 we track their influence and discern which signals are important in receiving citations. Three types of signals are examined: the author's reputation (as producer of the idea), the journal (as the broker of the idea), and the state of uncitedness (as an indication of the assessment by the scientific community of an idea). The empirical analysis points out that, first, the reputation of journals plays an overriding role in gaining attention in science. Second, in contrast to common wisdom, the state of uncitedness does not affect the future probability of being cited. And third, the reputation of a journal may help to get late recognition (so-called 'sleeping beauties') as well as generate so-called 'flash-in-the-pans': immediately noted articles but apparently not very influential in the long run.

1. Introduction

Never judge a book by its cover. Still, in everyday life scholars use all kinds of signals to distill the value of a book or an article. It has always been practice and will surely gain more weight in the future as the number of articles, papers and books exceeds by far the capacity of scholars to even browse or scan the most important research papers in their field. The 'cover' may reveal to a potential reader whether or not the article is worth reading. The author's name and reputation may be one tell-tale sign of quality, but other signals may be just as important. To name a few of the signals used by scholars: the reputation of the publisher or the editorial board of a journal and the number of citations an article has received. In this paper we will assess whether the most common signals used in science – the reputation of authors, journals and the state of uncitedness of an article – are relevant in explaining the attention an article receives in the long run. We will examine this question in depth for the discipline of demography.

The relevance of examining the use of signaling in science is without a doubt important at a number of levels. For the individual reader it may be important because of the simple fact that the reading time of a scholar is scarce. Spending this time in the most efficient manner implies that a scholar does not want to waste time reading a paper of no significance. Publishers and editors also do not want to waste their time, that of their referees or their potential readers. Hence reputations of authors may help to select papers in the decision whether or not to referee a paper. And finally decisions on firing and hiring in academia and the funding of departments are increasingly based on rankings, publication and citation records. The refined division of labor in academia has made it quite difficult to assess the quality of researchers or departments. The use of such signals is therefore primarily based on ignorance of subjects. Deans, administrators, sometimes even colleagues are unable to appreciate the content of research and the only lead to follow are external judgements like citations, publications in refereed journals and prizes. The number of signals has increased tremendously as internet technology allows publishers and others to generate numerous statistics to 'value' a contribution in science: the number of times an abstract or paper is downloaded, rankings of the most popular papers of a journal, or special reports on the hottest papers or authors in science.

In examining the question which signals are of importance in explaining the 'quality' of articles we will use as the unit of our analysis the *articles* published in a set of seventeen

demography journals (as marked by the ISI) in the years 1990-1992.[1] The quality of articles assessed by the community of social science scholars is approximated by the *impact* and *speed* with which knowledge is disseminated in the scientific community. The impact of an article boils down to the number of citations registered by the Web of Science (of ISI). The speed with which an article is disseminated in the scientific community is measured by the timing of the first citation. Of course, besides looking into the separate effects of timing and impact we will also try to see what role signals play when you combine both these quality dimensions. By making a distinction between slow and fast recognition and high and low impact we shed light on the issue touched upon by Glänzel et al. (2003) and Van Raan (2004) who show quite convincingly that publications that go unnoticed for a considerable number of years and then suddenly attract a lot of attention (so-called 'sleeping beauties') are extreme exceptions to the rule.

In this paper we will try to unravel what role signals play for both criteria of scientific quality and in our endeavor we will distinguish between three types of signals: (1) the author's reputation (as producer of the idea), (2) the reputation of the journal (as the broker of the idea), and (3) the state of uncitedness (as an assessment of an article's value by the scientific community).

The first two signals are quite common in scientometrics as much of the discipline revolves around measuring the impact authors and journals have upon the development of science. Author and journal reputation are generally felt to play a role in gaining attention in science. The question we are posing is not so much whether these reputations play a role, but *to what extent*? The third signal – the state of uncitedness - has not been thoroughly examined empirically but its use is quite common in everyday practice as most scholars would claim that the quality of an article could be deduced from its state of uncitedness of an article. The reason why the chances of a first citation decrease over time may be that observed uncitedness of articles signals to prospective readers that the article is of low quality. In other words, uncitedness may become a stigma and the longer an article remains uncited, the lower the perceived quality of the article. The decline over time may, however, also be a reflection of a selection process in which papers with certain characteristics are bound to get cited relatively early, whereas other papers need more time to be noticed and appreciated. In that case, the soc-called negative duration dependence may be at least partially attributed to a composition

1. In Van Dalen and Henkens (2001) we reported earlier on this data set. The present study can be seen as a follow-up study: the robustness of earlier conclusions, based on a five-year period, is tested but we explicitly extend the study by paying close attention to the timing of citations and the role signaling plays in science.

effect. The results of our exercise in first citations are quite illuminating as they suggest that the state of uncitedness does not signal that an article will not be cited in the future. This is in marked contrast with the plain observation of citation statistics, which suggest the existence of negative duration dependence: the longer an article remains uncited, the bigger the chance that the article will never be cited. However, our analysis shows that by taking account of the type of article and the journal in which it appears negative duration dependence is no longer a certainty.

Finally, every scholar hopes that his or her ideas will prove to be path-breaking. In that respect the sign of being noted immediately and being cited many times by the scientific community can be an informative sign and most of the influential ideas in many a science seem to conform to this type of pattern. Still, there are always the odd number of articles which are not noted early on but which gain a lot of attention late in life (so-called 'sleeping beauties'). Of course, the reverse case would need to be examined also because there are some articles which are noted immediately and cited a lot, but which do not seem to have a lasting impact and die early in life (which we call 'flash-in-the-pans'). It would be of some interest to see what role signals play in making such mistakes that are of course quite common in decision making under uncertainty. Our rudimentary assessment of these types of articles is that for both cases the influence of the journal reputation seems to be of considerable importance, whereas the reputation of authors is of negligible influence.

2. Signaling in Science

The importance of signals in the presence of quality uncertainty has been stressed by Akerlof (1970). The intuition for understanding why signals perform such an important role is that in the presence of asymmetric information – one side of a market knows what the quality of good represents, the other side has to guess – markets will fail to exist. Unless, of course, suppliers will be able to signal the quality of the goods supplied or the demand side of the market can profitably screen the goods. Under those circumstances, markets may be able to perform their function and demand and supply for a certain quality of goods can be met. Signals are therefore of utmost importance to show to 'buyers' whether they are dealing with a 'lemon' or not.

By way of analogy, this idea helps one to understand the 'market' for journal articles. The quality of articles differs enormously across the entire spectrum of scholars if one takes the number of citations as an approximation of the quality of an article (Klamer and Van Dalen, 2002). By using signals authors can make clear to their potential audience that they are

dealing with a quality idea and grab the attention of readers so they will invest time in reading the article. In this article we will focus on three types of signals to see how these signals affect timing and impact of an article.

The first signal is the *reputation of the author* or team of authors of an article. The track record of an author (approximated by his or her list of publications, the number of citations received, or prizes received), the academic department with which an author is affiliated or the country in which he resides are all signals tied to the author. The quintessential article stressing the role of reputations in receiving attention is Merton's (1968) article on the Matthew effect in science. The general claim posed by Merton is that there are increasing returns to fame. In other words, authors with high reputations received disproportionately more citations than authors with low reputations.

The second signal is the *reputation of the journal* in which the article is published. A prestigious journal signals to readers that the idea is of high quality. Of course, journals perform a double role as journals are the gatekeepers of the market for ideas. Screening the entire spectrum of articles for each and every individual would be an impossibility and part of the screening has therefore been delegated to journal editors and referees. The quality standards which journal editors uphold represent a screening device distinguishing between lemons and quality articles among the solicited and unsolicited manuscripts. But again here quality differences arise from two sides. First from the supply side, the distribution of high quality articles offered for reviewing differs per journal. Second, from the demand side, not every editor or referee will make the same decision on the initial choice of letting an article be refereed and subsequently in the choice of a suitable referee. By drawing on the editorial correspondence provided by authors of articles in top economics journals, Laband (1990) shows that referees' comments have a positive impact on the subsequent citations of papers. The main contribution of editors is in efficiently matching papers with reviewers. Editors of top journals take great care to match authors with suitable reviewers and the myth that authors with a high (low) reputation get matched with a reviewer of similar standing does not seem to hold up in practice (Hamermesh, 1994). Picking and making winners is not only a science but apparently also an art and so the quality of the editorial board - measured by the reputation of the editors and the past performance of the journal – will impinge on the choice of articles appearing in a journal.

The third signal is a signal that approximates the quality assessed by the academic community: *whether or not an article is cited*. There is a strong presumption in science to dub *uncited* papers as a 'failure' or at least a sign of 'inferiority'. This assessment is generally

made at an aggregate level. E.g., about ten years ago this presumption made headlines when the renown journal *Science* made the effort of collecting the statistics on citations and provoked a discussion on the value of science papers.[2] The gist of this bibliometric exercise was that about half the science papers was never cited within the 5 years time span after publication, thereby confirming the distrusting hunches of policy makers. *Newsweek* even made the bold claim that "nearly half the scientific work in this country is worthless" (April 2, 1991). Later on the figures were corrected for some anomalies but the blow could not be softened. Still, the issue keeps coming back as the force to publish and be cited has increased over time (Frey, 2003) and apparently resources to increase the quality of papers seems to be wasted, as Laband and Tullock (2003) deduce from the constancy of uncitedness over time and the simultaneous increase in time and money spent on academic research. These so-called 'dry holes' in research are a cause of concern for both science policy makers and scientists.

The ultimate question is, of course, whether 'dry holes' are really that dry or to rephrase this: do uncited articles really signal inferior quality? The tacit assumption made by many practitioners is that the chance that an article will be cited diminishes the longer it remains uncited. In short, common wisdom has it that *negative duration dependence* is a widespread phenomenon in the timing of first citations. If this experience rule is an informative signal then dry holes are a real cause of concern. Citations statistics concerning the first citations in *science* (see Glänzel et al., 2003) suggest that this common wisdom may be right. However, till date the empirical evidence of negative duration dependence in the timing of first citations is rather weak.

Finally, sometimes signals can send mixed information and the common statistical error judgements can occur: high quality articles may not immediately be noticed by the scientific community but in the long run may well be noticed. These types of articles – dubbed 'sleeping beauties' by Van Raan (2004) – are a rarity. The study by Glänzel *et al.* (2003) suggests for a large sample of science papers (450.000 articles) that the chance that a highly cited paper can be traced among the laggards is extremely small (0.00014 percent). But, of course, these 'maverick' papers do exist and what Glänzel *et al.* (2003) suggest is that these papers share the common property of being highly mathematical papers published in a 'foreign' (sub)discipline.

The other possibility may also arise: low quality articles may get noticed immediately based on the author's or journal's reputation but in the long run prove to be of little value.

2. See Hamilton (1990, 1991).

These types of statistical 'error' judgements in science (so-called errors I and II) have not been examined in a consistent manner. We make a distinction between four types of paper. The previous two error judgements are already two types. But we would like to also make a distinction between papers that are disregarded completely or cited a few times. And 'normal science' papers: the influence of these papers accords to a standard pattern, viz. that papers of authors or journals with a high reputation are noted faster and cited more often than others.

3. The citation data

In creating a database to test the various ideas we have used the *Web of Science* as published by the *Institute for Scientific Information* (ISI). To get an insight in the long-run impact of journals and their articles we have gathered data on the citation frequency and other characteristics of individual publications in demography journals in three consecutive years (1990-1992). For each article in our data set we established if, and how often they were cited in the ten years following their publication by other scholars who publish their work in the journals covered by the citation indexes published in the Web of Science. As we intend to measure knowledge dissemination in a scientific community, we exclude the number of self-citations by authors in our citation counts. The reason for choosing a ten-year exposure time and not a shorter period can be found in Glänzel and Schoepflin (1995) who report that it takes at least four to five years for articles to be well-accepted and cited in the social science literature (i.e. the highest impact of an article is attained in the fourth or fifth year after publication).

Demography is covered worldwide by some 330 population serials, according to *The Serials Directory* (1994), although a large number of these serials are bulletins of national statistics organizations.[3] Only 17 of the 330 journals have been selected by the SSCI as being important for the development of the discipline. The benefit of using the SSCI selection of demography journals is that it offers a wide variety of journals, not just the prestigious journals of large associations, but also the more specialized and less prestigious journals. The journals we have included in our sample are, in alphabetical order: *Demography*, the *European Journal of Population, Family Planning Perspectives, International Migration, International Migration Review, Journal of Biosocial Science, Journal of Family Welfare, Journal of Population Economics, Population, Population Bulletin, Population and Development Review, Population and Environment, Population Index, Population Research*

3 See for a more in-depth review of the demography journal literature Van de Kaa (2003).

and Policy Review, Population Studies, Social Biology and *Studies in Family Planning*. Book reviews, editorials and other so-called 'marginalia' are excluded in our sample as these types of articles do not contain research results. Data on circulation numbers of the different journals have been obtained from such established databases as *The Serials Directory* and *Ulrich's Plus - The Complete International Serials Database*.

The total sample size consists of 1,371 articles published in the years 1990-1992 in seventeen demography journals. The key characteristics of the consulted journals are summed up in the appendix to this paper. We have collected data at the level of individual articles by hand. In tracking down article content, we consulted all the issues of the journals in the years 1990-1992 by hand and used the electronic database POPLINE.

Descriptive statistics

In order to capture elements of quality and visibility in affecting the impact and the timing of citation we have collected a number of explanatory variables and the descriptive statistics of these variables are summed up in Table 1. The variables have been used before in Van Dalen and Henkens (2001) and are explained here again in brief.

Author characteristics.

An author's reputation is operationalized by the stock of citations accumulated by the author in the year 1990. Where there are two or more co-authors, individual reputations are used to generate an article-specific reputation variable: the reputation of the author with the best reputation. The Matthew effect suggests that the maximum score found among the authors is the best predictor of citation frequency, but in addition to the effect of reputation there should also be increasing returns to scale (read: fame). In order to control for possible non-linearity of the Matthew effect, we have also included a quadratic term. In line with the Matthew effect the coefficient for the quadratic term should be positive.

Besides the reputation of the author, other author variables used in this study are the number of authors and the presence of a US affiliation of at least one of the authors. This variable explicitly refers to the work location of the authors and not to US citizenship as it is the working conditions which matter when building a network. The 'US affiliation' variable is used primarily to test for the importance of connections with the leading country in demographic science, namely the US. As shown in Table 1, more than 50 percent of the articles have been written by an author who is affiliated with a US institution, or by a team of authors, one or more of whom are affiliated with a US institution.

Table 1: Descriptive statistics for variables in analysis (N = 1,371)

	Mean	Standard deviation	Min.	Max.
Citations per article after 10 years	7.12	13.67	0	158
Visibility variables				
Number of pages[a]	9.37	5.23	0.51	32.94
Presidential address	0.004	0.07	0	1
Order of articles in an issue[b]	3.93	1.87	1	6
Comment/reply/note	0.14	0.35	0	1
Content variables				
Historical content/focus of paper	0.05	0.22	0	1
Focus paper				
US/Canada	0.25	0.43	0	1
Europe	0.18	0.38	0	1
Africa	0.07	0.26	0	1
Asia/Australia	0.19	0.40	0	1
Latin America	0.05	0.21	0	1
Middle East	0.02	0.13	0	1
World	0.09	0.28	0	1
Non-empirical focus (e.g. theory, Essays, etc.)	0.15	0.36	0	1
Author variables				
Reputation of the most reputable author of a team (highest number of aggregate citations received 1990)	17.07	33.52	0	625
Number of authors	1.74	1.16	1	13
US connection authors	0.51	0.50	0	1
Journal variable				
Use of french language	0.13	0.34	0	1

(a) Pages are made equivalent to the size of pages of *Demography*, by standardizing for the number of characters on a full page of each journal to those of *Demography*.
(b) This variable has been censored from the right by assigning all articles from number six onward the value 6.

Article characteristics. The characteristics of the articles in question have been operationalized by focusing on indicators that capture the visibility and content of an article. The presidential address is a clear example of how visibility can affect the success of an article. The length of articles was operationalized by counting the number of words on a full-size page in each journal. To obtain a standardized measure, these figures were placed on an equal footing with the pages of *Demography* by taking the average number of words on a *Demography* page as the standard. The type of article (regular article = 0, comment/note/reply = 1) and the order in which an article appears in a journal issue are, in our view, variables that capture the idea of visibility in a journal issue. Because the journals differ considerably with respect to the number of articles appearing in an issue, we have put all articles that appear after the sixth position on a equal footing: all these back-of-the-journal articles receive a value of 6.

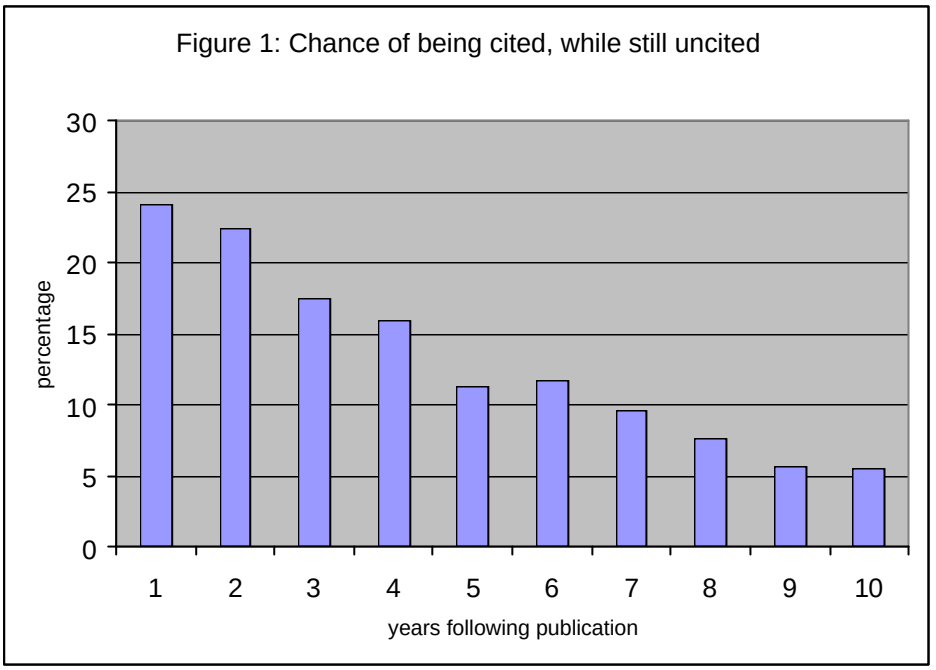
In examining the contents of articles we have constructed two types of dummy variables. First, a set dummy variables categorizes the regional empirical focus of the article in question, the articles with a US focus serving as the reference category. We distinguish the following regions: US/Canada, Europe, Asia/Australia, Africa, Latin America, Middle East, a global focus (hence no particular stress on one region in particular) and finally a non-empirical focus. The latter category includes essays, methodological articles, theoretical articles (either of a verbal nature or of a formal mathematical nature) and discussions. Second, a dummy variable indicates whether or not the article has a historical orientation. If the article contains data about, or an analysis focusing on the period preceding the second World War it has been classified as historical, otherwise not.

Journal characteristics. In order to examine the importance of journals in the allocation of citations we have used two different approaches. First, we followed Stewart's (1983) approach by using dummy variables for each journal in our sample (16 with the leading journal *Demography* as the reference category). Second, we used four distinct variables to operationalize journal differences. The demography journals are characterized by using the SSCI-impact factor of a journal in 1990, the reputation of the editorial board, the circulation numbers, and the use of the French language in communicating research (see the appendix for details). We have used the ISI impact factor to indicate the short-term impact of an article on the scientific literature. The ISI impact factor is based on citations of articles published in the last two years and this definition of impact may give a somewhat distorted picture of how knowledge is disseminated in the social sciences. Therefore, an additional indicator of journal

quality was obtained by calculating the average reputation of each journal's editorial board. The average number of citations received in 1990 was established for the editors and for the advisory editorial boards. The last two variables are straightforward. One dummy variable registers whether the article appeared in French. Two journals allow the French language to be used as a means of communication: the French-based journal *Population* and the *European Journal of Population*. The other variable concerns the circulation of the demography journals in question. The serials databases did not provide any information on circulation numbers of *Population and Environment* and the publisher was not prepared to disclose this information. For this particular case, the sample mean circulation value was imputed, computed from the non-missing values.

Timing first citations

A statistic not mentioned in Table 1 is the timing of first citations. We have presented the chance of being cited while still being uncited in Figure 1. This figure suggests that the chance of being cited declines the longer one remains uncited and this is exactly the image that lingers on in the minds of academics. However, appearances may be deceiving as these aggregate statistics cover up heterogeneity in the type of papers: differences in quality assessment by journal type, differences in author quality, different specializations and accompanying audiences, divergence in citation practices across sub-specializations within demography (e.g., sociology, biology, mathematics, anthropology, economics, gerontology, medicine), differences in content (by focusing on different regions in the world), and differences in the language of communication. In short, we need to examine this question at the micro-level in order discover whether uncitedness at a certain time signals 'inferiority'. Just because an article is uncited two or even five years later does not imply that it will not be noticed and used in the subsequent period(s). Figure 1 shows in detail that this is the case. Figure 1 depicts the probability of a previously uncited article being cited in the current year, by years after publication. Although an uncited article's chances decrease over the years, the decline is relatively modest. In the first year, an uncited article had a 24 percent chance of being cited. An article that did not receive a single citation in its nine years of existence still had a 5 percent chance of being cited in the 10th year. The reason why the chances of a first citation decrease over time may be that uncited articles signal to prospective users that the article is of low quality. In short, uncitedness becomes a stigma and the longer an article is uncited, the lower the quality and the less inclined researchers will be to cite it. The decline



over time may, however, also be a reflection of a selection process in which papers with certain characteristics are bound to get cited relatively early, whereas other papers need more time to be noticed and appreciated. In that case, the negative duration dependence may be at least partially attributed to a composition effect. One characteristic that may be important in this regard is the subfield in which demographers are active. Some demographers, like the family planning researchers, have different citation practices than other groups (migration, social biology, economics, mathematical demography). In Van Dalen and Henkens (1999), we showed how the balance of intellectual trade is structured within demography, but particularly with the outside disciplines. In addition, a number of other characteristics are known to play a role in allocating citations in demography (see Van Dalen and Henkens, 2001). Authors with distinguished reputations have an advantage when competing with 'rookie' authors for the attention of their fellow social scientists. A certain delay in being noticed might also arise due to the fact that some articles are not written in the lingua franca of science, namely English. All of these effects might help explain the decline in the chances of being cited as shown in Figure 1.

4. Explaining impact and the timing of first citations

In examining the various hypotheses we will explain the impact and timing of success by means of three definitions. Below we will introduce the three definitions and the accompanying methods to test the relevance of signaling in science.

4.1 Methods

Cumulative impact

The impact of ideas, i.e. the *frequency of citations* by other scholars, is the more common measuring rod of success in science. To make the split as clear as possible, most scientists would ascribe to the idea that what counts in science is being first in the race for priority (Stephan, 1996) and this race is rewarded by receiving numerous citations.

The ordinary least-squares method is not an adequate technique when the dependent variable represents a count or a binary indicator. Appropriate models for estimating the citation counts are the method of Poisson regression and its generalized version, i.e. negative binomial regression. In the negative binomial regression model, the individual units follow a Poisson regression model, but there is an omitted variable u_i such that e^{u}_{i} follows a gamma distribution with mean 1 and variance α. To see the encompassing character of the negative binomial regression model we can write this model down in general terms: $c_j \sim \text{Poisson}[\exp(\beta_0 + \beta_1 x_{1,j} + ... + \beta_i x_{i,j} + u_i)]$, where c_j is the rate at which an article is cited per time period and x_i (for $i = 1,..k$) are the explanatory variables, and $e^{u}_{i} \sim \text{gamma}(1/\alpha, 1/\alpha)$. An important reason for using the negative binomial regression model instead of the Poisson regression model is that the number of events tends not to follow a Poisson distribution as the Poisson distribution implies equality of mean and variance, which is rarely observed in social phenomena. In order to allow for overdispersion in the data, the Poisson regression model is generalized by invoking a gamma distribution. Of course, in estimating count models the scale parameter α (representing the degree of overdispersion) may be zero, which means that the underlying data are indeed Poisson-distributed. In order to account for the fact that citations per article are Poisson-distributed we test whether the restriction $\alpha = 0$ applies.

Timing of first citations

The other idea about what counts in science is more or less an Olympic ideal: the joy of participating in the race for fame. Winning the race is no longer an overriding criterion but

being able to compete (and satisfy some minimum criterion of quality). In science one can interpret this ideals being cited for one's ideas at least once. Being able to publish in an internationally refereed journal might be another participating ideal. For authors working in the backwaters of science the chance of being cited is a very important issue and the sooner one is cited for the first time, the better. The focus on first citations in our framework essentially implies that we focus on the *timing* of first citations and the role which reputations play in receiving the first citation.

The most appropriate method of evaluating the speed of dissemination of knowledge is duration analysis, which has its origins in what is typically called survival analysis. The sample of articles tracked by ten years of citation history offers the possibility of examining the moment of citation somewhat more closely. The idea behind duration analysis for this particular case is that all articles start their 'life' uncited, and based on the characteristics of the articles at the start of their life the central question in survival analysis is: *what determines the probability of leaving the initial state of uncitedness*? The hazard function $\lambda(t)$ is of prime interest in duration analysis as it approximates the probability of exiting the initial state within a short interval, conditional on having survived (i.e. still not cited) up to the starting time of the interval. The most simple form of a hazard function is that it is constant (the exponential distribution) in which case the duration of being uncited is *memoryless*, i.e. the probability of exit in the next interval does not depend on how much time has been spent in the initial state.

Outside the constant hazard functions there are two possibilities: positive and negative duration dependence. Positive (negative) duration dependence amounts to the case that the probability of leaving the state of uncitedness increases (decreases) the longer the article remains uncited. In estimating the hazard function we use the Gompertz distribution as it offers a suitable approximation of the distribution of the duration of uncitedness (as presented in Figure 1) and at the same time it offers us a test to see whether the constant hazard model (with the exponential form) is more appropriate. The proportional hazard model is described specifically by the Gompertz hazard function $h(t_j) = e^{gt_j}.e^{x_j b}$, where the concomitant covariates (x_j) have a multiplicative effect on the hazard function and γ is the ancillary parameter of interest measuring the presence of a duration effect. Furthermore, because the regressors are time-invariant - they only describe the characteristics of the article (type of authors, content, journal) at the time of publication - we will restrict our attention to proportional hazard models. And because the observation period per article is restricted to ten

years we have right censored the data. After year 10 we cannot determine whether an article is cited or not.

Sleeping beauties or flash-in-the-pans?

Finally, we examine the case for articles that follow the pattern of sleeping beauties or flash-in-the-pans. Of course, our database has certain limitations because the citation window is 10 years and the total number of articles is 1,371. If we would follow the criteria used by Van Raan (2004) or Glänzel et al. (20003) in operationalizing 'sleeping beauties' the chances of finding one would be very small. We hope to offer an provisional answer to the question of these rare type of articles by dividing the group of articles into four categories. Subsequently we test by means of multinomial logit whether or not reputations enhance the chance of becoming a hit in the long run even though the limelight does not fall immediately on an article, or alternatively, the chance of being noticed in the short run but forgotten in the long run. The four categories by crossing article by timing and impact: (1) the reference category being the class of articles that are cited little (less than 5 times in ten years) and late (i.e. after 3 years for the first time) also the articles that are never cited within the ten-year citation window are included in the reference category (N = 732 and average number of citations in ten years: 1.2 citations); (2) the sleeping beauties: those articles that are not cited in the first two years after publication but in spite of this delay in being noticed quite often (more than 5 citations) thereafter (N = 73; average citation count 10.8); (3) the flash-in-the-pans: those articles that are cited in the first two years, but with no subsequent citation success (N = 183; average citation count: 2.9); and (4) normal science articles: the articles that follow the pattern one would expect being cited early and subsequently many times (N = 383; average citation count: 19.8).[4] Perhaps these impact and timing categories may not accord to what one usually defines as early and late or little and many citations but within the discipline of demography these horizons and citation impact number are rough but adequate approximations (see Van Dalen and Henkens, 2004). To test whether reputations play a role in belonging to one of the four categories we will use the method of multinomial logit. This method is most appropriate as the outcomes are not binary and cannot be ordered. It is a simple extension of the logit model and for a more thorough introduction one can consult Wooldridge (2002). In a multinomial logit model we estimate a set of coefficients $\beta^{(i)}$ belonging to explanatory variables *X* and corresponding to one of the *i* outcomes (where *i* is four outcomes in our

4. Slight variations in categorizing articles does not affect the outcomes we present in section 4.2.

case), with the coefficient of one of the categories set to zero in order to identify the model, i.e. without this restriction the set of equations would generate more than one solution for the $\beta^{(i)}$**.** For instance, the probability that an article belongs to the category of 'sleeping beauties' (outcome 2) is, where $\beta^{(1)} = 0$:

$$\Pr(article = outcome2) = \frac{\exp[X.b^{(2)}]}{1 + \exp[X.b^{(2)}] + \exp[X.b^{(3)}] + \exp[X.b^{(4)}]}$$

Similar expressions apply for the other two outcomes.

4.2 Results: the impact of signals…

…on the cumulative number of citations

Table 2 presents the results for testing the case whether or not reputations of authors and journals enhance subsequent success. Two models are distinguished: model I tries to capture the characteristics of the journal by focusing on the initial reputation of the journal, the reputation of the editors (as an approximation of their role as screeners or gatekeepers) and the circulation of the journal within the scientific community. Model II replaces these variables by using journal dummies as each and every journal has specific characteristics (specialization, editorial policies, etc.) which are not captured by the journal variables of model I.

The results seem to confirm the common wisdom that the reputation of authors matter. However, one should be careful in exaggerating this effect because the marginal effect of a reputation is quite small: with a reputation of a 100 citations received in one year, the extra citation one will receive over a time horizon of ten years is 0.8 citation. Furthermore, as the squared reputation term suggests there are decreasing returns to fame. The results with respect to the reputation of journals is more robust: the estimation results show that in order to become influential one should publish a full-length article in one of the top journals. The top journals in demography generate average citation scores between 9 and 24 citations over a ten-year time horizon (in other words, an impact factor in the interval between 0.9 and 2.4), whereas the second-tier journals generate between 0.5 and 4.5 citations over ten years time (see Van Dalen and Henkens, 2004). A full-length article of say 30 pages would bring in an extra 2.1 citations.

Table 2: Explaining cumulative number of citations (after 10 years) in demography[a]

Explanatory variables:	Cumulative number of citations			
	Model 1		Model 2	
	Coefficient	t-value	Coef.	t-value
Author characteristics				
Max. reputation author ($x10^{-2}$)	0.89**	6.33	0.83**	6.16
Max. reputation author squared ($x10^{-4}$)	-0.13**	3.84	-0.11**	3.58
US affiliation authors	0.03	0.44	0.04	0.45
Number of authors	0.05*	2.05	0.04	1.49
Article characteristics:				
Visibility				
Presidential address	0.68	1.66	1.02**	2.62
Comment/reply/note	-0.34**	3.14	-0.38**	3.49
Number of pages	0.06**	7.74	0.07**	7.80
Order in a journal issue	-0.03	1.40	-0.04*	2.02
Content				
Historical orientation	-0.42**	2.68	-0.58**	3.68
Focus of article:				
US = base category	•	•	•	•
Europe	0.02	0.19	-0.22	1.87
Asia/Australia	-0.32**	3.17	-0.35**	3.17
Africa	-0.15	1.17	-0.40**	2.94
Latin America	-0.41**	2.58	-0.61**	3.89
Middle East	-0.54*	2.13	-0.76**	3.04
World	-0.12	0.97	-0.24*	1.98
Non-empirical focus	0.84	1.40	-0.22	1.93
Journal characteristics				
Demography = base category			•	•
Family Planning Perspectives	-	-	0.64**	4.35
Population & Development Review	-	-	0.57**	3.63
Population Studies	-	-	0.17	1.07
Studies in Family Planning	-	-	0.43**	2.68
Journal of Biosocial Science	-	-	-0.51**	3.15
International Migration Review	-	-	-0.49**	3.39
Social Biology	-	-	-0.75**	4.33
Population	-	-	-1.66**	10.06
Population Bulletin	-	-	-0.94**	3.12
Population and Environment	-	-	-1.28**	6.43
Population Research & Policy Review	-	-	-1.13**	5.71
European Journal of Population	-	-	-0.96**	4.44
International Migration	-	-	-1.32**	7.77
Journal of Family Welfare	-	-	-2.24**	9.56
Journal of Population Economics	-	-	-0.89**	4.69
Population Index	-	-	-0.39	1.13
Impact factor journal	0.74**	6.60	-	-
Reputation editorial board ($x\ 10^{-2}$)	0.01**	2.94	-	-
Circulation journal i (x 1000)	0.07**	4.05	-	-
Use of french language	-1.05**	7.87		
Constant	0.15	1.00	1.58**	7.37
α	0.90	19.23	0.79	1.26
LR χ^2(df)	931.1		1070.5	
Pseudo R^2	0.12		0.13	

(a) Estimation method: negative binomial regression. The symbol * denotes significance at $p < 0.05$; ** at $p < 0.01$. The sample size N is 1,371 articles.

With respect to the language used to communicate, it stands to reason that scientists who refrain from using the standard language in communicating their findings will receive less response to their ideas. The consequences of using the French language seem to be far-reaching. The citation frequency of articles written in French drops by 50 percent compared with English articles. Of course, there is always a possibility that it takes more time for French articles to be disseminated in social science literature than the five years used in this study. The fact remains, however, that French articles are at a considerable disadvantage in the race for priority.

…on the timing of first citation

To explain the timing of first citations and to test whether negative duration dependence really holds up once one pays attention to a set of article characteristics, we ran a hazard analysis. First, we estimated the hazard function that fitted Figure 1 best, which, appeared to be a Gompertz function. Next, we estimated three proportional hazard models controlling for article characteristics that were known at the time of publication, such as the reputation of the author(s), the size of the research team, the length of the article (in terms of number of pages), whether the article was a full-size article or a note, and finally the journal in which it appeared. The estimation results are presented in Table 3.

Both models make clear that characteristics of the communication *process* (visibility, language and reputation of authors and journals) are of prime importance in speeding up knowledge dissemination. There are, however, a number of notable differences in moving from model 1 to models 2 and 3. In model 1 the speed of obtaining a first citation is explained completely in terms of the producers of an article, without controlling for the quality of the journal in which the article appears. In models 2 and 3 we complement model 1 by controlling for journal quality: first by trying to capture the quality of each journal in terms of its initial status (measured by the impact factor), the reputation of the editorial board, the circulation and the use of the English language or not. In model 3 we replace these journal quality measures by journal dummies to take account of the idiosyncrasies of publishing and citation practices.

In moving from model 1 to model 2 one can see that if one does not control for journal quality too much weight is put on the individual author characteristics like reputation, the US affiliation and the spillover of collaboration. By controlling for journal quality, the reputation effect of authors is present but it is significantly smaller than in model 1 and there is no significant effect of being affiliated with a US institution and the benefits of collaboration

have also disappeared. Instead the results of model 2 suggest that the reputation (its history and the reputation of the editors) and editorial policy of journals (e.g. restricting to publishing English written articles) makes quite a difference.

Table 3: Explaining the timing of first citations in demography[a]

	Dependent variable: time when first cited					
Explanatory variables:	Model 1		Model 2		Model 3	
	Hazard ratio	t-value	Hazard ratio	t-value	Hazard ratio	t-value
Author characteristics						
Max. reputation author ($x10^{-2}$)	2.38**	4.83	1.68**	3.03	1.47*	2.38
Max. reputation author squared ($x10^{-4}$)	0.83*	2.15	0.91	1.28	0.94	0.92
US affiliation authors	1.45**	4.77	1.06	0.67	1.03	0.38
Number of authors	1.07*	2.43	1.04	1.33	1.03	0.85
Article characteristics:						
Visibility						
Presidential address	2.35*	2.06	1.79	1.40	2.33*	2.01
Comment/reply/note	0.91	0.86	0.81	1.78	0.73*	2.51
Number of pages	1.06**	8.90	1.04**	4.96	1.05**	5.13
Order in a journal issue	0.95**	2.83	0.95**	2.88	0.95*	2.55
Content						
Historical orientation	0.87	0.95	0.89	0.75	0.76*	1.72
Focus of article:						
US = base category	•	•	•	•	•	•
Europe	0.87	1.26	1.05	0.40	0.95	0.39
Asia/Australia	0.86	0.14	0.84	1.70	0.84	1.61
Africa	1.02	1.55	0.95	0.40	0.80	1.56
Latin America	0.67**	2.56	0.73*	1.97	0.57**	3.31
Middle East	0.84	0.66	0.83	0.70	0.68	1.43
World	0.93	0.58	0.97	0.26	0.96	0.36
Non-empirical focus	0.76*	2.53	0.84	1.50	0.84	1.43
Journal characteristics						
Demography = base category					•	•
Family Planning Perspectives	-	-	-	-	1.67**	3.10
Population & Development Review	-	-	-	-	1.95**	4.01
Population Studies	-	-	-	-	1.48*	2.39
Studies in Family Planning	-	-	-	-	1.36	1.74
Journal of Biosocial Science	-	-	-	-	0.79	1.36
International Migration Review	-	-	-	-	0.78	1.64
Social Biology	-	-	-	-	0.69	1.94
Population	-	-	-	-	0.36**	5.81
Population Bulletin	-	-	-	-	0.52	1.90
Population and Environment	-	-	-	-	0.40**	4.18
Population Research & Policy Review	-	-	-	-	0.51**	3.11
European Journal of Population	-	-	-	-	0.42**	3.70
International Migration	-	-	-	-	0.40**	5.12
Journal of Family Welfare	-	-	-	-	0.19**	6.41
Journal of Population Economics	-	-	-	-	0.62*	2.31
Population Index	-	-	-	-	0.75	0.76
Impact factor journal	-	-	1.65**	5.29	-	-
Reputation editorial board ($x\ 10^{-2}$)	-	-	1.01**	2.62	-	-
Circulation journal i (x 1000)	-	-	1.03	1.85	-	-
Use of french language	-	-	0.60**	3.76	-	-
γ	-0.06**	5.04	-0.03**	3.01	-0.01	1.26
Log Likelihood	-1891.2		-1818.7		-1765.5	

(b) Estimation method: parametric survival analysis with Gompertz distribution. The symbol * denotes significance at $p < 0.05$; ** at $p < 0.01$. The sample size N is 1,371 articles.

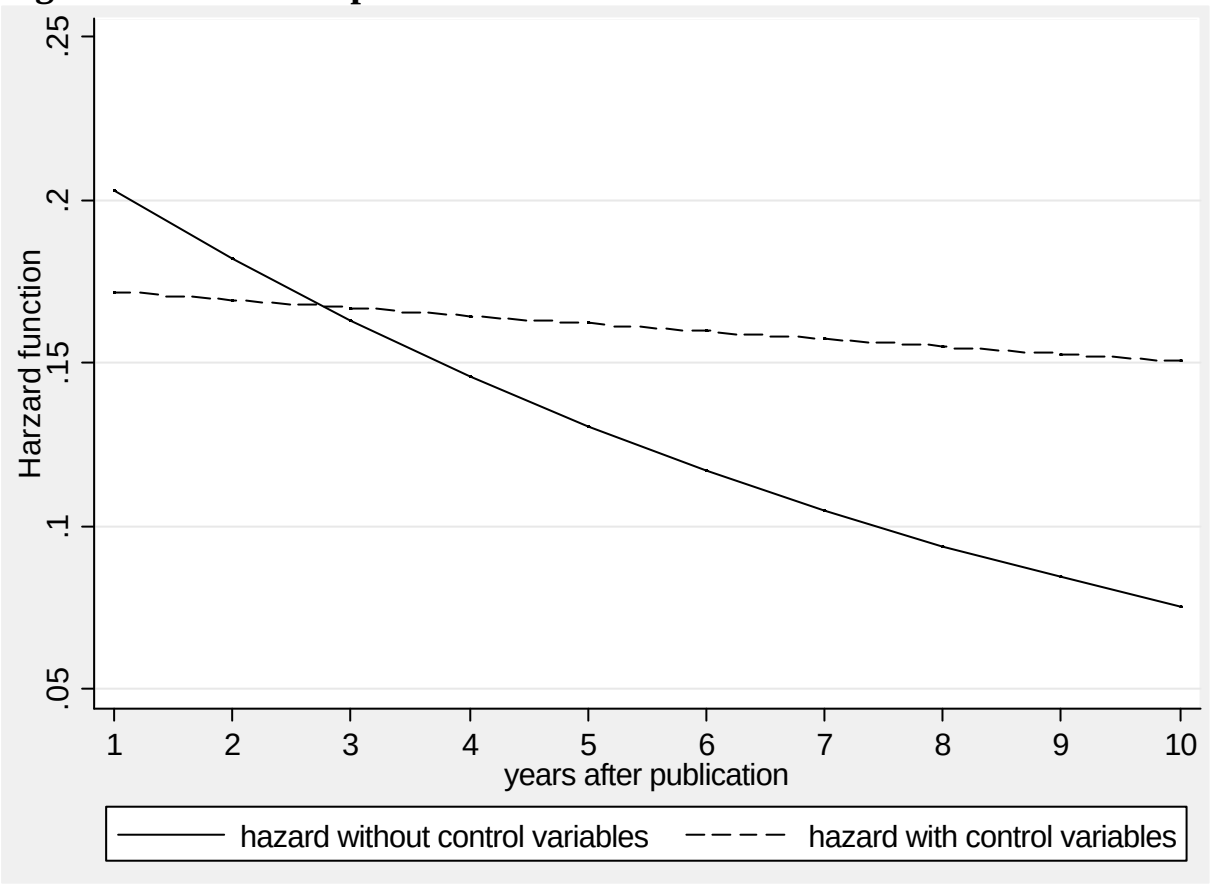
However, we also know that these journal quality measures are imperfect approximations of quality as the collection of demography journals are noted by different audiences with different citation practices (see Van Dalen and Henkens, 1999). The simplest way to correct for this in estimation is by representing each and every journal with a dummy variable, which has been done in model 3. The base category article is in this case an article that has appeared in the journal of the Population Association of America: *Demography*. The most notable findings of this model compared to the previous two models is the fact that all visibility characteristics of an article play a significant role in speeding up the timing of the first citation and the effect of the reputation of authors has decreased even further and is weakly significant.

Of course, we are primarily interested in estimating the parameter that indicates the presence of duration dependence in first citations and here we touch on something remarkable and novel: in the full model (model 3) there is *no sign of negative duration dependence* in the timing of first citations. In the first two models one could still trace the presence of negative duration dependence, although in model 2 the duration dependence effect is twice as small as that of model 1, which underscores the need to model citation processes as fully as possible. All the relevant elements – the producers (as reflected by the quality of authors and journals) and the consumers of articles (as reflected by the speed with which content of the article is used) – should be included, otherwise one could erroneously deduce the presence of duration dependence.

To get an idea of how strong the element of quality is, we have visualized two constructed hazard functions (see Figure 2). The dashed curve (the hazard function belonging to model 3 of Table 3) is constructed by setting each covariate at its mean value. The simple hazard model (solid line) is a reflection of duration dependence without controlling for any article characteristics (not shown in Table 3), whereas the dashed line reflects the 'corrected' duration dependence, i.e. by controlling for composition effects. As one can see, the slope of the dashed line is almost horizontal and is in marked contrast with the negative slope of the simple hazard model. Figure 2 suggests that the observed negative duration dependence is largely attributable to article characteristics known at the time of publication. In other words, the reasons why an article is not cited or cited relatively late, have to do with the journal in which the article appeared, certain visibility characteristics, and the reputation of the author(s). But perhaps the most important thing to notice is that the *absence of a duration effect* - after controlling for the above stated factors - indicates that a stigma of uncitedness

plays no role in the timing of the first citation. The conclusion that an article will never be cited because it remained uncited for quite some years therefore seems unwarranted.

Figure 2: Duration dependence in first citations



…in creating sleeping beauties and flash-in-the-pans

Finally, we want to examine the question what role signals play in creating sleeping beauties – so-called articles that are asleep for quite some years and suddenly get noticed (cf. Van Raan, 2004) and its antonym, flash-in-the-pans – articles that are noted almost immediately but that receive no attention whatsoever after the initial attention. Table 4 presents the findings for a limited number of explanatory variables (because the number of observations within the smallest categories is quite limited). The journals are split into two categories: top journals (*Demography*, *Population and Development Review*, *Family Planning Perspectives*, *Studies in Family Planning*, *Population Studies*, *Population Index* and *Population Bulletin*) and second—tier journals.

Table 4: The Role of Reputations in Creating Sleeping Beauties and Flash-in-the-pans (multinomial logit analysis)

	Dependent variable: Probability of belonging to quality category	
	Coefficient	t-value
Sleeping beauty*: Noted late and receiving many citations*		
Max. reputation author ($x10^{-2}$)	0.97	1.91
US affiliation	0.38	1.34
Number of authors	0.03	0.25
Comment	-0.69	1.44
Number of pages	0.07*	2.33
Order in a journal issue	0.01	0.17
Top journal (other journals=0)	1.43**	4.93
Constant	-3.63**	7.16
Flash-in-the-pan*: Noted early and receiving few citations*		
Max. reputation author ($x10^{-2}$)	0.91*	2.38
US affiliation	-0.02	0.12
Number of authors	-0.00	0.02
Comment	-0.26	0.93
Number of pages	0.05**	2.57
Order in a journal issue	-0.03	0.51
Top journal (other journals=0)	1.10**	5.25
Constant	-2.04**	5.98
Normal science*: Noted early and receiving many citations*		
Max. reputation author ($x10^{-2}$)	1.71**	5.16
US affiliation	0.25	1.40
Number of authors	0.08	1.12
Comment	-0.17	0.62
Number of pages	0.11**	6.16
Order in a journal issue	-0.16**	3.44
Top journal (other journals=0)	2.41**	13.13
Constant	-2.71**	8.46
Log Likelihood	-1241.78	
Pseudo R^2	0.19	

(a) The comparison category is the category of articles that are noted (and cited) late or never and that receive no or few citations. The symbol * denotes significance at $p < 0.05$; ** at $p < 0.01$. The sample size N is 1371articles.

The results are at first sight counterintuitive because the reputation of authors and journals are important in explaining the probability in being in one of the three categories compared to the reference category articles which receives little or no attention and if so quite late in the life of an article. The role of reputations in the 'normal science'-category is what one would expect: Table 2 has already shown for the entire sample how important reputations can be for gaining attention in science. But the two categories of interest – sleeping beauties and flash-in-the-pans – the result needs to be interpreted more carefully. The effect of author reputation in

explaining these two types of categories is far smaller than the coefficient for the 'normal science' category and furthermore the coefficients are only weakly significant. The most important factor in explaining the probability of belonging to one of these two article types is the reputation of the journal. In the case of the 'sleeping beauty' case it is better to publish your idea in a top journal than in a journal outside the core as it gives your idea a chance to a second life. To follow up on the metaphor of 'sleeping beauty': in order to be kissed alive by some prince at some future date it helps to be lying asleep in a quality bed. This result is in contrast to some anecdotal evidence that path-breaking work is more likely to be accepted in non-core journals (Gans and Shepherd, 1994) as these journals might be more open to heterodox approaches. The risk of publishing in such journals is, of course, that one's ideas will be noted somewhat later than in core journals because the readers of the non-core journals will not be drawn to consult the pages as the impact score of these journals signals that the average article will not be highly influential. One could also rephrase the conclusion in a more positive light: top journals apparently do not seem to be a barrier for non-standard work in demography.

The result of journal reputation for the case of flash-in-the-pans is more difficult to interpret. A possible interpretation is that the signals sent out by the reputation of journals give rise to misjudgements: initially articles are cited because of the seal of approval which they receive by being accepted in a top journal. However, being accepted does not mean that articles will have long-lasting influence on the discipline. An alternative interpretation is that top journals are also the journals where the academic debate of a discipline takes place. Some debates have a short life because the issues are fashionable or because the debate concerns topics that are easily settled or replaced by other interesting phenomena of the time. The fact that top journals consistently produce more flash-in-the-pans than second-tier journals is in that respect understandable. Second-tier journals are often more specialized and not in the habit of publishing debating points which interest an entire discipline.

5. Conclusions and Discussion

Nobel laureate Herbert Simon once made a trivial but far-reaching statement that "A wealth of information creates a poverty of attention" (Simon, 1971: 40). In many scientific disciplines this statement is becoming truer by the day. The number of articles, working papers, conference proceedings, books and newsletters is far too large for any capable scholar to absorb or even to scan. Signals have to be used in gaining attention and authors are satisfied to embrace the maxim of the Hollywood star Mae West: "It is better to be looked

over than overlooked". In that respect citations have come to acquire a different status over the course of time. Their original purpose was to give credit to the originators of an idea or finding, but with increasing competition and specialization in the academic field, the function of citations as an intellectual credit system is increasingly giving way to their being indicators of individual or departmental productivity (cf. Hargens and Schuman, 1990). Citations are now widely used to assess the viability of research programs and journals. Students use citation rankings to assess which university or department is worth paying large enrollment fees to (cf. Siow, 1997). Policy makers and (science) foundations use rankings to allocate funds in order to generate the 'biggest bang for their buck', in hiring and tenure decisions, citations are also directly or indirectly being used to assess individual scholars. And last but not least, scholars who are on the tenure track use citation rankings to decide which journals they should submit their papers to. The predominant use of citations in decision-making in academic life makes questions about the allocation of citations increasingly important (see Korobkin, 1999, and Frey 2003).

In this paper we have examined the role played by three types of signals in assessing quality: author reputations, journal reputations and the state of uncitedness of an article. The measurement of 'quality' is in this particular set-up: the cumulative number of citations after ten years and the timing of the first citation. Both measures are different but related dimensions and to assess them in combination we have also used four different categories of articles that differ by impact and the timing when they received their first citation.

To summarize our findings succinctly we can state the following four conclusions. First, the reputation of authors plays a role in gaining attention whether attention is measured by the cumulative impact of an idea or the speed with which an idea assimilates in the scientific community. Still the author reputation effect is small and therefore does the much cited Matthew effect of Robert Merton (1968) plays a relatively minor role in science (or to be more specific: in the science of demography). Though we do not find a very strong Matthew effect on citation counts there is still a possibility that reputations do matter but primarily at the stage when publications are refereed. For instance, if two articles of the same quality are submitted to the same journal, the article written by the more widely reputed author may be more likely to be accepted for publication than the article by a less established author. The extent to which this violation of the universalist rule occurs in the refereeing process by demography journals is not known.

Second, journals are the dominant force in allocating citations. Articles published in core journals receive considerably more citations than articles in second-tier journals and the

speed with which knowledge disseminates lies far higher in the core journals than in the journals with less visibility and less reputation.

Third, every scholar hopes that his or her ideas will prove to be path-breaking. In that respect the sign of being noted immediately and being cited many times by the scientific community is an informative signal and most of the influential ideas in many a science seem to conform to this type of pattern. Still, there are always the odd number of articles which are not noted early on but which gain a lot of attention late in life (so-called 'sleeping beauties'). And the reverse case applies also: there are some articles which are noted immediately and cited a lot, but which do not seem to have a lasting impact and die early in life (which we call 'flash-in-the-pans'). Our rudimentary assessment of these types of articles is that for both cases the influence of the journal reputation seems to be of considerable importance. Both types of articles are more likely to be found in top journals than in second-tier journals.

And finally we end with the finding which contradicts the myth that the chance of being cited for the first time declines with the age of an article. We find that the chance of previously uncited articles being cited *does not decline* as articles 'age'; the stigma of uncitedness does not play a role in the allocation of citations over time. The absence of a negative duration effect in the analysis of first citations may be seen as a sign that demography functions as an open science and an indication of substantial intellectual health (cf. Morgan and Lynch, 2001). The paradox is that this openness of demography may be jeopardized by the blessings of the so-called information age. Scholars use all kinds of signals to extract the quality of an article. When browsing through journals, attention is focused on the type of journal in which an article appears, who has written the article, whether it is a lead article or an article that is pushed to the back of a volume, etc. In using these signals, they are, however, unable to discern the exact impact of an article or, to put it bluntly: the stigma of *being uncited* is not emblazoned across an article. However, with the appearance of electronic journals, this unprejudiced attitude may disappear. The automatic registration of search behavior on the Internet and the use of rankings of journals (cf. the journals registered by Elsevier Science see www.sciencedirect.com or the working paper series of the Social Science Research Network: www.ssrn.com) by 'downloads' or 'abstract viewing' makes the stigma explicit and visible for anyone who searches for papers on the Internet. In short, having no *a priori* information about "who has cited who" is perhaps a state of blissful ignorance which will soon be a thing of the past.

References

Akerlof, G.A., 1970, The Market for Lemons: Quality Uncertainty and the Market Mechanism, *Quarterly Journal of Economics*, 89: 488-500.

Frey, B.S., 2003, Publishing as Prostitution? – Choosing Between One's Own Ideas and Academic Success, *Public Choice*, 116: 205-223.

Gans, J.S., and G.B. Shepherd, 1994, How Are the Mighty Fallen: Rejected Classic Articles by Leading Economists, *Journal of Economic Perspectives*, 8: 165-180.

Glänzel, W., and U. Schoepflin, 1995, A Bibliometric Study of Ageing and Reception Processes of Scientific Literature, *Journal of Information Science* 21: 37-53.

Glänzel, W., B. Schlemmer, and B. Thijs, 2003, Better Late Than Never? On the Chance to Become Highly Cited only Beyond the Standard Bibliometric Time Horizon, *Scientometrics*, 58: 571-586.

Hamermesh, D.S., 1994, Facts and Myths about Refereeing, *Journal of Economic Perspectives*, 8: 153-164.

Hamilton, D.P., 1990, Publishing by – and for? – the Numbers, *Science*, 250: 1331-1332

Hamilton, D.P., 1991, Research Papers – Who's Uncited Now?, *Science*, 251: 25.

Korobkin, R., 1999, Ranking Journals: Some Thoughts on Theory and Methodology, *Florida State University Law Review*, 26: 850-876.

Hargens, L.L., and H. Schuman, 1990, Citation Counts and Social Comparisons: Scientists' Use and Evaluation of Citation Index Data, *Social Science Research* 19: 205-221.

Klamer, A., and H.P. van Dalen, 2002, Attention and the Art of Scientific Publishing, *Journal of Economic Methodology*, 9: 289-315.

Laband, D.N., 1990, Is there Value-Added from the Review Process in Economics? Preliminary Evidence from Authors, *Quarterly Journal of Economics*, 103: 341-352.

Laband, D.N., and R.D. Tollison, 2003, Dry Holes in Economic Research, *Kyklos*, 56: 161-174.

Merton, R.K., 1968, The Matthew Effect in Science, *Science* 159: 56-63.

Morgan, S.P, and S. Lynch, 2001, Success and Future of Demography, *Annals of the New York Academy of Sciences*, 954: 35-51.

Simon, H.A., 1971, Designing Organizations for an Information-Rich World, in: M. Greenberger (ed.), *Computers, Communication and the Public Interest*, Johns Hopkins University, Baltimore, pp. 37-52.

Siow, A., 1997, Some Evidence of the Signalling Role of Research in Academia, *Economics Letters*, 54: 271-276.

Stephan, P.E, 1996, The Economics of Science, *Journal of Economic Literature*, 34: 1199-1235.

Stewart, J.A., 1983, Achievement and Ascriptive Processes in the Recognition of Scientific Articles, *Social Forces*, 62: 166-184.

Van Dalen, H.P. and K. Henkens, 1999, How Influential are Demography Journals?, *Population and Development Review*, 25: 229-251.

Van Dalen, H.P. and K. Henkens, 2001, What Makes a Scientific Article Influential? The Case of Demographers, *Scientometrics*, 50: 455-482.

Van Dalen, H.P. and K. Henkens, 2004,Who is Uncited After Ten Years? – Demographers and their Journals, *Population and Development Review*, 30: 489-506.

Van de Kaa, D.J., 2003, Population Journals, in: P. Demeny and G. McNicoll (eds.), *Encyclopedia of Population*, MacMillan, New York, pp. 555-557.

Van Raan, A.F.J., 2004, Sleeping Beauties in Science, *Scientometrics*, 59: 467-472.

Wooldridge, J.M., 2002, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA.