

**Is a picture worth a thousand words? The division of private copy remuneration
between authors of text and images¹**

Juan Santaló
IE Business School
Dept. of Strategy
Alvarez de Baena, 4
28006 Madrid, Spain
Tel: (34) 91-568 96 00
Fax: (34) 91 563 22 14
email: juan.santalo@ie.edu

ABSTRACT

I propose and implement a method to divide the collection of private copy remuneration between the authors of text and the authors of images. I basically propose that the division should be based on the economic value added by text and images in the original work. Using a unique dataset of books and magazines copied in Spain, I estimate the economic value of each item, text and image, according to different characteristics of the written creative work. My estimates indicate that the average economic value of the images is between 6.35% and 20.00% of the average economic value of the text. These numbers are close to estimates that simply use the proportion of space occupied by images to proxy their economic value.

¹ This article is based on the author's study, commissioned in 2014 by CEDRO (the Spanish collecting society for editors and authors of books, magazines, and other publications), of different methods to estimate how the collection of private copy remuneration should be divided between CEDRO and VEGAP (the Spanish collecting society for visual artists).

1. INTRODUCTION

In most intellectual property right (IPR) regulations authors relinquish some of their exclusive rights to their work so that under some conditions consumers can use copyrighted material without the author's consent and without payment (Towse, Handke, and Stepan, 2008). More specifically these IPR regulations allow private copying of protected work done by consumers in their private sphere. This private copying exception heavily enhances consumers' freedom to use IPR-protected goods, and in exchange, IPR regulations establish a scheme to compensate authors. According to this scheme users pay private copy remuneration charges on devices that allow or facilitate private copying, with the goal of compensating right-holders for the reproduction of their works without a license fee.

For private copy remuneration for written goods (basically books and magazines), IP legislation usually establishes an equipment levy collected on photocopy machines, regardless of the actual use and lifespan of the machine. The equipment levy payers are the manufacturers, importers, or buyers of these machines. In addition, there is an operator levy collected on the basis of the number of copies of protected works. The operator levy payers are people or organizations who make copies or make photocopy machines available to others, whether free of charge or for a fee. Examples of organizations that pay the operator levy are educational institutions, libraries, and copy shops. The total of operator and equipment levies is then divided between different copyright owners—basically publishers, authors of the written text, and authors of the images. The question then becomes what proportion should go to each party. This proportion is subject to controversy since different studies have found different estimates (CEDRO, 2012; VEGAP, 2012).

In this study I propose and implement a method to estimate how to divide private-copy remuneration between the authors of text and the authors of images according to its economic value added in the work being copied. Using a unique dataset of copied books and magazines provided by CEDRO (the Spanish collection agency for publishers and authors of books, magazines, and other publications), I run hedonic regressions that allow me to disentangle the average economic contribution of text and image, and I corroborate this finding by asking for direct cost data from publishing companies. My estimates indicate that the average economic value of the images is between 6.35% and 20.00% of the average economic value of the text. These numbers are close to estimates that simply use the proportion of space occupied by images to proxy their economic value.

My contribution is twofold. To my knowledge this is the first study that provides rigorous empirical analysis of how private copy proceeds for written work should be divided between authors of text and authors of images. Hence, this study should be of interest to public authorities in all countries that have established a mechanism for private copy compensation for written work. Additionally, from an academic point of view, I contribute to the literature about the economic analysis of IPR by answering Towse, Handke, and Stepan's (2008) call for more research about the sharing of remuneration between different author types.

2. THE LEGAL CONTEXT OF THE DISTRIBUTION OF PRIVATE COPY COLLECTION AMONG DISTINCT RIGHT HOLDERS

According to the legal system in many countries like Austria, Finland, Germany, Norway, Spain, and Switzerland, the total collection for private copy remuneration is shared between publishers and authors. The part corresponding to authors is divided

between the authors of the written text and the authors of the graphic content. For example, at the present time in Spain 55% of the total amount collected goes to authors, while the other 45% belongs to publishers.² The 55% assigned to the authors has to be divided between two distinct collective rights management associations: the Visual Entidad de Gestión de Artistas Plásticos (VEGAP) and the Centro Español de Derechos Reprográficos (CEDRO). VEGAP represents visual creators like painters, sculptors, photographers, illustrators, and designers, while CEDRO represents the authors of written text.

Table 1 displays the percentage of the total amount of private-copy remuneration collected that is distributed to the authors of images in different European countries. It basically oscillates between 6% and 10% of the total, while the rest is divided between publishers and text authors.

*****INSERT TABLE 1 ABOUT HERE*****

The distribution of resources between authors of text and image is a breeding ground for conflict. In Spain, VEGAP and CEDRO have been debating recurrently about what distribution is fair. VEGAP commissioned a study by the consulting company CIMEC that concluded that graphic authors should get 42.5% of the total allocated to authors (23.37% of the total revenues collected). In contrast, a study commissioned by CEDRO and done by the consulting company CONECTA concluded that graphic authors should get 16.75% of the authors' piece (9.21% of the total revenues collected). The most important discrepancy lies in the method used to measure the relative importance of the image. CONECTA measures the percentage of space occupied by the image in the overall work. CONECTA finds that images represent on average 16.75% of the total space

² Verdict of 21st of October 2010, C-467/08.

in a book or similar publication, and accordingly CEDRO argues that graphic authors should get 16.75% of the total private copy revenues allocated to authors. That is, it uses the following formula:

$$IEV = \frac{\textit{Number of equivalent pages occupied by illustrations}}{\textit{Total number of pages}}$$

where *IEV* represents the Image Economic Value. In Appendix 1 I describe specifically how CONECTA weighs the space occupied by images.

CIMEC (2012, p. 13) takes a different approach that considers that one image, regardless of the space it occupies, should be as important as one page of text: *“CIMEC does not use the concept of “page equivalents,” since we understand that one image in a photocopied page is a copied protected work, regardless of its size. In the same manner we understand that one page of text is a complete work (and not one-third or one-fourth of a work) regardless of whether it contains a lot of text or little text and regardless of whether the text format (for example the font size or how text is spaced out) might imply that the space occupied could be less than a page. Furthermore, in our study we count as a text page any page even if it includes just one image with a footnote. Those cases count for one page of text and one page of image.”* With this justification, CIMEC estimates the economic value of the image according to the following formula:

$$IEV = \frac{\textit{Number of images}}{\textit{Number of images} + \textit{Number of pages with text}}$$

A priori it is not clear which of the two methods to estimate the value of graphic content is better. The method based on space has more widespread adoption, being used in other European countries like Germany. Remuneration based on the space occupied by graphic content is commonly used not only by collective rights management organizations (CRMOs), but also in private market transactions. For instance, the

Association of Professional Illustrators in Spain advises professional illustrators about the tariffs they should charge for their work. This association explicitly proposes that illustrator fees should vary with the amount of space occupied by the image. According to its recommendations, graphic authors should charge for a black and white illustration that fills half a page of a paperback book only 60% of the price of the same illustration if it occupies a whole page in the same type of book (FEAIP, 2010).

However, the controversy between the two CRMOs is plainly lacking an economic foundation. In the next section I propose a method to estimate how authors should share the private copy collection based on the relative economic value contributed by each party to the total economic value of the work. In the conclusion I compare the results of this proposed new method with the results of the methods used by VEGAP and CEDRO.

3. METHODS

IPR regulations establish that private copy collections should compensate authors for the damage caused by private copying. I next describe a procedure for estimating the relative economic damage done to authors of images versus authors of text.

3.1. Basic Intuition. To capture this relative damage, ideally one should be able to estimate the revenue for each type of work in a world without private copying and in a world with private copying. The income gap between the two worlds would perfectly determine the weight of the image relative to the text in the distribution of compensation. I argue that compensation for private copying should be distributed between authors of text and images in the same ratio as the percentages of the final retail price that they are paid for their original contributions to the work.”

Before developing the argument more formally it is useful to explain the basic intuition with an example. Suppose that a book is worth €10, of which the author(s) of text get 7.5% and the author(s) of images get 1%. As a result, the payment to authors of text would be 7.5 times the payment to authors of images. If the book sells 1,000 copies, then text authors would get a total of €750 and image authors €100. Now suppose that because of private copying, the book sells only 700 copies instead of 1,000. Revenue for authors of text would be €525 and for authors of images €70. That is, in an ideal world the authorities would have to compensate authors of text with €225 and the authors of images with €30—again, a ratio of 7.5 to 1. That is, the relative value of the economic remuneration for the original work determines the relative level of compensation for private copy damages. One could argue that in a world with private copying the total percentage going to authors may decrease because publishers may need to retain a larger percentage of revenues to recover the fixed cost of production. Yet this does not necessarily change the conclusion. To see this let's revise the original example and assume that in a world with private copying text authors get only 3.75% of the retail price instead of 7.5% and image authors get just 0.5%. In the world with private copying, text authors would get €262.5 while graphic content authors would receive €35. This implies that now text authors should receive $750 - 262.5 = €487.5$ while image authors should get €65 ($100 - 35$)—again, a ratio of 7.5 to 1. A quick-witted reader may have noticed that this depends on the fact that the percentage received by each author type decreases in the same proportion. This is indeed the main assumption of the analysis. Below I explain why I do not think this is a bad assumption, but first I develop this intuition more formally.

3.2. Case of a Unique Representative Publication. I start by analyzing a representative publication (book, newspaper, or magazine) and for the moment ignore the fact that distinct publication types may differ in the relative importance of text versus image. I will make the more general analysis later. The subscript 1 will represent the hypothetical world with no private copying, while subscript 2 will identify a world with private copying. Then the compensation for private copy damages that text authors should receive, C_t , should be:

$$C_t = R_{t1}T_1 - R_{t2}T_2 \quad (1)$$

where R_{ti} represents the market price for a text unit in world i while T_i stands for the total quantity of text units in world 1 for the representative publication. Note that the unit of text that is used—line, word, or page—is irrelevant as long as R_{ti} and T_i employ the same unit of measure.

Similarly, private copy compensation for authors of graphic content, C_i , should be:

$$C_i = R_{i1}I_1 - R_{i2}I_2 \quad (2)$$

where R_i stands for the market price of an image unit, whatever that is, and I_i symbolizes the total number of image units used in the representative publication. As above, the unit that is chosen to measure the images (number of images or space occupied by them) is irrelevant as long as market prices and total quantity are measured in the same unit.

With this notation, the ratio of private copy compensation to authors of text and image is:

$$\frac{C_t}{C_i} = \frac{R_{t1}T_1 - R_{t2}T_2}{R_{i1}I_1 - R_{i2}I_2} \quad (3)$$

I assume that:

$$\frac{R_{t1}T_1}{R_{I1}I_1} = \frac{R_{t2}T_2}{R_{I2}I_2} = \bar{\alpha} \quad (4)$$

where $\bar{\alpha}$ is a constant parameter; then I can manipulate and simplify (3) to get:

$$\frac{C_t}{C_i} = \frac{R_{t1}T_1 - R_{t2}T_2}{R_{I1}I_1 - R_{I2}I_2} = \frac{\frac{R_{t1}T_1}{R_{I1}I_1 R_{I2}I_2} - \frac{R_{t2}T_2}{R_{I1}I_1 R_{I2}I_2}}{\frac{R_{I1}I_1}{R_{I1}I_1 R_{I2}I_2} - \frac{R_{I2}I_2}{R_{I1}I_1 R_{I2}I_2}} = \bar{\alpha} \quad (5)$$

The implications of putting together equations (4) and (5) are important for the identification strategy. Note that in reality I observe only the world with private copying (subindex 2), while the world without private copying is an unobserved counterfactual world. But with the assumption stated in (4), the ratio $\frac{R_{t2}T_2}{R_{I2}I_2}$ is enough to determine the ratio of compensation to authors of text versus authors of images. This is the strategy that I use in the empirical section.

Given how critical (4) is for the empirical strategy, it is worthwhile to reflect on the specific micro-foundations in which it could hold. If one assumes that publishers maximize profits in a world with competitive input factors, the remuneration paid to authors would be equal to its marginal productivity. This is:

$$\frac{R_{Ti}}{R_{Ii}} = \frac{\frac{\partial y_i}{\partial T_i}}{\frac{\partial y_i}{\partial I_i}} \quad (5)$$

where y_i represents the publishing output in world 1. I can use (5) and (4) to get:

$$\frac{R_{Ti}T_i}{R_{Ii}I_i} = \frac{\frac{\partial y_i}{\partial T_i} T_i}{\frac{\partial y_i}{\partial I_i} I_i} = \bar{\alpha} \quad (6)$$

Applying the natural logs to both sides of the equation (6) implies:

$$\ln\left(\frac{T_i}{I_i}\right) = \ln(\bar{\alpha}) + \left| \ln\left(\frac{\frac{\partial y_i}{\partial T_i}}{\frac{\partial y_i}{\partial I_i}}\right) \right| \quad (7)$$

According to Varian (1992), the elasticity of substitution between two factors of production, σ , is defined by:

$$\sigma = \frac{\partial \ln\left(\frac{T_i}{I_i}\right)}{\partial \ln\left(\frac{\frac{\partial y_i}{\partial T_i}}{\frac{\partial y_i}{\partial I_i}}\right)} \quad (8)$$

Hence (7) implies that $\sigma = 1$. In sum, the assumption needed for the identification strategy is that the elasticity of substitution between text and graphic content is equal to one. This would be true, for example, if the publishing business has a production function whose functional form is a Cobb-Douglas:

$$y_i = E_i^\theta (I_i^\alpha T_i^\beta)^\mu \quad (6)$$

where E symbolizes factors of production other than text and image (like capital) and both θ and μ are parameters.

With this assumption about the production function, it is straightforward to derive that if factor markets are competitive,³ the ratio of total compensation received by any author type to total revenue generated by the representative publication, $p_i y_i$, is constant, where p_i represent the unit price of the publication and y_i the total number of units sold. This is:

$$\frac{R_{Ii} I_i}{p_i y_i} = \alpha \mu; \quad \frac{R_{Ti} T_i}{p_i y_i} = \beta \mu \quad (7)$$

Hence (7) implies that (4) is satisfied, since it is straightforward to derive that

$$\frac{C_t}{C_i} = \frac{\beta \mu}{\alpha \mu} = \frac{\beta}{\alpha} \quad (8)$$

That is, the ratio of private-copying compensation to text authors to compensation to graphic content authors should be equal to the ratio of the percentage of total publication revenues paid by publishers to text authors to the percentage paid to

³ Although I require that the factor market be competitive, the publishing market need not be so. If publishers have some degree of market power, then

$$\frac{R_I}{p_y} = \left(1 + \frac{1}{\epsilon_d}\right) (\alpha \mu); \quad \frac{R_T}{p_y} = \left(1 + \frac{1}{\epsilon_d}\right) (\beta \mu)$$

where ϵ_d represents firm residual demand elasticity. Hence it would still be the case that $\frac{C_t}{C_i} = \frac{\beta}{\alpha} = \bar{\alpha}$.

authors of images. This means that if one can compile publishers' cost allocation and publication revenues, one gets a good estimate of how private copy collection should be distributed among different types of authors. I use this insight next to estimate how private copy collection should be divided between authors of text and images.

3.3. Distinct Types of Publications. Now I allow that there are distinct types of publication each using different proportions of text and graphic content—for example, architecture textbooks versus social science ones. It is worth considering this possibility because private copying could differently affect the demand for each type of publication. If so, the relative compensation to authors of text versus authors of images should also vary. Let's start by assuming that there are two publication types, A and B, with different relative importance of text versus graphics. More specifically, let's assume that:

$$y_{ij} = E_{ij}^{\theta_i} (I_{ij}^{\alpha_i} T_{ij}^{\beta_i})^{\mu_i} \quad (9)$$

where $j=2$ is a world with private copy and $j=1$ otherwise. The only modification with respect to (6) is that the parameters of the production function vary with the type of publication i . If there are two publication types, then:

$$\frac{C_t}{C_i} = \frac{R_t T_{A1} - R_t T_{A2} + R_t T_{B1} - R_t T_{B2}}{R_t I_{A1} - R_t I_{A2} + R_t I_{B1} - R_t I_{B2}} \quad (10)$$

Manipulating (10), I get:

$$\frac{\beta_A \mu (\Delta_A - 1) p_{A2} y_{A2} + \beta_B \mu (\Delta_B - 1) p_{B2} y_{B2}}{\alpha_A \mu (\Delta_A - 1) p_{A2} y_{A2} + \alpha_B \mu (\Delta_B - 1) p_{B2} y_{B2}} = \frac{\beta_A \mu (\Delta_A - 1) \omega_A + \beta_B \mu (\Delta_B - 1) \omega_B}{\alpha_A \mu (\Delta_A - 1) \omega_A + \alpha_B \mu (\Delta_B - 1) \omega_B} \quad (11)$$

where $\Delta_i = \frac{p_{i1} y_{i1}}{p_{i2} y_{i2}}$ measures the relative change in revenue for publication i due to private

copying and $\omega_i = \frac{p_{i2} y_{i2}}{p_{A2} y_{A2} + p_{B2} y_{B2}}$ stands for the relative revenue for each type of

publication as a proportion of total publishing industry revenue. (11) can easily be

generalized to the case of N publication types:

$$\frac{C_t}{C_i} = \frac{\sum_{i=1}^N \omega_i \beta_i (\Delta_i - 1)}{\sum_{i=1}^N \omega_i \alpha_i (\Delta_i - 1)} \quad (12)$$

Compared with the expression in (8), the expression in (12) presents two extra terms to be estimated. One is the relative weight of each industry type in the world with private copies. This is usually observable, and therefore its estimation does not represent a challenge. The second extra term is Δ_i . This is the change of demand for publication type i due to private copying. If this change were constant for all publication types, then I would not need to estimate it since I could cancel it from numerator and denominator. Since I am not sure this is the case, in the empirical procedure below I will estimate it.

4. EMPIRICAL ESTIMATION OF THE AVERAGE ECONOMIC VALUE OF GRAPHIC CONTENT

IN RELATION TO THE ECONOMIC VALUE OF TEXT

In this section I estimate the relative values of text versus image as inputs in the publishing production function, using the insights derived in the previous section. First I estimate the relative compensation of authors of text and images using the cost information provided by publishers. Second, I estimate this relative compensation with a statistical analysis that estimates how the image and the text explain the final prices of books, magazines, and similar publications.

4.1. Estimation Using Direct Cost Data from Spanish Publishing Companies. CEDRO has provided the author of this article with cost data representing the amounts paid for image rights in the process of producing a book. This information was provided to CEDRO by a number of publishing companies that in Spain jointly represent more than 25% of total turnover in the domestic book market. To preserve anonymity I have changed the real names of the publishing companies. Table 2 shows these estimates for

the year 2012, when image rights represented on average about 1.01% of the final book price.

*****INSERT TABLE 2 ABOUT HERE*****

Next, I estimate the annual amount paid for copyrighted text as a percentage of total turnover for the publishing industry in Spain. Table 3 shows the total payment for text author rights and the total turnover of the sector in Spain in 2012 by size of publisher as listed in the "Comercio Interior del Libro" (Domestic Book Trade 2012).

*****INSERT TABLE 3 ABOUT HERE*****

According to the information displayed in Table 3, the total publishing cost paid to the authors of text is about 6.90% of final retail price. However, the sample used to estimate the image costs described above comprises only very large companies. Therefore to make the comparison as accurate as possible, it may be convenient to use the figure of 7.46% that Table 3 reports for very large companies.

From the information displayed in Table 2 and Table 3 I can then compute an estimator of $\bar{\alpha}$:

$$\bar{\alpha} = \frac{7.46}{1.01} = 7.3867$$

Hence according to the expression in (5) the amount of the private collection that should be distributed to text authors should be around 7.39 times the amount assigned to the authors of graphic content.

4.2. Estimation Using Hedonic Regressions. An alternative way to find the percentage of the total economic value added of text versus image is to estimate how each item contributes to the total price of the book. That is, I estimate a production function in

which the combination of text, T , images, I , and other materials E gives rise to the production of a work of a certain quality which I denote by q .

4.2.1. *Sample.* The sample used to estimate the contribution of text and image to the value of the work comes from a survey commissioned by CEDRO in 2012 to estimate the number and prevalence of private copies in Spain. The sample was designed to be representative of Spain as a whole, and it was done by sending survey takers to places with a large volume of photocopying machines, including education centers, public administration, and private companies. The database includes information on the title of the work being copied, publisher, ISBN number, the estimated number of pages with graphic content, and the percentage of space that images occupied. It had information on 524 books and magazines. I have complemented these data with information published on Spain's Ministry of Culture (MCU) webpage.⁴ From this publicly available database, I obtained, for each book, price, year of first publication, number of pages, binding type, and height and length in centimeters. Since some magazines have no ISBN number and since some of the books in the MCU database do not have the final sale price, the number of observations used for the estimation is 379. Table 3 provides descriptive statistics of the variables used in the analysis.

4.2.2. *Variable Description.*

Price: The price in euros, obtained from the MCU database of books published in Spain.

Pages: The number of pages in the book, according to the MCU database of books published in Spain.

⁴ Available at:

http://www.mcu.es/webISBN/tituloSimpleFilter.do?cache=init&prev_layout=busquedaisbn&layout=busquedaisbn&language=es

Size: Page area in square centimeters. The MCU database provides the height and width of the page in centimeters

Years since Printing: The number of years that have elapsed since the printing of the book.

Years since 1st Edition: The number of years that have passed since the first edition.

Binding: Information on the type of binding, also extracted from the MCU database; it can be paperback, hard cover, leather, cart, Geltex, or other.

Image Page Equivalent: The database of CEDRO provides the space occupied by the images on the book pages that were copied. CEDRO also provides total pages copied. Image page equivalent is the result of multiplying the total number of book pages by the ratio of copied image equivalent pages to total number of copied pages. Please refer to the Appendix for all the details about how CEDRO computes the space occupied by images in any given publication.

Text page equivalent: The estimation of space occupied by the text in the book. It is the result of subtracting the number of image pages from the total number of pages in the book.

Finally I include dummy variables that indicate the general subject of the book. Table 4 displays the descriptive statistics of the sample used in the hedonic regressions. The average price of the books in the sample is €26.02. The graphic content occupies the equivalent of 16 pages, while the average number of text pages is 364. The subject most widely represented is social science, which accounts for around 27% of the observations in the sample.

*****INSERT TABLE 4 ABOUT HERE*****

4.3. **Estimation Assuming a Representative Publication.** As above, I start by assuming that the intensity of use of text versus image is the same across all publication types, and I estimate a Cobb-Douglas functional form:

$$q = E^\theta (I^\alpha T^\beta)^\mu \quad (13)$$

Note that in the notation employed in the previous section $q = py$. Since now I focus on the production process of a unique unit of one publication, I assume that $y=1$ and use book price, p , as a proxy of q . Hence I can rewrite (13) as:

$$\ln p = \alpha_1 \ln E + \alpha_2 \ln I + \alpha_3 \ln T \quad (14)$$

where $\alpha_1 = \theta$; $\alpha_2 = \alpha\mu$; $\alpha_3 = \beta\mu$. Note that with this production function, α_2 and α_3 represent the percentages of the publisher's revenue that would go to authors of images and text respectively. Hence, with this functional form assumption, the relative damage suffered by authors of image versus the damage suffered by authors of text could be approximated by the ratio $\frac{\alpha_2}{\alpha_3}$.

I start by estimating a linear regression of (14) above, including as controls the variables described above as well as dummies that indicate the subject of the book.

The results are included in the first column of Table 5.

Regarding the control variables, the results indicate that social science and humanities books have higher prices. The size of the page also explains the price of the book; a 10% increase in squared centimeters per page increases the price by 9.2%. Finally, while the time since printing has no significant effect on price, the years since the first edition have a significant negative effect: the price falls by around 3% for every year passed since the first edition. Both the space occupied by the images and

the space occupied by text increase the price of the book, but the *relative* effect of text pages is five times larger than that of image pages ($5 = 0.50 / 0.10$). This suggests that the economic value of the text is five times the economic value of the image.

In the second column of Table 5 I repeat the same linear regression but including fixed effects by binding type. These are important controls to add, since it may be that binding type is associated simultaneously with price and number of pages with graphic content. Hence not adding these controls could bias upwards the coefficient of number of image pages. However, with binding type fixed effects the essence of the results does not change; indeed, the ratio of added value of text to added value of images increases slightly to 5.85 ($= 0.47 / 0.08$).

Finally I add publisher fixed effects. With them I ensure that the above results are not due to different publishers selling books at higher prices while being more or less likely to use images. The third column of Table 5 shows the corresponding results. Once I control for publisher fixed effects, the importance of text pages for the book price increases, becoming 0.60, while the coefficient of equivalent image pages becomes 0.07. Hence, the estimated ratio of value of text to value of images is about 8.5 ($8.5 = 0.60 / 0.07$), much higher than the ratio found in the two previous regressions. However, as a cautionary note, one must take into account that in the database there are 285 publishers so that adding these fixed effects I am actually using information only from those publishers who have more work published in the sample with which these results are determined by a much smaller number of observations. In general, Table 5 shows that the contribution of the text to book prices oscillates between 5 times and 8.57 times the contribution of graphic content. Hence, in accord with the arguments developed in Section 3, the private copy compensation received by

text authors should be between 5 times and 8.57 times the compensation received by authors of images.

4.4. Estimation Allowing the Proportion of Graphic Content to Vary with Publication

Type. But suppose that the relative importance of text versus image varies by subject. Simple sample descriptive statistics indeed show some subject variance in the relative use of graphic content. Table 6 displays the average space occupied by images by topic. While in biomedical sciences the image content uses an average of just 2.62% of the overall publication space, in scientific-technical publications the percentage jumps to 22.62%.

*****INSERT TABLE 6 ABOUT HERE*****

I repeat the same hedonic regression estimated above but allowing the coefficient regressions to vary by book subject, and I run the following linear regression:

$$\ln p_i = \alpha_{1i} \ln E_i + \alpha_{2i} \ln I_i + \alpha_{3i} \ln T_i \quad (15)$$

where the subindex i identifies each of the seven subject types listed in the CEDRO database: humanities, social sciences, literature, biomedical sciences, scientific-technical, education, and other.

Table 7 displays the results of estimating (15) using both subject fixed effects and binding-type fixed effects. I do not estimate a regression with publisher fixed effects because given the low number of publishers that have more than one publication, regressions in which I also allow variation in the regression coefficients by book topic have a low number of degrees of freedom.

*****INSERT TABLE 7 ABOUT HERE*****

The Table 7 results show how the relative value contributed by the space occupied by images depends on the publication subject. While for scientific-technical books the

price-space image elasticity is statistically different than zero and with the largest point estimate equal to 0.131, in literature, biomedical sciences, and others the corresponding regression coefficient is negative, and statistically one cannot reject the null hypothesis that it is equal to zero. In humanities the impact of image space on price is positive and statistically significant, but its point estimate is about half that for scientific-technical books. The contribution of text space to book prices also depends on the book subject. Its value is highest in biomedical sciences and social sciences and quite low in others.

As I noted in the analytical section, once one relaxes the assumption that the relative importance of graphic content to text changes by publication topic, one must estimate the relative impact that private copying has on each subject type in order to be able to estimate the proper relative compensation for each author type. For this I assume that the change in demand due to the existence of private copying is proportional to the percentage of pages copied by subject. More formally, I assume that:

$$\Delta_i - 1 = \gamma F_i \tag{16}$$

where F_i symbolizes the percentage of pages copied in subject i and γ is a parameter that I assume is constant across publication topics. This parameter scales by how much book demand would increase when the percentage of copies increases. Its estimation is out of the scope of this paper, and since I assume it does not vary across book subjects it clearly will not affect the relative compensation of text versus graphic content authors.

With (16) I can rewrite (12):

$$\frac{C_t}{C_i} = \frac{\sum_{i=1}^N \omega_i \beta_i (\Delta_i - 1)}{\sum_{i=1}^N \omega_i \alpha_i (\Delta_i - 1)} = \frac{\sum_{i=1}^N \omega_i \beta_i \gamma F_i}{\sum_{i=1}^N \omega_i \alpha_i \gamma F_i} = \frac{\sum_{i=1}^N \omega_i \beta_i F_i}{\sum_{i=1}^N \omega_i \alpha_i F_i} \tag{17}$$

The CEDRO database includes the number of pages copied for each publication in the sample. Since I have compiled the total number of pages for each book, I can then compute the percentage of pages copied for each book in the sample. This is the proxy I use for F_i . The second column of Table 8 displays this percentage of pages copied by publication subject. It does not exhibit a large degree of subject fluctuation, and all percentages are contained in the interval of 1.11% to 1.36%. Finally, from the same CEDRO database I estimate how much weight each publication type has as a function of total book sales in Spain. This is shown in the first column of Table 8.

*****INSERT TABLE 8 ABOUT HERE*****

With the information collected in Tables 7 and 8 I can estimate the relative private compensation to text authors in relation to graphic content authors according to expression (17). Table 9 summarizes all the information I use to estimate (17). Note that I opt to assign an image coefficient equal to zero for all publication subjects in which I cannot reject the null hypothesis that it is equal to zero. With these numbers, then

$$\frac{C_t}{C_i} = \frac{\sum_{i=1}^N \omega_i \beta_i F_i}{\sum_{i=1}^N \omega_i \alpha_i F_i} = \frac{0.088}{0.005} = 15.75 \quad (18)$$

(18) implies that the private copy compensation received by text authors should be around 16 times the compensation received by authors of images. In other words, the private copy compensation received by graphic authors should be just 6.35% (=1/15.75) of the compensation assigned to text authors.

*****INSERT TABLE 9 ABOUT HERE*****

4.5. Estimation without Assuming That the Elasticity of Substitution between Text and Image Is Equal to One. The critical assumption in the identification strategy is that the elasticity of substitution between text and graphic content is equal to one. In all

the empirical specifications I have used a Cobb-Douglas functional form that by construction satisfies this condition. Next, I replicate the estimation procedure using a more general functional form. I employ a translog production function (Christensen, Jorgenson, and Lau, 1973) as follows:

$$\log p = \alpha_1 \ln T + \alpha_2 \ln I + \alpha_3 \ln T \ln I + \alpha_4 (\ln T)^2 + \alpha_5 (\ln I)^2 + Controls \quad (19)$$

If $\alpha_3 = \alpha_4 = \alpha_5 = 0$, then (19) would be identical to the expression (14) that I have estimated above. From (19), if I assume that R_T / R_I would be equal to the marginal productivity of text/image, I can derive that:

$$\bar{\alpha} = \frac{R_T T}{R_I I} = \frac{\alpha_1 + \alpha_3 \ln I + 2\alpha_4 \ln T}{\alpha_2 + \alpha_3 \ln T + 2\alpha_5 \ln I} \quad (20)$$

Using the CEDRO sample, I estimate the production coefficients in (19) using nonlinear regression techniques, and from (20) I determine the implied ratio of compensation to the distinct authors at the sample mean value of both text and image variables. Table 10 displays the results. In all three specifications, $\bar{\alpha}$ has a value close to six, implying that with this more flexible functional form the private copy compensation received by text authors should be six times that received by graphic content authors.

*****INSERT TABLE 10 ABOUT HERE*****

Finally we use the translog production function to estimate the empirical plausibility of having an elasticity of substitution equal to one. For this we use a linear approximation of the Constant Elasticity of Substitution production function (Hoff, 2004):

$$p = [\delta I^{-\rho} + (1 - \delta) T^{-\rho}]^{-\nu/\rho} \quad (21)$$

Where the elasticity of substitution, σ , equal to:

$$\sigma = \frac{1}{1 + \rho}$$

According to Hoff (2004) we can take a linear approximation of (21) that will follow the exact same functional form of (19) above with the following constraints:

$$\alpha_4 = \alpha_5 = \alpha_3/2 \tag{22}$$

$$\text{Where } \delta = \frac{\alpha_2}{\alpha_1 + \alpha_2}; v = \alpha_1 + \alpha_2; \rho = \alpha_3(\alpha_1 + \alpha_2)/(\alpha_1\alpha_2) \tag{23}$$

From (23) we also have an estimation of the elasticity of substitution between text and images. With this in mind we re-estimate (19) above imposing the constraints in (22) to estimate the elasticity of substitution between text and image. The results are displayed in Table 11. In none of the three specifications we can reject the null hypothesis that $\sigma = 1$

*****INSERT TABLE 11 AROUND HERE*****

5. DISCUSSION AND CONCLUSIONS

Table 12 summarizes the different estimates of the relative economic weight of text and images produced by the distinct methods used here. According to these estimates, the maximum private copy compensation that authors of graphic content should receive equals 20% of the compensation received by text authors; the lowest, 6.34%. In Spain, given that publishers receive 45% of the total amount of revenues collected for private copy compensation, my estimates suggest that the compensation assigned to authors of graphic content should range from 3.49% to 11%. These numbers

are relatively close to the numbers obtained if one simply computes the percentage of space allocated to image versus text for the average publication (9.21%). In general, this then suggests that the standard method of distributing private copy revenues according to the amount of space occupied by graphic content versus text provides a good estimator of damage compensation based on lost economic value.

In this study I have ignored the different payment mechanisms that distinguish authors of graphic content from text authors. While text authors receive a percentage of final retail price for each unit sold (with or without an advance payment), graphic content authors usually are paid a fixed amount of money regardless of book sales. However, for the purposes of this paper this difference should not change the conclusions. My goal is to estimate the income loss caused by private copying, and this income loss should not be affected by the payment method. To see this point, note that the estimation has to take into account not only the damage caused by a lower number of units sold, but also the loss of income from books that are never published because allowing private copying lowers expected sales. These author losses do not depend on the payment mechanism.

The method introduced here can be easily adapted to similar allocation conflicts in other situations in which the artistic work is composed of distinct contributions from different copyright holders. For example, a similar method could be employed to estimate how to divide radio royalties to compensate music composers versus music performers, or how to assign TV royalties between authors of visual work and music composers. Note, however, that this method relies on the assumption that the elasticity of substitution between the different inputs that constitute the final output is equal to one. It is hard to estimate the elasticity of substitution between inputs of production,

precisely because most empirical specifications of production functions assume it to be equal to one. However we provide some evidence in our sample consistent with an elasticity of substitution between text and image equal to one. Future research should further confirm the empirical validity of this assumption.

*****INSERT TABLE 12 ABOUT HERE*****

REFERENCES

CEDRO Report (2012), “Resumen método Econométrico idóneo para la determinación de los pesos de la imagen y del texto en la copia privada de libros o publicaciones asimiladas”. Report prepared by CONECTA consulting presented at the Intellectual Property Commission of the Spanish Ministry of Culture, first section, 19 September.

Christensen, L., D. Jorgenson, and L. Lau (1973), “Transcendental Logarithmic Production Frontiers”, *Review of Economics and Statistics*, 55; 28–45.

FEIP, Federación de Asociaciones de Ilustradores Profesionales (2010), Libro Blanco de la Ilustración Gráfica en España. Madrid. ISBN: 84-609-0681-7.

Rosen, S. (1974), “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”, *The Journal of Political Economy*, 82(1); 34–55.

Towse, R., C. Handke, and P. Stepan (2008), “The Economics of Copyright Law: A Stocktake of the Literature”, *Review of Economic Research on Copyright Issues*, 5(1); 1–22.

Varian, H. (1992), *Microeconomic Analysis*. New York: W.W. Norton & Company.

VEGAP Report (2012), “Peso de la imagen en la copia de productos editoriales y audiovisuales”. Report prepared by CIMEC consulting presented at the Intellectual Property Commission of the Spanish Ministry of Culture, first section, 19 September.

TABLES

Table 1: International Comparison of the Percentage of Total Private Copy Compensation Assigned to Authors of Images in 2011

Austria	Germany	Finland	Norway	Spain	Switzerland	United Kingdom
9.58%	9.66%	8.80%	9.40%	9.21%	6.90%	8.00%

Source: CEDRO (2012)

Table 2: Image Costs in the Publishing Production Process

Publishing Company	Turnover (in millions of euros) ⁱ	Image rights paid (in millions of euros)	%ⁱⁱ
Publishing Company 1	87.257	0.785	0.90%
Publishing Company 2	106.89	1.37	1.28%
Publishing Company 3	153.08	3.44	2.25%
Publishing Company 4	186.30	0.24	0.13%
Publishing Company 5	48.10	0.08	0.16%
Total	581.627	5.915	1.01%

ⁱ Information on turnover in 2012 from SAB database

ⁱⁱ Information provided by publishing companies

Table 3: Text Author Compensation by Type of Publishing Company

Firms	Text Author Payment (in millions of €)	Turnover (in millions of €)	%
Small	20.83	293.01	7.11%
Average	44.12	630.66	7.00%
Large	35.94	612.28	5.87%
Very Large	69.76	935.54	7.46%
Total	170.65	2,471.49	6.90%

Source: Comercio Interior del Libro 2012

Table 4: Sample Descriptive Statistics

Variable	Average	Standard Deviation	N
Price (€)	26.04	36.63	376
Size	394.36	162.08	376
Years printed	13.42	93.01	376
Years first edition	8.88	7.49	376
Image pages	16.54	49.04	376
Text pages	364.28	483.06	376
Humanities	0.19	0.39	376
Social Sciences	0.27	0.44	376
Literature	0.14	0.34	376
Scientific-Technical	0.14	0.35	376
Biomedical Sciences	0.06	0.24	376
Education	0.14	0.35	376
Other	0.05	0.22	376

Table 5: Hedonic Regressions with Final Book Price (in Logs) as Dependent Variable

Variable	Model (1)	Model (2)	Model (3)
Constant	-5.33*** (.65)	-5.06*** (.88)	-6.56*** (1.20)
Humanities	.26* (.14)	.19 (.14)	.20 (.19)
Social Sciences	.32** (.14)	.29** (.13)	-.01 (.18)
Literature	.15 (.15)	.05 (.15)	.15 (.21)
Technical	.06 (.15)	.09 (.14)	.22 (.26)
Biology & medicine	.19 (.17)	.25 (.17)	.06 (.24)
Education	-.01 (.15)	.04 (.14)	.11 (.21)
Size(logs of cm2)	.92*** (.10)	.77*** (.10)	1.04*** (.15)
Image Page Equivalents (logs)	.10*** (.02)	.08*** (.02)	.07** (.03)
Text Page Equivalents (logs)	.50*** (.03)	.47*** (.03)	.60*** (.06)
Years printed	.00 (.00)	.00 (.00)	.02 (.02)
Years first edition	-.03*** (.00)	-.03*** (.00)	-.04* (.02)
Observations	376	376	376
R2	.53	.60	.97
Binding fixed effects	NO	YES	YES
Publisher fixed effect	NO	NO	YES

Standard deviation is reported below each coefficient in parentheses. *** indicates significance at the 1% level. ** indicates significance at the 5% level. * indicates significance at the 10% level.

Table 6: Publication Space Occupied by Graphic Content by Subject

Subject	% of equivalent pages of graphic content
Literature	7.76%
Education	6.15%
Scientific-technical	22.64%
Biomedical Sciences	2.62%
Social Sciences	9.20%
Humanities	6.37%
Others	7.13%

Table 7: Estimation of Hedonic Regression Coefficients of Text and Image by Subject

	Image regression coefficient (1)	Text regression coefficient (2)	Ratio of (2) over (1)
Humanities	.067* (.036)	.380*** (.062)	5.671 (3.18)
Social Sciences	.062 (.060)	.730*** (.064)	11.774
Literature	-.127 (.099)	.145* (.088)	-1.141
Scientific-Technical	.131*** (.037)	.380*** (.081)	2,90
Biomedical Sciences	-.028 (.058)	.972*** (.146)	-34.71
Education	.062 (.045)	.660*** (.114)	10.64
Others	-.122 (.097)	.052 (.229)	-0.426

Only the regression coefficients of text and image space are displayed. The linear regression included the same control variables as in Table 5 including binding type and subject fixed effects but not publisher fixed effect.

*Standard deviation is reported below each coefficient in parentheses. *** indicates significance at the 1% level. ** indicates significance at the 5% level. * indicates significance at the 10% level.*

Table 8: Percentage of Total Sales and Percentage of Pages Copied by Subject

Subject	% Sales	% of pages copied
Humanities	18.17%	1.11
Social Sciences	38.58%	1.13
Literature	8.07%	1.23
Scientific-Technical	15.55%	1.21
Biomedical Sciences	5.78%	1.21
Education	10.96%	1.24
Others	2.89%	1.36

Table 9: Numbers Relevant to Estimating Relative Compensation for Private Copying

Subject	%Sales	Text regression coefficients	Image regression coefficients	% pages copied
Humanities	18.17%	0.38	0.067	1.11
Social Sciences	38.58%	0.730	0	1.13
Literature	8.07%	0.145	0	1.23
Scientific- Technical	15.55%	0.380	0.131	1.21
Biomedical Sciences	5.78%	0.972	0	1.21
Education	10.96%	0.660	0	1.24
Others	2.89%	0	0	1.36
TOTAL	100%			

Table 10: Estimation of a Translog Production Function

Variable	Model (1)	Model (2)	Model (3)
Constant	-3.13*** (.75)	-3.56*** (.74)	-3.59*** (.71)
Size(logs of cm2)	.78*** (.09)	.83*** (.09)	.85*** (.08)
Years printed	.00 (.00)	-.00 (.00)	-.00 (.00)
Years first edition	-.03*** (.00)	-.03*** (.00)	-.03*** (.00)
Image Page Equivalent(logs)	.12 (.11)	.14 (.11)	.13 (.11)
Text Pages Equivalent(logs)	-.07 (.17)	-.11 (.17)	-.12 (.17)
Text*Pages (both in logs)	-.04*** (.02)	-.04** (.02)	-.04** (.02)
Text pages (log) square	.07*** (.01)	.07*** (.01)	.07*** (.01)
Image pages(logs) square	.04*** (.01)	.04*** (.01)	.04*** (.01)
$\bar{\alpha}^i$	5.95*** (1.49)	6.05*** (1.47)	6.25*** (1.44)
Observations	376	376	376
R2	.68	.67	.66
Binding fixed effects	YES	NO	NO
Subject fixed effect	YES	YES	NO

Standard deviation is reported below each coefficient in parentheses. *** indicates significance at the 1% level. ** indicates significance at the 5% level. * indicates significance at the 10% level.

¹ $\bar{\alpha}$ stands for the implied ratio of private copy compensation to authors of text versus authors of graphic content evaluated at the mean sample values of both text and image.

Table 11 Estimation of a translog functional form imposing the constraints of a linearized constant elasticity of substitution

Variable	Model (1)	Model (2)	Model (3)
Constant	-3.91*** (.79)	-5.34*** (.74)	-4.94*** (.71)
Size(logs of cm2)	.77*** (.10)	.92*** (.10)	.87*** (.10)
Years printed	.00 (.00)	.00 (.00)	.00 (.00)
Years first edition	-.03*** (.00)	-.03*** (.00)	-.03*** (.00)
Image Pages(logs)	.06 (.13)	.10 (.13)	.09 (.13)
Text Pages (logs)	.45*** (.12)	.50 (.17)	.52*** (.12)
Text*Pages (both in logs)	.00 (.01)	.12** (.02)	-.00 (.02)
Text pages (log) square	.00 (.01)	-.00 (.01)	-.00 (.01)
Image pages(logs) square	.00 (.01)	-.00 (.01)	-.00 (.01)
$\bar{\alpha}^i$	5.70*** (1.62)	5.07*** (1.24)	6.06*** (1.67)
Elasticity of substitution	.96*** (.36)	1.00*** (.20)	1.01*** (.20)
Wald test elasticity of substitution equal to one	.01	.00	.00
Observations	376	376	376
R2	.58	.55	.53
Binding fixed effects	YES	NO	NO
Subject Fixed effect	YES	YES	NO

Table 12: Summary of the results

Methodology	Economic weight of image versus text (%)
Publishing costs	13.53%
Statistic contribution to the final book price, case of a representative book	11.67% - 20%
Statistic contribution to the final book price, case of multiple subjects with different text-image intensity	6.34%
Statistic contribution to the final book price, case of a representative book with a translog production function	16%-16.80%

Appendix 1

The CONECTA report measures the economic importance of images in book, newspapers, and magazines by estimating the ratio of the surface occupied by the image to the overall surface of the book or publication. This consulting company sends personnel to a representative sample of distinct photocopy centers, like photocopy shops or university libraries, to directly survey the characteristics of all material photocopied in those centers. How CONECTA weights the surface filled by graphic content is described on page 52 of CONECTA (2012):

“For each photocopied item we have obtained three distinct variables:

Number of pages with and without graphic content.

Number of illustrations on those pages with graphic content.

Percentage of space filled by the graphic content in those pages that have it, since a photocopy may have an image that fills the entire page or a combination of text and images of different relative importance. In this last case we compute the so-called number of equivalent pages of graphic content. This is the result of adding up the percentage of space occupied by graphic content in all photocopied pages.

We measure the surface occupied by the graphic content using two templates:

10 cell template

20 cell template

<p>Texto texto texto texto texto texto texto texto texto texto texto texto texto texto texto</p>	
	
<p>Texto texto texto texto texto texto texto texto texto texto texto texto texto texto texto</p>	

Image =60%

	<p>Texto texto texto texto texto texto texto texto texto texto texto texto texto texto texto</p>	
<p>Texto texto texto texto texto texto texto texto texto texto texto texto texto texto texto</p>		
<p>Texto texto texto texto texto texto texto texto texto texto texto texto texto texto texto</p>		
<p>Texto texto texto texto texto texto texto texto texto texto texto texto texto texto texto</p>		

Image=20%

In the case of direct photocopies from the original the questionnaire reports whether the front page contains graphic content independently of whether the front page is photocopied.”⁵

⁵ Translated to English from Spanish by the author of this article.