

Chapter 2

Some Fundamental Economics

2.1 The Economics of Copyright

2.1.1 *The Basic Concept of Copyright*

Understanding the Principles of Copyright

Copyright—lat. *copia*, plenty or to make plenty—generally provides its rightholder with “an exclusive right to copy, reproduce, distribute, adapt, perform or display” (Scotchmer and Maurer 2006, p. 76) her works of creative expression.¹ That is, the rightholder receives a bundle of exclusive rights which allow to exercise control over the use of her works for a specific period of time. In most states of the world and for the majority of existing works this period is set at author’s life plus 70 years.² An exception to this rule concerns works for hire,³ where the copyright term is limited

¹Note that for matters of simplicity we refrain from comparing all the differences between national copyright laws in detail. Instead, we will primarily refer to two models with the US Copyright and the German Urheberrecht, representing two different approaches (US versus Continental Europe)—especially regarding the scope of protection. References to other national features will only be made where appropriate. Accordingly, the features of the German Urheberrecht shall also represent similar characteristics as codified under other national legislation in continental Europe.

²Note that the terms of copyright have been subject to change by legislation, especially within the last two decades. See the next section on the history of copyright for a more detailed picture.

³Following section 101 of the US Copyright Act (U.S.C) a “work made for hire is (1) a work prepared by an employee within the scope of his or her employment; or (2) a work specially ordered or commissioned for use as a contribution to a collective work [...]” (17 U.S.C. §101). This distinction, however, is not applied in all copyright systems of the world. The German “Urheberrecht”, for instance, does not include such a distinction as the copyright is always assigned to the creator of a work. The creator can only grant certain rights of use to a third party (Eger and Scheufen 2012b, p. 39).

to either 95 years after publication or 120 years after creation. After expiration the work enters the public domain.⁴

Copyrightable subject matters are creative⁵ works of authorship in art, literature and science, fixed in any tangible medium of expression.⁶ In particular, copyright protection refers to any creative activity expressed in artforms, including for example literary works and programs for data processing (particularly computer software), musical works (including any accompanying words), pantomimes and choreographical works, dramatic works (including any accompanying music), graphic arts and architecture, pictorial and sculptural works, motion pictures and other audiovisual works as well as illustrations of scientific and technical form (Raskind 1998; Rehbinder 2006).⁷ The fact that copyright protection has been extended to cover not solely aesthetic subject matters, but also works of fact (like maps and directories) as well as data bases and computer programs shows that the sometimes mistakenly assumed difference between “copyright solely for creative works” and “patents for industrial products” has become somewhat blurred. As a matter of fact, for some products a creator may even choose between the two, as it is the case for computer programs.⁸ Also the combination of both rights is common in practice. Besides, some functional articles may be sought either under copyright or under the design patent regime, whereas in this context an attachment of a copyright to a product that already receives protection under the design patent is ineligible (Besen and Raskind 1991, p. 13).

For understanding the basic principles underlying the system of copyright law, however, it is important to emphasize its distinct features as compared to patent law and other forms of industrial property rights (like trademarks or industrial design).

⁴In contrast, the patent law provides with 20 years of exclusive use for a technical invention. However, it is worth to note that not all forms of intellectual property rights expire after a certain period of time. For example, trademark protection can last virtually forever.

⁵As such, at least a minimum degree of creativity is required for copyright protection. However, the term “creativity” is somewhat blurred as it is not explicitly defined by the law. Especially in the economic literature on copyright creativity is often neglected. For a discussion see Demsetz (2009) and Towse (2006).

⁶For an overview on the general subject matters of copyright see 17 U.S.C. §102 for creative works produced in the USA. In the German “Urheberrecht” the subject matters are listed under §1 and §2 UrhG.

⁷Here, Raskind (1998) emphasizes that even though copyright originates in the technology of printing, reforms in copyright law have been extending the subject matters of copyright protection as a response to advancements in the technologies to copy, reproduce and distribute copyrighted works. See Eger and Scheufen (2012a) on the changes and challenges of copyright law in the light of technological change.

⁸Note that there are differences between national copyright and patent laws regarding particular subject matters. As such, software may be a subject matter under US Patent Law (but also under US Copyright Law), while software is solely a subject matter under copyright in European memberstate’s IP laws. Only if a software-related invention exhibits a certain technical character, as is the case for the ABS brake system. On the “patent and/or copyright for software” debate see Watt (2007).

First and foremost, copyright law grants an exclusive right independently of any application or registration process. Notwithstanding the possibility of registering and depositing a work at the Copyright Office, the copyright generally arises “self-contained” by means of the creation of the work. In particular, protection to an original⁹ work of authorship¹⁰ is granted subject to its fixation in any tangible medium of expression. Following section 101 of the US Copyright Act this “fixation” requirement involves that a work is embodied in a copy or phonorecord, by or under the authority of the author, and “is sufficiently permanent or stable to permit it to be perceived, reproduced, or otherwise communicated for a period of more than transitory duration” (17 U.S.C. §101).¹¹ Thus, this requirement is already met when “the word processor causes the word order to be printed on paper” in the context of a literary work or when a song is being performed on a recording medium like a CD or MP3.¹² As a consequence, the definition of claims in copyright law and hence its scope is subject to litigation. While in patent and trademark law a comprehensive review of the file by the Patent and Trademark Office provides for the validity of a patent or trademark, respectively, the copyright leaves the claim for validity to the copyright holder, i.e. the author must provide evidence to verify validity of her copyright (Besen and Raskind 1991, pp. 11 et seq.). Second, section 102(b) of the US Copyright Law specifies that “in no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work.” (17 U.S.C. §102(b)). Thus, copyright only protects the expression within a work, but not the idea expressed by the author. In this respect, copyright protection is much narrower than a patent, since a patent grants a monopolistic right regarding a (complete) technology and hence the idea of an innovation. Third and somehow related is the feature of “independent creation”. Copyright law primarily focuses on actual copying, considering any use as an infringement of an author’s exclusive right if a person actually copies a protected work. That is, any independent creation of a similar or even identical work (duplicate) does not

⁹Originality implies that it owns its origin to the author and does not simply constitute a copy of some earlier work. However, the originality requirement is not as restrictive as the novelty requirement in patent law. A (derivative) work that incorporates other works can still be original and hence subject to copyright protection.

¹⁰Authorship refers to the requirement that a work needs to represent a modicum of intellectual activity. Besen and Raskind (1991) provide a simple example for understanding the basic idea. As such, if “a chimpanzee were to manipulate the keyboard of a personal computer with a graphics program” (Besen and Raskind 1991, p. 12), then protection would not be granted as the required modicum of intellectual activity is deemed lacking.

¹¹For reference to the US Copyright Law see henceforth <http://www.copyright.gov/title17/circ92.pdf> (last accessed on September 1, 2014).

¹²In this context, it is important to note the difference between the information good (e.g. the content of a book, the melody of a song etc.) and the information medium or carrier (e.g. the paper in the context of a book, the CD or MP3 medium for music etc.).

violate any of the right holder's exclusive rights and receives protection in equal measure (Gordon and Bone 1999, p. 190). As a consequence, there is no such principle as priority—which plays a major role in patent law (first-to-file versus first-to-invent).¹³ Finally, copyright law defines not only its subject matter and the granted rights, but also its limitations more specifically as compared to the patent statute (Besen and Raskind 1991, p. 12). As such, section 102 of the US Copyright Act explicitly lists the categories that are considered as a work of authorship. Once a work is subject to protection, copyright accords six basic rights to its owner: (1) the right to reproduce the copyrighted work in copies or phonorecords, (2) the right to prepare derivative works upon the copyrighted work, (3) the right to distribute the work in copies or phonorecords (including sale, rental, lending, leasing or other forms of ownership transfer), (4) the right to perform the protected work (in the case of literary, musical, dramatic and choreographic works, pantomimes, and motion pictures and other audiovisual works), (5) the right to display the protected work publicly (in the case of literary, musical, dramatic and choreographic works, pantomimes, and pictorial, graphic, or sculptural works) and (6) in the case of sound recordings the right to perform the copyrighted work publicly by means of a digital audio transmission (17 U.S.C. §106). However, as copyright law seeks to balance the rights of the owner of a protected work with the public interest of having access,¹⁴ the law provides with a number of limitations and exceptions that allow the use (with or without payment of compensation) of a copyrighted work without first seeking permission (authorization) from its rightholder(s).¹⁵ As copyright law is (still) a national law, however, these limitations and exceptions to copyright (and related rights) vary considerably between nation states. This diversity is acknowledged partially by international treaties which provide with general conditions for the application of such restrictions to the copyright. In this respect, the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) constrains national exceptions to the Berne Three-Step Test. Despite the efforts in harmonizing national IP laws (in the sense of providing minimum standards), article 10 of the World Intellectual Property Organization (WIPO) Copyright Treaty¹⁶ emphasizes still the will to leave it to the national legislation to decide on whether or not a certain limitation or exception is to be applied and if so, to determine its scope. Despite all these differences in the embodiment of limitations and exceptions in national copyright laws, copyright is mainly limited by the concepts of “fair

¹³For a discussion of “first-to-file versus first-to-invent” see Frost (1967).

¹⁴On the economics of this basic goal of copyright law see Sect. 2.1.2.

¹⁵Note that the “originality” requirement—with the integrated principle of independent creation, meaning copyright protection in equal measure for identical works that have been created independently—and the fact that copyright provides protection for the expression only (not the idea) may also be considered as limitations to copyright. The same also applies to the duration of copyright which limits copyright protection to a particular period of time.

¹⁶See http://www.wipo.int/export/sites/www/treaties/en/ip/wct/pdf/trtdocs_wo033.pdf (last accessed on September 1, 2014).

use” and independent creation (Watt 2004, p. 157). As previously mentioned, the latter concept of independent creation allows for equal protection of two identical works (duplicates) that have been created independently. Consequently, copyright protection is limited in so far, as it considers a use to be an infringement if a work has actually been copied. The “fair use” argument limits the scope of copyright protection as certain uses are considered to be “fair” in the sense that the benefits of such uses to the public are assumed to outweigh the negative effects on the rightholders. However, the conceptualization of this copyright limitation differs considerably between national copyright law systems. In general, two distinct traditions on “fair use” limitations have evolved in copyright history¹⁷: on the one hand, the “fair use” doctrine in US Copyright Law (17 U.S.C. §107) and “fair dealing” in UK Copyright Law (sections 29 and 30 of the UK Copyright, Designs and Patents Act)¹⁸ respectively, and on the other hand, the “catalogue of exceptions” (Schrankenatalog) in European Copyright Laws (in Germany, e.g. §§44 et seq. UrhG)¹⁹. The “fair use” doctrine (and the concept of fair dealing in UK Copyright Law) generally provides with a flexible instrument for assessing whether a use should be considered as “fair” and hence non infringing by means of a four-factor balancing test, explicitly taking into account the benefits and costs of certain uses of copyrighted works. In particular, section 107 of the US Copyright Law lists four statutory factors: First, the purpose and character of the use, raising two important issues to be evaluated: the commercial and the transformative character of the use. Second, the nature of the copyrighted work, explicitly taking into account the degree of creativity as “more creative” works are considered to be “closer to the core of intended copyright protection.” (Campbell 1994, p. 586). Third, the amount and substantiality of the portion used, reflecting on the degree to which a work has been copied and the relevance of the copied fraction in relation to the complete work. Obviously, the higher the amount of copied material, the less likely may be to justify such a use by a “fair use” argumentation. Fourth, the effect of the use upon the potential market for or the value of the copyrighted work, evaluating the costs imposed on the rightholder due to e.g. competition with a new product. In judging whether a use is conceived as “fair” in reference to section 107 of the US Copyright Law, finally, all four statutory factors have to be weighed against each other.²⁰ However, the list of section 107 is not exhaustive. In particular, the court shall

¹⁷These differences originate in the distinctive features of two opposite systems: common law and civil law.

¹⁸See <http://www.legislation.gov.uk/ukxi/2010/2694/made> (last accessed on September 1, 2014).

¹⁹See <http://www.gesetze-im-internet.de/bundesrecht/urhg/gesamt.pdf> (last accessed on September 1, 2014).

²⁰Note that in *Harper & Row v. Nation Enterprises* the US Supreme Court stressed the fourth factor as to be most important for assessing a fair use judgement (*Harper & Row v. Nation Enterprises* 1985, 471 U.S. 539). Notwithstanding, in *Campbell v. Acuff-Rose Music, Inc.* the Supreme Court more recently argued that “the four statutory factors are to be explored, and the results weighed together”, highlighting the need to weigh all statutory factors against each other (*Campbell, aka Skyywalker et al. v. Acuff-Rose Music, Inc.*, 510).

account for the user's intention in the particular case.²¹ In contrast, the European model of copyright law explicitly names all cases that are considered as limitations by means of a "catalogue of exceptions". Apparently, the "fair use" doctrine exhibits much more flexibility as it enables to account for specific circumstances by means of a case law approach. Notwithstanding, US and European copyright laws have certain common limitations that are generally accepted as exceptions in copyright protection. Sections 107 to 122 in US Copyright Law and e.g. paragraphs 44–63 in German Copyright Law (Urheberrechtsgesetz, UrhG) list several uses that are not conceived as an infringement of copyright. These limitations and exceptions—where a copyrighted work may generally be used without seeking authorization of the rightholder (with or without payment of compensation)—include especially uses for academic purposes—like research and teaching. In scientific research, for example, the right to cite and adopt expressions from other authors constitutes a crucial barrier of copyright protection, explicitly taking into account the cumulative character of scientific research and facilitating the creation of derivative works. Besides, exceptions regarding the personal use of copyright material (so-called private copy, e.g. in Germany §53 UrhG)²² abound as a common category in limiting copyright.²³ As such, copyright law allows for the making of single copies for private use if the source is unapparent illegal and follows a non-commercial purpose (Landes and Posner 1989, pp. 347 et seq.; Eger and Scheufen 2012a, pp. 157 et seq.). Other limitations refer for example to home recording of musical compositions.

Last but not least, a copyright may also be transferred to third parties. Section 201(d) of the US Copyright Law (Chap. 2) specifies that "the ownership of a copyright may be transferred in whole or in part by any means of conveyance or by operation of law, and may be bequeathed by will or pass as personal property by the applicable laws of intestate succession." (17 U.S.C. §201(d)). In particular, two different forms of transfer of copyright exist: (1) Copyright Assignment and (2) Copyright Licensing. The assignment of a copyright involves the transfer of rights in an exclusive and definite manner.²⁴ The licensing option, in contrast, provides the licensee only with the permission to use a certain right as contractually agreed upon. However, there are considerable differences between national copyright systems regarding the options for the transfer of copyright. The German "Urheberrecht" does not allow the assignment of a copyright, but the option to grant or licence

²¹See Mueller-Langer and Scheufen (2011b) for a recent "fair use" analysis regarding the Google Book Search Project.

²²Note that the US Copyright Law does not have a general "private copy" exception.

²³Other categories consider cases like parody, criticism, reviews, news reporting, archiving etc. The "fair use" doctrine (US Copyright Law) and the "catalogue of exceptions" (European Copyright Laws) follow various public goals such as freedom of speech, educational and equality of access as well as issues referring to market failures. See e.g. Hugenholtz (2001).

²⁴Notwithstanding, section 203 of the US Copyright Law specifies that a copyright owner has a non-revocable right for the termination of transfers. Consequently, the author of a book may reclaim her copyright. See section 203 U.S.C. on the conditions for the termination of copyright transfer.

certain rights of use, i.e. to grant a license for the right to distribute, reproduce, perform, or display the copyrighted work and obtain royalties for granting the right (Besen and Raskind 1991, p. 14).²⁵ A licence as such can be exclusive or non-exclusive. With the exclusive licence the licensor grants a certain right to use her work exclusively, i.e. without the option for the licensor to grant such rights to third parties. A common practice in this respect is the transfer of the right to distribute the copyrighted work to the publisher in exchange for participation in sales. The latter type—the non-exclusive licence—involves the option to grant multiple licences for the right to use the copyrighted work. The reasonability of such licences arises out of situations where the ratio of the value of a given piece of copyrighted material and the transaction costs of licensing is relatively small (Besen and Raskind 1991, p. 15). That is, in all cases where individual licensing agreements or vis-à-vis negotiations would involve prohibitively high transaction costs, so-called copyright collectives²⁶ enter as mediating parties between copyright owner and the consumers of a work. The Copyright Collective collects royalties by compulsory or individual licenses negotiated on behalf of the copyright holders, who in return are remunerated for the use of their works. Especially the compulsory license with statutorily determined license fees allows the remuneration of copyright holders whose works are used within the boundaries considered as limitations of copyright. In this regard, the private copying of copyrighted material is burdened with a fee on blank copying machinery or media (Besen and Raskind 1991, pp. 14 et seq.). However, while the idea of collective administration of copyright is widely shared, the statutory role of the Copyright Collective differs considerably between national legislations. As a consequence, the Anglo-American Copyright system is not as restrictive as the European system, where collective administration is mandatory.

Obviously, the system of copyright law is complex and needs to be seen in historical reflexion to understand the evolution of particular principles and the differences between national copyright legislations. As such, the following section reflects on the history of copyright and successively develops the pillars of the copyright system from its origins to modern copyright law.

The History of Copyright

The history of copyright, meanwhile, is inseparably linked to the development of technologies to produce, copy and distribute copyrightable works (Eger and Scheufen 2012a, pp. 160 et seq.). In particular, the invention of the “moveable type

²⁵See §29 UrhG. Accordingly, the rights of use can be granted im- or explicitly.

²⁶On the economics of Copyright Collective Societies (CCS) see e.g. Handke and Towse (2007). Well known Copyright Collectives are e.g. the American Society of Composers, Authors and Publishers (ASCAP) and the Broadcast Music Incorporated (BMI) in the USA. In Germany, the “Gesellschaft für musikalische Aufführungs- und mechanische Vervielfältigungsrechte” (GEMA) and the “Verwertungsgesellschaft Wort” (VG Wort) accordingly abound as examples of CCS.

printing press” by Johannes Gutenberg²⁷ around 1440—along with the invention of ink and paper sometime before—marks the starting point of actually discussing a need to regulate the production and copying of books.²⁸ Gutenberg’s printing press allowed for the first time the efficient mass production of books and eased the possibility to effectively reproduce books in multiple copies (Samuels 2000, p. 11). Prior to that time, the reproduction of books was extremely laborious and time consuming, since every single duplicate of a given work involved the production of handwritten copies by slaves and scribes (in ancient Greek and Roman civilisations), and monks and paid scribes (in medieval times). As a consequence of the introduction of the movable type to the Western World²⁹ by Gutenberg, however, the costs for producing copies of books were suddenly brought down to only a fraction of earlier copying practices. While the practice of handwritten copying involved high variable costs of production but rather negligible fixed costs, the new printing technology changed the environment to its opposite and hence allowed for the realization of cost advantages by means of unit cost degression.³⁰ Kapp and Goldfriedrich (1908) estimate that for a book edition of 500 copies the printing press enabled to reduce printing costs by a factor of 5 by the mid fifteenth century and even by a factor of eight at the end of the fifteenth century (as cited in Tietzel 1995, p. 42).³¹ As a result, the production of books exploded after this technological revolution with a massive shift in the number of books in libraries all over Europe.³² Thus, with the dispersion of the printing press in the fifteenth century an occasion was created asking for a regulation of the market for books in order to prevent from the free riding risk³³ for book publishers and authors.

The British Parliament was finally first to recognize a statutory right to copy by enacting the Statute of Anne³⁴ in March 1710 (Feather 1980, p. 19; Raskind

²⁷Originally Johannes Gensfleisch von Sorgenloch, referred to as Johannes Gutenberg. See also Venske (2000).

²⁸See chapter one in Samuels (2000) for a comprehensive review on “The Printing Press—The First Copyright Technology”.

²⁹In fact, in China the printing with carved wooden blocks as well as ink and paper had been invented around the sixth century.

³⁰See Eger and Scheufen (2012a) for a discussion.

³¹See also Eger and Scheufen (2012a) on the economic consequences of the printing press invention for the cost structure in the market for books.

³²Carl Sagan (1980) gives an impression on the effect of the “Gutenberg revolution” by highlighting that the number of books available in Europe increased rapidly from a few tens of thousands around 1450 to more than ten million in 1500. See also Samuels (2000) on page 13.

³³We will later reflect in detail on the economic reasoning and consequences of a free riding argumentation.

³⁴The full title is “An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Authors or Purchasers of such Copies, during the Times therein mentioned”. See <http://www.copyrighthistory.org/cgi-bin/kleioc/0010/exec/ausgabe/> (last accessed on September 1, 2014).

1998, p. 478; Samuels 2000, p. 12).³⁵ Notwithstanding, there had been efforts by the British Parliament to effectively regulate the printing prior to the Statute of Anne by granting a state-sanctioned monopoly over the printing of books.³⁶ After William Caxton had introduced Gutenberg's printing press to England in the fifteenth century, the economic threat of piracy led to the formation of a collective organization (the Stationers' Company) by the printers (known as stationers) in England. In 1557, finally, the Stationers' Company was chartered by royal decree and given the power to effectively control the printing and distribution of writings (Geller 2000, p. 216; Samuels 2000, pp. 11 et seq.). Only the members of the Stationers' Company were allowed to print and distribute books in England. Special licensing acts allowed publishers to receive rights to particular registered works, where only lawfully printed books entered a particular register which was again controlled by the Company (Geller 2000, p. 216). The printing of unauthorized books that were not registered was prohibited and sanctioned with imprisonment. Since members of the Stationers' Company were entitled to enter books into the Company's Register, the Company could virtually stop others from publishing copies (Geller 2000, p. 217). Thus, the charter of the Stationer's Company ultimately granted a monopoly right regarding the printing of books. However, when the last of these licensing acts expired in 1694 the British Parliament was reluctant—despite all lobbying by the Company—to renew the acts. Instead, the parliament enacted the Statute of Anne in 1710 that “for the first time protected the rights of authors rather than publishers of books” (Samuels 2000, p. 12). The rights created under the Statute of Anne provided authors with quite extensive protection of their works for a period of 14 years with an optional renewal for another 14 years and hence a maximum protection of 28 years. Nevertheless, the rights were also restrained under the statute by further specified limitations by which the statute followed other objectives like continued creation of useful literature or the advancement and spread of education. Consequently, granting a property right to the author was seen to follow a greater concern in the sense that the incentive provided to the author was also to serve the interest of society (Samuels 2000, p. 12; Raskind 1998, p. 478). Finally, the

³⁵Some authors date the Statute of Anne back to 1709. See for example Varian (2005). The confusion, however, roots in a change from the Julian to the Gregorian Calendar. The “British Calendar Act of 1750. Implemented Across The Years 1751, 1752, and 1753” (www.exit109.com/~ghealton/y2k/br1752a.html) induced a change in date marking the first day of a year. While before 1751 (Julian calendar) the turn of the year was fixed to March 25, the act changed this date to the first of January. So the date of the passing of the Statute of Anne on 25 March 1709 became finally 25 March 1710. See also Samuels (2000) on page 13.

³⁶Also in other European countries similar developments came to recognize the need to regulate the market of books. As such, the French crown, for example, entitled publishers with temporary privileges to print and sell designated books (Geller 2000, pp. 217 et seq.). In this regard, especially Italy should be stressed as being among the pioneers in these developments. Here, the city of Venice was first to provide printers and publishers with such privileges in 1469, which made Venice to be the first European publishing centre with more than eight million books being printed in the second half of the sixteenth century. See Grendler (1975).

British system became the role model of copyright law. However, with its design and emphasis varying especially between common versus civil law countries.

In the end of the eighteenth century copyright legislation started to spread to the European continent and Colonial America. In Colonial America a resolution was passed by the Congress in 1783,³⁷ recommending that each of the 13 states adopt a law regarding the right to copy (Samuels 2000, p. 13; Raskind 1998, p. 478). Shortly after the so-called Connecticut's Statute of 1783, 12 states (excluding Delaware) adopted the copyright statute (Joyce et al. 2010, §1.03 B). The US Constitution of 1787, finally, gave the Congress the power to grant a temporary monopoly to authors and inventors to their respective writings and discoveries to "promote the progress of science and useful arts" (*US Constitution: Article 1, Section 8*).³⁸ Before, the protection of "literary property" had been a matter for the states (Ginsburg 1990, p. 999). The first national copyright law was signed by George Washington on May 31, 1790, with "An Act for the Encouragement of Learning"³⁹ (Samuels 2000, pp. 13 et seq.). As a matter of fact, the US Copyright Act of 1790 was almost verbatim copied⁴⁰ from the Statute of Anne and granted to each author of a map, chart⁴¹ or book the right to exclude others from printing, reprinting, publishing or vending their work for two terms of 14 years (Samuels 2000, p. 12). In continental Europe, the evolution of national copyright law came to be distinguished from its Anglo-American counterpart.⁴² The French Copyright Laws⁴³ of 1791 and 1793 did not only emphasize the "economic rights" that were recognized by the statutes of the common law jurisdiction, but added a second "autonomous set of non-transferable prerogatives identified as 'moral rights'" (Raskind 1998, p. 478). In particular, the "moral right" (*droit morale*) in French Copyright Law strengthened the personal

³⁷The act was passed by the General Court of Connecticut under the title "An Act for the Encouragement of Literature and Genius" and is often referred to as the Connecticut's Statute. Other Pre-Constitutional statutes include for example the Massachusetts Statute among others. Several of these statutes prior to the US Constitution of 1787 particularly emphasize the "personal claims" of authors (Ginsburg 1990, p. 999; Joyce et al. 2010, §1.03 B).

³⁸See <http://www.law.cornell.edu/constitution/articlei#section8> (last accessed on September 1, 2014).

³⁹By sharing almost the same title with the Statute of Anne, the closeness of both copyright laws seems to be already at hand.

⁴⁰However, the US Copyright Law particularly emphasizes a purely utilitarian character for its copyright law. See also Sect. 2.1.2.

⁴¹Thus, the US Copyright Law added maps and charts as subject matters for copyright protection as compared to the Statute of Anne.

⁴²See Geiger (2010a) on the influence of the Statute of Anne in France.

⁴³Some historians see in the French decree of 1793 the world's first true Copyright Act. Bently (2010) argues that the Statute of Anne was referred to as a "right in copies" rather than a "right to copy" and hence applied rights to the printing and reprinting of books, but not copyright (Bently 2010, p. 9). Nevertheless, the influence of the British statute on other countries is uncontested (Cornish 2010; Geiger 2010a; Deene 2010). See the contributions of Bently and Kretschmer (2013) on <https://copyrighthistory.org> (last accessed on September 1, 2014) for a comprehensive collection of key documents and commentaries on the history of copyright.

claims of authors by emphasizing the intellectual and personal relationship of an author to her works. The French extension of copyright law included especially the right to the integrity of a work⁴⁴—meaning the right to object any alteration, distortion or mutilation of a work (Raskind 1998, p. 478). The rights of publishers and the society as a whole were conceived as derivative rights. Besides, the duration of the copyright was extended to last for the authors' life time plus 10 years. In Germany the history of copyright eventually begins with enacting a copyright law in Prussia in 1837. The duration was first set at 10 years. In 1845 copyright legislation extended the term to authors' life plus 30 years. It was not until 1870, though, that the North German Confederation (Norddeutscher Bund) enacted a formal copyright regarding literary works, illustrations, musical compositions and dramatic works. After the formation of the German Empire in 1871, the copyright law of the North German Confederation was adopted and extended by complementary laws in 1876 (Eger and Scheufen 2012a, p. 165). Similarly to the French Copyright Law, the German copyright legislation emphasizes besides a property right in intellectual products (Immaterialgüterrecht) a separate set of rights referring to aspects of "personality" (Persönlichkeitsrecht), where the latter corresponds largely to the French "droit morale" dogma. Most importantly, a work is perceived to be personal in the sense that a work is infused with the personality of the author (Gassaway 2002, p. 40). As a result, the Anglo-American Copyright Law (Common Law System) and the Copyright Law in Europe (Civil Law System) followed different paths in the development of a copyright legislation, since the European model more specifically addresses an author's right to her personal creation.⁴⁵ Accordingly, the European system is often referred to as the "authors' and neighbouring rights" model (Raskind 1998, p. 478; Towse 2006, p. 84).

By the end of the nineteenth century copyright legislations had been enacted by an increasing number of jurisdictions. By then, a demand had been developed to coordinate copyright legislation at the international level. This is for at least two reasons: First, large differences between national legislations regarding its subject matters and its scope induced an environment of legal uncertainty for authors and publishers, respectively. Second, the territorial character of copyright, i.e. protection for "natives" only, gave rise to economic uncertainty as soon as cross-border transactions of copyrighted works became more important during the industrial revolution and hence the development of global markets.⁴⁶ After a period of bilateral

⁴⁴Besides, the "moral rights" emphasize the right of attribution as well as the right to have a work published anonymously or pseudonymously.

⁴⁵This difference is closely related to the philosophical debate between the utilitarian and the natural right school. See Sect. 2.1.2 for a discussion.

⁴⁶In this regard, again technological change plays an important role. In particular, the invention of the steam engine by James Watt allowed for a reduction in transportation costs. In this context, Geller (2000) emphasizes that already during the eighteenth century "English novels crossed the Atlantic by steamship to be pirated in cheaper editions on the mass market in the United States" (Geller 2000, p. 233).

agreements between individual states, ten countries⁴⁷ formed a union for the protection of the right of authors in their literary and artistic works in 1886 in Berne, Switzerland—the so-called “Berne Convention”.⁴⁸ In essence, the Convention was supposed to foster the governing of copyright at the international level in two ways: First and foremost, by signing the agreement each member state agrees to give the citizens of other member states the very same level of copyright protection as it gives to its own citizens. This general condition is referred to as the principle of national treatment (Berne Convention, Article 3-5). Second, the Convention sets a framework of minimum standards for copyright protection that may be interpreted as a signal for the copyright legislation in the countries, e.g. the minimum level for the duration of copyright is set at authors’ life plus 50 years.⁴⁹ However, the “rule of shorter term” under Article 7(8) specifies that “unless the legislation of that country otherwise provides, the term shall not exceed the term fixed in the country of origin of the work” (Berne Convention, Article 7(8)), meaning that an author may not be granted longer copyright abroad than in his home country. Moreover, Article 9(2) of the Convention introduces a guideline for possible limitations or exceptions of copyright. In particular, the reproduction of a work is permitted “in special cases, provided that such reproduction does not conflict with a normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the author” (Berne Convention, Article 9(2)) and is referred to as the (Berne) Three-Step Test.⁵⁰ In the following years the Convention was subsequently re-negotiated in 1896 (Paris), 1908 (Berlin), 1928 (Rome), 1948 (Brussels), 1967 (Stockholm) and 1971 (Paris) and amended on September 28, 1979.⁵¹ Today 166 countries have signed the treaty, with a vast increase especially in the last two decades and the US joining on March 1, 1989.⁵² Despite its general vision to become a universal framework, the Berne Convention remained a rather European initiative until at

⁴⁷Belgium, France, Germany, Great Britain, Haiti, Italy, Liberia, Spain, Switzerland and Tunisia. Liberia was the only state that did not ratify the treaty in 1887.

⁴⁸The full title is “Berne Convention for the Protection of Literary and Artistic Works”. See http://www.wipo.int/treaties/en/ip/berne/trtdocs_wo001.html (last accessed on September 1, 2014).

⁴⁹Nevertheless, the treaty does not delegate legislation to a supranational body. In particular, Article 2 of the Convention specifies that “it shall be a matter for legislation in the countries of the Union to determine the protection to be granted” (Berne Convention, Article 2).

⁵⁰We will later reflect on the three-step test more carefully when analyzing the options for introducing OA publishing in the international copyright law framework. See Sect. 4.2.1.

⁵¹For instance, the three-step test under Article 9(2) of the Convention was introduced after the congress in Stockholm in 1967.

⁵²While in 1970 the Convention counted 58 contracting parties, the number increased steadily to 70 in 1980, 83 in 1990, 147 in 2000 and 166 in 2013 (last accessed on September 1, 2014). See http://www.wipo.int/treaties/en/statistics/StatsResults.jsp?treaty_id=15&lang=en (last accessed on September 1, 2014). See http://www.wipo.int/treaties/en/ShowResults.jsp?lang=en&treaty_id=15 (last accessed on September 1, 2014) for a detailed list of the contracting parties of the Berne Convention.

least the mid twentieth century.⁵³ As a matter of fact, two distinct systems coexisted: The Berne Convention (European System) and the Buenos Aires Convention (Inter-American System).⁵⁴ Signed by the US and nineteen Latin-American countries⁵⁵ in 1910, the Buenos Aires Convention (Article 3) introduced the obligation to a state that all rights are reserved as a necessary condition for maintaining copyrights (Lipszyc 2010, p. 386). Under the Convention copyright protection was granted for the shorter of the terms of either the protecting country or the source country of the work—referred to as the rule of the shorter term, stated in the Articles 6 and 7 of the Convention. However, besides the vague formulation of several provisions of international copyright law, the Convention particularly suffered from the absence of Argentina and Chile until the 1950s (Lipszyc 2010, p. 386). By then the UNESCO had developed a Universal Copyright Convention that was adopted in Geneva in 1952, giving those countries some form of multilateral copyright protection that still disagreed with the “European model” of the Berne Convention. In particular, the US followed a totally different path than the Berne Convention, since US copyright protection was subject to the registration of a work at the Copyright Office and granted for a fixed and renewable term. The Berne Convention instead asked for the very opposite. The need for several structural modifications in their copyright law and economic interests made the US to refuse to join the Berne Convention (Eger and Scheufen 2012a, pp. 165 et seq.). Besides, several other countries—especially the developing world and the Soviet Union—believed that the Berne Convention and the extensiveness of the copyright system would primarily benefit the western world. The conflict of the different copyright systems, finally, lead to the foundation of the World Intellectual Property Organization (WIPO) in 1967.⁵⁶ The WIPO was generally thought to mediate between the conflicting interests and quickly turned to a specialized agency of the United Nations in 1970. The fact that the US finally refrained from their “registration constraint”, joining the Berne Convention in 1989, may be seen as a success of the WIPO initiative. With the advent of the World Trade Organization (WTO) in 1994, the basic aspects of the Berne Convention were integrated to Article 9 of the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) and became binding to all member states of the WTO. The TRIPS-Agreement as such does not only cover copyright issues, but sets standards

⁵³The fact that a French statesman and literate, named *Victor Hugo*, initiated the development of the Berne Convention emphasizes the French influence to formulate the principles of the agreement in Berne. As such, the “moral rights” principle is explicitly accounted for under Article 6 of the Convention.

⁵⁴See Lipszyc (2010) for a review.

⁵⁵Argentine Republic, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, Guatemala, Haiti, Honduras, Mexiko, Nicaragua, Panama, Paraguay, Peru, Salvador, Uruguay and Venezuela. For the Convention see http://ipmall.info/hosted_resources/lipa/copyrights/The (last accessed on September 1, 2014).

⁵⁶The WIPO is based in many respects on the United International Bureaux for the Protection of Intellectual Property that was set up in 1893 when the bureaus that administered the Berne and Paris Convention had merged. See May (2009) on page 17.

and requirements for all forms of (formal) intellectual property rights, including patents, copyrights, trademarks etc. In particular, the minimum standard for the duration of copyright protection is set at author's life time plus 50 years, where no registration is required. Most importantly, the TRIPS-Agreement ensures that no discrimination of states with respect to non citizens is possible, as each state has to treat its citizens and the citizens of other TRIPS countries likewise.⁵⁷ Despite all efforts in harmonizing standards for copyright protection, national legislations remain to be responsible for its design. As a consequence, there is no such thing as an international copyright law.

In the twentieth century the terms of copyright was steadily extended by national (and international) reforms, regarding both the duration and the scope of copyright. In particular, two general reasons abound for these changes in copyright law: (1) technological change and (2) rent-seeking. First and foremost, the development of new technologies to produce, copy and distribute copyrightable works induced a new environment for competition between the original and the copy. The technology to print books remained rather unchanged until the twentieth century. The rotary press had mechanized the process of automatic paper delivery. However, most of the known printing technologies still required laborious workflows, like the type setting and particularly the transfer of ink to paper by pressing the paper against a plate (Samuels 2000, p. 17). In 1938, Chester F. Carlson introduced a fully new printing technology with the dubbed xerography (from Greek *xeros* "dry" and *graphia* "writing"), an electrostatic dry-printing process.⁵⁸ Nevertheless, the new technology was first rather thought of as an alternative printing process until the advent of the Xerox 914 copier in 1959, easing the making of copies of printed material dramatically and triggering a series of significant revisions in the copyright laws. While the impact of the photocopier turned out to be rather small for traditional entertainment works,⁵⁹ like books (novels etc.), newspapers, magazines etc., it revolutionized the dissemination of scholarly journals by copying en masse from library collections.⁶⁰ The conflict between journal publishers and

⁵⁷See http://www.wto.org/english/tratop_e/trips_e/t_agm0_e.htm (last accessed on September 1, 2014) for the TRIPS-Agreement.

⁵⁸In particular, Carlson combined electrostatic printing with photography. As such, the dark parts of a picture are negatively charged, whereas the light parts lose their charge by exposure to light. Consequently, the positively charged toner powder sticks to the darker parts of the picture. In a final step, a heater seals the toner on the paper (Samuels 2000, p. 18). See Owen (2004) for a comprehensive analysis of the impact of the Xerox copier for the printing industry.

⁵⁹The reason is rather easy to grasp. First, the copying of a complete novel requires labor and time and hence is not necessarily cheaper than its original. Second, the quality of a copy of a novel and the convenience for its use make it to be no perfect substitute for the original.

⁶⁰Sung et al. (2009) provide with general facts and a description of the impact of photocopying on the copyright industries as well as with links for further reading via <http://blogs.ischool.berkeley.edu/i103su09/structure-projects-assignments/research-project/projects-and-presentations/copyright-and-the-advent-of-xerox-machines/> (last accessed on September 1, 2014). In Sect. 2.1.2 we will further elaborate on this issue by introducing to the economics of copying.

libraries, finally, led to a number of court cases⁶¹ and especially induced the codification of the “fair use” principle in the US Copyright Act of 1976.⁶² In particular, the Copyright Act of 1976 specified exemptions for the photocopying by libraries and archives and allowed the copying of copyrighted works for teaching purposes. Besides, the US adopted a unitary term based on the date of the author’s death.⁶³ Subsequent amendments to US Copyright Law—like the Semiconductor Chip Protection Act of 1984 and the Vessel Hull Design Protection Act—further broadened the scope of copyright protection to new subject matters as new product technologies were developed.⁶⁴ The advent of the internet (more particular the World Wide Web) in 1989, finally, somehow revolutionized the making of copies and hence the conditions in the markets of copyrighted works. In particular, two parameters in the copyright “equation” changed: On the one hand, with digital technologies the costs for making copies are almost zero.⁶⁵ On the other hand, the quality loss of earlier copying technologies is eliminated with significant consequences for competition, since original and digital copy are perfect substitutes (Wiebe 2010, p. 323; Eger and Scheufen 2012b, p. 50). The new challenges in the digital environment quickly induced a debate in the international forum. In 1996, the WIPO adopted two treaties, the “WIPO Copyright Treaty”⁶⁶ and the “WIPO Performance and Phonogram Treaty”,⁶⁷ clarifying copyright on the internet (Samuels 2000, p. 111). On the national level, the US Digital Millennium Copyright Act of 1998 and the EU Copyright Directive of 2001 followed. In particular, the changes made in copyright legislation were related to the development of “Digital Rights Measurement” (DRM) technologies (sometimes referred to as technological protection measurements), i.e. software which virtually enables the controlling of

⁶¹These include e.g. *Williams & Wilkins Co. v. United States* (1973), *Encyclopedia Britannica Educ. Corp. v. Crooks* (1978), *Basic Books Inc. v. Gnomon Corp.* (1980) and *Addison-Wesley Publishing Co. v. New York Univ.* (1983).

⁶²The “fair use” doctrine was not unknown before its codification under section 107 in 1976, as the federal courts had applied the common law form of this doctrine since the 1840s.

⁶³As previously mentioned, the US Copyright Law was until then based on a fixed initial and renewal term.

⁶⁴The development of new technologies also affected other fields of the copyright industry. As such, the invention of the video recorder (Betamax) and later DVD players induced significant changes for the film industry. See for example Gordon (1982) for an analysis of the Betamax Case or Mortimer (2007) on DVDs. Similarly, the impact of new tape recording technologies (records, tapes, CDs and later MP3) for the music industry. In this context, the US Congress introduced a new amendment in 1972, involving two separate copyrights for music (for the composer) and sound recordings (for the record company). See also Samuels (2000).

⁶⁵This aspect is closely connected to a third new environmental condition in the digital era, i.e. digital information can be distributed without the need for any physical medium such as paper or a CD (Eger and Scheufen 2012b, p. 49).

⁶⁶See http://www.wipo.int/export/sites/www/treaties/en/ip/wct/pdf/trtdocs_wo033.pdf (last accessed on September 1, 2014).

⁶⁷See http://www.wipo.int/export/sites/www/treaties/en/ip/wppt/pdf/trtdocs_wo034.pdf (last accessed on September 1, 2014).

access, use and trading of electronic content (like movies, music etc.). In this regard, the changes made to national copyright strengthened the positions of DRM advocates and made it illegal to bypass DRM technologies, even “if the person doing so would otherwise have the legal right to access the information behind the digital fence” (Corrigan and Rogers 2005, p. 168). This “fencing off” of information, finally, induced a debate upon the “information commons” and countermovements to the traditional copyright system with the “Creative Commons” (CC) movement introduced by Lawrence Lessig or the Google Book Search Project.⁶⁸ Second, also rent-seeking motives have significant explanatory power for the gradual strengthening of copyright legislation. As such, the extension of the copyright term to 70 years after the author’s death with the Sonny Bono Copyright Term Extension Act (CTEA) in 1998, is a classical example of how economic interests can shape the law. In particular, the Sonny Bono Act was decisively lobbied by Walt Disney, as the first drawings of Mickey Mouse would have entered the public domain in 2003 without an extension of the copyright term (Lévêque and Ménière 2004, p. 68; Corrigan and Rogers 2005, p. 164).⁶⁹

In the end, this series of revisions on the copyright laws—due to new media and copying technologies as well as lobbying activities by the parties who are exploiting the copyrights of the creator (publishers etc.)—have extended the copyright in terms of its duration and its scope, especially regarding the capable subject matters. After extending the copyright term in 1831 to 28 years, renewable for 14 years, and again in 1909 to 28 years, renewable for another 28 years, the US Congress extended the term for 1 year each year after 1962. From 1976, the renewal term for “old” works was set at 47 years and extended to life of the author plus 50 years for new copyrighted works (Corrigan and Rogers 2005, p. 156). With the Council Directive 93/98/EEC of October 29, 1993,⁷⁰ the EU harmonized the term to author’s life time plus 70 years, with the US following with the Sonny Bono Act in 1998. At the same time new subject matters had to be defined due to the development of new media technologies. As the principle of copyright originates in the technology of the printing press, the “right to copy” was first considered as a right in books only. The enactment of the US Copyright in 1790, finally, broadened the scope to include also maps and charts. Throughout the nineteenth century several new subject matters were added by the US Congress, broadening the scope e.g. to musical compositions (1831), dramatic works (1856) and photography (1865). With the Copyright Act of 1976, finally, the Congress included a phrase to more abstractly define copyright’s subject matter as “works of authorship” (17

⁶⁸See Sect. 2.1.3 for a discussion of the counter movements as reactions to the broadening of copyright protection. For a seminal work on the idea of “Creative Commons” see Lessig (2004). A law and economics analysis of the Google Book Search Project is Mueller-Langer and Scheufen (2011b).

⁶⁹On rent-seeking in the UK copyright legislation see Mitra-Kahn (2011).

⁷⁰See <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31993L0098:EN:HTML> (last accessed on September 1, 2014).

U.S.C. §102(a)) regardless of a work's medium of fixation (Hardy 1999, pp. 663 et seq.). By this generalization of copyright's subject matter, the Congress strived for a more flexible framework to respond to and accommodate the development of new media technologies. As a matter of fact, section 102(a) of the US Copyright does not specifically name computer programmes as a subject matter of copyright. The German "Urheberrecht", in contrast, explicitly lists computer programmes under §2(1) UrhG. Thus, the copyright has always been a matter of change as new technological developments and economic interests kept driving the system to continuously adapt to changing conditions. Economic theory, meanwhile, has been providing with the necessary tools for evaluating the impact of possible changes for the effectiveness of the copyright system. The following section shall introduce to the analysis of copyright from an economic perspective.

2.1.2 The Economics of Copyright and Copying

Legal and Economic Philosophy

The conceptual origins of the system of intellectual property rights in general and copyright in particular can be traced back to the debate of primarily two distinctive philosophical schools: The utilitarian school (Jeremy Bentham, John Stuart Mill, Thomas Jefferson) and the natural rights school (John Locke and Jean-Jaques Rousseau).⁷¹ While utilitarianism emphasizes the relevance of the interplay between individuals—standing on the shoulders of one another—and hence perceiving an invention as a product by society in a cumulative process of producing intellectual assets, the natural rights approach strengthens the position of each individual's contribution in advancing the knowledge base of the society. As such, the utilitarian notion of intellectual property argues that intellectual creations are creations by society and as such should serve the interests of all members at large (Granstrand 1999, p. 24; Menell 1999, pp. 130 et seq.; Eger and Scheufen 2012a, p. 153). Or as Thomas Jefferson frames it: "Inventions then cannot, in nature, be a subject of property. Society may give an exclusive right to the profits arising from them, as an encouragement to men to pursue ideas which may produce utility, but this may or may not be done, according to the will and convenience of the society, without claim or complaint from anybody." (cited in David 1993, p. 26).⁷² Following the line of argumentation by the natural rights school, in contrast, each individual is perceived of having a natural claim to the results of his or her physical or mental labour (Granstrand 1999, p. 23). And hence, there is essentially no difference

⁷¹The list may be even broadened to contributions by David Hume, Immanuel Kant and Friedrich Hegel among others, who indirectly influenced our present notion of intellectual property (Granstrand 1999, pp. 23 et seq.). See Palmer (1990) for further reading.

⁷²See also Granstrand (1999) on pages 34 et seq.

between property in physical or intellectual creations. In particular, the natural rights school argues that any result from individual's intellectual labour may and should be perceived as an extension of that individual's identity, "an extension of which the individual could not be deprived by others, and especially not by societal institutions" (Granstrand 1999, p. 23). Or more generally speaking and following Locke (1689) it is the notion that all humans possess an unquestionable property in their own person (Menell 1999, p. 157).

In the historical evolution of the copyright system—as well as of the system of intellectual property rights in general—the influence of natural rights arguments eventually declined. As a consequence, the "modern" twentieth century copyright system can be seen rather in the tradition of the utilitarian argumentation, while considerable differences regarding its perception can be found in comparing the Anglo-American and European copyright system. As such, especially for the development of copyright in the United States the utilitarian notion of intellectual property became an essential principle. Menell (1999) cites a report by the Congressional Committee on the 1909 US Copyright Act which thoroughly highlights the utilitarian position by the Congress. Here it says: "The enactment of copyright legislation by Congress under the terms of the Constitution is not based upon any natural right that the author has in her writing [...] but upon the ground that the welfare of the public will be served [...] by securing to authors for limited periods the exclusive right to their writings." (*US Copyright Act (1909)*). The copyright systems in Continental Europe, in contrast, explicitly account for the natural rights argument as they eventually distinguish between economic and moral rights.⁷³ As previously outlined, the German copyright legislation, for instance, distinguishes between a proprietary right in the intellectual product—*Immateriälgüterrecht*—and a separate form of rights regarding the creator's personality—*Persönlichkeitsrecht* (Raskind 1998, p. 478).

Theoretical and Normative Foundation

A systematic analysis of copyright from an economic theory perspective, meanwhile, had long been a neglected topic in economics and is much younger as the system of copyright itself. While first comments on copyright issues may be traced back to early economists like Adam Smith,⁷⁴ it was not until the seminal work by

⁷³Throughout the process of harmonization by means of international treaties the differences have clearly diminished over time (Reichman and Okediji 2012, p. 1377). In this regard, Goldstein (2001) finds that the differences between the two traditions are more in terms of emphasis than outcome. Ginsburg (1990) highlights that both the French and US Copyright law exhibit a mix of both traditions.

⁷⁴Smith approaches the issue of intellectual property by making a case for temporary monopolies that are justified and reasonable (in contrast to "unjust" monopolies). In his "Lectures on Jurisprudence 11" Smith already distinguishes the exclusive privileges, like patents and copyrights, from other forms of property. In this context, Smith postulates that "the author of a new book

Plant (1934) to establish a research field that may be called “the economics of copyright” (Towse et al. 2008, p. 1; Watt 2004, p. 153). As such, Plant’s article marks the launch of a broad scientific literature on the economics of copyright, whereas subsequent works primarily approached the issue of copying.⁷⁵ However, only after the early 1970s—where the work by Breyer (1970) may be perceived as most influential—economists started to regularly contribute to the advancement of the literature in the economics of copyright and copying, respectively (Gordon and Bone 1999, p. 192).⁷⁶ As most important for laying the foundation of modern copyright analysis can be seen the work by Landes and Posner (1989). The authors provide a first comprehensive analysis regarding the various doctrines in copyright law, making reference to explicit aspects like the duration and the scope of copyright protection as well as on the fair use doctrine in US Copyright Law (17 U.S.C. §107). Furthermore, Landes and Posner (1989) were first to distinguish analytically between the fields of the “Economics of Copyright” and the “Economics of Copying”. A distinction that has since been frequently adapted and has become standard in the economic literature.⁷⁷ The usefulness of distinguishing between copying and copyright issues is based upon a difference in the analytical focus of each field. While the economics of copyright primarily focuses on problems referring to the legal framework and hence a legal question, the economics of copying analyze problems which are related to advances or changes in the technologies for reproduction (Towse et al. 2008, p. 9).⁷⁸ Despite the breadth of topics being analyzed in the literature on the economics of copyright and copying, two general approaches can be highlighted to form the basis of the traditional analytical framework in law and economics: (1) The Public-Goods Approach and (2) The Property-Rights Approach.

In this regard, the economic rationale for providing legal protection for works in arts, literature and science by means of copyright—as an exclusive right—is

has an exclusive privilege of publishing and selling his book [...] as an encouragement to the labours of learned men.” (Meek et al. 1762, p. 83). See Hadfield (1992) on pages 11 et seq. for a comprehensive review on the history of the economics of copyright.

⁷⁵Obviously, copying is closely related to copyright, since copyright is designed to control copying (Watt 2004, pp. 159 et seq.). Seminal works on the economics of copying can be traced back to articles by Hurt and Schuchman (1966), Breyer (1970), Novos and Waldman (1984) and Johnson (1985). A more recent article on this topic is Varian (2005). See Towse et al. (2008) for a comprehensive literature review.

⁷⁶Note that by then relevant contributions on the economics of public goods by Samuelson (1954) and information economics by Arrow (1962) and Stigler (1961) provided with the necessary tool kit for approaching the problems in the market for information goods. See also Gordon and Bone (1999) on page 192.

⁷⁷See for example Varian (2005).

⁷⁸Obviously, copyright and copying are often interrelated and can, as such, not always be strictly distinguished. Nevertheless, a distinction between the two fields seems useful and reasonable as they differ with respect to the problem that is being analyzed. As such, we will proceed by also distinguishing between the “Economics of Copyright” and the “Economics of Copying”.

ultimately determined by the nature of the market for information goods. From an economics perspective information goods exhibit the classical characteristics of a public good,⁷⁹ characterized by non-rivalry and non-excludibility (Gordon and Bone 1999, pp. 191 et seq.; Koboldt 1995, p. 133; Eger and Scheufen 2012a, p. 154). Accordingly, book contents or musical compositions, respectively, can be copied or used at will without affecting the amount and quality being consumed by others (non-rivalry). In addition, information goods impose high transaction costs for identifying and excluding other market participants from consuming the good (non-excludibility).⁸⁰ That is, the public good character finally induces an incentive to consume the information good without bearing the (sunk) costs of production—in the sense of the opportunity and risk costs by the creator. As a consequence, a fundamental free-riding problem (prisoner’s dilemma)⁸¹ emerges with “copying” as the dominant strategy, inducing a divergence between private and social incentives to innovate and hence an underprovision of information goods, as the creator has to fear not being able to recoup her sunk costs of actually producing the good (Arrow 1962). The concept of copyright law offers a solution to eventually overcome this free-riding problem by providing the creator of a work with an exclusive right that enables her to exercise control over the use of her works and hence providing ex ante with an expectation to internalize an economic surplus—or at least to recoup her sunk costs.⁸² That is, the Public-Goods Approach provides a general argument to justify the “privatization” of intellectual creations by referring to the need to create incentives for creative activities (Demsetz 1970; Gordon and Bone 1999, pp. 192 et seq.).

The Property-Rights Approach complements the Public-Goods Approach by providing a tool box for designing the principles of copyright law in view of the criteria of economic efficiency. As conceptually most important in this context may be seen the works by Ronald H. Coase and particularly the Coasian bargaining

⁷⁹The public good character of information (goods) was first to be acknowledged by Saint Augustine, sometime between 391 and 426 (Wills 1999) and later by Thomas Jefferson in 1813 (Dalrymple 2003).

⁸⁰The notion “public good” in this context may sometimes also be perceived somewhat misleading. In this regard, Landes and Posner (2003) emphasize: “It sounds like a good produced by the government as opposed to the private sector. That is true of public goods that people cannot be excluded from having the benefit of even if they don’t contribute to the cost of supplying the goods. The clearest example is national defense. Many public goods, however, including intellectual property, are excludable in the sense that it is possible to condition access to them on payment.” (Landes and Posner 2003, p. 14).

⁸¹The prisoner’s dilemma game in the market for information goods is analyzed by Gordon (1992a). She shows that in a game where the players may choose between creating a work on their own or simply copying the work of another, the strategy of copying will strictly dominate the creation. As a result, no work is being created. See also Gordon and Bone (1999) on pages 192 et seq.

⁸²Obviously, the economic arguments for justifying copyright protection are closely related to the theory of externalities (Heide 2004). In this respect, the public good theory is only part of a much broader theory of externalities.

solution. Following Coase (1960), an efficient allocation will be reached by a bargaining between the players (creator and user) regardless of whether player 1 or player 2 is entitled with a property right. As a necessary condition for this market solution, however, Coase (1960) emphasizes the relevance of well-defined property rights and the absence of transaction costs. But, especially in the market for information goods transaction costs are by nature rather large. The reason and inevitable consequences are rather easy to grasp. Imagine a situation in which every market participant would be entitled with a natural right to copy a work. Then, every creator would have to contract with everyone who had access to the work. Obviously, a situation that would not lead to a Pareto-efficient outcome due to high information costs (Gordon and Bone 1999, pp. 193 et seq.; Gordon 1992b, 1989). However, the same may be true when multiple and dispersed users of a work would have to contract with the creator to receive the permission to use a particular work. A condition that may explain why in copyright law the tradition of enforcing a “property rule” is frequently replaced by the concept of a “liability rule”, where a user may use the work without permission and instead pays a fair price that is set by a third party (collecting society). In this context, Calabresi and Melamed (1972) showed that with increasing transaction costs it may be reasonable to shift from a “property rule” toward the more flexible concept of a “liability rule”.⁸³ The principle of “flat tax compensation” (pauschale Vergütung) can be viewed in the tradition of this basic economic insight. The Property-Rights Approach, finally, provides with the instruments to specify the property rights between the rightholder and the user, giving reasoning to a welfare maximizing design of the copyright system. As such, the economic analysis of copyright law (as well as the system of intellectual property rights in general)⁸⁴ manifests in a classical maximization problem from a social welfare perspective (Eger and Scheufen 2012a, pp. 155 et seq.). The optimal design of copyright law will be addressed in the following section.

The Economics of Copyright

The basic economic intuition behind copyright is somewhat twofold. First, a state may choose to grant an exclusive right to the author of a work to overcome the market failure associated with the provision of a public good. Thus, copyright is designed to control copying activities to prevent from an underprovision of information goods in the context of the free-riding environment. The exclusivity in this regard gives the rightholder the market power (in the sense of a temporary

⁸³Especially due the technological revolution ushered in by the internet and the vast increase in opportunities in the digital environment, a debate on a shift in paradigm in favor of a “liability rule” has been induced (Eger and Scheufen 2012a, p.155). See also Lemley and Weiser (2007).

⁸⁴For an overview on the economics of intellectual property rights see for example Besen and Raskind (1991) besides Posner (2005). A brief overview in German language is provided by Mueller-Langer and Scheufen (2011a).

monopoly) to increase the product prices above marginal costs and hence to generate profits which serve as an incentive to engage in creative activities *ex ante*. Second, as copyright restricts access to information goods, it creates economic costs (dead-weight loss). Copyright therefore should be limited in so far, as to guarantee that creative works are created at minimum costs (Landes and Posner 1989; Besen and Raskind 1991, p. 5). In particular, easing access to basic or prior creations in a cumulative environment of knowledge production—i.e. in an environment where each creator is “standing on the shoulders of giants” (Turnbull 1959)—may be relevant to foster not only innovations today, but also tomorrow. In so far, copyright should be limited to foster the dissemination of new ideas, building the foundation of creative works tomorrow (Eger and Scheufen 2012b).⁸⁵

The economic analysis of copyright, finally, reflects on this basic trade-off between creating incentives to innovate (*benefit argument*) and restricting access to information (*cost argument*).⁸⁶ In this context, Landes and Posner (1989) describe copyright law as an attempt to balance a conflict of interest between the parties on the supply side (creators and publishers) and the demand side (users) of the market. While the creator and publisher of a copyrightable work seek for an extension of copyright protection to force an internalization of the economic surplus, the consuming entities aspire for a cheap or free (open) access to the good. From economics perspective this conflict of interests manifests in a classical maximization problem, maximizing the difference between the benefits of providing incentives for authors to create a work and the costs associated with a limitation of its access (Landes and Posner 1989, p. 326; Mueller-Langer and Scheufen 2011a, p. 140; Hilty and Peukert 2004). In designing an optimal copyright law, a social welfare analysis generally distinguishes three dimension of copyright protection: (1) Duration, (2) Depth or Height and (3) Breadth (Varian 2005, p. 124; Watt 2004, p. 157; Eger and Scheufen 2012a, p.157).⁸⁷

The first dimension, the duration of copyright refers to the time horizon and hence the amount of years copyright protection is enforced for (Watt 2004, p. 157). In this respect, most (developed) states of the world grant copyright protection for 70 years after the death of the author (Eger and Scheufen 2012b). Recent reforms in copyright law—e.g. the US Sonny Bono Copyright Term Extension Act (CTEA) in 1998 or the guidelines 93/98 EWG for a harmonization of EU copyright law to 70 years—have extended the length of copyright protection. The minimum standard

⁸⁵As such, copyright can be seen as a form of social contract, where the public agrees to a law to ensure creative development for the benefit of following generations, but the costs of having to pay a higher price today.

⁸⁶Accordingly, the economics of copyright joins the tradition of the IPR argumentation and hence trades off between underproduction and underconsumption of information goods due to the divergence between private and social incentives to innovate (Arrow 1962).

⁸⁷There are several works on the optimal structure of copyright, where some papers even predate the seminal paper by Landes and Posner (1989) by a few years. See, for example, Novos and Waldman (1984), Pethig (1988) or Besen and Raskind (1989). See also Liebowitz and Watt (2006) on pages 516 et seq. for a review.

has been fixed by international agreements (e.g. TRIPS, WIPO Copyright Treaty) to author's life time plus 50 years. Landes and Posner (1989) provide with insights to the intertemporal trade-off coinciding with a lengthening of copyright protection. As copyright grants an exclusive right to the author—thereby preventing others from copying—copyrightable works are removed from the public domain for the duration of copyright, increasing the costs for subsequent or derivative works. Accordingly, an extension of its duration is only reasonable if its discounted positive effects for creations today (static efficiency) manage to compensate its negative effects on the incentive to create (subsequent) works tomorrow (dynamic efficiency), marking a social equilibrium where marginal costs equal marginal benefits.⁸⁸

Both latter dimensions refer to the depth and breadth of the protection, clarifying the aspects of a work that are protected and giving advice to uses being considered as an infringement of copyright law (Watt 2004, p. 157; Landes and Posner 1989, pp. 347 et seq.). As the depth of copyright is concerned, the law provides only protection to the expression, but not the idea expressed by the author. We have seen that copyright protection is much narrower than a patent, since a patent grants a monopolistic right regarding a (complete) technology and hence the idea of an innovation. In contrast, following section 102(b) U.S.C. the same idea may be expressed differently without infringing copyright law. With other words: The ideas expressed in this chapter on the economics of copyright in my own words does not state an infringement of the copyright by the pioneers in this field.⁸⁹ Whereas copying a portion or the whole article without reference to the pioneering authors would turn me to an infringer.⁹⁰ The breadth of copyright is mainly limited by fair use in US Copyright Law (17 U.S.C. §107) or in the European Copyright Laws (in Germany: §§44-63 UrhG) by a catalogue of exceptions.⁹¹ As previously outlined these limitations and exceptions refer to certain special cases which allow to use a work without the permission of the rightholder.⁹² We have seen that especially

⁸⁸As such, the considerations on the optimal duration of copyright just follows the basic arguments expressed in the model of Fig. 2.1.

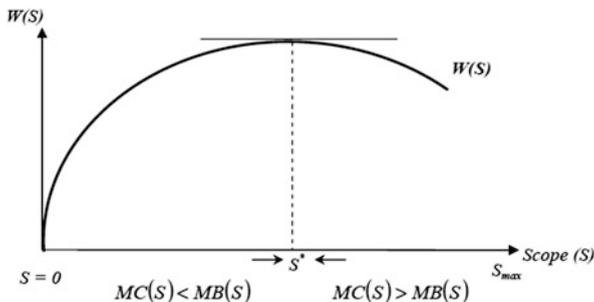
⁸⁹Obviously, especially in science, where priority to a discovery, eponymy and hence citations are considered as the currency of the market, a reference to prior and seminal works is essential and still required. However, this is rather an issue of the moral or personal rights of the author but copyright. See Sect. 2.2 on the economics of science. A seminal work on the economics of science is Merton (1973) (earlier works include Merton (1948) and Merton (1968)) besides Freedman (1960).

⁹⁰See Landes and Posner (1989) on pages 349 et seq. for a discussion on “distinguishing ideas from expression”.

⁹¹The fair use doctrine (US Copyright Law) and the catalogue of exceptions (European Copyright Laws) follow various public goals such as freedom of speech, educational and equality of access as well as issues referring to market failures. See e.g. Hugenholtz (2001).

⁹²Following the Fair Use Doctrine (17 U.S.C. §107) four statutory factors help courts to assess whether a use is considered to be “fair” and hence non infringing: (1) the purpose and character of the use, (2) the nature of the copyrighted work, (3) the amount and substantiality of the portion used and (4) the effect of the use upon the potential market for and value of the copyrighted work. See Gordon (1982) on the economics of fair use. Mueller-Langer and Scheufen (2011b) provide a recent analysis of the Google Book Search Settlement under fair use considerations.

Fig. 2.1 The optimal scope of copyright (Source: Mueller-Langer and Scheufen (2011b))



for educational purposes—like research and teaching—a general “fair use” defense is implemented. In scientific research, especially the rights to cite and adopt expressions from other authors constitutes a crucial barrier of copyright protection, explicitly taking into account the cumulative character of scientific research and facilitating the creation of derivative works. Other categories refer to issues like the private copy, home recording of musical compositions, parody, criticism, comment, news reporting, archiving etc. In the literature the dimensions of depth and breadth are often bundled under the title “copyright scope” (Watt 2004, p. 157).

In this respect, some simple economics may eventually explain the underlying rationale regarding the optimal scope of copyright protection.⁹³ Landes and Posner (1989) already emphasize the relevance of a minimum standard of copyright protection. However, a “too much” of protection—in the extreme case considering every use as an infringement of copyright—forces a situation where an extension in the scope of copyright would lead to marginal costs (MC) exceeding marginal benefits (MB). The economics of this rationale are just straightforward. To see this, consider the following simplified argument. Let $W(S)$ denote the social welfare subject to the scope (S) of copyright, where $\frac{dW(S)}{dS} \geq 0 \forall S \in [0, S^*] \wedge \frac{dW(S)}{dS} < 0 \forall S \in]S^*, S_{max}]$, with S^* revealing the optimal scope satisfying for a maximization of the social welfare $W(S)$. Accordingly, in S^* we satisfy the condition that $\frac{dW(S)}{dS} = 0$ and hence that $MB(S^*) = MC(S^*)$. Geometrically this optimal scope is displayed by the horizontal tangent to $W(S)$, leaving us with a maximum in S^* since $W(S)$ exhibits the characteristics of an inverted u-shape. Figure 2.1 illustrates this rationale in the optimal scope of copyright (Mueller-Langer and Scheufen 2011b, pp. 13 et seq.).⁹⁴

Obviously, there is an optimal scope of copyright, since the social welfare is not strictly increasing in S , such that $S^* < S_{max}$. Consequently, a use will be expected to infringe copyright as long as a protection of this particular use comes along with a positive net gain in social welfare. In contrast, the scope is not broadened to uses that are to the right of S^* , since an extension would cause a decrease in social welfare

⁹³The following (simplified) model is taken from Mueller-Langer and Scheufen (2011b).

⁹⁴Note that the graphic as well as the technical arguments break things down to an ideal type of the relationship between scope and economic returns.

from copyright. Accordingly, an adjustment and optimal definition of the scope of copyright will have to balance the costs and benefits following the economics of this simple model. A closer look at the rationale already emphasizes the differences in the flexibility regarding the two competing copyright principles. While the fair use doctrine (17 U.S.C. §107) in the US Copyright law explicitly accounts for a balancing of the costs and benefits of a single use of copyright material in the context of a case by case analysis, the European tradition explicitly lists the exceptions which are considered as “fair uses”. Thus, the US fair use doctrine is much more flexible. A condition that particularly gains relevance in an environment of vast structural changes in the presence of technological change.⁹⁵

However, consciousness is needed regarding the calculus of an optimal design of copyright as each of the three dimensions may not be viewed separately. In particular, a comprehensive analysis will have to consider all possible interdependencies between these dimensions. An optimal mix of all dimensions involved will then create another dilemma in copyright law due to high information needs as well as a very fine line between the relative optimality of the three dimensions (Watt 2004, p. 157).⁹⁶

The Economics of Copying

While the Economics of Copyright approaches the problems associated with the supply of information goods by analyzing the effects of public intervention (copyright law) only, the Economics of Copying more generally addresses the effect of copying in an environment of technological change, providing with a framework to specifically discuss alternative mechanisms to copyright law (Handke 2010, p. 31; Towse et al. 2008, p. 9). As such, the Economics of Copying builds a bridge to the changing environment for copying as new technologies are developed. The previous reflection on the history copyright clearly emphasizes that the development of new copying technologies—since the advent of the Xerox copier in 1959—have eased the copying of copyright material in terms of costs, quality and time effort. Given this framework, the Economics of Copying analyzes the relation between the right holders’ costs for generating copies of a given work, and the costs of unauthorized copying (Handke 2010, p. 31).⁹⁷ In this regard, Besen (1986) shows that unauthorized copying may lead to a loss of social welfare, if economies of

⁹⁵See Sect. 2.1.3 for a closer look.

⁹⁶Watt (2004) (on pages 157 et seq.) provides a formal model on the optimal design of copyright, clearly emphasizing a dilemma with respect to the optimal mix of the three dimensions in copyright law. In particular, he shows that there is no unique solution for minimizing the deadweight loss subject to the participation constraint of the author.

⁹⁷Another trade-off that is being analyzed in the literature on the economics of copying can be traced back to the nature of the cost structure of information goods, i.e. copyrighted works are characterized by high fixed costs of creating the work in the first place but rather low marginal costs of generating multiple copies of a given work. See particularly O’Hare (1985) and Pethig

scale can be observed in the copying of a copyrighted work. However, this trade off may be less important in an environment of digital technologies, where the costs of copying are reduced to virtually zero.

In particular, the works by Stan Liebowitz in the 1980s led to a somehow revolutionary notion of copying activities.⁹⁸ While until that time copyright was rather perceived as the battle against (unlawful) copying of works of authorship,⁹⁹ Stan Liebowitz forwarded an idea that was entitled indirect appropriability and seen as an alternative to the public intervention in the market for information goods (Handke 2010, p. 33).¹⁰⁰ Historically seen, the papers by Liebowitz and his followers may be seen as a reaction on the vast increase of court cases after the advent of the Xerox 914 copier in 1959. The new copying technology most significantly affected the market for academic journals as it became common practice to copy single journal articles or complete issues en masse from (university) library collections.¹⁰¹ Liebowitz (1985) argued that particularly in the case of academic journals¹⁰² the ability of photocopying increased the value of a given piece of work and hence the willingness to pay. As a matter of fact, he showed that unauthorized copying could eventually allow sellers of originals to capture this additional value by the simple

(1988). A more recent analysis is Pollock (2007). See also Handke (2010) on pages 31 et seq. and Towse et al. (2008) on pages 9 et seq.

⁹⁸Liebowitz (1981) published a booklet in 1981 that created the fundament for a series of papers, analyzing the impact of reprography on the copyright system. The concept of indirect appropriability follows particularly from Liebowitz (1985, 1986). See also Liebowitz (1983).

⁹⁹Liebowitz (1985) provides with a quote by the publishing house “Williams and Wilkins Company” to illustrate the notion of publishing houses towards the impact of photocopying on the journal publishing market. They argue that “uncontrolled photocopying is largely responsible for the deaths of two journals [...] and if the condition is allowed to continue, many more will either go out of business or be published under government subsidy.” (Liebowitz 1985, p. 956; as quoted in Thatcher 1978, p. 324). As we will see later, this topic has not lost its topicality in the digital (Open Access) environment.

¹⁰⁰Liebowitz’s famous paper on indirect appropriability was published in the *Journal of Political Economy* in 1985. Together with the papers by Novos and Waldman (1984) and Johnson (1985) in the same journal and around the same time, the paper induced a flood of new papers reflecting on the economics of copyright and copying (Watt 2005, p. 1). In 2005, the *Review of Economic Research on Copyright Issues* published a series of papers in a symposium for the twentieth anniversary of the concept of indirect appropriability. Here, Liebowitz (2005), Johnson and Waldman (2005), Johnson (2005) and Boldrine and Levine (2005) reflect critically upon indirect appropriability in the context of technological change.

¹⁰¹See Sect. 2.1.1 for short review. In particular, see Liebowitz (1981, 1983) besides chapter one in Samuels (2000).

¹⁰²Liebowitz (1985) was rather sceptical about a generalization of the concept of indirect appropriability for other markets but the academic journal market. He argues: “The copying of other media may or may not have impacts similar to those found for photocopying. Only case-by-case empirical investigations of institutions and markets can discover the impacts of these other forms of copying.” (Liebowitz 1985, p. 956; as cited in Watt 2005, p. 1). As such, Johnson and Waldman (2005) give examples of markets in which the concept of indirect appropriability is probably not applicable.

means of price discrimination. Price discrimination is possible (and profitable) whenever suppliers can ask for different prices from different types of consumers. However, only if different groups of consumers can be distinguished according to their varying willingness to pay for a single product (Liebowitz 1985; Watt 2005, p. 1).¹⁰³ In this regard, journal publishers typically apply a dual pricing strategy as they distinguish between two sub-markets—institutional and individual subscribers. As the elasticity of demand varies significantly between the two markets, price discrimination¹⁰⁴ is feasible and profitable for journal publishers, where institutional subscription prices are significantly higher than those for individual subscriptions (Joyce and Merz 1985, p. 274; Rosenbaum and Ye 1997, p. 1611). This is for at least two reasons: First, the fraction of the budget spent for journal subscription is much higher for an individual. Second, the hypothesis of an inelastic demand for libraries is supported by a tendency of libraries to be complete rather than selective, while individuals may revert to the sharing or renting option as a close substitute (Joyce 1990, p. 1127). In addition, and this is the point Liebowitz (1985) made, the copying option of single journal articles may induce an increase in the value a library assesses to a journal and hence creates an argument for a higher pricing scheme for institutional subscribers.¹⁰⁵ Furthermore, also tying and bundling strategies abound as common pricing schemes in the academic journal market. Here, major publishers offer bundles of different journals to libraries, bundling across journals and also across print and electronic versions (Edlin and Rubinfeld 2005, p. 441). In this regard, a membership to the American Economic Association (AEA), for example, includes a subscription to a bundle of seven journals, e.g. *The American Economic Review*. Furthermore, the annual subscription fee for the bundle of journals is \$420 for only print subscription, \$840 for print subscription and an electronic site license and \$665 for only electronic site license, respectively (Mueller-Langer and Watt 2010, p. 54). The advent of digital copying eventually even created a “better” fundament for such price discriminating practices.¹⁰⁶ All

¹⁰³For illustrating his point, Liebowitz (1985) uses a simple analogy, referring to the automobile company “Ford” that sells cars to both private individuals and automobile rental companies (e.g. Hertz, Sixt etc.). He shows that by accounting for the resale value of used cars (or authorized copies) when they purchase them, Ford (or the copyright owner) can indirectly profit from later uses of their sold cars (or unauthorized copying).

¹⁰⁴Following a classification by Pigou (1920), economists typically distinguish between three types of price discrimination: (1) first-degree or perfect price discrimination, (2) second-degree price discrimination or non-linear pricing and (3) third-degree price discrimination. As such, dual pricing falls under the headline of third-degree price discrimination as the supplier chooses a pricing scheme in accordance to the Amoroso-Robinson relation. See Robinson (1932).

¹⁰⁵Phillips and Phillips (2002) note that individuals will treat a single journal as a private good in economic terms, while the journal will take on attributes of a public good for institutional subscribers who will make the journal available to a pool of users.

¹⁰⁶This builds the bridge to the serials crisis that was outlined in our introduction and hence stresses one of the reasons for rethinking the general institutional structure of the academic journal publishing market. On price discrimination in the digital economy see Fudenberg and Villas-Boas (2012). On bundling strategies for information goods see Choi (2012).

of this reveals that copying may not have a detrimental effect on publishers, since indirect appropriability by means of price discrimination as well as exposure and network effects may eventually create a baseline to indirectly appropriate revenues from unauthorized copying activities.

In 2000 the concept of indirect appropriability was revived by Hal Varian, who applied the concept to the environmental conditions in the presence of digital copying (Towse et al. 2008, p. 10). By analyzing the market for information goods—including not only journals but also books, computer software, music and videos—Varian (2000) identifies three general circumstances where the opportunities for sharing may increase the profits of a content producer: (1) when the marginal cost of producing a piece of work are above the transaction costs of sharing, (2) when in a situation of low transaction costs of sharing the work is rarely used, and (3) when the producer can identify different types of consumers (high versus low value users) to segment the sharing market. By modelling the different cases identified by Besen (1986), Varian (2005) analyzes the impact of digital technology copying on the price setting for a content supplier (Towse et al. 2008, p. 10). More recent papers have analyzed the relationship between optimal copyright duration and price discriminating practices in a digital environment. As such, Meurer (2001) and Gordon (1998) analyze the effect of copyright law on the ability of suppliers of copyrighted works to price discriminate. Yuan (2010) studies the opposite case, simulating the optimal copyright duration subject to price discrimination.

As a consequence, the broad discussion on the relationship between copyright law (plus alternative mechanisms) and new developments of copying technologies already reveals the importance of research to understand the interdependencies of various different issues. Nevertheless, recent reforms in copyright law still rather suggest that an extension of copyright law is required to motivate authors to be creative in an environment of digital copying. These developments and the reactions by academia shall be discussed in the following subsection.

2.1.3 Copyright in Transition: Challenges in the Information Age

Copyright and Technological Change

Obviously, it is the development of new technologies for the production, copying and distribution of copyrightable works that has been driving major changes, especially in the last quarter of the history of copyright law. The series of reforms—starting with the “WIPO Copyright Treaty” and the “WIPO Performance and Phonogram Treaty” in 1996 that were implemented into national legislation shortly after—have decisively extended the rights of the owners of copyrightable works. All efforts said to adapt copyright law to the challenges of an environment that is characterized by digital media technologies. In particular, with the Digital Millennium Copyright Act (DMCA) in the US and the EU Copyright Directive

of 2001 a new section was added to copyright that specifically addresses the use of Digital Rights Management (DRM) technologies, i.e. a class of access control technologies that may be used by sellers of digital content or devices to effectively control the access, use and distribution of digital content (e.g. sound recordings, movies etc.). As such, the DMCA¹⁰⁷ adds a new chapter 12 to title 17 of the US Code, clarifying that (any)¹⁰⁸ circumvention of technological measures is prohibited by the law and criminalizes the production and dissemination of technologies whose primary function aims at circumventing technological measures that protect copyrighted content. These developments in the “right to copy” can generally be understood as a reaction towards the new digital environment. As previously outlined, this digital shift induced considerable changes in the cost structure and the quality of copying technologies (Eger and Scheufen 2012a, pp. 169 et seq.).

In academia, these developments have induced a heated controversy regarding the effectiveness and general role of intellectual property rights.¹⁰⁹ In particular, several economists have expressed their rather critical perspective on the attempts to further strengthen copyright law. First and foremost, on 20 May 2002 seventeen leading economists¹¹⁰ (including five Nobel laureates) presented an *amicus curiae* brief as a response to the Sonny Bono Copyright Term Extension Act (CTEA), revealing doubt on the reasonability of the 20-year copyright term extension for existing and future works of authorship. Akerlof et al. (2002) stress that it is highly unlikely that the economic benefits from extending copyright under the CTEA outweigh its costs. In particular, the authors argue that the CTEA provides no significant effect for creating new works, while significantly increasing the social cost of the temporary monopoly. In addition, they stress that the copyright term extension increase the cost regarding the production of new creative works that make use of existing materials. Liebowitz and Margolis (2005) respond to the arguments forwarded by Akerlof et al. (2002), highlighting that a more comprehensive analysis of the effects of copyright reforms still require “an examination of empirical magni-

¹⁰⁷See DMCA (1998) on pages 3 et seq.

¹⁰⁸Generally the new chapter of the DMCA (1998) “divides technological measures into two categories: measure that prevent unauthorized access to a copyrighted work and measures that prevent unauthorized copying of a copyrighted work.” (DMCA 1998, pp. 3 et seq.) Accordingly, only the circumvention of the first category is generally prohibited, while the second category may be subject under fair use. Furthermore, section 1201 specifies certain saving clauses and exceptions (DMCA 1998, section 1201(a)(1)(B)-(E)).

¹⁰⁹Note that there have always been critics of a system of intellectual property rights as a means to overcome the market failure associated with information goods. As such, already Plant (1934) questions the need for any type of a legal protection system, pointing to first mover advantages as means of appropriation for the creator. Also Hurt and Schuchman (1966) and Breyer (1970) question the effectiveness of copyright law.

¹¹⁰Including *George A. Akerlof, Kenneth J. Arrow, Timothy F. Bresnahan, James M. Buchanan, Ronald H. Coase, Linda R. Cohen, Milton Friedman, Jerry R. Green, Robert W. Hahn, Thomas W. Hazlett, Scott Hemphill, Robert E. Litan, Roger G. Noll, Richard Schmalensee, Steven Shavell, Hal R. Varian and Richard J. Zeckhauser*. Henceforth Akerlof et al. (2002).

tudes that no one has fully undertaken.” (Liebowitz and Margolis 2005, p. 457).¹¹¹ Png and Wang (2009) and Hui and Png (2002) provide empirical evidence for the impact of the extension of the copyright term and the European Rental Directive, finding no statistical robust effect of either reform on the production of new movies. Other critics even advocate the abolition of copyright. Among the strongest advocates of an abolition of copyright law in the information age are Boldrine and Levine (2002)¹¹² and Lessig (2001).¹¹³ Besides, Ku (2002) suggests—by analyzing the Napster case—that intellectual property may even be counterproductive with respect to digital products. As a consequence, there has been a series of important papers that have been analyzing possible alternatives to a system of copyright law.¹¹⁴ Especially rewards and prizes as a means of public funding have been discussed as a reasonable alternative for creating incentives for innovative endeavour. In this context, Shavell and van Ypersele (2002) show that a system of intellectual property rights has actually no fundamental social advantage over a reward system. Furthermore, the authors advocate an optimal reward system, i.e. a system where the creator may choose between rewards and intellectual property rights, over a pure IPR regime.

Other scholars, in contrast, have still been advocating a strong copyright protection. Some authors have even raised arguments to support a copyright that virtually lasts forever (Landes and Posner 2002, 2003; Turnbull 1998). In particular, Landes and Posner (2002) argue in favour of an infinitely renewable copyright that would just follow the rationale applied in the system of trademark law.¹¹⁵ The authors provide with empirical evidence for the fact that for the vast majority of existing works a renewal of the copyright term would eventually not be valuable. This is particularly true for works of low average commercial value. With a short but renewable copyright term these kind of works would enter the public domain at an early stage, minimizing access, transaction and administration costs. Those few works that are of high average commercial value for its right holders would, however, retain their value by remaining in copyright protection forever (Landes

¹¹¹On the need for further empirical research on the impact of copyright law for the supply of creative works see Png (2006) as well as Handke (2010).

¹¹²See also Boldrine and Levine (2005) and Boldrine and Levine (2008). For a discussion see McManis (2009).

¹¹³Shavell (2010) very recently raised the question of an removal of copyright law concerning academic works. We will elaborate further on Shavell (2010) in Sect. 2.3. At its heart, this thesis extends on Shavell’s work, seeking to give answers to the question whether an abolishment of copyright for academic works is socially desirable.

¹¹⁴Varian (2005) gives a short overview on the alternatives to a legal copyright protection. Gallini and Scotchmer (2002) discuss different scenarios where legal mechanisms are not superior to its alternatives. See also Liebowitz and Watt (2006) for a discussion of the alternatives to copyright in the music industry.

¹¹⁵A trademark is generally granted for a term of 10 years with the option to indefinitely renew the term. That is, a trademark can virtually last forever. See Besen and Raskind (1991) or Mueller-Langer and Scheufen (2011a) on the economics of trademark law.

and Posner 2002, p. 41). Thus, Walt Disney would have had the chance to retain their rights for Mickey Mouse without affecting all other existing works and especially without any means of rent-seeking. Such a system would, however, be especially valuable with respect to out-of-print and orphan works, i.e. works for which locating the copyright owner has become prohibitive or even impossible. Orphans would enter public domain after expiration of the shorter copyright term as there would be nobody to renew its term. Obviously, an indefinitely renewable copyright could eventually provide with a solution to a still prevalent dilemma that has been repeatedly tried but failed to be solved by the US Congress in the history of copyright legislation (Eger and Scheufen 2012b).¹¹⁶

Recent Developments in the Economics of Copyright

Recent developments in the economics of copyright have taken a more critical position regarding the relevance of copyright for creating incentives for creative activities. In this context, Towse (2001) shows that there is eventually no empirical evidence that copyright actually increases the earnings of the creators in the creative industries. Moreover, she emphasizes the inequality in the level of earnings by particularly pointing to two observations: First, due to market power and better opportunities for rent-seeking by publishers we can observe a distortion in the distribution of income for the benefit of the publishers. Second, several examples of industries abound where monetary rewards seem to be rather negligible for stimulating creativity. As such, Towse (2001) shows that only a small minority of superstars generates a considerable income from copyright,¹¹⁷ while the vast majority of creators can hardly earn their living (Eger and Scheufen 2012a, pp. 171 et seq.).

Obviously, there must be other factors but financial gains that explain why somebody engages in creative activities. In this regard, especially the role of so-called intrinsic motivation or other motivational factors (like reputation or social recognition) has been emphasized to better explain certain behavioural patterns in some creative industries or branches. Most important in creating a new notion of knowledge production has been the emergence of so-called open peer-production models, like the movement of Open Source Software (OSS) or Wikipedia.¹¹⁸ The

¹¹⁶See also Mueller-Langer and Scheufen (2011b) who discuss the orphan works dilemma in the light of the Google Book Search Project.

¹¹⁷On the economics of superstars see Rosen (1981).

¹¹⁸With the organization model “the bazaar” as opposed to the traditional model of “the cathedral” Raymond (1999) turned the notion of intellectual property on its head. Raymond (1999) argued that the development of a new software should proceed on the basis of an open peer production process (“the bazaar”) where everybody can contribute voluntarily by writing a part of the source code. The traditional model, in contrast, favoured an organization in small isolated teams on the basis of secrecy—very much like the organization in the building of a cathedral. Raymond (1999) somehow revolutionized the way of thinking organizational procedures by adding a third model to

general idea of the OSS movement is that large and informal groups of volunteers contribute freely to the development of a software by directly affecting the source code of the software and providing with incremental improvements or solving specific problems, i.e. they provide with certain modules (pieces of code) of the software. Recalling the rationale of a system of intellectual property rights, finally, Lerner and Tirole (2002) frame the revolutionary character of the OSS model by asking: “Why should thousands of top-notch programmers contribute freely to the provision of a public good?” (Lerner and Tirole 2002, p. 198). The literature on the economics of OSS finds three basic arguments for answering this intriguing question. Accordingly, programmers may be motivated by means of (1) simple enjoyment, (2) career concerns and (3) ego gratification (Lerner and Tirole 2002, pp. 212 et seq.; Rossi 2006, pp. 16 et seq.).¹¹⁹ Accordingly, a programmer may be motivated to add a new algorithm to the source code of a software as she simply enjoys her doing. In particular, the open character of OSS may give her the freedom to choose a project that matches her skills and interests. An aspect that is certainly more difficult to satisfy in the hierarchy of a software company. Lerner and Tirole (2005) emphasize that “a ‘cool’ open source project might be more fun than a routine task.” (Lerner and Tirole 2005, p. 58). Both later aspects—career concerns and ego gratification—may be summarized under the single heading of “signaling incentives” (Lerner and Tirole 2002, p. 214). This signaling may take on two dimensions: On the one hand, a contributor may have the chance to demonstrate her ability in solving specific problems in software engineering and hence may profit from participating by means of better labour market opportunities. On the other hand, a programmer may seek to just be a member in the group of peers and hence seeks a “socialisation” in the community of software experts. The assignment of the term “hacker” may then be understood as a form of honor within the OSS community (Lakhani and Wolf 2005, p. 5; Scheufen 2011, p. 5). As a matter of fact, the “peer-production” model may create a baseline for questioning the role of intellectual property rights (copyright) in stimulating creative endeavour. Nevertheless, in evaluating the role of copyright in a certain market, like the market of science, an understanding of the characteristics and motivational conditions of this particular market is crucial. Consequently, the following section shall introduce to the economics of science and create a fundament to further assess the role of academic copyright as an incentive to do research.¹²⁰

the traditional models of the market and hierarchy as distinguished by Coase. This “paradox” is also referred to as the Coase’s Penguin. Here, the penguin refers to the mascot of Linux as one of the pioneers in the OSS movement. See Benkler (2002).

¹¹⁹We will later see that the motivation of a programmer to contribute to the production of an Open Source Software correlates considerably with the motivational structure for a scientist. See Scheufen (2011) on the analogy between open source and open access.

¹²⁰Obviously, this book sees itself in the tradition of this literature stream, questioning the role of copyright law in the market for academic publishing.

2.2 The Economics of Science

2.2.1 *The Concept of Science and Scientific Research*

Understanding the Nature and Principles in Science

Science—lat. *scientia*, knowledge or information—can generally be described as the process of human activity to develop or generate an accurate knowledge and understanding of nature, regarding the past, the present and the future (Freedman 1960; Coccia 2006, p. 11). The Webster’s New Collegiate Dictionary defines science as “systemized knowledge as an object of study” or “knowledge covering general truths or the operation of general laws esp. as obtained and tested through scientific method” (Mish 1985, p. 1051). Obviously, science is far from being a static concept, but is rather a matter of continuous modification of or adaption to new environmental conditions. Furthermore, the development of new scientific methods may eventually revert a current state of knowledge as our perception of the world (a theory) may be falsified.

Conceptually, the term science needs to be distinguished from the rather closely related terms of “research” and “scientific research”. In fact, science and scientific research are frequently being used synonymously (Coccia 2006, pp. 9 et seq.). The term research refers to the general process of gathering information. Freedman (1960) describes research as a continued search for knowledge and understanding. It is to be distinguished from science, as research does not necessarily include scientific information and the application of scientific methods. As a matter of fact, the term research is today rather used to mean anything from surfing the web for good sale offers to reading the daily newspaper. In contrast, scientific research adds “scientific” to the expression as it refers to the continuous process of scientific knowledge accumulation by the application of scientific methods. Research as such is then only the first step to scientific research, as it more generally describes the gathering of data. A person reading a book to receive a better understanding of the world is hence not (yet) engaged in scientific research. Recalling the definition of science in general reveals that scientific research is much broader and can be seen as the combination of both science and research. Figure 2.2 shall summarize the conceptual differences of the terms science, research and scientific research.

Obviously, the application of scientific methods is crucial to go the additional step towards scientific research. In this regard, we can generally distinguish between two approaches or types of methods: the deductive approach and the inductive approach.¹²¹ The deductive approach starts at a more general conception of a certain problem and seeks to deduce a more specific information. As such, a general

¹²¹This general differentiation can be traced back as far as to Aristotele (384 BC–322 BC) who was among the pioneers to describe the deductive approach in methodology. Among the first to develop the inductive reasoning was Francis Bacon (1561–1626), where Galileo Galilei (1564–1624) later added the mathematical formalisation. See also Coccia (2006) on pages 11 et seq.

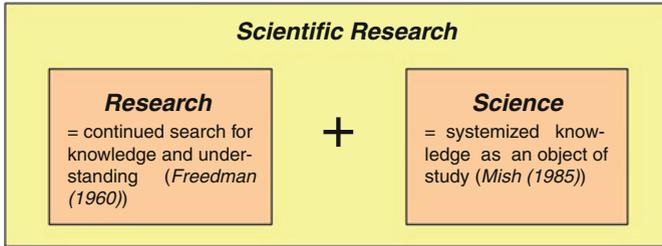


Fig. 2.2 A conceptualization of scientific research

theory states the starting point for deriving hypotheses which are tested against the background of real life observations. The specific information derived from testing our hypotheses is the evidence derived for a falsification or verification, which allows conclusions on the reasonability of the theory being tested. As a consequence, the deductive approach is often referred to as a “top-down” approach. The inductive method, in contrast, follows a “bottom up” approach and hence moves from a specific observation towards a generalization or definition of a theory. As scientific research requires the application of either deductive or inductive reasoning, there are two fields of scientific research that follow the lines of either of the two approaches: basic research and applied research (Godin 2001; Coccia 2006, p. 11). While basic or pure research is conducted solely for the purpose of accumulating and extending on existing knowledge, applied research seeks the resolution of a particular problem. It is important to note that basic research often builds the foundation for applied research, as it provides with a more general understanding of the functioning of the world on which applied researchers can extend on. Notwithstanding the fact that the lines between these two fields can be blurred in some respect, a distinction is reasonable and useful for the purpose of our investigation.

As we seek to understand why scientists do science and whether a certain publishing model matches with the norms and organizational structure in scientific research, it is necessary to clarify the antetype of a scientist that builds the core of our analysis. Most importantly, our analysis will focus on such activities that are not primarily addressed at providing any commercial purpose. That is, we will not be interested in the commercialization of science, like it is often the case for example in the field of biotechnologies. Here patents play an important role to protect one’s intellectual assets.¹²² Consequently, our analysis will primarily address the field of basic research, where a publication rather than a patent application is the output of research. As already outlined in our introduction, we will focus on academic works to which Peter Suber refers to as royalty-free literature. This has two important implications: First, the publisher receives the work from the authors at no costs.

¹²²See e.g. Scotchmer (2011) on the role of patents in university research.

Second, the author should (*ceteris paribus*) be open to the publishing mode (open or closed access) as she is not losing any revenue (Suber 2012, p. 9). It is this type of literature that we have in mind when analyzing the impact of copyright versus open access for the scholarly system as a whole. Before we move on, however, an understanding of the incentives of authors of this type of literature is important.

The following chapter shall first give an overview of the fields of study in the analysis of scientific research before focussing on the reward structure in science and hence the motivational forces of the researcher that we have in mind.

The Analysis of Scientific Research: Fields of Study

The analysis of scientific research covers several fields of study and as such ranges from economics—including pure economics, managerial economics, political economy and economic history—to sociology and philosophy. In “modern” economics eventually three different fields of study can be distinguished.¹²³

First, economists have come to realize the importance of science for the advancement of technologies and hence for economic growth. In fact, modern growth theory emphasizes the role of science and technological innovations for productivity. Especially the works by Romer (1986, 1990) increased the attention to science as the major factor for technological innovations.¹²⁴ In particular, economists in this field have analyzed the relationship between science and technology as well as the role of knowledge spill-overs from science for economic growth.¹²⁵ Also the role of scientists in the industry has been addressed to understand the scientist’s input for creating the capacity of firms in the development of innovations. As such, Cohen and Levinthal (1989) highlight that scientific knowledge is crucial for both

¹²³Coccia (2006) follows the structure of Stephan and Audretsch (2000) and distinguishes primarily ten different fields of study in the economics of science, ranging from the public nature of scientific knowledge to the studies of scientific research and economic growth. In particular, see Coccia (2006) on page 11. Nevertheless, all ten fields may be summarized to cover primarily three different fields of study in the analysis of scientific research.

¹²⁴Earlier models in growth theory especially by Solow (1956) and Swan (1956) were only able to explain the impact of technological progress on the long-run rate of growth by means of a “residual” (Solow residual) as growth was exogenously determined by new technological developments. Endogenous growth theory in contrast endogenizes technological progress by particularly emphasizing the relevance of so called spill-over effects. On the origins of endogenous growth see Romer (1994).

¹²⁵On the relationship between science and technology see in particular Rosenberg (1974), Scherer (1982) and Gibbons and Johnston (1974). Extending on Romer (1990) especially Griliches (1992), besides Jaffe (1989), Audretsch and Feldman (1996) and Acs et al. (1994) analyzed the importance of knowledge spillovers from science for economic growth. Diamond (2004) provides with an overview on Zvi Grichile’s contributions for understanding the economics of technology and growth.

the production of new knowledge and the adoption of external knowledge developed outside of the firm—so-called absorptive capacity.¹²⁶

Second, economists have analyzed the scientific labour market and the human capital embodied in scientists for understanding the labor market conditions in science. Ehrenberg (1992) provides a comprehensive analysis on the main characteristics in the labor market for scientists. By analyzing the determinants that explain the supply and demand in the scientific labor market, Leslie and Oaxaca (1993) review the process of forecasting labor market conditions that help to understand the failure and success of particular research careers.

The third field in the analysis of scientific research addresses the appropriation problem associated with the production of a pure public good. The origins of these studies revert back to the analysis of the nature and the conditions for the production of scientific knowledge. Obviously, the most fundamental characteristic of scientific knowledge is its public good nature. The starting point for analyzing the consequences of the public good nature of scientific knowledge are the seminal works by Samuelson (1954) on the theory of public goods and Arrow (1962) and Stigler (1961) on information economics, laying the foundation for the analysis and consequences coinciding with the provision of public goods. Arrow (1962) particularly emphasizes the consequence of an underprovision of the public good as private and social incentives differ considerably.¹²⁷ More recent works by Callon (1994) and Dasgupta and David (1994) provide with a more differentiated picture on the public good problem, highlighting the role of tacit knowledge (Callon 1994) and arguing that scientific knowledge may only satisfy the characteristics of a public good if they are codified in a manner that can be understood by others (Dasgupta and David 1994). Furthermore, several economists have tried to measure scientific contributions, i.e. the added value a certain publication provides for the progress of science,¹²⁸ and to explain productivity differences among scientists (Allison and Stewart 1974) or over the life cycle of a scientist's career (Levin and Stephan 1991; Diamond 1986). A matter of particular interest in this field of study is to understand the incentives structure in the market of science. This latter issue is of particular relevance for the purpose of our investigation, as we seek to understand what impact a regime change—from a closed access (copyright) to an open access regime—would have on the incentives of scientists to contribute to the advancement of science. As a consequence, the following section provides a brief review on the

¹²⁶See also Mansfield (1995), Nelson (1962), Rosenberg (1990) and Lichtenberg (1988) on the role of scientists and basic research for industrial innovations.

¹²⁷See Sect. 2.1.2 for a review of Arrow's argument and the consequences for a justification of copyright law.

¹²⁸A prevailing method for measuring the impact of a certain publication to the advancement of science is to revert back to citation counts. Nowadays individual contributions are being evaluated by the Social Science Citation Index of Thomson Reuters. See Lindsey (1989) on the use of citation counts for measuring scientific output. See also Garfield (1955) on the foundations. A literature review on bibliometric measures of productivity is provided by Diamond (2000).

general insights gained from the economics of science for understanding the motives of scientists by particularly addressing the reward structure of scientific research.

2.2.2 *The Reward Structure of Science*

Some General Insights

Since our analysis shall later reflect more carefully on the norms, incentives and organizational structure in the market of science, we will first need to understand more about the mechanisms or primarily about the rewards that explain motivational patterns in this market. This section shall provide an understanding of the motives of a scientist to contribute to the advancement of science and hence seeks to answer a compelling question: Why do scientists do science?

In this regard, we may learn from the literature in psychology which generally distinguishes between two types of motivation that drive individual behaviour: intrinsic and extrinsic motivation.¹²⁹ While intrinsic motivation refers to an action that is driven by an inherent satisfaction or simple enjoyment, extrinsic motivation explains any action by means of attaining a certain outcome or reward. As a result, an individual that is intrinsically motivated may engage in a certain activity due to the “fun factor” or the challenge she assigns to her doing. Thus, it is rather the inherent interest in a certain activity than the external prods, rewards or pressure that may explain certain decisions in human behavior (Ryan and Deci 2000, p. 56). As a simple example we may assume that the reader of this section is curious of getting to know about a scientist’s motives. In contrast, any performance that can be summarized as being extrinsically motivated is done to attain some separable outcome (Ryan and Deci 2000, p. 60). In fact, Ryan and Deci (2000) emphasize that eventually the vast majority of all human activities is driven by extrinsic motivation. The authors stress that after the early childhood “the freedom to be intrinsically motivated becomes increasingly curtailed by social demands and roles that require individuals to assume responsibility for nonintrinsically interesting tasks.” (Ryan and Deci 2000, p. 60). Experimental studies eventually show that the relevance of extrinsic motivation increases with each advancing grade in school.¹³⁰ The nature of the rewards that trigger such behaviors may be direct (money, prizes) or indirect (CV-effects, future income). Recalling the earlier example of our reader, an extrinsic motivation to continue with the reading of this section may be to learn new skills because he or she understands the potential value or utility of these skills for

¹²⁹On the theory of motivation in general see Cofer and Apply (1967). For an analysis of extrinsic and intrinsic motivation and their interdependencies in human behavior see Deci et al. (1999) and Deci and Ryan (1985).

¹³⁰See the literature cited in Ryan and Deci (2000).

generating present or future net income, e.g. by an application of these insights to his or her own research or in case of a student to achieve a good grade in her exams.

We apply this more general framework for assessing a typology on the rewards that drive the motivation pattern of a scientist. Thus, we will first look at possible extrinsic motives for doing science before turning to other explanatory factors beyond (monetary) rewards that may be rather explained by intrinsic motivation. After all, a typology of the rewards will summarize the different motives in science and highlight the relationship between and relevance of the intrinsic and extrinsic rationale for the decision to engage in scientific research.

Priority to Discovery: The “Ribbon” and the “Gold”

At the core of any scientist’s endeavour to “do science” is primarily the goal to establish priority to discovery (Merton 1957, 1973). In this context, particularly two “motivational fields” can be distinguished: (1) peer-recognition (“the ribbon”) and (2) monetary rewards (“the gold”).¹³¹ We will see that the boundaries between both types of motives are somewhat blurred and to a large extent build on each other. Accordingly, the rewards scientists may gain from being first to communicate a discovery is the recognition that is assigned to the researcher by her peer group. In its core is the pursuit to accumulate reputation in the eyes of her peers (Stephan and Levin 1992, p. 18). The reputation capital accumulated over a scientific career, finally, seeks two particular purposes.

On the one hand, it may serve the simple means of ego gratification in the sense of a community-based intrinsic motivation.¹³² Here, several forms of recognition can be distinguished. First and foremost, priority to discovery allows for the attachment of the scientist’s name to her discovery and hence leads to the reward of eponymy. In fact, in practice several examples of eponymy abound, such as the Gaussian Curve (Carl Friedrich Gauss, 1777–1855), the Pasteurization (*Louis Pasteur*, 1822–1895) or the Otto engine (*Nicolaus August Otto*, 1832–1891). In economics, the Nash-Equilibrium (*John Forbes Nash Jr.*, 1928–) or the Pareto-Criterion (*Vilfredo Federico Pareto*, 1848–1923) state famous examples for the association of the founder with his or her discovery. Second, prizes and awards state a special form of reward which seek to recognize the role of particular scientists for the advancement in certain disciplines.¹³³ Zuckerman (1992) estimates more than 3,000 different

¹³¹The notion of “the ribbon” and “the gold” follows from Stephan and Levin (1992) and Stephan and Everhart (1998).

¹³²Following Lindenberg (2001), intrinsic motivation can be distinguished in enjoyment-based and community-based intrinsic motivation. We will elaborate on the first one in the next section.

¹³³An award or prize may have different levels of properties. Besides social prestige and recognition, Frey and Neckermann (2009) also highlight that (1) winning an award generally provides its recipient with a “warm glow” or good feeling, (2) awards are conferred by principals whose opinion the agent values and (3) awards provide with monetary compensation or other material benefits. In addition, it is the enjoyment derived from being in competition with other

prizes in the early 1990s in North America alone. Among the variety of different prizes, the Nobel Prize¹³⁴ provides the most prestige to its laureates. In mathematics, for instance, the so-called Fields Medal for mathematicians not older than 40 and the Wolf Prize provide with a similar prestige. Third, a scientist may be rewarded by means of titles and fellowships. As such, the achievements of one's contributions may be honored by the degree of a *doctor honoris causa* (honorary doctorate). Or a scientist may be elected to national or international academic societies—e.g. the Royal Society, the European Academy of Sciences and the Bavarian Academy of Sciences—as a particular form of recognition. Also invitations to keynotes, board memberships and editorial positions at prestigious journals are important aspects.

On the other hand, the reputational capital accumulated over time also acts in pursuance with career concerns. That is, scientists are interested in an impressive curriculum vitae to signal one's status within the scientific community and to receive appointments for professorships at prestigious departments or to be considered for consulting. A necessary premise for establishing priority and hence to gain recognition is by signaling one's contributions and hence skills on the basis of (high quality) publications.¹³⁵ A frequently applied output measurement for assessing and operationalizing a scientist's output is the so-called Social Science Citation Index (SSCI) by Thomson Reuters.¹³⁶ Accordingly, it is not the number of publications but the number of citations assigned to the papers of an individual researcher that counts for indirect benefits which accrue through CV-effects. In the end, this instrument finally helps to distinguish between different types of researchers (good or bad type/high or low quality) and provides the researcher with the opportunity to signal her skills for future job opportunities.

However, the behavioural patterns of our researcher may not only be intrinsically motivated. In particular, the reputational capital or the "reputation-building-claims" (Dasgupta and David 1994, p. 498) accumulated by different means may as well

peers that may provide the participant with pleasure irrespective of outcome (Frey and Neckermann 2009, pp. 76 et seq.).

¹³⁴The Nobel Prize is each year awarded for achievements in physics, chemistry, medicine or physiology, literature and peace. Since 1969 the "Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel" extended the scope of Nobel Prizes with respect to the economic discipline. Each laureate receives a Nobel Prize diploma, a Nobel Prize medal and is endowed with the prize money. Since 2001 the prize is set at (SEK) 10 million per full Nobel Prize. The announcement of the laureates is each year in October. The awards ceremony is held in Stockholm and Oslo (only the Nobel Peace Prize) each year within the week of December 10, which is the day when Alfred Nobel died in 1896. See <http://www.nobelprize.org> (last accessed on September 1, 2014) for more information.

¹³⁵Stephan (1996) highlights that while eponymy or the Nobel Prize is most often beyond the reach for the majority of researchers, recognition by means of publications or citations is in reach for almost every scientist.

¹³⁶The SSCI generally provides with bibliographic and citation information to find research data and to analyze trends, journals and individual researchers. In present (September 2014) the SSCI lists 3,000 of the world's leading social science journals and covers 50 disciplines. See <http://scientific.thomsonreuters.com/products/ssci> (last accessed on September 1, 2014).

be expressed in monetary terms, since the payment schedule in science allows for material rewards like a higher salary or access to research facilities subject to a scientist's reputational standing. In this regard, the nature of science as a "winner-takes-it-all" contest positively defines a payment schedule that consists of two components. First, a fixed flat¹³⁷ salary as a compensation for the risk of not being the winner of the "priority game" and hence as an incentive to actually enter the game. Second, a reward to the winner of the scientific competition in terms of a bonus that is granted subject to the scientist's reputational status. Most important is to understand that the most productive scientists may enjoy substantial salary premiums as a good performance makes the scientist attractive for other institutions (Stephan and Levin 1992, p. 21). That is, salary is at least indirectly related to productivity as more productive scientists are more likely promoted. Nevertheless, there is also evidence for "pay-for-performance" models that are applied at some universities. An extreme example of such a model is the Vienna University of Economics and Business Administration, where faculty members receive a reward payment of 1,000 € for a publication in an "A journal" and 3,000 € for a publication in an "A+ journal" (Frey and Neckermann 2008, p. 2). Even though these "pay-for-performance" models are applied in a far less rigid way at most universities, more successful researchers may still be able to increase their salary by means of bargaining or by receiving outside research funding.¹³⁸ Moreover, Stephan and Everhart (1998) emphasize the opportunity to act as consultant¹³⁹ for private enterprises or as an employed researcher in the industry.¹⁴⁰ As a consequence, every publication and more importantly citations support the scientist in increasing his own market value—not only in reputational but also in financial terms.¹⁴¹ Last but not least, also prizes or awards generally provide its recipients with material rewards. As such, the Nobel Prize is endowed with a monetary reward of (SEK) 10 million (i.e. approx. 1.1 million Euros) per full Nobel Prize.

¹³⁷Dasgupta and David (1987) argue that a discovery that is made a second or third time does not add any value to the social surplus resulting from its first discovery (Stephan 1996, p. 1202). Thus, only a flat salary as an "entrance royalty" allows for an efficient allocation of resources, as only this payment schedule helps in bearing the risks of being a loser in the "priority game".

¹³⁸See e.g. Frey and Neckermann (2008) on the various (monetary) rewards in academia.

¹³⁹A famous example in the economics discipline is *Hal Ronald Varian* who was professor of microeconomics at the University of California at Berkeley before he joined Google Inc. as a consultant in 2002.

¹⁴⁰Of course, the job market opportunities for scientists are much broader and an analysis is in need for a deep understanding about the conditions in the scientific labour markets. As this is beyond the scope of this work, we would like to point to some literature for further reading. Especially the works by *George Stigler* are important to recognize. See Diamond (2005) for an overview on Stigler's contributions. A general overview on the (labour) market for scientists is given by Stephan (1996) on pages 1211 et seq.

¹⁴¹In this respect, Diamond (1986) estimates that the net present value of a 35-year old scientist (in 1994 dollars) ranges from \$2,225 (for a physicist) to \$6,750 (for a mathematician).

Intellectual Satisfaction: Solving the Puzzle

Nevertheless, science is obviously more than just money and prestige. The sociologist Warren Hagstrom explains that “research is in many ways a kind of game, a puzzle solving operation in which the solution of the puzzle is its own reward” (Hagstrom 1965, p. 16; also cited in Stephan and Levin 1992, p. 18). As such, a scientist may have an inherent interest in doing scientific research that may best be explained by intellectual satisfaction. Just like pure intrinsic motivation is mostly a phenomenon observed in childhood, a scientist may be somehow special as her interest is purely driven by the excitement of being at the forefront of a particular research field. In this regard, Hull (1988) understands science as “play behavior carried to adulthood” (Hull 1988, p. 306). Most important is to highlight a form of enjoyment-based intrinsic motivation that drives the decision to contribute in a certain field of interest (Stephan and Levin 1992, p. 18).

In this context, Csikszentmihalyi (1974)—as one of the pioneers in studying the dimension of enjoyment—highlights the importance of reaching a so-called “state of flow” that maximizes the enjoyment derived from solving a particular problem or puzzle.¹⁴² A state of flow is reached if the skills of a particular person are matched to the challenges of a task. A matter of particular relevance to ensure such a matching is the concept of academic freedom, as it provides the scientist with the freedom to choose the optimal zone of activity in which her state of flow is maximized. Thus, puzzling at the forefront of an exciting research field may then provide with positive net benefits or an additional satisfaction far from monetary terms (Scheufen 2011, pp. 3 et seq.).

The Rewards in Science: A Typology

All of the aforementioned aspects greatly reveal the complexity of the mechanisms in and organization of scientific research. The immanent rules to foster scientific progress and to implement measurements for a selection process in the “scientists game” show that our researcher’s motivation is far from being monocausal. In fact, several motivational factors may explain behavioral patterns of researchers in the market of science. Figure 2.3 gives an overview on the main motivational factors (dark grey) and the various rewards (light grey) in science.

In conclusion, a typology on the reward structure reveals three aspects to play an important role regarding the incentives structure in science. First and foremost, the recognition (“the ribbon”) awarded to priority and hence the gained scholarly esteem from being first to acknowledge a certain issue. In this context, scientists

¹⁴²This goes in line with the above mentioned typology provided by Lindenberg (2001) who distinguishes between enjoyment-based and community-based intrinsic motivation. See also Lakhani and Wolf (2005) on pages 4 et seq. On this basis, Scheufen (2011) looks at the parallels in the motivation of scientists and software engineers/ programmers.

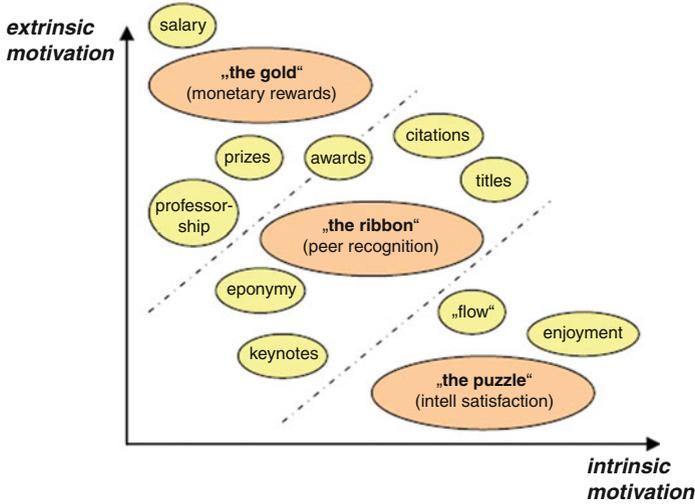


Fig. 2.3 The reward structure in science: a typology

seek to accumulate individual reputation to gain both peer-recognition (community-based intrinsic motivation) and future job opportunities by means of a performance signaling. Second, the monetary rewards (“the gold”) gained in terms of salary bonuses or prize money also reveal extrinsic motivational forces that explain certain behavioural patterns in the scientific labour market. Third, the pure enjoyment or intellectual satisfaction from solving a puzzle or being at the forefront of the “scientific game” are important to note as a pure (enjoyment-based) intrinsic motivational force.

2.2.3 Some Implications: Why do Scientists do Science?

All of the above shows that the truth of why scientists do science is somewhere in the middle. It is neither the money, nor the pure satisfaction or altruistic motives that drive a researcher to write and publish academic works. Scholars write scientific papers because advances in the knowledge within their field advances their careers (Suber 2012, p. xx). Thus, a scholar writes for the impact, not for a direct monetary reward. What is important for the researcher is the reputation and hence the value a journal publication adds to her bibliography. An analysis of the effectiveness of alternative publishing models will have to take these fundamental insights into account. An optimal publishing regime should hence solve a fundamental trade-off: On the one hand, scholars should receive credit for their writings according to their impact for the advancement of science. The performance of a scientist is measured by the sum of academic works weighed with their impact, i.e. the

reputation/ impact factor that is assigned to a particular journal. The environment for individual career advancement is competitive. On the other hand, new findings should be made available immediately and at minimal costs to (1) enable for priority to discovery and (2) maximize knowledge diffusion.

2.3 A Comparison: The Shavell Model and Beyond

In comparing the economic rationale of copyright with the insights from an economic analysis of the reward structure in science, a discrepancy with respect to both mechanisms seems already at hand. While the public good nature of knowledge or information induces a fundamental market failure with respect to rewards that are market based, in science a non-market reward mechanism has evolved, providing incentives for scientists to behave in socially responsible ways (Stephan 1996, p. 1201). In particular, we have seen that copyright seeks to award the creator of a work with a means of appropriating financial returns as an incentive for creative endeavour. In science, however, these monetary rewards seem to be rather negligible as scientists typically earn hardly any royalties from publishing their research results. In science the authors are rather indirectly rewarded as an impressive publication list or citation rate induces “reputation building claims” for future career concerns. As a matter of fact, with reputation rather than royalties stating the relevant currency in the market of science it may be asked whether copyright for academic works is at all reasonable.

This intriguing question, however, has only recently aroused interest in academia. The development of new technologies to produce, copy and distribute academic works in a more and more digital environment has courted resentment in academia with respect to the rights and duties of academic publishers. While copyright seemed reasonable in the past as a lever for the emergence of (commercial) publishers, recent price discriminating practices by commercial publishers may have already become too expensive to bear (Litman 2006, p. 104). Only with the digital revolution ushered in by the internet publishers gained new price discriminating strategies, such as bundling and versioning options with different print and online choices. As already outlined in the introduction,¹⁴³ this development lead to an increase of serial expenditures (serial unit costs) by 273 % (more than 188 %) from 1986 to 2004 (Ramello 2010). As a consequence of this so-called serial crisis, finally, significant cuttings of (university) library subscription portfolios has induced critique in academia. At the same time, however, the advent of the internet gave also birth to a new publishing mode and movement—the Open Access movement—which revived a discussion between publishers, scientific associations and scientists as to whether

¹⁴³See Chap. 1.

the traditional copyright/closed access or the open access model may fit better to the norms, incentives and organizational structure in the market of science.¹⁴⁴

Despite a lively public debate for now more than a decade, a first comprehensive economic analysis on the effectiveness of both regimes was not done until very recently.¹⁴⁵ In December 2009, Steven Shavell published a working paper analyzing the reasonability of a possible shift towards an OA regime in publishing academic works.¹⁴⁶ In a nutshell, Shavell's arguments can be summarized as follows: (1) scientists seek to accumulate reputation which is increasing in readership, (2) readership is higher under open access and hence scholarly esteem, (3) the publication costs due to a shift towards the "author-pays" principle under open access will be covered by most universities, and (4) there are several reasons why a shift towards an open access publishing model will not be smooth without legislative steps (Shavell 2010; Eger and Scheufen 2012b, p. 55). In July 2010, the Society of Economic Research on Copyright Issues (SERCI)¹⁴⁷ published a special issue as a response to Shavell's primer, putting some of Shavell's conclusions up for discussion. In this context, Ramello (2010) provides with an empirical insight to the market structure of the academic publishing market. Obviously, the academic publishing market has exhibited a trend towards a high market concentration, where only a handful of large journal publishers (especially Elsevier, Springer and Wiley Blackwell) have substantial market power. The previously discussed serial crisis may also be seen as a consequence of this development.¹⁴⁸ Furthermore, Mueller-Langer and Watt (2010) re-consider some of Shavell's modelling assumptions which may significantly change the perspective on the conclusions made from the Shavell model. First, the authors discuss Shavell's assumption that scholarly esteem can be proxied by the simple means of readership alone. This would be true if reputation as a function of readership would be strictly increasing for all values of readership (Mueller-Langer and Watt 2010, p. 46). However, a simple consideration may cause doubt on this belief as it is not the readership but the reputation or ranking of the target journal that matters. An author would more likely submit to a well-esteemed journal with a small number of readers than to a low-esteemed journal with a higher audience. Mueller-Langer and Watt (2010) conclude that it is important to analyze the impact of quality-adjusted readership on scholarly esteem. Extending on McCabe and Snyder (2005) they also raise the question whether copyright may be important to establish a certain level of reputation for a journal in the first

¹⁴⁴We will further elaborate on the characteristics and evolution of the Open Access movement in Chap. 3. For a general overview on most recently discussed issues see Eger and Scheufen (2012b).

¹⁴⁵Before the topic had been of interest especially among lawyers. See e.g. the works of Reto Hilty (2006a,b, 2007) as well as Hansen (2005) and Gienas (2008). See Mueller-Langer and Scheufen (2013) for a broad literature review.

¹⁴⁶The paper was later published in the *Journal of Legal Analysis*. Henceforth Shavell (2010).

¹⁴⁷See <http://www.serci.org/> (last accessed on September 1, 2014).

¹⁴⁸We will further elaborate on the market conditions in the market for academic publishing in Sect. 3.1.

place. Taking both consideration into account, Mueller-Langer and Watt (2010) are more sceptical regarding a shift towards a universal open access regime in academic publishing. Second, they argue that an abolishment of copyright may be particularly detrimental for top institutions (like Harvard etc.) with a relatively high publication output since the total publication fees (open access regime) may more than outweigh the savings in total subscription fees (CA regime). Looking at eight top-tier journals in economics and 517 institutions they suggest that publication fees under open access would “punish” any institution above a certain level of research output. As a result, the authors propose a modification of the open access regime which may provide for the best of both the copyright and the open access regimes (Eger and Scheufen 2012b, pp. 55 et seq.). In addition, Watt (2010) emphasizes that an analysis of the possible impact of an elimination of copyright for academic works will have to take into account both the total quantity of publications but also the average quality of the works. In this regard, McCabe and Snyder (2004, 2005) provide with a simple (two-sided market) model to provide evidence for a possible quality degradation of journal content. The authors argue that if journal publishers charge a fee per publication, this is likely to result in a situation of accepting papers that would otherwise not have been accepted. As such, a publisher will have the incentive to accept additional papers to internalize the fees paid by the authors. Obviously, each publisher would accept additional papers as long as the marginal benefits would outweigh its costs. As a result, the set of accepted papers would likely increase and hence cause a degradation of the minimum quality threshold for paper acceptance.¹⁴⁹

Despite the increasing interest in the topic and a flood of various papers analyzing the effects of an abolishment of copyright for academic works ever since Shavell (2010), several questions still remain unresolved. Most importantly, it is still questionable whether online access to journal content has created an environment that supports the demand of scientists for reputation or higher citation counts.¹⁵⁰ In fact, there is still very little and even contradicting empirical findings for these important questions. In this regard, Lawrence and Giles (2000) provide with evidence for a three times higher citation rate on average for open access content. Besides, Eysenbach (2006) applies a longitudinal bibliometric analysis of a cohort of open access and non open access articles, analyzing “The Proceedings of the National Academy of Science” (PNAS).¹⁵¹ He highlights that open access articles are more immediately recognized and cited by a factor of two. Also Norris

¹⁴⁹Similarly Jeon and Rochet (2010) stress that open access induces profit maximizing publishers to set socially inefficiently low quality standards. Also in a two-sided market model, Mueller-Langer and Watt (2012) identify countervailing effects and conclude that also high-quality journals may provide open access.

¹⁵⁰No doubt, access to academic works via the internet has significantly changed the way scientists search and use journal content with important implications for the productivity of both research and economic development (McCabe 2011, pp. 21 et seq.).

¹⁵¹See <http://www.pnas.org/> (last accessed on September 1, 2014).

et al. (2008) find significantly higher citation rates for OA articles, investigating journals in four different disciplines—ecology, applied mathematics, sociology and economics. Similar findings are reported by Hajjem and Gingras (2005) for ten different disciplines, where the advantage of OA ranges between 25 and 250 % by discipline and year. Bernius and Hanauske (2009) extend the scope by investigating the impact of OA on the reputation of scientists in their peer-group, and show that researchers gain in reputation when shifting to the OA mode. In a similar vein, Bernius (2010) reports significant cost advantages of OA. Moreover, he stresses that OA would accelerate the creation and dissemination of scientific knowledge.¹⁵² There are, however, also a few sceptical studies on the readership and citation advantage of open access. Davis (2009) estimates that the citation advantage of open access journals is eventually declining by about 7 % per year and is only 17 % taken all journals together. In contrast, there are also empirical investigations that doubt any citation advantage of open access journals. In this context, Davis et al. (2008) show that open access articles reveal indeed significantly higher download numbers, but with respect to the number of citations there is no such difference between open access and non open access articles. Davis (2011) finds that not the research community may benefit from open access to scientific literature but communities of practice that consume but rarely contribute to the journal content.¹⁵³ McCabe and Snyder (2011) confirm the impression of no significant impact of online access on the number of citations. Using panel data on citations to economics and business journals and controlling for article quality by adding fixed effects, they show that JSTOR may help to boost citations by only 10 % whereas ScienceDirect has no effect at all. Nevertheless, in the light of many prevailing methodological weaknesses, such as the problem of selection bias (Bosch 2009), there is room for further and more comprehensive empirical investigations.

Finally, a last branch of literature has been investigating researchers attitudes towards OA publishing. Bernius et al. (2009) consider all stakeholders (scholars, publishers, libraries and funding organizations) and identify reasons why authors may still prefer the “closed access” mode over OA publishing. Imagining a world of two coexisting regimes (closed and open access), the literature emphasizes that researchers may be locked-in to an inefficient Nash Equilibrium (closed access) due

¹⁵²In a study for the Joint Information Systems Committee in the UK, Houghton and Oppenheim (2010) argue that, in the long run, both OA journals and self-archiving platforms will show positive net benefits.

¹⁵³Interesting in this regard is the potential role of free online access for advances in the standard of living in developing countries. For example, access to biomedical and health literature may allow doctors in those countries to improve important health standards. In this context, the “Research4Life” programme has been launched to provide free or reduced fee access to the literature in agriculture (AGORA), health (HINARI), environmental science (OARE) as well as development and innovation (ARDI). We will assess the potential role of this programme in Sect. 4.2.2.

to the reputation advantage of established closed access journals.¹⁵⁴ Accordingly, a survey analyzing 481 scientists from different disciplines shows that researchers tend to exhibit a “wait and see” attitude toward OA publishing (Mann et al. 2008). Eger et al. (2013, 2014) highlight that this “wait and see” attitude may differ considerably between disciplines and countries, not only depending on aspects that find their origin in the publishing culture but also policy issues that are more or less successful in promoting gold or green OA.

All of the above shows a great avenue for further research—ranging from theoretical approaches for assessing social welfare implications of copyright versus open access in academia to empirical investigations accounting for the effect of online and free online access to the academic literature.¹⁵⁵ Before approaching some of these intriguing questions, however, we will first look at the scientific journal market and the open access movement in science to understand important characteristics and mechanisms that lay the foundations for our analysis.

¹⁵⁴We will further elaborate on the reputation advantage of closed access journals in Chap. 3 by analyzing the impact factor distribution of closed access versus open access venues in different academic disciplines. See also the discussion in Chap. 5.

¹⁵⁵Feess and Scheufen (2013) identify three different lines of research in the OA debate: (i) studies on the economic impacts of alternative publishing models, (ii) studies assessing the effects of open access on readership and citations, and (iii) studies investigating researchers’ attitude and behavior towards open access.



<http://www.springer.com/978-3-319-12738-5>

Copyright Versus Open Access
On the Organisation and International Political Economy
of Access to Scientific Knowledge
Scheufen, M.
2015, XVI, 191 p. 39 illus., Hardcover
ISBN: 978-3-319-12738-5