

An explanation of the relation between the fraction of multinational publications and the fractional score of a country.

Peer-reviewed author version

EGGHE, Leo (1999) An explanation of the relation between the fraction of multinational publications and the fractional score of a country.. In: Scientometrics, 45(2). p. 291-310.

DOI: 10.1007/BF02458438

Handle: <http://hdl.handle.net/1942/791>

# **AN EXPLANATION OF THE RELATION BETWEEN THE FRACTION OF MULTINA- TIONAL PUBLICATIONS AND THE FRACTIONAL SCORE OF A COUNTRY**

by

L. EGGHE

LUC, Universitaire Campus, B-3590 Diepenbeek, Belgium<sup>1</sup>

and

UIA, Universiteitsplein 1, B-2610 Wilrijk, Belgium

---

## **ABSTRACT**

Consider a country's national output, measured by counting the number of authors from country  $c$  that collaborate in every paper in a bibliography. Depending or not that country  $c$  appears at least once in every paper, we are able to deduce the corresponding relationship between  $c$ 's fractional score and its fraction of multinational papers to which  $c$  belongs.

One of these models, a slowly decreasing concave function is completely similar to the observed relation in [Nederhof and Moed, *Scientometrics* 27(1), 39-52, 1993] between the fractionated score of a country  $c$  and its fraction of multinational papers.

The proof of the models developed here uses a stochastic property of weighting schemes, namely that the average fractional score of a country equals its total score.

**Keywords and phrases** : fractional score, total score, multinational publication.

---

<sup>1</sup>Permanent address

## **I. Introduction**

In measuring the national scientific output of a country, in a certain fixed domain, several measures need to be used. It is indeed a complicated situation since each paper in the studied set can have multiple authors and these authors can belong to several countries, some of these appearing more than once.

One class of measures describes the share of a country  $c$ , in the totality of all countries appearing in the bibliography. We have here several possible, non-equivalent measures. One way is to count every appearance of each country and to weight it with full credit 1. This is called the total scoring system (for formulae, see further). Another way is to credit each appearance in a paper with  $a$  authors by the number  $\frac{1}{a}$ . This is called the fractional scoring system. One could even give scores dependent on the rank of each author in a multi-authored paper in the sense that the smaller the rank, the higher the score (hence giving the highest score to the first author - or his/her country). This is called a proportional scoring system. An extreme case of the latter is the first author scoring system, where only the first author (or his/her country) receives a score (of 1). Since the other authors (and hence their countries) receive a score of 0, the system does not describe the collaboration in the bibliography very well. We will not use it further on and refer to Cole and Cole (1973) and to Schubert, Glänzel and Braun (1989) for its use and for criticism on this use. For more on the usability of total, fractional or proportional scores we refer the reader to Egghe and Rousseau (1990), Price (1981), Van Hooydonk (1997) and to Egghe, Rousseau and Van Hooydonk (1998). In the latter paper the differences between these scoring methods are studied.

In Nederhof and Moed (1993) another scoring method for countries is used, the so-called fractionated scoring method. Here a country  $c$  receives a score  $\frac{1}{b}$  in a paper if  $b$  is the total number of different countries in this paper and if  $b \neq 1$ . If  $b=1$  the fractionated score is 0. They further discuss the pros and cons of this method and investigate the relation between a country's fractionated score and its fraction of multinational papers to which this country belongs. We say that country  $c$  has a multinational paper if there exists at least one country

$c' \neq c$  such that this paper is co-authored by authors from  $c$  and  $c'$  (at least). The relation obtained by Nederhof and Moed (1993) is expressed in their Figure 1 (p. 45) and seems to be decreasing very slowly and in a concave way (but close to linear). The figure is reproduced here, with permission (Fig. 1)

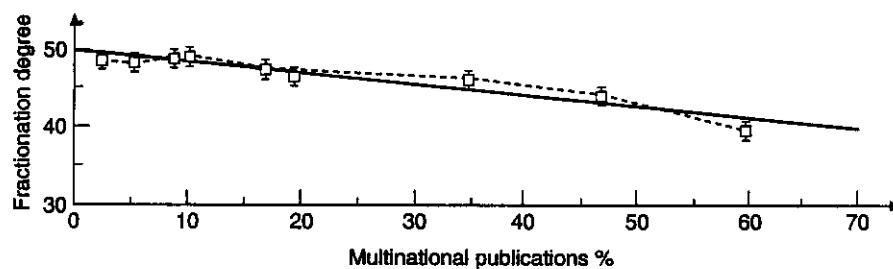


Fig. 1 The Nederhof-Moed experimental relation between the fractionated score and the fraction of multinational publications of a country.

The object of our study is to find an explanation of the graph in this Figure. We will, however, study a slightly different problem as follows : instead of studying fractionated scores we will study fractional scores. The reason is simple : we were not able to explain Fig. 1 as such but we reached a result where the vertical axis denotes fractional degree instead of fractionated degree. It is all right to do so, first of all since fractional scores are also interesting to compare with fractions of multinational publications and, secondly we conjecture that the result will be similar, i.e. the obtained curve could be very similar to the one of Fig. 1.

One must, however, bear in mind that fractional scores in case only one country  $c$  appears have a value 1 while, according to the definition above, the fractionated score can be - at most -  $\frac{1}{2}$ . So we will have an ordinate axis with values between 0 and 1 (Fig. 1 has ordinate (possible) values between 0% and 50% (i.e.  $\frac{1}{2}$ ). Another remark must be made : in Nederhof and Moed (1993) one fixes a country  $c$  and one only looks at papers in which this country  $c$  appears at least once. In a general framework of a bibliography, however, one can also have that  $c$  does not belong to a paper (i.e. there are no co-authors from this

country  $c$  in this paper). We will work out both models : first the general case, where papers exist in which  $c$  does not appear : in this case we will find a model that is different from the Nederhof-Moed graph. Then we will use conditional expectations to see how this model changes if we presuppose that  $c$  belongs to each paper at least once. In this case we will recover the Nederhof-Moed graph completely. These results will be given in section III.

In section II we prove some preliminary results on total, fractional and proportional scores that will enable us to use, henceforth in section III, only total scores instead of fractional ones. This will simplify the arguments given in section III.

## **II. Total, fractional and proportional scores and their expectations.**

We will fix our framework now. We suppose we have a general bibliography where we consider authors per article and the countries to which they belong. We assume that we can determine unambiguously, for each author, the corresponding country. We henceforth will only consider "countries" since this is more general than studying "authors" : the author-case can be deduced from the country-case by assuming - as a special case - that each "country" appears (maximally) only once in each article.

We denote by  $N$  the number of articles in the bibliography and for each  $i \in \{1, \dots, N\}$  by  $a_i$  the number of authors in article  $i$ , amongst  $a_i(c)$  are from country  $c$ .

The total scoring system (T) gives a score 1 for every appearance, hence country  $c$  scores  $a_i(c)$  in paper  $i$ , yielding a total score of  $\sum_{i=1}^N a_i(c)$ . For reasons of comparing this with other scoring devices (see further) we are only interested in relative scores. Since, in total, there are  $\sum_{i=1}^N a_i$  scores to receive we have that the relative total score of  $c$  equals

$$Q_T(c) = \frac{\sum_{i=1}^N a_i(c)}{\sum_{i=1}^N a_i} \quad (1)$$

The fractional scoring system (F) gives a score  $\frac{1}{a_i}$  for every author of country  $c$  in paper  $i$ , hence a score  $\frac{a_i(c)}{a_i}$  for country  $c$  in paper  $i$ .

Since the total score of the system obviously is  $N$  we have that the relative fractional score of  $c$  equals

$$Q_F(c) = \frac{1}{N} \sum_{i=1}^N \frac{a_i(c)}{a_i} \quad (2)$$

It is well-known (see Egghe, Rousseau and Van Hooydonk (1998)) that country-rankings according to (1) are very different from country-rankings according to (2). Yet, in this paper we will be able to relate (1) and (2) in a way we can use it in our last section. This relationship will be given in this section.

The proportional scoring system (P) is - scientometrically speaking - less important, although it has been defined in Van Hooydonk (1997). We include it here because of its mathematical interest and its relation with the total and fractional scoring systems.

The proportional scoring system - as the fractional one - gives a weight 1 per paper but divides this "1" in a different way : instead of giving each author in paper  $i$  a score  $\frac{1}{a_i}$  (as is so in the fractional case) we look at the rank  $R$  of each author, thereby giving an author on rank  $R \in \{1, \dots, a_i\}$  a score

$$\frac{2}{a_i} \left( 1 - \frac{R}{a_i + 1} \right). \quad (3)$$

It can readily be seen (cf. also Egghe, Rousseau and Van Hooydonk (1998)) that the relative proportional score of country  $c$  equals

$$Q_P(c) = \frac{1}{N} \sum_{i=1}^N \frac{2}{a_i} \left( a_i(c) - \frac{R(i,c)}{a_i + 1} \right) \quad (4)$$

$$Q_P(c) = 2Q_F(c) - \frac{1}{N} \sum_{i=1}^N \frac{2R(i,c)}{a_i(a_i + 1)} \quad (5)$$

where  $R(i,c)$  denotes the sum of all ranks occupied by country  $c$  in paper  $i$ .

System (T) is the simplest from a probabilistic point of view : given  $N, a_1, \dots, a_N$ , we only need to know  $\sum_{i=1}^N a_i(c)$ , while for sytem (F) we need to know  $a_i(c)$  (here we fix the country  $c$ ). In other words for (T) we only need to know the total appearance of  $c$  in the bibliography, while for (F) we need to know how many times  $c$  appears in each article. (F) implies (T) uniquely but, per (T), there are many different (F)-situations, and hence also different values of  $Q_F(c)$ , per fixed  $Q_T(c)$ . We can now ask the question : what is the average of all the possible  $Q_F(c)$ -values, given the fixed  $Q_T(c)$ -value. We will denote this (conditional) expectation by  $E_T(Q_F(c))$ . We have the following result.

**Theorem 1** :  $E_T(Q_F(c)) = Q_T(c)$ , for every country  $c$ .

**Proof** : We have (using linearity of conditional expectations)

$$\begin{aligned} E_T(Q_F(c)) &= E_T\left(\frac{1}{N} \sum_{i=1}^N \frac{a_i(c)}{a_i}\right) \\ &= \frac{1}{N} \sum_{i=1}^N E_T\left(\frac{a_i(c)}{a_i}\right), \end{aligned}$$

where  $E_T\left(\frac{a_i(c)}{a_i}\right)$  denotes the average of all values  $\frac{a_i(c)}{a_i}$ , given the fixed (T)-situation (and hence the fixed  $Q_T(c)$ ). This average - of course - is weighted according to the different probabilities of occurrence of the  $\frac{a_i(c)}{a_i}$ . Given  $Q_T(c)$  as the overall probability for  $c$  to occupy one of the  $\sum_{i=1}^N a_i$  positions, the probability to have  $a_i(c)$  occurrences of  $c$  in  $a_i$  places of publication  $i$ , is

$$\binom{a_i}{a_i(c)} Q_T(c)^{a_i(c)} (1 - Q_T(c))^{a_i - a_i(c)} \quad (6)$$

Hence (since  $0 \leq a_i(c) \leq a_i$ )

$$\begin{aligned}
& E_T \left( \frac{a_i(c)}{a_i} \right) \\
&= \sum_{a_i(c)=0}^{a_i} \frac{a_i(c)}{a_i} \binom{a_i}{a_i(c)} Q_T(c)^{a_i(c)} (1-Q_T(c))^{a_i-a_i(c)}
\end{aligned} \tag{7}$$

But, as is easily seen :

$$a_i(c) \binom{a_i}{a_i(c)} = a_i \binom{a_i-1}{a_i(c)-1}.$$

So (remarking that in (7) the term for  $a_i(c)=0$  is zero)

$$\begin{aligned}
& E_T(Q_F(c)) \\
&= \frac{1}{N} \sum_{i=1}^N Q_T(c) \sum_{a_i(c)=1}^{a_i} \binom{a_i-1}{a_i(c)-1} Q_T(c)^{a_i(c)-1} (1-Q_T(c))^{(a_i-1)-(a_i(c)-1)},
\end{aligned}$$

the last sum equalling 1, being the total chance of a binomial distribution.

Consequently :

$$E_T(Q_F(c)) = Q_T(c). \quad \square \tag{8}$$

Now comparing the fractional and proportional systems we can remark that the latter one is finer than the former. Indeed : given the  $a_i(c)$  for all  $i$  ( $c$  fixed) there are several possibilities for the ranks that  $c$  in article  $i$  occupies, all leading to different (P)-scores. Conversely, given the ranks of  $c$  in every paper  $i$ , we can easily determine  $a_i(c)$  and hence the (F)-situation.

So we can ask the following question, similar to the one asked above : given a fixed (F)-situation, hence a fixed  $Q_F(c)$ , what is the average of the  $Q_P(c)$ -values that agree with  $Q_F(c)$ ? We denote this by  $E_F(Q_P(c))$ . We have the following result.



**Theorem 2** :  $E_F(Q_P(c)) = Q_F(c)$  for every country  $c$ .

**Proof** : Equation (5) gives

$$E_F(Q_P(c)) = 2Q_F(c) - \frac{1}{N} \sum_{i=1}^N \frac{2E_F(R(i,c))}{a_i(a_i+1)}, \quad (9)$$

using linearity of conditional expectations and the fact that  $Q_F(c)$  is fixed here. Of course,  $E_F(R(i,c))$  denotes the average of the  $R(i,c)$ , given the fixed (F)-situation. For this it suffices to prove that, for every situation yielding  $R(i,c)$ ,  $\exists!$  situation  $R'(i,c)$  such that

$$\frac{R(i,c) + R'(i,c)}{2} = a_i(c) \frac{a_i+1}{2}. \quad (10)$$

Since the right hand side of (10) is fixed here, we have that the overall average

$$E_F(R(i,c)) = a_i(c) \frac{a_i+1}{2} \quad (11)$$

from which (9) yields

$$E_F(Q_P(c)) = 2Q_F(c) - \frac{1}{N} \sum_{i=1}^N \frac{a_i(c)}{a_i} \quad (12)$$

So all that needs to be proved is that (10) is true. So take any (P)-situation. We construct the following second (P)-situation : we mirror the first one in the following way : if a country  $c'$  occupies a place  $r \in \{1, \dots, a_i\}$  in the first situation, we let  $c'$  occupy rank  $a_i - r + 1$  in the second situation. This also goes for  $c' = c$ . Denoting by  $R'(i,c)$  the sum of the ranks occupied by  $c$  in the second situation, we have

$$R'(i,c) = \sum_{r \in A(i,c)} (a_i - r + 1),$$

where  $A(i,c) \subset \{1, \dots, a_i\}$  is the set of ranks that  $c$  occupies in the first situation. Hence

$$\begin{aligned} R'(i,c) &= (a_i+1) \sum_{r \in A(i,c)} 1 - \sum_{r \in A(i,c)} r \\ &= (a_i+1)a_i(c) - R(i,c), \end{aligned}$$

from which (10) follows.  $\square$

Corollary :  $E_T(Q_P(c)) = Q_T(c)$  (13)

for every country  $c$ .

Proof : By the above theorems we have

$$\begin{aligned} E_T(Q_P(c)) &= E_T(E_F(Q_P(c))) \\ &= E_T(Q_F(c)) \\ &= Q_T(c). \end{aligned}$$

Here we use the fact that  $E_T \circ E_F = E_T$  since  $(T)$  is a rougher situation than  $(F)$  and by the fact that conditional expectations are unique (see e.g. Chow and Teicher (1978), p. 200, theorem 1).  $\square$

Equations (8) and (13) allow us to use  $Q_T(c)$  when we want to prove regularities involving  $E_T(Q_F(c))$  or  $E_T(Q_P(c))$ . More concretely, in the next section we will investigate the relationship between the fraction of multinational publications of a country and its average relative fractional score. For the latter we will hence use  $Q_T(c)$ , which will simplify the

arguments considerably. We are convinced that the above remark will also be applicable in other (scientometric) situations.

Note : It is easy to see that nor the first author count system, neither the fractionated author count system have hierarchical relations with (T), (F) or (P)-systems, in the sense discussed above.

### **III. The relation between the fraction of multinational publications and the fractional score of a country.**

As mentioned in the introduction we will study two cases

- (i) General case : a general bibliography where a fixed country  $c$  might or might not appear as citizenship of one of the authors of an article
- (ii) Conditional version of (i) : we only look at the subcollection of the bibliography for which a fixed country  $c$  appears at least once in every article (Nederhof-Moed case).

#### **III.1 General case.**

It is underlined that this subsection is devoted to the case that a country  $c$  might or might not appear as citizenship of one of the authors of an article, i.e. a case not considered by Nederhof and Moed (In Nederhof and Moed (1993) one only considers case (ii) above, with which we will deal in subsection III.2).

We fix the bibliography and a country  $c$ . Hence  $Q_T(c)$  is the overall probability that an author from country  $c$  occupies one of the available places in article  $i$

$$\underbrace{x \ x \ x \ \dots \ x}_{a_i}$$

(for each  $i \in \{1, \dots, N\}$ , where paper  $i$  has  $a_i$  authors).

The fraction  $f_c$  of multinational publications of  $c$  (i.e. the fraction of papers in which  $c$  and at least one other country  $c' \neq c$  appears) is one minus the fraction  $s_c$  of papers in which only  $c$  appears minus the fraction  $n_c$  of papers in which  $c$  does not appear :

$$f_c = 1 - s_c - n_c \quad (14)$$

It is now clear that paper  $i$  (with  $a_i$  co-authors) has a probability  $Q_T(c)^{a_i}$  to have only authors from country  $c$  and has a probability  $(1 - Q_T(c))^{a_i}$  to have no authors from country  $c$ . Hence, over the complete bibliography, (14) yields

$$f_c = 1 - \frac{1}{N} \sum_{i=1}^N [Q_T(c)^{a_i} + (1 - Q_T(c))^{a_i}] \quad (15)$$

$$f_c = \frac{1}{N} \sum_{i=1}^N [1 - Q_T(c)^{a_i} - (1 - Q_T(c))^{a_i}] \quad (16)$$

This function is hence an average of functions of the type

$$y = 1 - x^a - (1 - x)^a, \quad (17)$$

where  $x \in [0, 1]$  and  $a \in \mathbb{N}$ ,  $a$  being the number of authors in a paper. The graphs of (17) all have the same form as in Fig. 2, hence this is also the graph of (16). Note that we graphed  $Q_T(c)$  versus  $f_c$  since this was also the case in the Nederhof-Moed graph (Fig. 1). The higher the  $a$ 's, the higher the value of the top of  $y(x)$ , but  $y(x)$ , hence also  $f_c$ , will never be 1.

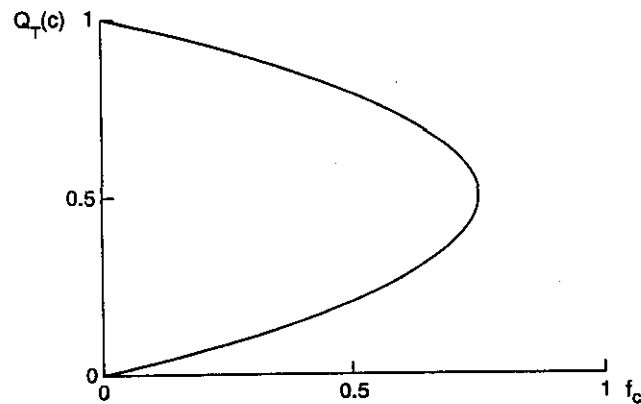


Fig. 2. Relation between  $Q_T(c)$  and  $f_c$  in the general case.

The model (16) reveals that  $f_c$  is low for low and high values of  $Q_T(c)$ . This is easily understood : if  $Q_T(c)$  is low,  $c$  does not belong to many papers (a case that we will exclude in the next part) and for this reason the overall fraction of multinational papers of  $c$  is low. In case  $Q_T(c)$  is high, many papers are written by authors from  $c$  only so that, again, the fraction of multinational papers of  $f_c$  is low. In a way we could already predict that, excluding papers where  $c$  does not appear will lead to the upper half of the graph of Fig. 2 and this is a slowly decreasing concave function, resembling Fig. 1. We will study this case in a more exact way in the next subsection.

The maximum of  $f_c$  is found for  $Q_T(c) = \frac{1}{2}$  as is easily seen. This maximal value is

$$f_c \left( \frac{1}{2} \right) = 1 - \frac{2}{N} \sum_{i=1}^N \left( \frac{1}{2} \right)^{a_i} < 1 \quad (18)$$

The higher the  $a_i$ , the higher  $f_c \left( \frac{1}{2} \right)$  but this is true for every value of  $Q_T(c)$  (see formula (16)), a logical fact : the more co-authors that are available, given fixed  $Q_T(c)$ , the higher chance to have multinational papers.

Note again that  $Q_T(c) = E_T(Q_F(c))$  so that Fig. 2 gives also the relation between  $f_c$  and the expected fractional score of country  $c$ .

### III.2. The case that country $c$ appears at least once in every paper.

Since  $Q_T(c)$  is the overall probability for country  $c$  to appear as citizenship of one of the authors of one of the papers in the entire bibliography, we cannot exclude that  $c$  does not appear as such in some papers. But we can look at the part of the bibliography where every paper has at least one author from  $c$ . In this part we can apply conditional probabilities to study then the fraction of multinational papers of  $c$  (see Fig. 3).

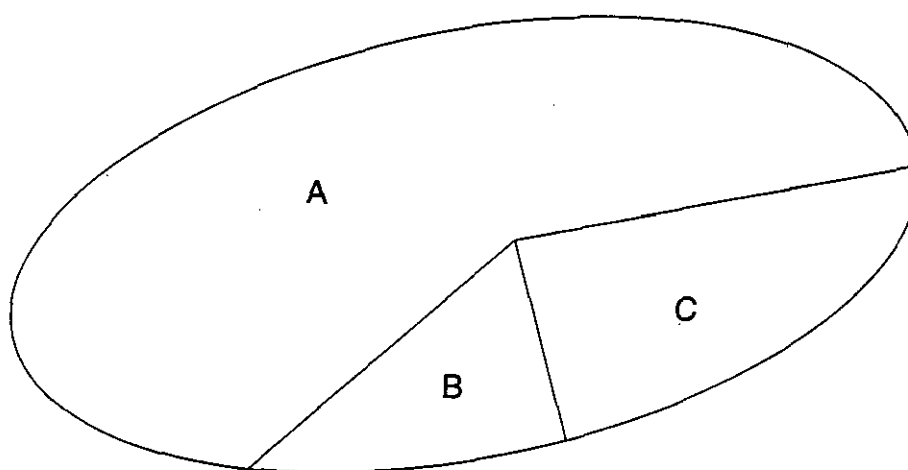


Fig. 3. Three parts of a bibliography

A : Every paper in A is a multinational paper of  $c$  :  $c$  appears as citizenship of one of the authors and at least one author belongs to another country.

B : Every author in every paper of B belongs to country  $c$ .

C : No author from a paper in C belongs to country  $c$ .

In the previous part  $f_c$  was  $P(A)$ , now we study

$$f_c^* = P(A|A \cup B), \quad (19)$$

the conditional probability of A, given that we are in  $A \cup B$ , hence (equivalently) not in C.

We hence have now

$$f_c^* = \frac{P(A)}{P(A \cup B)}. \quad (20)$$

By using the same argument as in subsection III.1, we now have (since  $P(A \cup B) = 1 - P(C)$ ) :

$$f_c^* = \frac{1 - \frac{1}{N} \sum_{i=1}^N [Q_T(c)^{a_i} + (1-Q_T(c))^{a_i}]}{1 - \frac{1}{N} \sum_{i=1}^N (1-Q_T(c))^{a_i}}$$

$$f_c^* = \frac{\sum_{i=1}^N [1-Q_T(c)^{a_i} - (1-Q_T(c))^{a_i}]}{\sum_{i=1}^N [1-(1-Q_T(c))^{a_i}]} \quad (21)$$

Of course, the exact shape of  $f_c^*$  in function of  $Q_T(c)$  depends on the values of  $a_1, \dots, a_N$ . We have checked several cases and we always find the following relationship : slowly decreasing in a concave way. In fact all shapes are similar to the simple functions

$$y = \frac{1 - x^a - (1-x)^a}{1 - (1-x)^a}$$

$$y = 1 - \frac{x^a}{1-(1-x)^a} \quad (22)$$

for  $a \in \mathbb{N}$  and  $x \in [0,1]$ . They all have graphs as in Fig. 4 (all concavely slowly decreasing) (except if  $a=1$  ; then  $y=0$ , obviously). If  $a$  increases, then the decrease of  $y$  becomes slower.

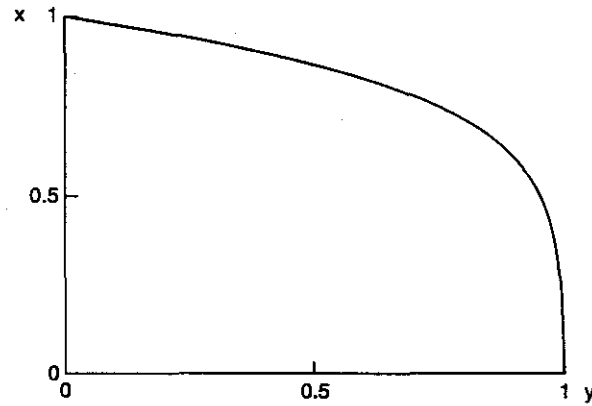


Fig. 4. Graph of the functions of the type (21) ( $a > 1$ ).

The only difference between the graphs of (21) and the type of graph depicted by Fig. 4 is the fact that  $f_c^*$  does not go to 1 if  $Q_T(c) \rightarrow 0$ . This is due to the fact that in (21), some  $a_i$ s can be 1.

In fact we can easily prove the following result :

$$\lim_{Q_T(c) \rightarrow 0} f_c^* = \frac{\sum_{i=1}^N a_i}{\sum_{i=1}^N a_i} \leq 1, \quad (23)$$

a logical result. Note also that this value increases if the  $a_i$  increase. Only if no  $a_i=1$  then this limiting value is 1, again a logical result.

Formula (23) represents the value of the absolute maximum of the fraction of multinational publications of a country  $c$ , given  $a_1, \dots, a_N$ . In the appendix we present some concrete graphs of  $f_c^*$  versus  $Q_T(c)$  as given by formula (21). They all are similar to the graph in Fig. 5.

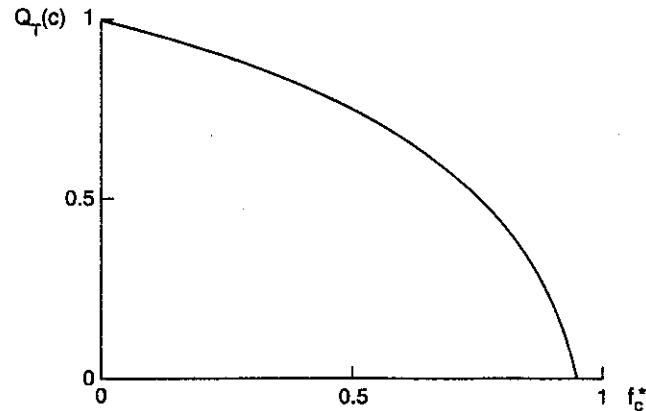


Fig. 5. General form of the graph of  $Q_T(c)$  in function of  $f_c^*$ .

Note again that  $E_T(Q_F(c)) = Q_T(c)$  so that (20) and Fig. 4 also shows the relation between the expected fractional score of a country  $c$  and its fraction of multinational publications. The



resemblance with Fig. 1 is so high that we can say that we have provided a (first) explanation of the Nederhof-Moed findings (albeit we could not work with fractionated scores - this is left as an open problem).

**Remark** (as suggested by one of the referees)

The model, as derived in section III.2 deals with the situation  $Q_T(c)$  versus  $f_c^*$  (in words : the total score of a country versus its fraction of multinational publications). As explained in theorem 1,  $Q_T(c)$  is the average of all fractional scores (with the same fixed  $Q_T(c)$ -score). This means that graphs as in Fig. 5 are "averaged" over clouds of points where the relation between  $Q_F(c)$  and  $f_c^*$  is exhibited. It is our point that such averaged curves resemble very much the original ones (although in the latter, much more irregularity can occur).

Finally, as agreed by the referee, we assume that graphs of the relationship between the fractional score of a country and its fraction of multinational publications resembles the graphs of the similar relation, where we replace "fractional score" by "fractionated score".

For these reasons we say that model (21) is a (first, simplified) explanation of the Nederhof-Moed regularity as shown in Fig. 1. Hereby, of course, we leave open the full explanation of Fig. 1 itself.

## **References**

- Chow, Y.S. and Teicher, H. (1978). Probability Theory. Independence, Interchangeability, Martingales. New York : Springer Verlag.
- Cole, J.R. and Cole, S. (1973). Social Stratification in Science. Chicago : the University of Chicago Press.
- Egghe, L. and Rousseau, R. (1990). Introduction to Informetrics. Quantitative Methods in Library, Documentation and Information Science. Amserdam : Elsevier.
- Egghe, L., Rousseau, R. and Van Hooydonk, G. (1998). Methods for accrediting publications to authors or countries : consequences for evaluation studies. Preprint.
- Nederhof, A.J. and Moed, H.F. (1993). Modeling multinational publication : Development of an on-line fractionation approach to measure national scientific output. *Scientometrics* 27(1), 39-52.
- Price, D. De Solla (1981). Letter to the editor. *Science*, 212, 987.
- Schubert, A., Glänzel, W. and Braun, T. (1989). Scientometric datafiles. A comprehensive set of indicators on 2649 journals and 96 countries in all major fields and subfields 1981-1985. *Scientometrics* 16, 3-478.
- Van Hooydonk, G. (1997). Fractional counting of multi-authored publications : consequences for the impact of authors. *Journal of the American Society for Information Science* 48, 944-945.

## Appendix

Below we present several cases of application of formula (21). It suffices to know the proportional values of the number of times an article has 1,2,3,... authors. We present several cases, applicable to the different sciences (less or more co-authors). From the examples below it will be clear how to proceed with other coefficients. Note that they are easily determined, given a concrete set of data.

We restrict our attention to  $a_i \leq 6$  although other extensions are possible of course.

### Case I

$a_i$	weight
1	2
2	1
3	0.5
4	0.4
5	0.3
6	0.2

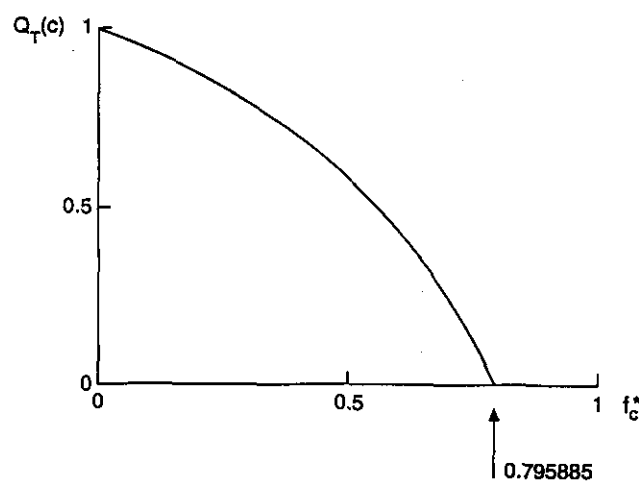
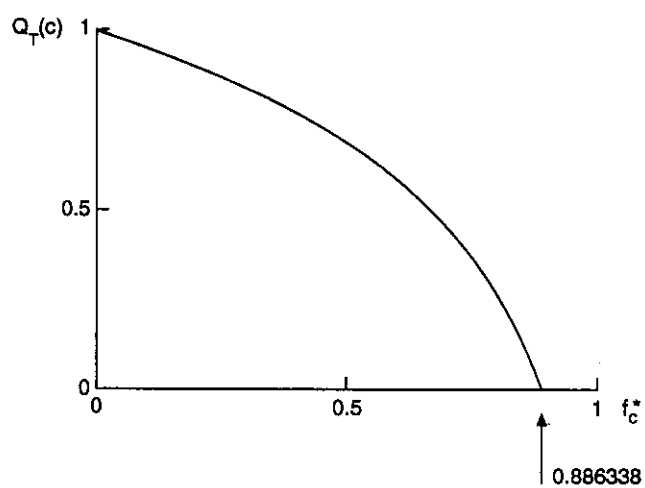


Fig. 6.  $Q_T(c)$  versus  $f_c^*$  in case I.

**Case II**

$a_i$	weight
1	1
2	1
3	0.5
4	0.4
5	0.3
6	0.2

Fig. 7.  $Q_T(c)$  versus  $f_c^*$  in case II.**Case III**

$a_i$	weight
1	0.5
2	1
3	0.5
4	0.4
5	0.3
6	0.2

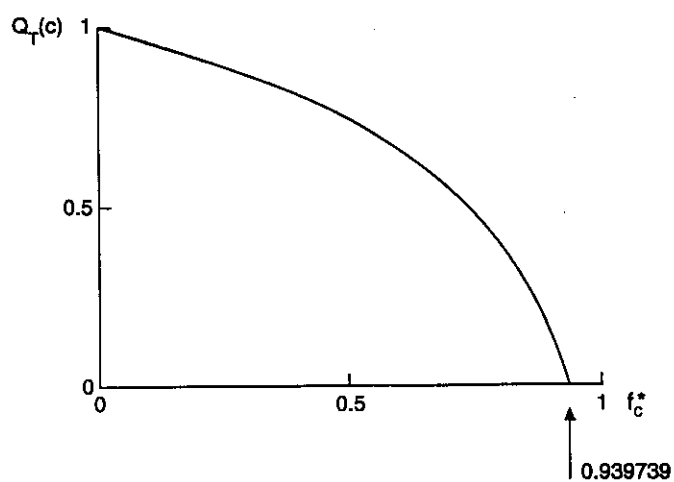


Fig. 8.  $Q_T(c)$  versus  $f_c^*$  in case III.

#### Case IV

$a_i$	weight
1	0.5
2	1
3	1
4	0.4
5	0.3
6	0.2

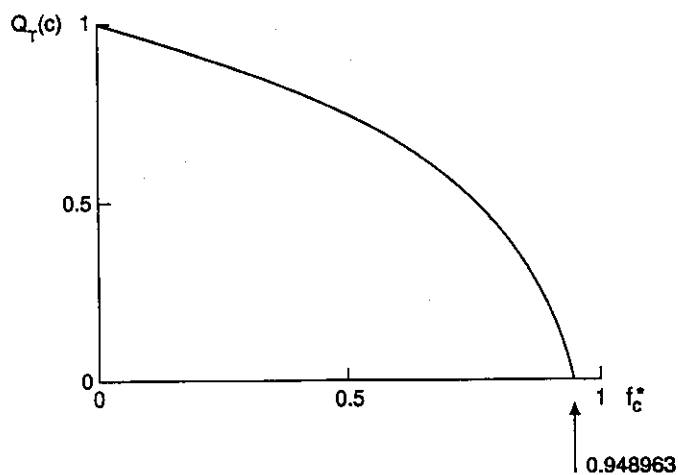
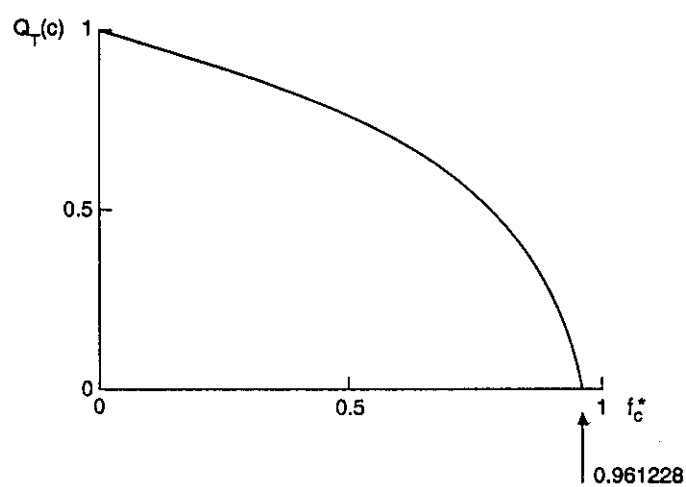


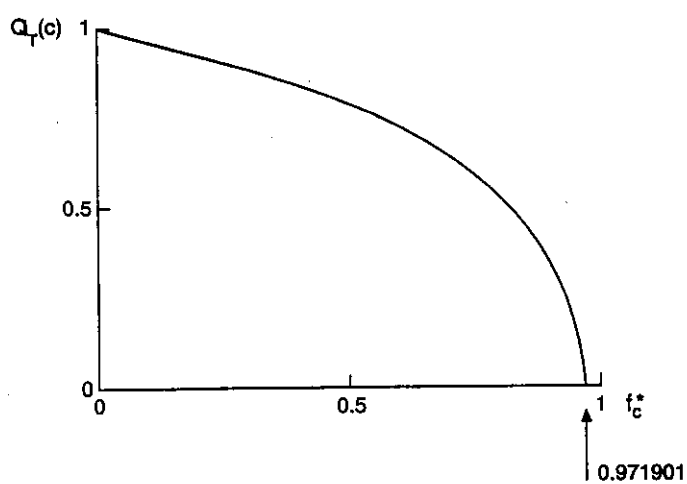
Fig. 9.  $Q_T(c)$  versus  $f_c^*$  in case IV.

Case V

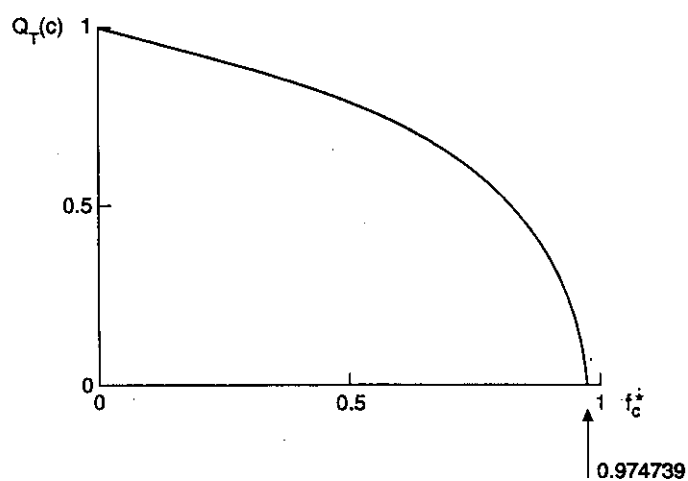
$a_i$	weight
1	0.5
2	1
3	1.5
4	0.8
5	0.3
6	0.2

Fig. 10.  $Q_T(c)$  versus  $f_c^*$  in case V.Case VI

$a_i$	weight
1	0.5
2	1
3	1.5
4	1.5
5	0.6
6	0.3

Fig. 11.  $Q_T(c)$  versus  $f_c^*$  in case VI.**Case VII**

$a_i$	weight
1	0.5
2	1
3	2
4	1.5
5	0.7
6	0.3

Fig. 12.  $Q_T(c)$  versus  $f_c^*$  in case VII.