Journal of Informetrics 9 (2015) 16–24

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

The value of experience in research

Thor-Erik Sandberg Hanssen*, Finn Jørgensen

Bodø Graduate School of Business, University of Nordland, Universitetsalleen 11, No.-8049 Bodø, Norway

ARTICLE INFO

Article history:
Received 20 June 2014
Received in revised form 22 October 2014
Accepted 3 November 2014
Available online 23 November 2014

Keywords:
Human capital
Experience
Citation count
Publications
Negative binomial model

ABSTRACT

This article examines how different factors influence the number of times articles in the five most recognized transportation journals are cited. The effects of most of the explanatory variables indicating the characteristics of articles, authors and journals correspond with earlier studies of citation counts. Special focus in this study is placed on estimating the relationship between researchers' human capital or skills and their experience. For the purpose of this study, human capital is defined as a scientist's ability to conduct research at the frontier of his or her discipline and is measured by how frequently his or her research is cited. Experience is measured by counting the number of their previous scientific articles. Using negative binomial regression, we find that experience offers a statistically significant positive effect on the human capital of scientists. However, this effect diminishes rapidly with the level of experience. This suggests that young researchers relatively quickly learn the skills and gain the knowledge necessary to produce high-quality research.

1. Introduction

Human capital is the accumulated skill and knowledge of human beings (Bade & Parkin, 2011) and an important factor in individual, firm and national economic growth (Blundell, Dearden, Meghir, & Sianesi, 1999). This is because individuals with more or higher quality human capital achieve higher performance (Dimov & Shepherd, 2005) and produce higher quality goods (Stokey, 1991). In Norway, for example, estimates from National Statistics and The Ministry of Finance suggest that the present value of human capital amounts to between 71% and 81% of national wealth (Flåten, 2013).

Human capital is similar to physical means of production (Becker, 1964); as in the case of investments in physical means, human capital investments can improve output. The most important investments to be made in human capital are education and on-the-job training (Becker, 1964). The majority of students experience negative net earnings while in school. This can be considered the cost of investing in human capital through schooling, and the monetary returns on such investments have been found to be significant at each stage of the educational process (Blundell, Dearden, & Sianesi, 2005). Training also has an important effect on the relationship between earnings and age, and because the income of trained individuals tends to lie below marginal productivity during training and becomes equal afterwards, these individuals' income increases sharply at the end of their training period before levelling off (Becker, 1964).

As far as researchers are concerned, the total number of times their papers have been cited is a good proxy for their human capital; i.e., their ability to do quality research at the frontiers of their discipline (Diamond, 1986). Such an operational measure of a scientist's skills or human capital has gradually become more important because the number of citations is now frequently examined when decisions are made with regard to funding, promotion, and career advancement in academia

* Corresponding author. Tel.: +47 75 51 76 56.
E-mail address: thh@uin.no (T.-E.S. Hanssen).

http://dx.doi.org/10.1016/j.joi.2014.11.003
1751-1577/

(Dries, Pepermans, & Carlier, 2008). Moreover, the number of citations is commonly used when evaluating journals, departments and research institutions (see Bornmann & Daniel, 2008). Thus, the widespread use of citation counts at all levels in academia indicates that scientists and decision-makers perceive that those institutions in which scientists in total have many citations are also institutions with plenty of human capital.

This article aims to analyze how different factors influence the number of times an article is being cited using data from the five most recognized transportation journals. We look at the articles published in these journals between 2000 and 2005 and count how many citations each of them received in late 2012. This paper extends the work done by Hanssen and Jørgensen (2014) in two key ways. First, it takes into account the authors' total experience, or on-the-job-training, before they wrote the article in question, operationalized by way of the total number of articles they have published previously in peer reviewed scientific journals. Second, the total number of times their previous works has been cited is introduced as an explanatory variable indicating the authors' skills or total human capital at the time they published the article in question. As mentioned previously, the number of times an author is cited signifies his or her ability to conduct quality research at the frontiers of his or her discipline. Summing up, compared with Hanssen and Jørgensen (2014) special focus is placed here on analysing how authors' experience and skills both influence the number of citations their articles will achieve.

The remainder of this article is organized as follows. Section 2 presents literature on human capital and the authors' hypotheses on the influence of authors' experience and skills on citation counts. In Section 3, a review of the literature on factors influencing citation counts is presented along with our hypotheses related to the influence of selected author, article and journal characteristics on citation counts. The methodology applied is presented in Section 4, followed by the estimation results in Section 5. Finally, in Section 6, the findings are discussed and summarized.

2. The influence of author experience and author skills on citation count

The relationship between productivity and experience has been investigated in sectors as diverse as baseball (Krohn, 1983) and venture capital firms (Dimov & Shepherd, 2005). In science the relationship has been addressed using age as a proxy for experience (Bonaccorsi & Daraio, 2003). However, past performance is a much better predictor of scientific productivity than age (Stroebe, 2010). Thus, researchers who are productive at a young age tend to be more productive when they are older, as opposed to researchers who are less productive at a young age. For this reason, instead of age, this article uses the number of previous publications to provide a measure of experience. Based on the above, experience can be considered an investment in human capital as such, offering increased productivity through higher citation rates. However, there are diminishing returns to experience (Berman, Down, & Hill, 2002; Loughran, Nguyen, Piquero, & Fagan, 2013), meaning that an additional unit of experience yields less and less additional human capital.

Whilst the number of previous publications by authors is an indicator of the authors' experience, their total number of previous citations describes their skills to a larger extent. The argument for including the number of previous citations as an explanatory variable is twofold. First, better known scientists tend to receive more credit than less well-known scientists, even if their work is similar (Merton, 1968). Frequently cited researchers generally have higher status than researchers who are cited less frequently. Because status influences perceptions of quality, those with high status often receive favourable assessments (Azoulay, Stuart, & Wang, 2014), which in turn can increase their odds of being cited. Higher status does not only influence perceptions of quality. Scientists with high status are also more likely to attract tangible resources, such as research funding and outstanding graduate students, which can result in research of higher quality (Merton, 1968). Second, many citations indicate that previous articles written by the author(s) have been at the frontier of research, i.e., of high quality. However, empirical studies cast doubts on previous number of citations as measure of authors' skills because an article citation record also seems to be affected by factors that are not directly linked to the scientific quality of it, for example its number of words in the title, place in the alphabet of the first author etc. (Hanssen & Jørgensen, 2014). Despite the last caveats, we expect to find a positive association between citation count and the number of previous citations (Stroebe, 2010). Based on the above discussion, we derive the following two hypotheses:

H1. Author experience affects the number of times the article is cited, in that the relationship between the numbers of times an article is cited and the author's total number of previous publications increases concavely.

H2. Author skills affect the number of times the article is cited, in that articles written by authors who have been cited many times previously are more frequently cited.

3. Other factors associated with citation counts

The main objective of this study is to investigate the influence of authors' experience on citation counts. This is done by testing H1 and H2. To isolate the effect of the number of previous publications and citations on citation counts, we control for a set of variables identified in previous research to correlate with citation counts. Broadly speaking, these factors can be related to author characteristics, article characteristics or journal characteristics.

3.1. Authors' characteristics beyond experience and skills

A number of studies have found that articles written by authors from more than one country are more frequently cited than articles where all authors are from the same country (e.g. Aksnes, 2003; Hanssen & Jørgensen, 2014; Katz & Hicks, 1997). Authors of multinational articles are more likely to belong, or have access to, separate research networks than authors from just one country, which suggests that such articles can be expected to be spread to a wider audience of researchers faster when they have been written by authors from more than one country (Hanssen & Jørgensen, 2014). Moreover, because the start-up cost of international research projects is likely to exceed the start-up cost of a national project, it is to be expected that such projects are only initiated when a really good idea exists. Research projects based on good ideas may in turn be expected to generate more citations.

Several studies conclude that the number of times an article is cited increases significantly by its number of authors (Beaver, 2004; Borsuk, Budden, Leimu, Aarssen, & Lortie, 2009; Katz & Hicks, 1997). It has been suggested that this association is the result of the larger scientific network of $n + 1$ authors, than that of n authors only (Hanssen & Jørgensen, 2014). Moreover, articles written by more than one author may benefit from a division of labour.

Articles written by authors whose surname begins with a letter close to the beginning of the alphabet are more frequently cited than articles written by authors whose surname is at the end of the alphabet (Tregenza, 1997). Such an association could be related to the function that allows for search results in abstract and citation databases to be sorted according to the alphabetical position of the first author's surname.

Based on this discussion, we derive the following hypothesis with regard to the above author characteristics:

H3. (a) Articles written by authors from more than one country are more frequently cited than those written by authors from just one country, (b) articles written by several authors are more frequently cited than articles written by one author only (c) articles written by authors in which the first author's surname is early in the alphabet are more frequently cited than when his/her surname is at the end of the alphabet.

3.2. Article characteristics

A number of studies have found a positive association between the number of times an article is cited and the length of its reference list (e.g. Hanssen & Jørgensen, 2014; Lokker, McKibbon, McKinlay, Wilczynski, & Haynes, 2008; Vieira & Gomes, 2010). It has been suggested that this is because references make an article more visible (Didegah & Thelwall, 2013). Because researchers might ask to be notified electronically when their own articles are cited, it is reasonable to assume that, all other things being equal, an article with many references will generate more attention than an article with fewer references. Moreover, researchers may feel indebted to researchers who have cited their work and may be inclined to return the citation (Stremersch, Verniers, & Verhoef, 2007; Webster, Jonason, & Schember, 2009).

The title is considered the most important element of scientific articles (Jamali & Nikzad, 2011). Its primary function is to draw readers' attention and indicate the article's content in a short glimpse, thus leading to initial selection or rejection of the article (Yitzhaki, 2002). Consequently, titles are an important marketing tool as hundreds of articles are published every year and compete with each other to be read (Jamali & Nikzad, 2011). It has been argued that shorter titles might be more frequently cited because long or confusing titles may act as deterrents to further reading (Vintzileos & Ananth, 2010). Moreover, articles with short titles make it less laborious to refer to them. These arguments gain support from empirical studies (e.g. Hanssen & Jørgensen, 2014; Jamali & Nikzad, 2011; Paiva, Lima, & Paiva, 2012). However, other works indicate that a higher number of citations is positively correlated the length of the article's title (Habibzadeh & Yadollahie, 2010) because a longer title is more likely to contain any given search term in electronic databases (Subotic & Mukherjee, 2014), i.e., they become better marketed.

Finally, older articles are, all else equal, more frequently cited than ones published more recently (Bergh, Perry, & Hanke, 2006; Judge, Cable, Colbert, & Rynes, 2007). This is reasonable because citations accumulate over time.

On the basis of this discussion, we derive the following hypothesis with regard to article characteristics:

H4. (a) The number of times an article is cited increases with the number of works it refers to, (b) from previous research it is ambiguous how the length of an article's title affects its number of citations and (c) the number of times an article is cited increases with the number of years since publication.

3.3. Journal characteristics

The average citation rate of the journal in which an article is published has been found to be the single most important factor driving citations to an article (Judge et al., 2007; Vanclay, 2013). This is mainly because the academic threshold level for an article to be accepted varies a lot between journals. Based on this, we derive the following hypothesis:

H5. The journal in which an article is published affects the number of times it is cited. All other things being equal, articles published in recognized journals with high impact factors are more frequently cited than articles published in more unknown journals. As far as our selected five journals are concerned, impact factor surveys and results from earlier studies showed the following citation ranking: Transportation Part B > Transportation Science > Transportation > Transportation Part A > Journal

Table 1
Summary statistics of the dataset.

Variable code	Definition	Mean	S.D.	Min	Max
Dependent variable					
CI	Citations per article by 27 December 2012	28.08	29.58	0	294
Author characteristic					
PP	Experience, measured by the number of prior publications by the authors	44.78	52.78	0	537
PC	Authors' skills, measured by total number of prior citations by the authors	159.97	339.39	0	7 417
AN[a]	Place in alphabet of first author's surname	10.96	7.18	1	26
NA	Number of authors (*NA* = 1 more than one, *NA* = 0 if one)	0.77	0.42	0	1
AC	Number of countries the authors are from (*AC* = 1more than one, *AC* = 0 if one)	0.20	0.40	0	1
Article characteristic					
RF	Number of references	25.96	16.73	0	167
WT	Words in title	10.75	3.66	1	28
YP	Years since publication	9.94	1.38	8	12
Name of journals					
TR	Transportation, *TR* = 1 if published here, *TR* = 0 otherwise	0.13	0.34	0	1
PA	Transportation Research Part A, *PA* = 1 if published here, *PA* = 0 otherwise	0.27	0.44	0	1
PB	Transportation Research Part B, *PB* = 1 if published here, *PB* = 0 otherwise	0.28	0.45	0	1
TS	Transportation Science, *TS* = 1 if published here, *TS* = 0 otherwise	0.19	0.39	0	1
JT	Journal of Transport Economics and Policy, *JT* = 1 if published here, *JT* = 0 otherwise	0.13	0.33	0	1

[a] Articles written by first authors whose surname begin with the letter *A* are given the value 1, a first author whose surname begins with the letter *B* is given the value 2, etc.

Table 2
Pairwise correlation matrix of the explanatory variables.

	PP	*PC*	*AN*	*NA*	*AC*	*RF*	*WT*	*YP*	*TR*	*PA*	*PB*	*TS*	*JT*
PP	–												
PC	63*	–											
AN	10*	08	–										
NA	28*	17*	−05	–									
AC	22*	17*	03	27*	–								
RF	02	15*	−01	−05	03	–							
WT	03	03	00	01	06	05	–						
YP	−10*	−16*	−01	−08*	−03	−12*	−03	–					
TR	−04	−04	03	01	−01	05	10*	00	–				
PA	−03	−02	−01	04	02	−05	10*	02	−23*	–			
PB	06	01	02	00	11*	−05	−03	−02	−25*	−38*	–		
TS	08*	11*	−07*	10*	−05	−01	−12*	−01	−19*	−30*	−31*	–	
JT	−09*	−08*	04	−19*	−11*	11*	−05	03	−15*	−23*	−24*	−18*	–

Note: Decimal points were removed. Abbreviations are defined in Table 1.
* Indicates correlation coefficients statistically significant at a 5 percent level or better.

of Transport Economics and Policy (Hanssen & Jørgensen, 2014). Our hypothesis is that the same ranking will apply with our extended dataset.

4. Method

4.1. The data

The data presented in this article were collected from Scopus, the world's largest abstract and citation database of peer-reviewed literature (www.scopus.com). The database was searched in order to obtain data from the articles published in the following five internationally most highly recognized peer-reviewed transportation journals: Transportation (*TR*), Transportation Research Part A (*PA*), Transportation Research Part B (*PB*), Transportation Science (*TS*) and Journal of Transport Economics and Policy (*JT*). Data on all articles published in the above-mentioned journals in the five-year period from January 1, 2000, to December 31, 2004, are analyzed as of December 27, 2012; i.e., 8–12 years after they were published. These particular time periods were chosen because they allow sufficient time for articles to reveal their scientific impact. In total, the dataset includes 779 articles, which on average have been cited 28.1 times as of December 27, 2012.

Table 1 summarizes the variables explored in this study. The dependent variable, i.e., the number of article citations (*CI*), refers to the number of times each article in our dataset has been cited in Scopus by December 27, 2012. The explanatory variables are grouped in variables indicating (1) authors' characteristics; (2) articles' characteristics and (3) the name of the journal in which the article is published.

The pairwise correlation matrix of the explanatory variables is shown in Table 2. As expected, the highest correlation coefficient (0.63) is between authors' total previous number of publications (*PP*) and their previous number of total citations (*PC*).

Although Scopus is the world's largest abstract and citation database of peer-reviewed literature, a disadvantage relates to older articles. It has been reported that the database is not complete for the years prior to 1995 (www.scopus.com). As a result, for some of the most experienced researchers, i.e., those who published articles prior to 1995, our dataset might underestimate their values of *PP* and *PC*. Moreover, both *CI* and *PC* include self-citations. Despite these caveats we believe that we have collected a good and unique dataset for the purposes of our study.

4.2. Statistical procedure

The dependent variable in our study, i.e., the number of citations, is a non-negative integer. This suggests the application of count-data regression methods that can account for the integer nature of the data. The benchmark parametric model for count data is the Poisson distribution (Cameron & Trivedi, 2013). The Poisson model is defined by the following equation:

$$P(n_i) = \frac{\lambda_i^{\mathrm{ni}} \exp(\lambda_i)}{n_i!} \tag{1}$$

In the above equation, $P(n_i)$ is the probability of article i having n citations, and λ_i is the expected number of citations for article i. Poisson regression models are estimated by specifying the Poisson parameter λ_i, i.e., the expected number of times article i is cited as a function of explanatory variables. The explanatory variables in our study are listed in Table 1. The most common relationship between explanatory variables and the Poisson parameter is the log-linear model (Washington, Mannering, & Karlaftis, 2010).

$$\lambda_i = \exp\left(\beta X_i\right) \text{ which is the equivalent of } \ln(\lambda_i) = \beta X_i \tag{2}$$

In the above equation, X_i is a vector of the explanatory variables of article i and β is a vector representing the parameters to be estimated.

However, a limitation of the Poisson distribution is that the variance of the data is restrained to be equal to the mean. Overdispersion (variance greater than the mean) or underdispersion (variance less than the mean) of data violates this constraint and leads to biased coefficient estimates (Shankar, Mannering, & Barfield, 1995). Because the variance of our dependent variable is greater than its mean, we had to reject the Poisson distribution. Instead, a negative binomial model was estimated because it allows the variance of the dependent variable to be larger than the mean (Washington et al., 2010). Negative binomial regression has been applied in several studies to assess the influence of various factors on citation counts (e.g. Bornmann, Schier, Marx, & Daniel, 2012; Davis, Lewenstein, Simon, Booth, & Connolly, 2008; Walters, 2006)

The negative binomial model is derived by adding an error term (ε_i) to the expected number of citations (λ_i) in Eq. (2). We then have that for each observation i

$$\lambda_i = \exp\left(\beta X_i + \varepsilon_i\right) \tag{3}$$

where $\exp(\varepsilon)$ in (3) has a gamma distribution with mean one and variance τ. The parameter τ is often referred to as the overdispersion parameter because it allows the variance of the dependent variable to be greater than its mean.

4.3. Model specification

To evaluate the impact of the different explanatory variables on an article's citation count the following negative binomial regression model is employed:

$$\begin{aligned}\lambda = CI = exp\big(&\beta_0 + \beta_{PP} PP^{\alpha} + \beta_{PC} PC + \beta_{AN} AN + \beta_{NA} NA + \beta_{AC} AC \\ &+ \beta_{RF} RF + \beta_{WT} WT + \beta_{YP} YP + \beta_{TR} TR + \beta_{PA} PA + \beta_{PB} PB + \beta_{TS} TS + \varepsilon\big)\end{aligned} \tag{4}$$

The analysis is made at article level and the symbols have the same definitions as in Table 1.

It follows from the hypotheses presented in Section 2 that articles written by experienced researchers, as measured by the number of prior publications (*PP*), will more frequently be cited than articles written by less experienced researchers. However, the return of one additional unit of experience on *CI* is expected to be higher for unexperienced researchers, i.e., with few prior publications, than for experienced researchers. Hence, we assume that *CI* increases concavely with PP; that is $(\partial CI/\partial PP) > 0$ and $(\partial^2 CI/\partial PP^2) < 0$. From (4) follows that $(\partial CI/\partial PP) = CI \times \alpha\beta_{PP} PP^{\alpha-1}$ and $(\partial^2 CI/\partial PP^2) = \alpha CI\left(\beta_{PP}(\alpha-1) PP^{\alpha-2} + \alpha\beta_{PP}^2 X_{PP}^{2\alpha-1}\right)$. Consequently, in order to meet the conditions placed on the signs of the first and second derivatives of *CI* with respect to *PP*, $\beta_{PP} > 0$ and $0 < \alpha < 1$ in Eq. (4)[1].

[1] A necessary but not sufficient condition for $(\partial^2 CI/\partial PP^2) < 0$ is that $0 < \alpha < 1$. If $\alpha > 1$, $(\partial^2 CI/\partial PP^2) > 0$.

Table 3
Results of the negative binomial regression analysis (dependent variable: citations per article, $n = 779$, $AIC = 6650$)[a].

	Coefficient	(t)
$PP^{0.261}$	0.1585***	(3.45)
PC	0.0001	(0.78)
AN	−0.0071*	(−1.68)
NA	0.0176	(0.21)
AC	0.1773**	(2.15)
RF	0.0140***	(6.78)
WT	−0.0188**	(−2.19)
YP	0.0602**	(2.55)
TR	0.3919***	(3.12)
PA	0.4915***	(4.46)
PB	0.7493***	(6.88)
TS	0.5579***	(4.76)
Constant	1.6377***	(5.39)

[a] Level of significance: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.10$ (two-tailed).

Moreover, equation (4) implies that the relationship between *CI* and *j*, $j = \{PC, AN, NA, AC, RF, WT, YP, TR, PA, PB, TS\}$ is semi-loglinear and the elasticity of *CI* with respect to *j*, EL_jCI can be written as $EL_jCI = (\partial CI/\partial j) \times (j/CI) = \beta_j \times j$, $j = \{PC, AN, NA, AC, RF, WT, YP, TR, PA, PB, TS\}$

Hence, a marginal change in *j* by one unit changes *CI* by $(\beta_j \times CI)$ units and by $(100 \times \beta_j)$%. Our earlier hypotheses imply that $\beta_{AN}, \beta_{WT}, < 0$. The remaining parameters are expected to have a positive impact on citation count (*CI*), i.e., $\beta_{PC}, \beta_{RF}, \beta_{NA}, \beta_{AC}, \beta_{YP}, \beta_{TR}, \beta_{PA}, \beta_{PB}, \beta_{TS} > 0$. The previously referred to ranking of the actual journals themselves also implies that we would expect $\beta_{PB} > \beta_{TS} > \beta_{TR} > \beta_{PA}$.

5. Estimation results

Table 3 provides results for the negative binomial regression model, i.e., of Eq. (4), using 779 articles published between 2000 and 2004 in five transportation journals. The value of α in Eq. (4) was found by letting its value run over small intervals, and then choosing the value that minimizes the value of Akaike Information Criterion (*AIC*) (Akaike, 1973). This resulted in $\alpha = 0.261$. Hence, the value of α is in line with the assumptions that we made in Section 4.3.

All estimated parameters have signs in the hypothesized directions. An increase in the number of previous publications by the authors (*PP*), the number of references in the articles (*RF*), years since publication (*YP*) and the number of countries the authors are from (*AC*) significantly increases the number of times an article is cited (*CI*). More precisely, when *RF*, *YP* and *AC* increase by one unit, the number of citations will increase by 1.4%, 6.0% and 17.7%, respectively. Moreover, the place of the first author in the alphabet (*AN*) and the number of words in the title (*WT*) reduce how frequently an article is cited; an increase in *AN* and *WT* by one unit decreases the number of times an article is cited by 0.7% and 1.9%, respectively. Finally, Table 3 shows that the articles in Transportation Research Part B (*PB*) are most frequently cited whilst those in Journal of Transport Economics and Policy (*JT*) are least frequently cited. Articles in *PB* are, on average, cited 75% more frequently than articles in *JT*. However, the previous number of times authors are cited (*PC*) does not significantly influence how often their coming articles are referred to. This may partly come from the high correlation between the number of previous publications (*PP*) and *PC* (see Table 2), and partly from the fact that many factors that only loosely signal authors skills (for example *AN* and *WT*), affect how many times an article is cited.

Let us have a closer look at the relationship between the number of times an article is cited (*CI*) and the authors' experience, as measured by their total number of previous publications (*PP*). After controlling for article, other author and journal characteristics, the results of this study confirm that the ability to conduct quality research at the frontier of a discipline increases concavely with the authors' experience. The illustrations in Figs. 1 and 2 are based on the results in Table 3. Fig. 1 shows the relationship between *CI* and *PP* and Figure shows 2 the relationship between marginal increase in *CI* $\left((\partial CI/\partial PP)\left(= \alpha\beta_{PP} PP^{\alpha-1} \times CI\right)\right)$ and *PP* when all other explanatory variables have their average values as denoted in Table 1.

Fig. 1 illustrates that approximately 10 years after being published[2], an article written by authors with no previous publications will, on average, have been cited nearly 18 times. An article written by authors who have 10 prior publications in total, can, during the same period of time, be expected to receive 24 citations, i.e., 34% more citations than articles written by only inexperienced researchers. It is evident from the figure that the predicted number of citations increases sharply at the beginning of a researcher's career before it levels off; articles written by authors with a total of 20 prior publications can, for example, expect only one more citation than articles from authors with 10 prior publications.

[2] The average age of an article in our dataset is 9.94 years (YP = 9.94).

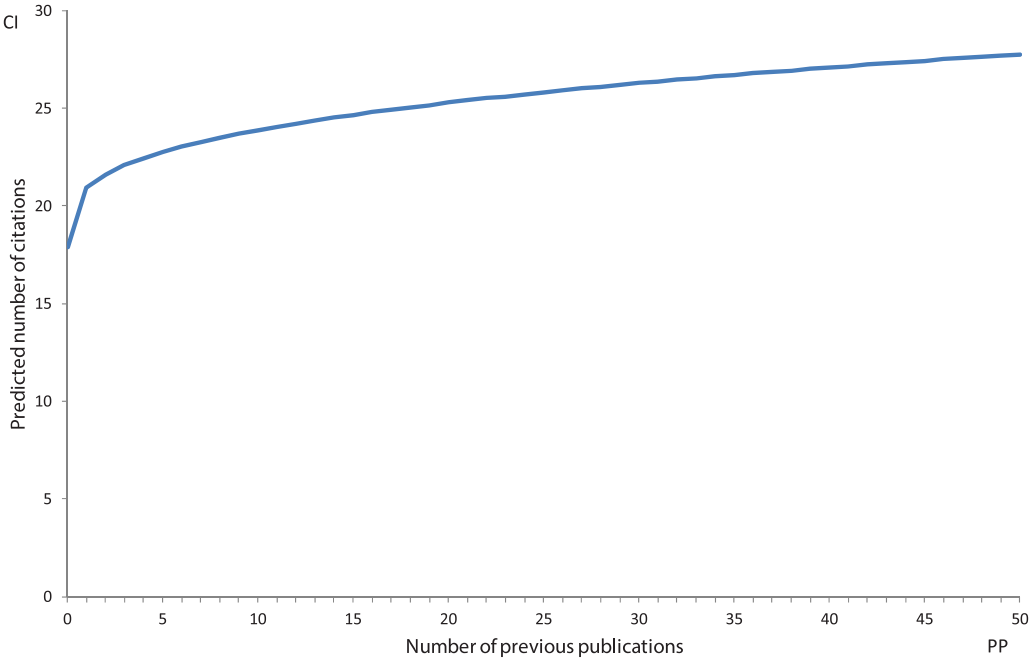

Fig. 1. The relationship between the expected number of citations and the number of previous publications, when all explanatory variables except *PP* have their average values.

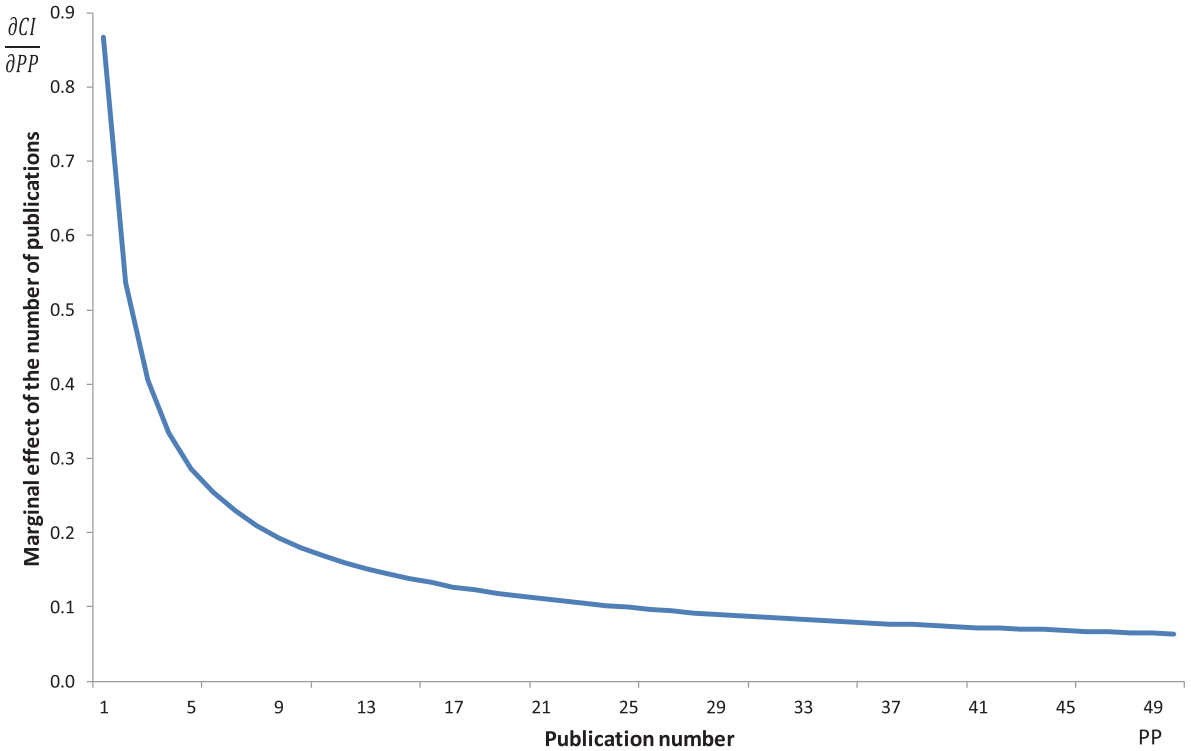

Fig. 2. Diminishing returns to scale from increased experience when all explanatory variables except *PP* have their average values.

The same thing is shown in Fig. 2 by the marginal effect on *CI* of increasing *PP*. The return to scale on researchers' skills or human capital from experience seems to diminish rapidly. The marginal effect on the number of expected citations from the tenth publication is, for example, only 21% of the marginal effect of the first publication. This suggests that researchers tend to learn the craft of research fairly quickly.

6. Concluding remarks

This study employs data from research published in five highly regarded transportation journals in order to examine how different factors influence the number of times an article is cited. The estimated results correspond, in general, to previous studies with regard to the direction and extent to which the various explanatory variables influence an article's citation count.

In this study, special focus is placed on analysing how researchers' skills, as measured by the number of times their articles are cited, are influenced by their previous experience or human capital, as measured by their total number of publications prior to the articles in question. We find that research by scientists with many prior scientific publications tends to be more frequently cited than research by inexperienced researchers. This suggests that there is a positive association between experience and the human capital of scientists. However, the effect of experience on human capital diminishes rapidly. Thus, investments in scientists' human capital through education can reduce the relatively steep learning curve currently observed among inexperienced scientists.

The above findings offer practical advice to researchers, research institutions and funders of research. Investments in human capital are primarily done through education and training (Becker, 1964). The effect of training on the human capital of scientists is confirmed in our study, which suggests that experienced scientists, to a greater degree than less experienced ones, do possess the skills and knowledge needed to conduct high quality research. However, in order to somewhat smoothen the learning curve most scientists experience in writing their first articles intended for publication in scientific journals, they could attend courses and seminars on topics such as academic writing, critical reading of academic texts and the handling of relevant ethical issues. Such training could help young researchers attain skills and knowledge that would enable them to better conduct research at the frontier of their discipline. By investing more in scientists' human capital through education in the craft of research, young researchers would not have to use their first scientific articles as part of their "education" or research training.

Research institutions benefit by having a high stock of human capital in their organizations. It is therefore in their interests to educate and train their own employees. This could be done by offering courses in the craft of research and by financing the cost of having experienced scientists review articles written by less-experienced researchers prior to submission to journals. Moreover, research institutions ought to require, or at least encourage, their PhD candidates to write their thesis as a collection of articles instead of as monographs. This provides the candidates with vital training in how scientific articles are written.

Funders of research ought to favour research teams with a high level of experience, measured by the number of prior publications, when allocating their resources. The rationale for this is that projects with participants who have a high number of publications will be more likely to produce high quality research.

Finally, it should be observed that our study – in line with all empirical studies – does have weaknesses due to erroneous data reports, as well as the validities of the variables and omitted variables. In particular, the use of citation counts as a proxy for human capital is a debatable issue. Moreover, this study has focused on articles published purely in transportation journals. It may therefore be difficult to generalize our results to other fields of research.

Despite the above limitations, this paper nevertheless offers an attempt to determine how a scientist's ability to conduct research at the frontier of a discipline, i.e., the scientist's human capital, is related to his or her of experience in writing and having published scientific articles previously.

Acknowledgements

The authors are grateful to Berner Larsen at Bodø Graduate School of Business, University of Nordland for his assistance during the preparation of this article.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. Petrov, & F. Czakil (Eds.), *Proceedings of the second international symposium on information theory* Budapest, (pp. 267–281).
Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, *12*(3), 159–170.
Azoulay, P., Stuart, T., & Wang, Y. (2014). Matthew: Effect or fable? *Management Science*, *60*(1), 92–109.
Bade, R., & Parkin, M. (2011). *Foundations of economics*. Boston: Pearson Addison-Wesley.
Beaver, D. D. (2004). Does collaborative research have greater epistemic authority? *Scientometrics*, *60*(3), 399–408.
Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis with special reference to education*. Chicago, IL: The University of Chicago Press.
Bergh, D. D., Perry, J., & Hanke, R. (2006). Some predictors of SMJ article impact. *Strategic Management Journal*, *27*(1), 81–100.
Berman, S. L., Down, J., & Hill, C. W. L. (2002). Tacit knowledge as a source of competitive advantage in the national basketball association. *Academy of Management Journal*, *45*(1), 13–31.

Blundell, R., Dearden, L., Meghir, C., & Sianesi, B. (1999). Human capital investment: The returns from education and training to the individual, the firm and the economy. *Fiscal Studies*, *20*(1), 1–23.
Blundell, R., Dearden, L., & Sianesi, B. (2005). Evaluating the effect of education on earnings: Models, methods and results from the national child development survey. *Journal of the Royal Statistical Society, Series A: Statistics in Society*, *168*(3), 473–512.
Bonaccorsi, A., & Daraio, C. (2003). Age effects in scientific productivity: The case of the Italian National Research Council (CNR). *Scientometrics*, *58*(1), 49–90.
Bornmann, L., & Daniel, H. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, *64*(1), 45–80.
Bornmann, L., Schier, H., Marx, W., & Daniel, H. D. (2012). What factors determine citation counts of publications in chemistry besides their quality? *Journal of Informetrics*, *6*(1), 11–18.
Borsuk, R. M., Budden, A. E., Leimu, R., Aarssen, L. W., & Lortie, C. J. (2009). The influence of author gender, national language and number of authors on citation rate in ecology. *Open Ecology Journal*, *2*, 25–28.
Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data*. Cambridge: Cambridge University Press.
Davis, P. M., Lewenstein, B. V., Simon, D. H., Booth, J. G., & Connolly, M. J. (2008). Open access publishing, article downloads, and citations: Randomised controlled trial. *BMJ*, *337* (Clinical research ed.)
Diamond, A. M. (1986). What is a citation worth? *The Journal of Human Resources*, *21*(2), 200–215.
Didegah, F., & Thelwall, M. (2013). Determinants of research citation impact in nanoscience and nanotechnology. *Journal of the American Society for Information Science and Technology*, *64*(5), 1055–1064.
Dimov, D. P., & Shepherd, D. A. (2005). Human capital theory and venture capital firms: Exploring home runs and strike outs. *Journal of Business Venturing*, *20*(1), 1–21.
Dries, N., Pepermans, R., & Carlier, O. (2008). Career success: Constructing a multidimensional model. *Journal of Vocational Behavior*, *73*(2), 254–267.
Flåten, O. (2013). Overvurdert humankapital og undervurdert naturkapital. *Samfunnsøkonomen*, *4*, 4–10.
Habibzadeh, F., & Yadollahie, M. (2010). Are shorter article titles more attractive for citations? Crosssectional study of 22 scientific journals. *Croatian Medical Journal*, *51*(2), 165–170.
Hanssen, T. E. S., & Jørgensen, F. (2014). Citation counts in transportation research. *European Transport Research Review*, *6*(2), 205–212.
Jamali, H. R., & Nikzad, M. (2011). Article title type and its relation with the number of downloads and citations. *Scientometrics*, *88*(2), 653–661.
Judge, T. A., Cable, D. M., Colbert, A. E., & Rynes, S. L. (2007). What causes a management article to be cited-article author or journal? *Academy of Management Journal*, *50*(3), 491–506.
Katz, J. S., & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, *40*(3), 541–554.
Krohn, G. A. (1983). Measuring the experience-productivity relationship: The case of Major League Baseball. *Journal of Business & Economic Statistics*, *1*(4), 273–279.
Lokker, C., McKibbon, K. A., McKinlay, R. J., Wilczynski, N. L., & Haynes, R. B. (2008). Prediction of citation counts for clinical articles at two years using data available within three weeks of publication: Retrospective cohort study. *BMJ*, *336*(7645), 655–657.
Loughran, T. A., Nguyen, H., Piquero, A. R., & Fagan, J. (2013). The returns to criminal capital. *American Sociological Review*, *78*(6), 925–948.
Merton, R. K. (1968). The matthew effect in science. *Science*, *159*(3810), 56–63.
Paiva, C. E., Lima, J. P. S. N., & Paiva, B. S. R. (2012). Articles with short titles describing the results are cited more often. *Clinics*, *67*(5), 509–513.
Shankar, V., Mannering, F., & Barfield, W. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis & Prevention*, *27*(3), 371–389.
Stokey, N. L. (1991). Human capital, product quality, and growth. *The Quarterly Journal of Economics*, *106*(2), 587–616.
Stremersch, S., Verniers, I., & Verhoef, P. C. (2007). The quest for citations: Drivers of article impact. *Journal of Marketing*, *71*(3), 171–193.
Stroebe, W. (2010). The Graying of Academia: Will it reduce scientific productivity? *American Psychologist*, *65*(7), 660–673.
Subotic, S., & Mukherjee, B. (2014). Short and amusing: The relationship between title characteristics, downloads, and citations in psychology articles. *Journal of Information Science*, *40*(1), 115–124.
Tregenza, T. (1997). Darwin a better name than Wallace? *Nature*, *385*(6616), 480.
Vanclay, J. K. (2013). Factors affecting citation rates in environmental science. *Journal of Informetrics*, *7*(2), 265–271.
Vieira, E. S., & Gomes, J. A. N. F. (2010). Citations to scientific articles: Its distribution and dependence on the article features. *Journal of Informetrics*, *4*(1), 1–13.
Vintzileos, A. M., & Ananth, C. V. (2010). How to write and publish an original research article. *American Journal of Obstetrics and Gynecology*, *202*(4), 344.e341–344.e346.
Walters, G. D. (2006). Predicting subsequent citations to articles published in twelve crime-psychology journals: Author impact versus journal impact. *Scientometrics*, *69*(3), 499–510.
Washington, S. P., Mannering, F. L., & Karlaftis, M. G. (2010). *Statistical and econometric methods for transportation data analysis*. Boca Raton, FL: Chapman & Hall/CRC.
Webster, G. D., Jonason, P. K., & Schember, T. O. (2009). Hot topics and popular papers in evolutionary psychology: Analyses of title words and citation counts in evolution and human behavior, 1979–2008. *Evolutionary Psychology*, *7*(3), 348–362.
Yitzhaki, M. (2002). Relation of the title length of a journal article to the length of the article. *Scientometrics*, *54*(3), 435–447.