TI 2000-032/1
Tinbergen Institute Discussion Paper

What Makes a Scientific article Influential?

Hendrik P. Van Dalen
Kène Henkens

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
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

WHAT MAKES A SCIENTIFIC ARTICLE INFLUENTIAL?

A CITATION ANALYSIS

Hendrik P. van Dalen* and Kène Henkens**

March 8, 2000

Abstract:

In this paper we examine, by means of a citation analysis, which factors influence the impact of articles published in demography journals between 1990 and 1992. Several quantifiable characteristics of the articles (characteristics with respect to authors, visibility, content and journals) are strongly related to their subsequent impact in the social sciences. Articles are most frequently cited when they deal with empirical, ahistorical research focusing on populations in the developed world, when they are prominently placed in a journal issue, when they are written in English and when they appear in core demography journals. Furthermore, although eminent scholars are likely to be cited on the basis of their reputation, the effect of reputation appears to be small in demography.

* Erasmus University Rotterdam, Tinbergen Institute and OCFEB, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands; and Scientific Council for Government Policy, P.O. Box 20004, NL-2500 EA The Hague. At the time of writing this paper Van Dalen was associated with the NIDI.
** Netherlands Interdisciplinary Demographic Institute (NIDI).

Acknowledgments
Comments by Aat Liefbroer, Frits Tazelaar, Evert van Imhoff and Frans van Poppel are gratefully acknowledged. Research was carried out with financial support from the Population Studies Program of the Netherlands Organisation for Scientific Research (NWO), grant # 18.071.

1. Introduction

What makes a scientific article influential in the social sciences? Is it purely scientific merit which determines whether an article will be well-cited, or do other, non-scientific factors come into play? This bibliometric study tries to shed some light on the forces that systematically affect the number of times a demographic article is cited by fellow scientists. This question is relevant for a number of reasons. First of all, by showing how the citation frequency of articles is affected by factors that are thought to dominate the communication process we can reveal whether the principles of open scientific communication apply to demography as a science. Open scientific communication is a prime requirement for scientific progress. Whether this principle applies to demography is an open question. Several members of the demographic community (Greenhalgh 1996; McNicoll 1992; Demeny 1988) have raised concerns about the level of scientific communication among its practitioners. To date there is, however, limited empirical information regarding the way in which demographers communicate. Citations reveal which articles have been influential in the literature. Needless to say, citations are not entirely accurate measures of intellectual influence, as numerous commentators have remarked (see for an overview of this literature: Cronin 1984) but they are nevertheless indicators that provide some insight into scientific communication. Van Dalen and Henkens (1999) give some information by showing that communication among demography *journals* with different specializations is weak and the majority of articles in the second-tier journals remain uncited in the five years following their publication. Unfortunately, no studies are available on what makes *articles* published in demography journals influential. By using a citation analysis this paper tries to fill this void.

A second reason for examining citation behavior in the social sciences is related to the observation that citations and impact scores of journals are increasingly becoming the central indicators used to evaluate scholarly publications and their producers. A survey by Hargens and Schuman (1990) revealed that approximately 60 percent of graduate sociology departments at US universities use citation counts when making decisions about hiring, promotion and tenure. One of the consequences of this practice is that the tacit reward system in science has increasingly become an explicit reward system as citations translate into dollars.[1] For this reason a critical assessment of the

[1]. For instance, Diamond (1986) shows that the marginal value of a citation, starting from a level of citations of zero, varies between $50 and $1,300 on a yearly basis. Differences in citation practices are probably the most

processes that govern the allocation of citations is not only of importance to sociologists of science and bibliometricians but is also relevant to science policy in general. If extra-scientific attributes affect the citation frequency systematically and significantly there is reason to question the current emphasis on citation numbers.

In detecting the forces that explain citation frequency we will use as the unit of our analysis the articles published in a set of demography journals in the years 1990-1992. The approach of this paper may perhaps be simple in nature, but the findings are novel and thought-provoking as they provide some insight into how influential certain demographic articles are in the social sciences. Before we introduce the data, the statistical methodology and a number of tests in sections 4 to 6, we will first review (section 2) theories of communication in science and subsequently introduce, in section 3, the elements that are potentially useful in explaining the intensity with which articles are cited.

2. Theories of scientific communication

In disentangling the myriad of forces behind the allocation of citations, there are essentially two views on this process of communication: the universalist and the social constructivist view. The universalistic view on the structure and development of science assumes that the reward structure of the science community has been set up in such a way that it elicits compliance of scholars with the norms of 'openness': speedy disclosure of new findings and their submission to others for verification, application, and extension (Merton 1957, 1973). Citations in this process represent the intellectual influence of the scientific literature up to a particular date. The universalists claim that cognitive content and article quality dominate the development of science. Social elements may perhaps play a minor role but they certainly do not affect the final outcome in science. In their seminal contribution, Cole and Cole (1973) stressed that stratification as found in science is normal and in line with the principles of meritocracy.

If citations are indeed only based on scientific merit and not correlated with extra-scientific attributes there are no grounds to question the reward structure of science. If, however, extra-scientific attributes play a significant role in the allocation of citations, the universalistic norm may not

---

important determinants of these pecuniary differences. Disciplines such as economics and mathematics have relatively low publication and citation rates and therefore the marginal value of a citation tends to be rather high, contrary to disciplines such as physics and chemistry where citation and publication rates are relatively high.

be an accurate reflection of how science evolves. Social constructivists like Latour (1987), Lindsey (1978) and Gilbert (1977) believe that the universalist norm is being systematically violated. They argue that scholars use citations as tools of persuasion, instead of as indicators of intellectual merit. In their view, it is not the content of an article that is of primary importance, but the position of the author in the scientific stratification system. They assert that influence is distributed on the basis of *who one is* rather than *what one says*. If two scientists contribute the same idea in different articles, most of the recognition goes to the more established of the two scientists. So, the extent to which an article or author is cited is not the outcome of an achievement process as reflected in the content and quality of the article, but of an ascriptive process in which the reputation and extra-scientific character traits of the authors matter. In this process, functionally irrelevant author characteristics such as eminence, sex, rank and institutional affiliation are assumed to be the main determinants of the allocation of citations and rewards.

Most research that addresses the relative importance of so-called achievement and ascriptive processes in the allocation of citations and rewards, use individual *scholars* as the unit of analysis (Allison and Long 1990; Bayer and Folger 1966; Hargens and Hagstrom 1982). This approach has been criticized for several reasons. First, leading citation databases only report the first author, which can introduce substantial bias in the dependent variable. Second, the studies use aggregated citation counts over all the articles written by an individual and do not examine how article characteristics affect the distribution of influence within science. This is unfortunate as these characteristics may be important determinants of the number of citations. Articles - which nowadays are the basic units in the communication system - do not only differ in importance, but also in length and topic. Scientists may differ in the average characteristics of their articles, which may lead to misspecified models. In line with Stewart (1983) and Baldi (1998), this study assumes that an analysis of articles is the most appropriate way to understand the distribution of influence in science. Stewart (1983) has examined the number of citations received by a selection of articles in geophysical research published in selected journals in 1968. He concluded that although universalism was the most important underlying principle in the allocation of recognition, there was some evidence of reputation effects. The additional effect on citation frequency was, however, relatively small. In his study on the astrophysics research area, Baldi (1998) included not only the characteristics of *cited* articles but also the characteristics of the *citing* articles. He discovered that authors “are most likely to cite

articles that are relevant to their work in terms of subject, recency of knowledge, theoretical orientation, and seem to have little concern with the characteristics of the authors who write them." (Baldi 1998: 843). However, both studies analyze a relatively small number of articles, and only articles in a narrow and clearly defined subfield of the natural sciences. The natural sciences are generally highly codified: there is consensus on what constitutes quality research and which problems are central to the development of the discipline in question. The social sciences are, however, less codified and delimitating subfields in, for example, demography may not prove to be very useful if one wants to assess the *impact* of a demographic article because most citations in demography come from outside the narrowly defined field of demography (see Van Dalen and Henkens 1999). Because the social sciences are different from the natural sciences, Baldi (1998) suggests that conclusions concerning the natural sciences may not hold in the social sciences, as "citations in the social sciences may be influenced less by the intellectual content of articles and more by the characteristics of their authors." (1998: 844). Moreover, it is well established in the scientometric literature that there are big differences in citation practices between the natural and the social sciences. For instance, Hamilton (1990, 1991) and subsequently Pendlebury (1991) showed that 22 percent of all articles in the natural sciences were never cited in the five years following publication. The status of the social sciences and the arts and humanities was very different: 48 percent and 93 percent, respectively, of the articles did not receive a single citation within five years. These figures suggest that there are big differences in the level of communication among scientists in different scientific disciplines, which substantially restricts the generalizability of the results concerning the natural sciences to other scientific disciplines. To our knowledge, no study has as yet detected the factors that underlie the distribution of citations in the social sciences, and whether the principles of meritocracy or principles brought forward by the social constructivists are dominant in this process.

3. Inequality in Science

The outcome of the communication process in science is reflected in the frequency distribution of citations received in the academic debate. It has long been noted that inequality in attention given to ideas is highly skewed toward 'giants' in a particular discipline. The registration of inequality in science was essentially started by Lotka (1926) who formulated the following law of scientific productivity: if n_1 is the number of scientists who publish one article, then the number n_k of scientists

publishing k articles (for $k > 1$) in the same field can be gauged on the basis of the following equation: $n_k = n_1/k^2$. Hence, if 1,000 scientists publish 1 article during their lifetime then 250 will have produced 2 articles, 111 will have written 3 articles, etc. Later, formulations of inequality distributions in scientific productivity were made by Price (1976) who formulated a variant of Lotka's Law, viz. one half of the total output of articles published by a population of N scientists will be the work of $\sqrt{N}$ most productive members of the population. Holub *et al.* (1991) made a variant of Price's claim by stating that the number of important publications in a science is the square root of the total number of publications in a research field. All these statistical theories of inequality distributions share the common trait that they lack a behavioral theory of citation, which stands to reason as most inventors of these distributions believed that the making of science was ruled by an iron law. Under these circumstances the main task of a 'scientometrician' is then to discover the natural constants of the process of publication and citation. David (1994) has argued that sociologists and economists of science still have some explanatory work to do if they want to understand more clearly why the large majority of articles in social science receive so little attention and why just a small percentage of articles make the grade in terms of citations.

In this paper we aim to contribute to this literature by disentangling the potential sources of citation frequency. Ideally, an overall assessment of the quality of an article would enable us to examine whether citation frequency and quality are strongly related. However, a fundamental methodological problem is that it is difficult, if not impossible, to define quality criteria in an objective manner. A common way of dealing with these methodological complications is to focus on the manifest characteristics of an article (cf. Stewart 1983; Baldi 1998) and to trace the influence an article generates. We make a broad distinction between three groups of determinants of citations received: (1) the characteristics of the journal in which the article appears; (2) the characteristics of the article itself, such as the visibility and the content of the article; and (3) the characteristics of the author(s) of the article in question. In order to understand each of these factors more fully, we will elaborate on them below.

Journal characteristics

The social dimension of science is most clearly seen at the stage of acceptance of ideas when the gatekeepers of science, i.e. referees and editors, decide whether or not an idea or a point of view has sufficient quality to be of interest to the community of scientists. If an article has insufficient quality to merit publication in the eyes of the referee or the editor, it will not survive the screening process. The importance of journal editors and referees in assessing quality should therefore not be underestimated. As Keyfitz (1993: 539) notes: "in the criteria a journal uses for determining what is 'best' it can be a powerful means for shaping the discipline." Each and every journal has a different editorial policy with respect to standards of research, the language in which it wants to publish ideas, the extent of specialization it allows, and last but not least, each journal has a different circulation and therefore a different visibility to the scientific community (cf. Baldi 1998), which may be the result of the fact that a journal is affiliated with an association or scientific society. The characteristics of journals are therefore pivotal to understanding why articles in some journals gain more attention than articles dealing with a similar topic or with similar authors in a different journal.

Article characteristics

The previous remarks concerned the quality of the article as determined on the basis of the journal characteristics. However, the role of the editors does not stop at the process of accepting and rejecting papers. They also influence the *visibility* of articles within a particular journal issue. Editors decide which articles should lead an issue and the subsequent sequence of articles, usually closed by notes, comments, replies and journal trivia. Smart and Waldfogel (1996) have shown, for a number of core economics journals, how important this factor can be in explaining citation frequency. Articles that lead an issue are generally more frequently cited than articles that appear in the back of an issue. The order in which articles appear may reflect the editors' perception of the range of quality. Another reason why front articles are noticed more often than articles at the back, may have to do with the way in which scientists browse through journals: leading articles are more visible than those in the back.

Besides the order of appearance in journals, each and every article has other characteristics which may increase or decrease its visibility. It stands to reason that research notes, comments and replies receive less attention than the regular articles of a journal issue as these contributions usually make small points which do not warrant a full-size article. Among full-size articles, there is also some

differentiation which may affect the visibility of an article, such as the length of the article. Again, like the order of an article in a journal issue, the length of an article may also contain a quality element. It is at the journal editors' discretion to allow longer papers if they think a paper of a particular size warrants publication and at the same time they also have full discretion to shorten papers if they think the subject is not worth the submitted length or they may even relegate the paper to the status of note. The potential explanatory power of article length is stressed in a citation analysis by Bayer (1982: 531) of a selection of articles from the *Journal of Marriage and the Family*. He finds that the length of the published articles is positively correlated with subsequent citations. A drawback of his conclusion is that this correlation is based on bivariate analysis and not the more appropriate technique of multivariate analysis.

Among the full-size articles we will also draw attention to articles that are made prominent by the scientific societies and associations themselves, viz. presidential addresses and invited lectures. Presidential addresses are usually seen as a reward for those who have become established figures within the hierarchy of a scientific society. These lectures are generally given by eminent scholars who review developments in their field. In order to give such reviews extra visibility, they tend to be the lead article of a journal. These articles will probably be more heavily cited than regular articles as these types of articles sum up and offer a perspective on a particular research field, providing a focus on the major issues at stake within a discipline.[2]

The second group of article characteristics concerns the *content* of the article. Especially in sciences that are empirically oriented, we expect that the empirical focus of an article can have a significant impact on the number of citations it receives. On a global scale, the US dominates every scientific discipline, in particular the social sciences. For instance, in the years 1995-1996, US scientists published 60 percent of all social science articles (Tijssen et al. 1998).[3] The situation in the field of demography is much the same, as the top journals are dominated by US authors and as we will show half of all demographic articles are produced by US scientists.[4] Dominance or scientific

---

[2]. An extra reason for setting the presidential addresses apart is that, contrary to regular articles, these addresses are not (heavily) refereed.

[3]. May (1997) shows that for articles published in science, medicine and engineering, influence measured by the number of citations is even more skewed.

[4]. Furthermore, as has been argued by demographers such as Demeny (1988), Hodgson (1991), and Greenhalgh (1996), the history of demographic thought has been heavily influenced by US foreign policy and the US preoccupation with family planning.

leadership implies that not only can US-researchers dominate production standards but they can also shape the direction of research. Scientists check their results in terms of plausibility by using the findings of other researchers. We may assume that, in order to obtain a satisfactory check, they will primarily consult results which relate to the same country or region. As the US is the technological leader in the social sciences and since the most prolific researchers are based in the US, it stands to reason that in an empirical science such as demography a preference for US-oriented research will show up in the citation frequency of articles. In small countries, opportunities for specialization are often limited, as are other publications that can serve as reference points. Researchers in these countries may well use the US as a reference point for their research because most of the demographic research is being carried out there. Hence, an asymmetry may be expected in the use of demographic literature that favors the scientific leader. Frey and Eichenberger (1993, 1997) stress this aspect for the study of economics by stating that only those scholars who do not aspire beyond the local market for economists can afford to work on topics with an emphasis on local data.

Another aspect related to article content concerns their historical orientation. Several studies have shown that citation practices across the discipline of history and other social sciences differ widely (Hamilton 1990; Pendlebury 1991; Schwartz 1997). The most recent study mentions that 91 percent of articles in history remain uncited, a percentage which by far exceeds the uncitedness rate of sociology (49 percent). On the basis of this finding, together with Caldwell's (1996) general impression that US academics have a relatively weak orientation towards historical analyses of populations, we assume that articles with a historical orientation will be less frequently cited.

Author characteristics

The last factor which needs to be discussed is the importance of personal elements, such as reputation and networks. To understand the importance of reputation in knowledge dissemination, one needs to consult the work of the father of the sociology of science, Merton (1968), who was the first to point to the existence of reputation effects, which function as a signal of quality. Zuckerman (1977) comes to a similar conclusion in her extensive study of Nobel laureates. Both point to the disproportionate amount of credit that flows to established names in science at the expense of the scientific rank-and-file. The phenomenon of the skewed distribution of influence has been aptly

coined the Matthew effect by Merton (1968). According Matthew's gospel "For unto everyone that hath, shall be given, and he shall have abundance: but from him that have not shall be taken away even that which he hath." Merton suggested that scholars might be rationally allocating their time by paying special attention to studying the works of colleagues on the basis of the preceding reputations of the articles' authors, simply because they are not able to read everything that is published in their field. As such, the Matthew effect may be functional for science because it increases the visibility of a contribution that might be overlooked if contributed by a lesser known scientist. Merton, however, also stated that "when the Matthew effect is transformed into an idol of authority it violates the norm of universalism embodied in the institution of science and curbs the advancement of knowledge."(1968: 457).

Social constructivists like Lindsey (1978) and Latour (1987) emphasize this last remark by pointing out that reputations are a violation of the universalist principles underlying communication in science. Reputations serve as feedback mechanisms which perpetuate the influence incumbents have over outsiders. As such, reputation prevents all authors from having an equal chance at being noticed. This social constructivist explanation of the role of reputations underlines that scientists use citations strategically. Scientists are believed to refer to an article of an author with high prestige merely to make their article more credible.

The second author characteristic frequently believed to affect the likelihood of citation is the number of authors involved in the writing of an article (cf. Stewart 1983; Bayer 1982). Two reasons for examining this characteristic may be put forward. First, the most straightforward reason for using a team of researchers is that collaboration is potentially an efficient organization of complementary capabilities in research. The gains from specialization and division of labor are the prime determinants of the accumulation of citations as such collaborative work generates benefits which could not have been attained if he researchers had worked on their own. Hence, the citations received for collaborative work are primarily a reward on the basis of scientific merit. The second reason for paying attention to the number of authors is that each author brings in his or her own network of scientific relations and scholars inside the network are more likely to be knowledgeable about the author's work and cite it than scientists outside the network. This aspect is particularly relevant to understanding the dissemination of knowledge as most sciences are characterized by a concentration

of prolific authors in just a few institutes (see Clemens et al. 1995; Teachman, Paasch and Carver 1993).

4. Citation Data

4.1 Data Sources

We have used the *Social Science Citation Index* (SSCI) as published by the *Institute for Scientific Information* (ISI) to gather data on the citation frequency and other characteristics of individual publications in the selected demography journals in three consecutive years (1990-1992). 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. 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.[5] 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 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. In tracking

[5]. Most bibliometric studies make a selection of top journals only (e.g. Clemens et al. 1995; Smart and Waldfogel 1996) which offers a biased insight, at least if one wants to obtain an overview of an entire research field or discipline.

down article content, we consulted all the issues of the journals in the years 1990-1992 by hand and used the electronic database POPLINE.

4.2 Operationalization

Citation counts. For each article in our data set we established if, and how often they were cited in the five years following their publication by authors in the journals covered by the SSCI. As we intend to measure knowledge dissemination in a scientific community, we exclude the number of self citations by authors in our citation counts.[6] The reason for choosing a five-year exposure time and not a shorter period can be found in Glänzel and Schoepflin (1995) who report that it takes 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).

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.

[6]. Longer citation counting intervals would, of course, give more reliable measurements of the propensity to be cited. In order to test the reliability of 1-year interval variability of citation counts, Allison (1978: 244) suggested a measure of the reliablity of citation counts. In our study, the reliability amounted to 0.92, which is comparable with the reliability found by Stewart (1983) who uses a six-year time interval.

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

	Mean	Standard deviation	Min.	Max.
Citations per article after 5 years	3.57	6.78	0	70
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.

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 (Anderson, Baselevsky and Hume 1983: 456).

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.

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. In order to control for possible non-linearity of the Matthew effect, we have also included a quadratic term.

The 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.

5. Methodology

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.

In estimating the citation counts in demography, we assume that an exposure time of five years is sufficient to obtain an adequate impression of the influence of articles. Hence, each and every article is subject to the same exposure size and the influence of the exposure time is absorbed in the constant term. The negative binomial regression model also allows us to examine with relative ease the rates of individual explanatory effects, holding all the other *x*'s constant, except one. This so-called incidence rate ratio (IRR) is defined as the relative incidence rate of an event resulting from a change in x_i: $\exp(\beta_i \Delta x_i)$.

The tacit assumption underlying the negative binomial regression model is that differences in the influence of articles are reflected in the number of citations gathered in the years following publication. An article with ten citations could be viewed ten times as influential as an article with just one citation. However, there are two good reasons for paying close attention to the *probability* that an article will be cited. First of all, the ones who cite an article may use an article over and over again during the course of their careers. The ten citations gathered by an article may well stem from one author who is very productive and who uses the idea ten times during the exposure time. Another reason for being cautious with citation counts is that citation practices may differ across specializations or across sciences. Being cited in the natural sciences is not very special, whereas being cited in the social sciences is quite a feat. In order to complement the negative binomial regression analysis, we will therefore also address the question as to whether or not an article is cited with the aid of a logistic regression analysis.

6. Results

6.1 The distribution of citations

In Table 1 (see section 4) we established that the total number of citations received (excluding self-citations) by the average demography article after five years is 3.6, with a minimum of zero citations and a maximum of 70 citations. However, what interests us at this point is the extent to which the citations are distributed equally (or unequally) over all the articles. Table 2 provides insight into the distribution of citations by article and the market share of influence by article.

Table 2: Distribution of citations received by articles published in 1990-1992 in the five years following publication

Number of citations after 5 years	Articles *N*	Percentage articles	Cumulated percentage	Percentage of total citations[a]	Cumulated percentage total citations
0	497	36.3	36.3	0.0	0.0
1	252	18.4	54.6	5.1	5.1
2	139	10.1	64.8	5.8	10.9
3	105	7.7	72.4	6.4	17.3
4	77	5.6	78.0	6.3	23.6
5	47	3.4	81.5	4.8	28.4
6	41	3.0	84.5	5.0	33.4
7	32	2.3	86.8	4.6	38.0
8	27	2.0	88.8	4.4	42.4
9	15	1.1	89.9	2.8	45.2
10-15	69	5.1	95.0	16.6	61.8
> 15	68	5.0	100.0	38.3	100.0
Total	1,371	100.0		100.0	

(a) Total value of citations (i.e. the number of articles with *i* citations multiplied by the value *i*) amounts to 4,897.

As shown, 55 percent of the articles received 1 or less than 1 citation in the five years following publication and no more than 5 percent of the articles received 16 or more citations in this period. The last two columns of Table 2 show the market share of attention by taking the total value of citations (i.e. the sum 4,897) as the measure of influence. As shown in these columns, the top 10 percent of the articles (those with 10 citations or more) attracted 55 percent of total attention. These findings for the science of demography are similar to the findings presented by Price (1963), Cole and Cole (1973) or Cox and Chung (1991) for other disciplines. In short, demography is not very different from other social or natural sciences in this respect.

6.2 Explaining Citation Behavior

The results of the negative binomial regression to explain the number of citations articles received in the five years following publication are presented in Table 3. The results of the logit analysis to explain the probability of being cited within that period are presented in Table 4. In both tables, two models have been estimated. Model 1 gives the relationship between our dependent variable and variables relating to author, visibility and content characteristics as well as journal dummy variables to control for journal differences. In model 2 the journal dummies are replaced by journal characteristics such as the impact factor, the reputation of the editorial board, journal circulation and the use of the French language in articles.[7]

Explaining the number of citations

The results of model 1 in Table 3 suggest that the number of citations that articles have received is related to all the characteristics previously reviewed in section 3. The model shows that the journal in which an article is published is strongly related to the number of citations. To start with the journal dummies first, compared with standard articles in *Demography* one can say that articles published in general journals with a small circulation (e.g. *European Journal of Population, Population Research and Policy Review*) and in specialized journals (e.g. *International Migration Review, Journal of Population Economics*) are cited much less frequently than articles in top-rated general journals (e.g. *Population and Development Review, Population Studies*) and family planning journals such as *Family Planning Perspectives*.

As shown in the column with the IRRs, the average number of citations received by articles in the French journal *Population* is 17 percent of the number of citations received by standard articles in *Demography*. In model 2 of Table 3 it is assumed that differences between journals reflect not only quality differences, but also aspects such as circulation and the use of French. The model suggests that variations in journal quality (reflected by the SSCI impact factor and reputation of the editorial board in 1990) are the most important determinants of differences in citations received by articles.

Table 3: Negative binomial regression of the total number of citations after 5 years[a]

[7]. In addition to testing the models in Tables 3 and 4, we have also tested all the models for the appearance of a trend in citation behavior over time as the articles stem from three different years. Because there appeared to be no significant trend, we have presented all the models without the year dummies.

Explanatory variables	Dependent variable: Number of citations within 5 years					
	model 1			model 2		
	Coefficient	IRR	t-statistic	Coefficient	IRR	t-statistic
Author characteristics						
Max. reputation author ($x10^{-2}$)	0.56**	1.75	3.60	0.62**	1.87	3.29
Max. reputation author squared ($x10^{-4}$)	-0.06	0.94	1.76	-0.09*	0.91	2.42
US affiliation authors	0.18*	1.20	1.98	0.22*	1.25	2.52
Number of authors	0.07*	1.07	2.37	0.05	1.05	1.83
Article characteristics:						
Visibility						
Presidential address	0.95*	2.59	2.33	0.79	2.20	1.86
Comment/reply/note	-0.60**	0.55	5.08	-0.55**	0.58	4.68
Number of pages	0.01	1.01	1.15	0.02*	1.02	2.55
Order in a journal issue	-0.07**	0.93	3.57	-0.06**	0.94	3.14
Content						
Historical orientation	-0.61**	0.54	3.50	-0.42*	0.66	2.42
Focus of article:						
US excluded	•	•	•	•	•	•
Europe	0.03	1.03	0.19	0.11	1.11	0.81
Asia/Australia	-0.27*	0.77	2.24	-0.36**	0.70	3.28
Africa	-0.50**	0.60	3.29	-0.50**	0.61	3.39
Latin America	-0.78**	0.46	4.28	-0.72**	0.49	3.87
Middle East	-0.72*	0.49	2.49	-0.68*	0.51	2.29
World	-0.03	0.97	0.25	-0.01	0.99	0.10
Non-empirical focus	-0.10	0.90	0.85	-0.18	0.83	1.51
Journal characteristics						
Demography excluded	•	•	•			
Family Planning Perspectives	0.61**	1.84	2.63			
Population & Development Review	0.47*	1.60	1.98			
Population Studies	0.07	1.07	0.27			
Studies in Family Planning	-0.02	0.98	0.10			
Journal of Biosocial Science	-1.12**	0.33	4.74			
International Migration Review	-0.78**	0.46	3.36			
Social Biology	-1.36**	0.26	5.33			
Population	-1.75**	0.17	7.08			
Population Bulletin	-0.12	0.89	0.38			
Population and Environment	-1.62**	0.20	5.97			
Population Research & Policy Review	-1.71**	0.18	6.03			
European Journal of Population	-1.57**	0.21	5.16			
International Migration	-1.35**	0.26	5.16			
Journal of Family Welfare	-2.59**	0.07	8.39			
Journal of Population Economics	-1.29**	0.27	4.75			
Population Index	-0.25	0.78	0.66			
Impact factor journal				0.72**	2.06	7.92
Reputation editorial board ($x\ 10^{-2}$)				2.53**	1.03	5.78
Circulation journal i (x 1000)				0.06**	1.06	3.04
Use of french language				-0.92**	0.40	6.14
Constant	1.81**		7.07	0.13		0.90
ln á	-0.18**		2.68	-0.07		1.08
LR test á = 0 : $\div^2$ (1)		2047.1			2221.3	
degrees of freedom (= df)		32			20	

$\div^2$ (df)	857.1	767.4
Log Likelihood	-2702.2	-2747.1

(a) The symbol * denotes significance at $p < 0.05$; ** at $p < 0.01$. IRR denotes the incidence rate ratios. The sample size N is 1,371 articles.

The results seem to indicate that the circulation has an additional positive effect on citation frequencies. An extra thousand subscribers to a journal will increase the number of citations received within the exposure time by 6 percent. Circulation numbers may be a reflection of the field size and as such this finding compares well with Archibald and Finifter's conclusion (1990) that journals specializing in larger fields will be cited more frequently. The use of French in communicating scientific results is substantially penalized by fellow scientists as citations received by French articles amount to no more than 40 percent of those received by English articles of comparable quality and focus.

With respect to article characteristics, the *visibility* of the article has a major impact on the number of citations accumulated. For instance, a presidential address tends to generate far more citations than regular journal contributions. The opposite applies to notes and comments. As shown in the column with the incidence rate ratios (IRR), a note tends to receive 55 to 58 percent of the citations which a regular article would receive. In model 1 the number of pages does not seem to affect citation frequency. This finding is not robust, however, as the number of pages does affect citation frequency in model 2.

The last visibility characteristic concerns the order in which articles appear in a journal issue. As shown in Table 3, this variable is significantly related to an article's citation frequency. Articles appearing first in a journal issue are more frequently cited than articles appearing at the back of an issue.

With regard to the *content* of the articles, the results suggest that articles in which non-US data are analyzed are less frequently cited than articles with an empirical focus on the US. However, this rule does not apply to articles using European data or articles with a global perspective or a non-empirical focus. Hence, writing an article based on, for instance, African data will receive 60 percent of the citations which a comparable article based on US data would have received. This particular effect of article content is more or less stable across the two regression models. Far less stable is the effect of articles that have a historical focus. Still, the sign of the effect itself is unambiguously

negative: historical articles receive 54 percent (model 1) to 66 percent (model 2) of the citations received by regular articles focusing on post-World War II data. This particular finding is in line with the assertion by Greenhalgh (1996) that modern demography is dominated by an ahistorical focus. Furthermore, the results refine the conclusion of Bayer's (1982: 531) citation analysis of a selection of articles from the *Journal of Marriage and the Family* that the degree of empirical orientation has no relation with citation frequency. The reason why he fails to find an effect is perhaps easy to trace: in (family) sociology most articles are empirical and under those circumstances, an empirical-nonempirical dichotomy is not the most distinguishing feature. His conclusion may no longer be valid if one refines the categorization of 'empirical research', as we have done in this paper.

With respect to the author characteristics, models 1 and 2 both show that articles produced by authors with a good citation record are more frequently cited than articles written by authors with a weak citation record. Authors with a record of 100 citations obtain an average of 1.75 (model 1) to 1.87 (model 2) more citations than authors with no citation to their name. However, the effect is non-linear as may be deduced from the second row, but this effect is extremely small and has a negligible impact on the reputation effect of citation records. In addition to the reputation effect, the number of authors and a US affiliation of at least one of the authors increases the number of citations.[8] An additional co-author increases the number of citations by 7 percent and having a US affiliation increases the citation count by 20 percent.

Finally, the test statistics at the bottom of Table 3 also show that the negative binomial regression model is a more appropriate model for estimating citation counts than the Poisson model. On the basis of the likelihood ratio statistics, which test whether the dispersion parameter α is equal to zero, one can safely reject the hypothesis that the citation data are Poisson- distributed.

To be or not to be cited

[8]. Other personal variables include the authors' sex or race. There is a vast literature that has focussed on sex discrimination. In general, the results either point to a negative effect on scientific productivity of being female in a male-dominated science (Baldi 1998; Xie and Schauman 1998) or else they indicate a complete lack of discrimination effects (Clemente 1973; Blank 1991; Smart and Waldfogel 1996). Departmental effects and individual reputations are found to be far more important than sex or race (Clemente 1973; Allison and Long 1990). For a limited set of articles for which we could establish the sex of the authors, we analyzed whether sex is a predictor of citation frequency in demography. We did not find significant traces of citation discrimination. In fact, women even seem to be slightly more influential than men. This may be attributed for a part to their professional involvement in the specialization of family planning, an area characterized by higher citation practices than other specializations in demography.

A question that remains unsolved in the previous regression analysis is whether each and every article receives a fair chance of being heard and subsequently used. As mentioned in the methodology section of this paper, we can also test the openness of demography as a science by analyzing a binary dependent variable: cited = 1 and not cited = 0.

We tested two models with the same explanatory variables as those described in the previous section.[9] Model 1 in Table 4 shows that with respect to the influence of the journal in which an article is published, the results are more or less in line with the earlier results concerning the citation counts in Table 3. Differences between *Population and Development Review, Family Planning Perspectives* and *Demography* are no longer significant, and here too the probability of being cited in one of the journals outside the core is small: compared with an article published in *Demography*, the probability of being cited in a second-tier journal varies from 2 to 16 percent. Model 2 reemphasizes the influence of journal quality and language on the probability of being cited. As mentioned in the previous subsection, writing an article in French is severely penalized: the probability of being cited is 50 percent of that of a comparable English-language article. In view of the fact that 64 percent of demography articles in the sample period were cited within five years, this estimate gives us a rough idea of why the French journal *Population* (only 38 percent of its articles are cited) plays such a minor role in demography. Our results suggest that much would be gained if English were to become the official language of this journal. For some time now, the editors of *Population* have realized that the French language limits its readership, as reflected by the fact that they now publish an English edition of *Population*, which contains translations of the best French articles of the preceding year.

[9]. Compared with the previous estimation procedures we have left out the variable 'presidential addresses' as this variable predicts the occurrence of citation perfectly.

Table 4: Logit regression of the chance of being cited in the five years following publication

	Dependent variable: chance of being cited within 5 years					
Explanatory variables	model 1			model 2		
	Coefficient	Odds ratio	t-statistic	Coefficient	Odds ratio	t-statistic
Author characteristics						
Max. reputation author ($x10^{-2}$)	0.96*	2.61	2.07	1.09*	2.98	2.38
Max. reputation author squared ($x10^{-4}$)	-0.12	0.89	0.97	-0.15	0.86	1.31
US affiliation authors	0.08	1.08	0.44	0.22	1.24	1.32
Number of authors	0.03	1.03	0.47	0.00	1.00	0.03
Article characteristics						
Visibility						
Comment/reply/note	-0.83**	0.44	3.75	-0.66**	0.52	3.17
Number of pages	-0.05	0.95	1.53	-0.02	0.98	0.90
Order in a journal issue	-0.14**	0.87	3.35	-0.11**	0.90	2.91
Content						
Historical orientation	-0.44	0.64	1.34	-0.27	0.76	0.90
Focus of article:						
US excluded	•	•	•	•	•	•
Europe	0.10	1.10	0.38	0.06	1.06	0.26
Asia/Australia	-0.03	0.97	0.13	-0.38	0.68	1.74
Africa	-0.07	0.93	0.23	-0.17	0.84	0.59
Latin America	-0.65*	0.52	1.98	-0.58	0.56	1.81
Middle East	-0.27	0.77	0.52	-0.20	0.82	0.40
World	0.05	1.05	0.16	0.05	1.05	0.19
Non-empirical focus	-0.17	0.84	0.71	-0.33	0.72	1.44
Journal characteristics						
Demography excluded	•	•	•			
Family Planning Perspectives	0.01	1.01	0.01			
Population & Development Review	0.82	2.27	1.01			
Population Studies	0.28	1.33	0.36			
Studies in Family Planning	-1.09	0.34	1.65			
Journal of Biosocial Science	-2.43**	0.09	3.99			
International Migration Review	-1.86**	0.16	3.07			
Social Biology	-2.30**	0.10	3.62			
Population	-2.85**	0.06	4.71			
Population Bulletin	-0.05	0.95	0.04			
Population and Environment	-2.93**	0.05	4.58			
Population Research & Policy Review	-2.87**	0.06	4.30			
European Journal of Population	-3.09**	0.05	4.47			
International Migration	-2.65**	0.07	4.17			
Journal of Family Welfare	-4.02**	0.02	6.01			
Journal of Population Economics	-2.06**	0.13	3.12			
Population Index	-1.25	0.29	1.04			
Impact factor journal				1.42**	4.14	6.44
Reputation editorial board ($x\ 10^{-2}$)				3.97**	1.04	4.33
Circulation journal (x 1000)				-0.01	0.99	0.17
Use of french language				-0.70**	0.50	2.94
Constant	3.26**		4.90	0.20		0.74

Degrees of freedom (= df)	31	19
$\times^2$ (df)	409.7	357.7
Log Likelihood	-692.9	-718.9

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

With respect to article characteristics, the only remaining strong effects are the visibility aspects of an article. A note or comment is roughly half as likely to be cited compared with a regular full-size article and the probability that a second article in a journal issue will be cited is approximately 90 percent of the citation probability of the first article. Contrary to the results found in the model to explain citation frequencies, the regional focus does not significantly affect the probability of remaining uncited in the five years following publication. The only exception to this observation is an article focusing on Latin American data. Compared with a US focus, the probability of being cited drops by almost 50 percent. The lack of influence of regional focus in Table 4 and the observation that we do not find a significant effect of journal circulation is remarkable. It may indicate that market size considerations primarily affect citation frequencies rather than occurrences of citation.

The only author characteristic that remains important in explaining citation frequency is the reputation of the most prominent author. To give an extreme example, an author with 100 citations to his or her name increases the probability of being cited within five years by a factor 2.6 (model 1) to 3.0 (model 2), compared with an author with no citations at all to his or her name. Although this may seem large, it remains a low multiplier effect as demographers with 100 citations to their name are the 'big shots' in their profession.

Finally, we have to pay attention to the relative importance of the various groups of factors in our models. We have summarized some diagnostic test statistics in Table 5 for various models and various estimation techniques underlying model 1.

Table 5: Relative importance of variable categories

	Negative binomial regression				Logit regression			
Model specification:	Pseudo R^2	df	$\div^2$(df)	LogL	Pseudo R^2	df	$\times^2$(df)	LogL
1. Journal characteristics	0.115	16	719.6	-2771.0	0.198	16	355.2	-720.2
2. Article characteristics	0.038	12	238.5	-3011.5	0.057	11	102.1	-846.7

3. Author characteristics	0.042	4	265.4	-2998.1	0.066	4	117.6	-839.0
4. Journal + Article	0.132	28	827.1	-2717.2	0.224	27	403.0	-696.3
5. Journal + Author	0.120	20	751.1	-2755.2	0.202	20	363.4	-716.1
6. Journal + Author + Article	0.137	32	857.1	-2702.2	0.228	31	409.7	-692.9

The table shows that journal characteristics are the most important factors in our model. Although article and author characteristics provide some additional explanatory power, we should add that the contribution of author characteristics is rather small. In other words, if articles were published anonymously, our ability to explain the variance in citation frequencies would hardly be affected. The much-discussed Matthew effect in science does not seem to play as dominant a role as one would expect on the basis of the attention it receives in sociology of science literature.

7. Conclusions and discussion

The central question of this paper was: what makes a demographic article influential? From the outset it should be stressed that both the universalist view and the social constructivist view carry some weight in explaining citation frequencies. There is some evidence that author reputations do effect citation counts, but their additional explanatory power is weak. The reputation effect of authors on citation counts is in line with a result found in a similar study by Stewart (1983) on articles in geophysical research. The minor impact of social forces is in contradiction with Baldi's suggestion (1998) that in the social sciences citations may be influenced less by the intellectual content of articles and more by the characteristics of their authors than in the natural sciences.

Journals are a dominant force in allocating citations. Articles published in core journals receive considerably more citations than articles in second-tier journals. With respect to the important role played by journals, a few comments are due here. First, the reputation of journals (and their editors) for being original and influential can attract high-quality papers. Merton's Matthew effect (1968) is therefore not only visible at the level of individual authors but perhaps also at journal level. Potentially influential papers will therefore be submitted to, and perhaps published in, journals that enjoy a reputation for being influential. Hence, this element of the scientific communication process reinforces a journal's reputation. This makes the entry of second-tier journals into the group of core journals very difficult. Editors of first-tier journals tend to have access to a pool of original and high-quality papers, and by choosing and publishing high-impact articles they perpetuate the status a journal enjoys. Second, the results suggest that besides quality, a large circulation and the use of the English language make a journal and the articles it publishes more influential.

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. However, empirical evidence of the review process at the *American Economic Review* suggests that author reputations do not affect acceptance rates (Blank 1991).

In our efforts to detect the dominant forces underlying citation behavior we came across another element of science that has not been widely examined in sociological studies of science, namely leadership in science. This element does not concur perfectly with the social constructivist view in which position and affiliation are seen as dominant factors, nor does it concur with the universalist view in which regional leadership is not fully acknowledged as a force of any importance. In our view, the topic of debate is set by the technological leader or leaders. Our study shows that publishing empirical knowledge focusing on the US or Europe increases the citation frequency significantly compared with researchers who focus on developing countries. Two explanations for this finding spring to mind. First, articles with a US or European focus may simply be better articles. Second, scientists pay more attention to American and European articles whilst ignoring equally good articles with a different empirical focus. If the latter assumption is true, science may be in trouble as it suggests that there is reason to believe that the observation regarding the Americanization of economic science made by Frey and Eichenberger (1993, 1997) also applies to demography. They claim that incentives for ambitious scholars in economics are geared toward studying US data and US problems. The citation data in this paper show that it is perhaps a trend not so much toward Americanization but rather toward Westernization that distorts scientific development as articles focusing on developing countries are not as highly rewarded as articles focusing on 'developed' countries. In that respect, Greenhalgh (1996) seems to be right in her characterization of demography as being 'Eurocentric', which Greenhalgh (1996: 27) defines as "the notion that Europe and its offshoots are superior to the rest of the world and the source of all significant change."

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. Only top-quality articles may reach the English-speaking community of scientists, if they have been translated, and even then with a considerable time-lag. A recent example of this phenomenon is mentioned by Portes (1998: 3) who notes how the French sociologist Pierre Bourdieu's groundbreaking analysis of social capital did not gain immediate attention in the English-speaking world, because his initial treatment of the concept of social capital appeared in a provisional note 'Le capital social', published in the *Actes de la Recherche en Sciences Sociales* in 1980. Only the translations, first in German in 1983 and then in English in 1985, made the dissemination of his ideas possible, but of course with a considerable time-lag.

When evaluating the results presented here, we need to draw attention to several limitations. By choosing citation counts as a dependent variable we restricted our study to establishing the role of the *cited* article in explaining citation behavior. We did not consider the characteristics of the *citing* articles as determinants of citations. The fact that information on the citing articles is diffuse and hard to obtain restricts our ability to assess the importance of relational variables such as US affiliation. As such, our study takes us part of the way to understanding citation behavior in demography. Further research may attempt to include characteristics of citing articles.

Despite these limitations, the results of this study suggest that to a large extent the principles of open scientific communication apply to demography as a science. Such openness is essential to guarantee the cumulation of scientific knowledge. However, in view of the fact that staffing and financial decisions are made increasingly on the basis of citations, we must also be aware of the pitfalls of using citations as the yardstick of 'quality'. If we fail to do so, we risk making mistakes that could ultimately be detrimental to the quality of the academic debate.

References

Allison, P.D. 1978. "The Reliability of Variables Measured as the Number of Events in an Interval of -253 in *Sociological Methodology*, edited by K.F. Schuessler. San Francisco: Jossey-Bass Publishers.

Allison, P.D., and J.S. Long. 1990. "Departmental Effects on Scientific Productivity." *American Sociological Review* 55: 469-478.

Anderson, A.B., A. Baselevsky, and D.P.J. Hume. 1983. "Missing Data: A Review of the *Handbook of Survey Research*, edited by P. H. Rossi, J.D. Wright, and A.B. Anderson. New York: Academic Press.

Archibald, R.B., and D.H. Finifter. 1990. "Multivariate Citation Functions and Journal Rankings." *Eastern Economic Journal* 16: 151-158.

Baldi, S. 1998. "Normative Versus Social Constructivist Processes in the Allocation of Citations: A Network-Analytic Model." *American Sociological Review* 63: 829-846.

Bayer, A.E. 1982. "A Bibliometric Analysis of Marriage and Family Literature." *Journal of Marriage and the Family* 44: 527-538.

Bayer, A.E., and J.F. Folger. 1966. "Some Correlates of a Citation Measure of Productivity in *Sociology of Education* 39: 381-389.

Blank, R. 1991. "The Effects of Double-Blind Versus Single-Blind Reviewing: Experimental Evidence from the American Economic Review." *American Economic Review* 81: 1041-1067.

Caldwell, D.C. 1996. "Demography and Social Science." *Population Studies* 50: 305-333.

Clemens, E.S., W.W. Powell, K. McIlwaine, and D. Okamoto. 1995. "Careers in Print: Books, Journals and Scholarly Publications." *American Journal of Sociology* 101: 433-494.

Clemente, F. 1973. "Early Career Determinants of Research Productivity." *American Journal of Sociology* 79: 409-419.

Cole, J.R., and S.Cole. 1973. *Social Stratification in Science*. Chicago: University of Chicago Press.

Cox, R.A.K., and K.H. Chung. 1991. "Patterns of Research Output and Author Concentration in *Review of Economics and Statistics* :740-747.

Cronin, B. 1984. *The Citation Process - The Role and Significance of Citations in Scientific Communication*. London: Taylor Graham.

David, P.A. 1994. "Positive Feedbacks and Research Productivity in Science: Reopening Another -89 in *Economics of Technology*, edited by O. Granstrand. Amsterdam: Elsevier.

Demeny, P. 1988. "Social Science and Population Policy." *Population and Development Review* 14: 451-479.

Diamond, A.M. 1986. "What is a Citation Worth?" *Journal of Human Resources* 21: 200-215.

Frey, B.S., and R. Eichenberger. 1993. "American and European Economics and Economists." *Journal of Economic Perspectives* 9: 185-194.

_____. 1997. "Economists: First Semester, High Flyers and UFOs." Pp. 15-48 in *Economic Science and Practice*, edited by P.A.G. van Bergeijk et al. Cheltenham: Edward Elgar.

Gilbert, N. 1977. "Referencing as Persuasi *Social Studies of Science* 7: 112-122.

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

Greenhalgh, S. 1996. "The Social Construction of Population Science: An Intellectual, Institutional, and Political History of Twentieth-Century Demography." *Comparative Studies of Society and History* 38: 26-66.

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

Hamilton, D.P. 1991. "Research Papers: Who is Cited Now?." *Science* 251: 25.

Hargens, L.L., and W.O. Hagstrom. 1982. "Scientific Consensus and Academic Status Attainment *Sociology of Education* 55: 183-196.

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.

Hodgson, D. 1991. "The Ideological Origins of the Population Association of America." *Population and Development Review* 17: 1-34.

Holub, H.W., G. Tappeiner, and V. Eberharter. 1991. "The Iron Law of Important Articles." *Southern Economic Journal* 58: 317-328.

Keyfitz, N. 1993. "Thirty Years of Demography and *Demography*." *Demography* 30: 533-549.

Latour, B. 1987. *Science in Action - How to Follow Scientists and Engineers Through Society.* Milton Keynes: Open University Press.

Lindsey, D. 1978. *The Scientific Publication System in Social Science*. San Francisco: Jossey-Bass Publishers.

Lotka, A.J. 1926. "The Frequency Distribution of Scientific Productivity." *Journal of the Washington Academy of Sciences* 16: 317-323.

May, R.M. 1997. "The Scientific Wealth of Nations." *Science* 275: 793-796.

Merton, R.K. 1957. "Priorities in Scientific Discovery." *American Sociological Review* 22: 635-659.

_____.1968. "The Matthew Effect in Science." *Science* 159: 56-63.

_____.1973. *The Sociology of Science*. Chicago: University of Chicago Press.

Pendlebury, D.A. 1991. "Letter." *Science* 251: 1410-1411.

Portes, A. 1998. "Social Capital: Its Origins and Applications in Modern Sociology." *Annual Review of Sociology* 24: 1-24.

Price, D.J. de Solla. 1963. *Little Science, Big Science*. New York: Columbia University Press.

_____. 1976. "A General Theory of Bibliometric and Other Cumulative Advantage Processes." *Journal of the American Society for Information Science* 27: 292-306.

Schwartz, C.A. 1997. "The Rise and Fall of Uncitedness." *College & Research Libraries* 58: 19-29.

Smart, S., and J. Waldfogel. 1996. "A Citation-Based Test for Discrimination at Economics and Finance Journals." NBER Working Paper, no. 5460. Cambridge MA.

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

Teachman, J.D., K. Paasch, and K.P. Carver. 1993. "Thirty Years of De *Demography* 30: 523-532.

Tijssen, R.J.W., T.N. van Leeuwen, B. Verspagen, and H. Hollanders.1998. *Wetenschaps- en technologie-indicatoren 1998* (transl. 'Science and Technology Indicators 1998'). Leiden: DSWO Press.

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

Xie, Y., and K.A. Schauman. 1998. “Sex Differences in Research Productivity: New Evidence *American Sociological Review* 63: 847-870.
Zuckerman, H. 1977. *The Scientific Elite - Nobel Laureates in the United States*. New York: Free Press.

Appendix

Table A1: Key figures for demography journal, 1990-1992

Name journal:	# Authors	% US connection	Impact factor 1990	Average reputation editorial board[a]	Reputation author[b]	Circulation[d]
1. Demography	1.8	92.5	1.80	23.8	33.6	4,000
2. Family Planning Perspectives	2.3	96.7	1.30	27.8	34.9	10,000
3. Population Development Review	1.6	64.0	1.33	33.5	34.4	5,000
4. Population Studies	1.6	53.2	0.91	14.2	18.5	2,800
5. Studies in Family Planning	3.0	70.7	1.08	25.8	19.6	6,000
6. Journal of Biosocial Science	2.1	31.4	0.41	13.1	14.8	600
7. International Migration Review	1.4	54.6	0.40	6.5	9.7	2,500
8. Social Biology	1.9	74.7	0.39	0.5	16.6	1,650
9. Population	1.2	0.6	0.24	9.9	5.9	4,500
10. Population Bulletin	2.2	83.3	0.62	6.0	13.2	6,000
11. Population and Environment	1.3	86.7	0.17	26.2	10.7	n.a.
12. Population Research Policy Review	1.8	90.9	0.58	8.1	20.6	700
13. European Journal of Population	1.3	26.2	0.30	19.6	8.6	400
14. International Migration	1.3	28.7	0.12	0.0	5.3	2,000
15. Journal of Family Welfare	1.7	0.0	0.05	0.0	1.2	3,000
16. Journal of Population Economics	1.5	48.1	0.40[c]	23.8	32.8	500
17. Population Index	1.4	88.9	0.43	29.8	49.9	4,400
Average	1.7	50.8	0.64	14.7	17.1	3,432

(a) Measured by the average number of citations accumulated by the editorial board (editor(s) and advisory editorial board) in 1990.
(b) Measured by the total number of citations accumulated by the author with the highest reputation (citation stock) of an article in 1990.
(c) The *Journal of Population Economics* was not yet recorded in 1990 by SSCI and 0.40 is the figure applicable to the first recorded years 1994-1995.
(d) Sources: *The Serials Directory - An International Reference Book*, Birmingham, Alabama; and *Ulrich's Plus - The Complete International Serials Database*, New York, Bowker Electronic Publishers.