

## A COMMENT: THE “COPYRIGHT FACTORS”

RICHARD WATT

### 1. INTRODUCTION

Economists (and policy makers) have shown much interest in attempting to calculate the fraction of total GDP that can be directly attributed to one particular aspect of the economy. For example, this symposium is concerned with estimating the contribution to GDP that can be attributed to copyright. Typically, such studies are based on the following methodology. Let  $x_i$  be the total value added of industrial sector, or *activity*,  $i$ . Assuming that there are  $n$  activities in all, total GDP is then given by  $X = \sum_{i=1}^n x_i$ . Now, define a vector of “copyright factors”, or weights,  $w = (w_1, w_2, \dots, w_n)$ , where the weight  $w_i$  is intended to represent that fraction of value added of activity  $i$  that is directly dependent upon copyright. Note that we must have  $0 \leq w_i \leq 1$  for all  $i$ , and that the total sum of the weights has no specific meaning. Given this, the total contribution to GDP of copyright (the total value added in the economy that is directly dependent upon copyright) is given by  $A \equiv \sum_{i=1}^n w_i x_i$ . In order that studies be comparable, both over countries and over time, this contribution is often expressed in relative terms, that is, the objective variable is  $Y \equiv \frac{A}{X}$ .

Clearly, this methodology is heavily dependent upon the subjective estimates that are present in the vector of weights. The greater are these weights, the greater will be the final contribution value that is calculated. Aside from pure data gathering, I would suggest that the task of assigning the appropriate weights to each activity is far and away the most important one that must be faced by any researcher who attempts to calculate the national contribution of copyright. In spite of this, this particular issue has not received a great deal of attention in the final reports of the studies that have been done to date, or in the recent WIPO Guide book on this subject.

### 2. GROUPINGS OF ACTIVITIES

In all of the studies on the contribution of copyright done up to now, the total set of activities is split into a reduced set of groups; the “core” activities, “related” activities (often sub-divided further into groups like “partial”, “dependent”, or “distribution” etc.), and “unrelated” activities. In terms of the weights, the activities in the core are assigned a weight of 1, and the unrelated activities are assigned a weight of 0. The weight to be assigned to each related activity is strictly between 0 and 1, and the exact values of these weights are subject to the assumptions, and guesswork, of the researcher. The final result of the national contribution of copyright does not really depend upon how activities are grouped, but rather upon the weights that are assigned to each activity within each group. However, the efforts that have been made by previous studies and in the WIPO *Guide* in suggesting exactly how each particular activity should be grouped, are invaluable for researchers

since it helps out considerably with the final assignment of weights to each activity, at least as far as the extremes go. In particular, we can safely assume that activities that should be assigned to groups of lower dependence should get a lower copyright factor.

It would seem that the choice of exactly which activities should receive a weighting of 1 (i.e. which activities constitute the core) is a relatively uncomplicated matter, especially with the help of previous studies and the *WIPO Guide*. The problems of weight assignment really begin when one considers the exact choice of where related activities stop, and where unrelated activities begin. In reality, it is extremely hard (perhaps impossible) to think of an economic activity that is entirely unrelated to copyrighted products. If one existed, it would have to have no dependence at all on any form of software or computing in general, digital information management, the written word, and publicity.<sup>1</sup> However, surely some activities exist that should be assigned a copyright factor that may be arbitrarily close to 0. Thus the unrelated group of activities would be those whose dependence upon copyright is so close to 0 that no significant imprecisions are introduced by assuming that it is indeed 0.<sup>2</sup> Thus, it would seem that perhaps the most important task is in assigning the weights to the related activities.

All of the national studies that have been made, as well as the *WIPO Guide*, take meticulous steps in identifying the activities to be included within each group, normally according to some sort of description of their dependence upon copyrighted products that groups them naturally. However, they then go on to establish a one-to-one correspondence between the groups and the size of the copyright factors that are assigned, and so each group could also just as easily be defined according to the copyright factor, or weighting, that is to be associated with each. The group of core activities, for example, usually defined as being those activities whose primary, or only, activity is directly related to copyrighted products, could just as easily be defined as “those activities whose copyright factor is equal to 1”. Unrelated activities would be defined as those with a copyright factor of 0, and related activities would be those that have a copyright factor that is strictly between 0 and 1.

In spite of economists’ natural aversion to counter-factual analyses, we can imagine our activity groupings in the following way. Firstly, we define a complete list of all products that are copyright protected. Then we imagine that all of these products somehow disappear. The core activities can be thought of as those activities that would disappear completely along with copyright products. The unrelated activities would be those that are unaffected by the disappearance of copyright products, and the related activities would be those that are negatively affected, but not completely annihilated.

In this comment, I will suggest that the validity of the final number that is calculated as the total national contribution of copyright to GDP, in terms of exactly how reasonable is the choice of weights for the related activities, can be checked with a

---

<sup>1</sup>At one time, I thought that an activity like “shepherding” could qualify as being unrelated to copyright. However, recently I saw a television documentary in which a shepherd stated that unless he could spend his time out on the pastures reading novels, he would certainly abandon the profession due to outright boredom. Given that testimony, one can formulate a good argument for shepherding to be included in the copyright core!

<sup>2</sup>In any case, if a marginal activity is counted as unrelated rather than related but with a tiny copyright factor, the effect upon the final national contribution figure will be insignificant.

very simple calculation that gives a first approximation to the correct value of contribution, although it does imply a particular definition of exactly what constitutes the related activities.

### 3. A SIMPLE LOWER BOUND ON THE OBJECTIVE VARIABLE

As has been mentioned above, in order to reduce the tedious task of estimating exact weights, researchers have often used the simplified methodology of dividing all activities into 3 mutually exclusive groups; activities that are totally dependent upon copyright, activities that are partially dependent upon, or related to, copyright, and activities that are independent of, or unrelated to, copyright. Then, each activity in the first group is assigned a weight of 1, and each activity in the third group is assigned a weight of 0. In this way, only the groupings themselves and the weights to be assigned to the related activity groups need to be dealt with.

Using this methodology, if we imagine that the  $n$  activities are ordered according to the estimate of their true relative dependence upon copyright, then the  $n$  activities in total can be partitioned as follows; activities 1 to  $h$  are fully dependent (or core activities), activities  $h + 1$  to  $k$  are partially dependent (or related), and activities  $k + 1$  to  $n$  are independent (or unrelated). In this way, any given partition into the three groups of activities is fully defined by the twosome  $(h, k)$ . The vector of weights is defined by

$$w = (1, 1, \dots, 1, w_{h+1}, w_{h+2}, \dots, w_k, 0, 0, \dots, 0)$$

where with no loss in generality we assume that the related activities are ordered such that  $w_i \geq w_{i+1}$  for  $i = h + 1, \dots, k$ . Once again, it is important to stress that setting the weights of the core activities to 1 and those of the unrelated activities to 0 is only meant to be an approximation to the truth. In all reality, the core activities would each carry a weight that is arbitrarily close to 1, and the unrelated activities would each carry a weight that is arbitrarily close to 0.

Furthermore, define the total value added of each of these groups of activities as

$$\begin{aligned} \sum_1^h x_i &\equiv C \\ \sum_{h+1}^k x_i &\equiv R \\ \sum_{k+1}^n x_i &\equiv U \end{aligned}$$

In this way, total GDP is simply  $T = C + R + U$ , and the total value of the contribution to GDP of copyright is given by

$$A = C + \sum_{h+1}^k w_i x_i$$

In relative terms, the contribution to GDP of copyright is given by

$$Y = \frac{C + \sum_{h+1}^k w_i x_i}{T} \quad (1)$$

Now, let us assume that the researcher is able to fully define the core activities with no problems, that is, the parameter  $h$  is given, and the only problems are that of defining the parameter  $k$ , and assigning weights to each of the related activities.

Now, let us postulate the following:

**Criterion 1.** *The weights assigned to the related activities and the division between related and unrelated activities ( $k$ ) should satisfy the condition*

$$Z \equiv \frac{\sum_{h+1}^k w_i x_i}{R} \geq Y$$

Criterion 1 can be seen to be nothing more than a certain definition of exactly what we mean by the “related” activity sector. It defines the related activities in such a way that *the average dependence of the activities included is at least as great as the average dependence of the entire economy*. This seems quite reasonable, simply since this seems to be the type of idea that is intended to be captured by the related activity sector of the economy.

Before going any further, we should check that, given the pre-established value of  $h$  (that is, the activities included in the core), we can always find a partition defined by  $k$  such that criterion 1 is satisfied. Firstly, note that if  $k = n$ , that is all activities outside of the core are included as related (i.e. there are no unrelated activities), then we would have  $R = \sum_{h+1}^n x_i$ , and

$$Y = \frac{C + \sum_{h+1}^n x_i w_i}{C + R}$$

But the copyright dependence of the related sector is just

$$Z = \frac{\sum_{h+1}^n x_i w_i}{R}$$

Now, in general it is true that

$$\frac{\partial \left( \frac{a+x}{b+x} \right)}{\partial x} = \frac{b-a}{(b+x)^2}$$

which is strictly positive if  $b > a$ . But writing

$$W(x) \equiv \frac{\sum_{h+1}^n x_i w_i + x}{R + x}$$

then  $Z = W(0)$  and  $Y = W(C)$ , and since  $C > 0$ , we clearly have  $Y > Z$  in this case.

Thus if the partition is defined such that there are no unrelated activities, then the objective variable ( $Y$ ) will be strictly greater than the average dependence of the related sector ( $Z$ ).

Secondly, say the partition is defined such that  $k = h + 1$ , that is, there is only one activity in the related group. In that case, the objective variable is valued at

$$Y = \frac{C + x_{h+1} w_{h+1}}{C + x_{h+1} + \sum_{h+2}^n x_i}$$

and the average dependence of the related sector is simply

$$Z = \frac{x_{h+1}w_{h+1}}{x_{h+1}} = w_{h+1}$$

Thus, we would have  $Z > Y$  if

$$w_{h+1} > \frac{C + x_{h+1}w_{h+1}}{C + x_{h+1} + \sum_{h+2}^n x_i}$$

Operating on this, it is a simple matter to reduce it to

$$w_{h+1} > \frac{C}{T - x_{h+1}}$$

However, since we know from previous studies that  $C$  is only a small fraction of  $T$  (most studies place it at between 3 and 5 percent), and since  $x_{h+1}$  is only one single activity in the economy, it will be almost insignificant in relation to  $T$ , thus the right-hand-side of this equation cannot be much greater than about 0.05. On the other hand  $w_{h+1}$  must be rather close to 1 (since the activity ranked at position  $h + 1$  is the most copyright dependent one outside of the core). Thus we can quite safely assume that such a partition does indeed yield  $Z > Y$ .

Finally, note that, whatever is  $k$ , we have

$$Z_k = \frac{\sum_{h+1}^k x_i w_i}{\sum_{h+1}^k x_i} \quad \text{and} \quad Z_{k+1} = \frac{\sum_{h+1}^k x_i w_i + x_{k+1} w_{k+1}}{\sum_{h+1}^k x_i + x_{k+1}}$$

and

$$Y_k = \frac{C + \sum_{h+1}^k x_i w_i}{T} \quad \text{and} \quad Y_{k+1} = \frac{C + \sum_{h+1}^k x_i w_i + x_{k+1} w_{k+1}}{T}$$

Clearly,  $Y$  is increasing in  $k$ . But it turns out that  $Z$  is decreasing in  $k$ . To see why, note that indeed  $Z$  would decrease with  $k$  if

$$Z_{k+1} = \frac{\sum_{h+1}^k x_i w_i + x_{k+1} w_{k+1}}{\sum_{h+1}^k x_i + x_{k+1}} < \frac{\sum_{h+1}^k x_i w_i}{\sum_{h+1}^k x_i} = Z_k$$

Simple operations on this inequality yield

$$w_{k+1} \sum_{h+1}^k x_i < \sum_{h+1}^k x_i$$

which is obviously true.

Thus, under the (very reasonable) assumption that  $w_{h+1} > \frac{C}{T - x_{h+1}}$ , it turns out that as  $k$  is increased,  $Z$  starts out above  $Y$  and ends up below  $Y$ , but  $Z$  is decreasing as  $k$  increases, while  $Y$  is increasing. Clearly, there must be a single  $k$  (not located at either extreme) for which  $Z = Y$ .

Now, go back to the objective, which is given by equation (1). Under criterion 1, we get  $\sum_{h+1}^k w_i x_i \geq YR$ , and so

$$Y = \frac{C + \sum_{i=h+1}^k w_i x_i}{T} \geq \frac{C + YR}{T}$$

Rearranging, this reduces directly to

$$Y \geq \frac{C}{T - R} = \frac{C}{C + U} \quad (2)$$

In particular, the lower bound given by (2) does not depend upon the exact weights that are set on the related activity group, but only the exact split that is made between core, related and unrelated activities. And, making use of the previous studies and the WIPO *Guide*, this split could be rather simple to calculate.

Naturally, it is interesting to consider exactly how useful this bound is by comparing it to the results of the studies that have already been done, although none of the reports actually define the related group as in criterion 1 here. In the final published reports, none of the studies actually report the (unweighted) values of the related group of activities, however from the reports we can get a minimum of intuition on these numbers.

To begin with, most of the studies suggest that the total added value of the core is between about 3% and 5% of total GDP. Thus, let us take  $C = 0.04T$ . On the other hand, given criterion 1, clearly the related group of activities would be very important in the economy. Let me simply assume that it constitutes one half of the unweighted total value added, that is,  $R = 0.5T$ .

With these figures, our lower bound is

$$\begin{aligned} \frac{C}{T - R} &= \frac{0.04T}{T - 0.5T} \\ &= \frac{0.04}{0.5} \\ &= 0.08 \end{aligned}$$

This figure, 8%, compares to the figure given in the latest US study (7.75%), but is greater than for most other countries, which are normally around 5%, give or take a percentage point. However, in the Australian study (for example), the total core value added is only about 1.5% of total GDP, and so our criterion would yield a final value of about 3%, which does compare very well with the final figure obtained in that study (3.3%).

Let us compare this with one of the real-world studies – the path-breaking study corresponding to the US Economy for 1989.<sup>3</sup> In that study, the total GDP was (all figures are expressed in billions of dollars) \$5,250.8. The total GDP attributed to the core industries was \$173.7146, or approximately 3.3% of GDP. Taking directly the 50% rule of thumb indicated above for the total (unweighted) GDP of the related activity division would give us a final figure for the total contribution of about 6.6%. In the study, the end result was a figure of 5.78%. In particular, the 6.6% is a lower bound, and yet the study returns a figure that is even more conservative, indicating that the choices of weights for the related activities, and/or

---

<sup>3</sup>I thank Stephen Siwek for kindly supplying me with the relevant data.

the choices of activities that are excluded as non-related, were rather conservative<sup>4</sup> – probably a good thing for empirical work. In that particular study, the total sum of all of the unweighted GDP of the related activities<sup>5</sup> was \$1,121.7232, or about 21.4% of total GDP. This is indeed conservative, at least in my opinion. The total weighted contribution of the related activities was taken as \$128.9775, thus the average weight applied there was about  $\frac{128.9775}{1121.7232} = 0.1149$ , i.e. about 11.4%. Now, in that study, about 75.3% (that is,  $100 - 3.3 - 21.4$ ) of all GDP was taken as being unrelated to copyright, or in terms of what was set out above, the copyright factor was considered to be so small that it can be taken as being insignificant. However, even if we were to say that the average factor for those activities was as low as 1% (i.e. less than a tenth of what is attributed, on average, to the related activities), then we would get 0.75%, i.e. close to a full percentage point, added to the total contribution figure, bringing us very close to the rough estimate of about 6.6%.

#### 4. CONCLUSION

In this paper, I have suggested the possibility of a simple calculation that returns a lower bound on the total contribution of copyright to GDP, once the groupings between the core, and unrelated activities has been made, but independently of the exact weights that should be assigned to the activities that are not in either of these two groups (i.e. those that remain in the related group). On the other hand, in order to do this it was necessary to take a particular definition of exactly what particular activities should be included in the related group (activities that, without having a copyright factor of 1, are on average at least as dependent upon copyright as is the economy as a whole). Thus, with a relatively low level of effort, one can get what appears to be quite an accurate, but still only intended as a rough estimate, answer to the question of exactly what is the total contribution of copyright to GDP.

In particular, if the final figure that is calculated in a study exceeds the lower bound considerably, then we can assume that the researcher has been quite liberal (as compares Criterion 1) as far as the selection of activities that are to be included in the related, rather than unrelated, group is concerned, and/or the choice of weights for the related group has been liberal. On the other hand, a figure that does not reach the lower bound suggests that the choice of weights (or of activities included in the related rather than the unrelated group) has been somewhat conservative.

In short, it is my contention that, at least as a first working approximation to the final contribution figure that one is searching for in this type of study, the short-cut suggested here may be useful. One would firstly decide on exactly which activities are to be included in the core, add up their respective added values, express this sub-total of added value as a percentage of total GDP, and then multiply by 2 (on the assumption that about 50% of total GDP is represented by related activities). The 50% rule does not seem to be all that much more *ad hoc* than the final decisions that need to be made when assigning final copyright factors, and it would also give a figure that is far more comparable across different countries. The methodology also implies that the type of definition as to what is a related industry and what is

---

<sup>4</sup>The authors themselves recognise that their choices of weights are indeed conservative estimates.

<sup>5</sup>There were three groupings; “partial” (which included 16 industry sub-groupings), “dependent” (included 7 industry sub-groupings), and “related” (a further 2 industry sub-groupings).

not (here given as Criterion 1) would be common across different studies. Finally, if the studies that have been done so far are any indication, the short-cut method suggested here does not err too far from what is finally calculated anyway.

RICHARD WATT, PROFESSOR OF ECONOMICS, UNIVERSIDAD AUTÓNOMA DE MADRID