1

Introduction

1.1 Patenting: an early history

The word *patent* is derived from the Latin *pateo*, meaning to be open or exposed. During the medieval period, 'letters patent' were frequently used by monarchs and other sovereign powers to confer various awards such as lands and titles, or to create new rights and privileges, such as the right to incorporate. These awards were a matter of public record, hence the need for them to be 'patent', that is, open to scrutiny. In time, letters patent were also found to be a convenient way of awarding privileges to foreign craftsmen who were willing to share craft technology. These form the earliest precursors to the modern patent grant; in England the first such examples are the letters of protection awarded in 1331 to Flemish weaver John Kempe, on the condition that he instruct native apprentices.[1] The patent was confirmed by a 1337 statute that accorded safe conduct and the king's protection to any other foreign weavers who wished to settle in England.[2]

These early grants, however, did not stipulate that their holders could *exclude* others from practising a designated craft or technology; the first example of such a patent grant in England occurred over a century later in 1449, when John Utynam (also from Flanders) was granted the exclusive right to use his 'art' of making all colours of glass. The text of Utynam's patent, which was ordered on the direct authority of Henry VI, set out the reasons for this award:

> [T]he said art has never been used in England and John intends to instruct divers lieges of the king in many other arts never used in the realm beside the said art of making glass, the king retains him therefore for life at his wages and fees and grants that no liege of the king learned

[1] Edward Hulme, 'The history of the patent system under the prerogative and at common law', *LQR*, 12 (1896), 142.

[2] Ibid., 143. 11 Edw 3, *c*. 5 (1337).

> in such arts shall use them for a term of twenty years against the will and assent of John, under a penalty of £200.[3]

As with Kempe's patent, the rationale for awarding this privilege was to encourage the importation of new technology. At around the same time in Venice, a second rationale for patents was also beginning to emerge – that by awarding exclusive rights to inventors, patents could encourage the *development* of new technology. The world's first patent law, passed in Venice in 1474, makes this logic explicit:

> There are men in this city, and also there come other persons every day from different places by reason of its greatness and goodness, who have most clever minds, capable of devising and inventing all kinds of ingenious contrivances. And should it be legislated that the works and contrivances invented by them could not be copied and made by others so that they are deprived of their honour, men of such kind would exert their minds, invent and make things that would be of no small utility and benefit to our State.[4]

A century later, the same rationale starts to appear in English patents as well: in 1562, a patent was awarded to George Cobham in the hope that it would 'encourage others to discover like good engines and devices'.[5]

1.2 Patents in economic theory and practice

Since this early period, a variety of other reasons have been adduced for awarding exclusive rights to inventors. In particular, some have reasoned that there must exist a moral imperative for recognising and protecting

[3] Henry Maxwell-Lyte, *Calendar of the Patent Rolls, Henry VI, 1446–1452* (London: Anthony Brothers, 1909), 255. The patent is signed 'per ipsum regnum', indicating that the decision to grant the petition had been made by the king personally. A. L. Brown, 'The authorisation of letters under the great seal', *Bulletin of the Institute of Historical Research*, 37 (1964), 127, 142.

[4] Before continuing: 'the decision has been made that, by authority of this Council, any person in this city who makes any new and ingenious contrivances not made heretofore in our Dominion, shall, as soon as it is perfected so that it can be used and exercised, give notice of the same to the office of our Proveditori di Comun, having been forbidden up to ten years to any other person in any territory and place of ours to make a contrivance in the form and resemblance of that one without the consent and license of the author'. Lionel Bently and Martin Kretschmer, 'Venetian Statute on Industrial Brevets', Primary Sources on Copyright (1450–1900), http://copy.law.cam.ac.uk/cam/tools/request/showRepresentation?id=representation_i_1474, accessed 6 October 2013.

[5] Jeremy Phillips, 'The English patent as a reward for invention: the importation of an idea', *JLH*, 3 (1982), 74.

the contribution of an inventor to the common weal. John Stuart Mill, for example, argued that 'it would be a gross immorality of the law to set everybody free to use a person's work without his consent, and without giving him an equivalent'.[6] Sometimes this argument appears with a Lockean gloss: that it is the labour and effort involved with creating a hitherto unknown invention which entitles the inventor to its dominion.[7] In essence, though, the principal and practical reasons for patents have remained essentially unchanged: they encourage the development and diffusion of technology.[8] In the formal economic literature, Kenneth Arrow provides the classic exposition of the argument that awarding patents to inventors encourages inventive activity.[9] Arrow begins by making two assumptions about the nature of inventions. First, that the costs of invention are infinitely greater than the costs of imitation (which are assumed to be zero). Second, because they are nonrivalrous, one's use of an invention will not diminish another's capacity to use it as well. Consequently, without patent protection, an inventor cannot appropriate a return above the market rate and recoup the costs of invention as well as the costs of production. However, a patent, by conferring the (temporary) right to exclude other parties from using the invention, permits the inventor to earn monopolistic profits and to recover the costs of invention. Alternatively, the patent holder can recover costs by licensing the use of the invention to other parties in return for royalties.

Nobel laureate Douglass North adopts Arrow's model of invention and patenting for his analysis of institutions in economic development. North argues that the social rate of return to inventive activities has always been high, but that historically, most societies have failed to achieve an optimal amount of inventive activity. This he attributes (as does Arrow) to the nonrivalrous nature of invention. Although inventions are costly to

[6] Quoted in Fritz Machlup and Edith Penrose, 'The patent controversy in the nineteenth century', *JEH*, 10 (1950), 18.

[7] For an interesting discussion on Locke and intellectual property rights, see Robert Merges, *Justifying Intellectual Property* (London: Harvard University Press, 2011), chap. 2.

[8] This statement does overlook one important development. Some economists and lawyers argue that granting broad patents facilitates the orderly development of technological opportunities. This argument rests on the premise that unless there is a controlling patent, many will perceive and pursue the same opportunity, leading to wasteful duplication of effort and an over investment in R&D. Edmund Kitch, 'The nature and function of the patent system', *The Journal of Law and Economics*, 20 (1977), 276.

[9] In particular, Kenneth Arrow, 'Economic welfare and the allocation of resources for invention', in NBER, *The Rate and Direction of Inventive Activity: Economic and Social Factors* (Princeton, NJ: Princeton University Press, 1962), 609–26.

develop, they can be duplicated at no cost to the copier. As a result, the inventors' private rate of return will be negligible compared to the social rate of return and so inventive activities are spurned: '[T]he failure to develop systematic property rights in innovation up until fairly modern times was a major source of the slow pace of technological change.'[10] North suggests that it is only with the introduction of patents that the inventor is reliably assured a return commensurate with the value of their inventive output. In England, this is supposed to have occurred in 1624 with the passing of the Statute of Monopolies, which, by incorporating the patent system into common law, guaranteed the intellectual property rights (IPRs) of inventors. This, argues North, gave England a comparative advantage in the development of technology, which ultimately led to industrialisation.[11]

A second reason for awarding patents is they promote the diffusion and commercialisation of new technology. Broadly speaking, this occurs in three ways. First, if patent protection is unavailable, inventors will need to work their invention in secret to prevent competitors from using it as well. Working in secret, however, imposes significant public and private costs. To be effective, it necessarily involves preventing the flow of technical information to other potential users, leading to duplication of inventive efforts and stymieing sequential technological development. In addition, efforts at maintaining secrecy impose limits on production.[12] By allowing an inventor to eschew secret working, patents mitigate these economic costs – although monopoly pricing of a patented invention will still lead to lower production and higher prices than would be the case in a free market.

Second, to obtain the patent, the inventor is normally obliged to author a detailed description of the invention (a specification), and these are often made available to the public through patent offices and public libraries. Properly enforced, the requirement to publish the invention in a patent specification provides other inventors and manufacturers with

[10] Douglass North, *Structure and Change in Economic History* (London: Norton, 1981), 164. More recently, North has referred to the 'enormous importance in history of property rights as providing incentives for innovation and enrichment'. Douglass North, 'A recommendation on how to intelligently approach emerging problems in Intellectual Property systems', *Review of Law and Economics*, 5 (2009), 1132.

[11] Douglass North and Robert Thomas, *The Rise of the Western World: A New Economic History* (Cambridge: Cambridge University Press, 1973), 154.

[12] Michele Boldrin and David Levine, *Against Intellectual Monopoly* (Cambridge: Cambridge University Press, 2008), 167.

easy access to the latest technical information.[13] Last, by defining and delineating property in technology, patents facilitate the transfer of usage rights to other parties. Conversely, trading secret technology is often prohibitively difficult. For the agreement to be effective, the seller must be bound by the terms of the agreement from publishing or selling the secret again. This may not always be possible (the legality of such agreements was uncertain into the second quarter of the nineteenth century), and even if the agreement can take effect, it is difficult to ensure that third parties have not acquired knowledge of the secret already or will not discover the secret subsequently. Moreover, in the absence of any record of title, the buyer cannot ensure that the technology has not been fraudulently obtained from someone else. In contrast, a record of patents and their holders are normally maintained by patent offices, allowing the buyer to confirm the title of the seller and legally exclude other parties from using the technology.

Over the past thirty years, policy initiatives in many developed countries have concentrated on strengthening and augmenting patent rights in an attempt to meliorate their economic and technological benefits. During the 1980s and 90s, the United States placed IPR at the centre of its trade negotiation agenda, culminating in the 1994 'Agreement on Trade Related Aspects of Intellectual Property Rights' (TRIPS). TRIPS established 'minimum' standards of intellectual property protection among World Trade Organization members, instituting, for example, a minimum term of twenty years for patents.[14] Empirical research, though, has begun to query the efficacy of recent policy changes and the reasoning behind them. In 1988, for example, the Japanese government enacted new patent reforms, prompted (in part) by pressure from the United States.[15] These reforms were supposed to improve the security of patent rights by allowing for multiple claims in a single application and by extending the patent term for some technologies. However, no statistically significant increase in research and development (R&D)

[13] Surveys indicate that many firms chose not to patent inventions for precisely this reason. Wesley Cohen, Richard Nelson and John Walsh, 'Protecting their intellectual assets: appropriability conditions and why US manufacturing firms patent (or not)', NBER Working Paper 7552 (2000), 15.

[14] Lionel Bently and Brad Sherman, *Intellectual Property Law*, 3rd edn (Oxford: Oxford University Press, 2009), 353.

[15] Mariko Sakakibara and Lee Branstetter, 'Do stronger patents induce more innovation? Evidence from the 1988 Japanese patent law reforms', *RAND Journal of Economics*, 32 (2001), 80.

spending, or inventive output, could be attributed to these reforms. In a similar vein, another study has collated and analysed the effects of 271 separate patent reforms, enacted over the past 150 years, in sixty different countries. It found that wherever patent protection was already reasonably strong, additional patent protection never had a positive impact on innovation and could even have negative consequences for inventive output.[16] These negative consequences are partly attributable to overly strong and broadly defined patent rights, allowing their holders to exclude other parties from pursuing related or 'downstream' inventions in the protected technology. The problem is exacerbated when these patents overlap with each other and such a scenario has arisen in nanotechnology, where the United States Patent and Trademark Office has awarded more than 1,600 patents that make reference to carbon nanotubes, dozens of which claim essentially the same thing.[17] For the researcher or inventor venturing into this 'patent thicket', an enormous range of licences have to be obtained before even the most basic work can be undertaken.

Arguments for the strengthening of patent rights have been also undermined by the revelation that in many technological sectors, patents are not the primary means of appropriating returns from invention. In 2000, Wesley Cohen surveyed the appropriation methods of large American firms, delineating three broad strategies: (1) exploitation of complementary capabilities and lead time, (2) secret working and (3) legal mechanisms such as patenting. Although these strategies are not mutually exclusive, Cohen concludes that in the majority of sectors, exploiting lead time and protecting secrecy are more important appropriation strategies than patenting and other legal mechanisms. Patenting, however, is found to be important in a minority of industries, particularly pharmaceuticals and certain classes of machinery and computing.[18] This is not some peculiarity exclusive to American industry; a major study of large Japanese corporations estimates that the cessation of patent

[16] Josh Lerner, '150 years of patent protection', NBER Working Paper 8977 (2002), 2.

[17] This discussion is derived from Joshua Pearce, 'Physics: make nanotechnology research open-source', *Nature*, 491 (2012), 519–21. For a prescient discussion of this problem from a legal perspective, see Mark Lemley, 'Patents and nanotechnology', *Stanford Law Review*, 58 (2005), 601–30.

[18] In contrast, a similar Japanese survey in 1999 found legal mechanisms to be more important than the other two appropriation strategies. Cohen, Nelson and Walsh, 'Protecting their intellectual assets', 13–14.

protection would result in a 60 per cent reduction in R&D spending by chemical firms, but only a 5 per cent reduction in mechanical firms.[19]

Pronounced sectoral differences in appropriation strategies also existed 150 years ago. Petra Moser has undertaken an empirical study of the exhibits in four World Exhibitions in the second half of the nineteenth century, beginning with the 1851 Great Exhibition at the Crystal Palace. In this Exhibition, only 11.1 per cent of British and 15.3 per cent of American exhibits were patented.[20] There was, however, significant variation in the propensity to patent between different industrial sectors.[21] Moser categorises the exhibits into ten different sectors. The two sectors with the highest patenting rates were manufacturing machinery (29.8 per cent) and engines (24.6 per cent).[22] In contrast, only 5 per cent of mining and metallurgy exhibits and 5.1 per cent of chemical exhibits had been patented. These results should be treated with some caution. It is likely that many exhibits were not patented simply because they were ineligible. Zorina Khan notes in a similar study of nineteenth-century fairs that many exhibits were examples of high-quality workmanship rather than new inventions, and so lacked sufficient novelty for patent protection.[23] Because Moser does not exclude nonpatentable exhibits, the propensity to patent exhibits that *were* patentable is higher than indicated by her figures.

Nonetheless, there is no reason to suppose that this influences the apparent variation in the intersectoral propensity to patent. Moser's findings are supported by the work of O'Brien, Griffiths and Hunt on British textile inventions in the eighteenth century. They traced 174 key

[19] Ove Granstrand, 'Innovation and intellectual property rights', in Jan Fagerberg, David Mowery, and Richard Nelson (eds.), *The Oxford Handbook on Innovation* (Oxford: Oxford University Press, 2005), 282.

[20] Some of the difference in patenting rates can probably be attributed to the higher overall quality of American exhibits. They would be of a higher quality because the effort for an American to exhibit in Britain would be greater than for a British inventor, meaning it was only worth the additional trouble with an invention they perceived to be of a higher value than their British counterpart. Although Moser suggests that this effect would have been minor because of the uniform rules of selection adopted by the central commissions, this does not preclude the possibility of self-selection by Americans prior to application. Petra Moser, 'Why don't inventors patent?', NBER Working Paper 13294 (2007), 37.

[21] These sectors are 'mining and metallurgy', 'chemicals', 'food processing', 'engines', 'manufacturing machinery', 'civil, military and naval engineering', 'agricultural machinery', 'scientific instruments', 'manufactures' and 'textiles'. Ibid.

[22] Ibid.

[23] Zorina Khan, 'Going for gold: industrial fairs and innovation in the nineteenth century', mimeo (2010), 12.

textile inventions (which as novel advances would have been patentable), of which 44 per cent were patented.[24] They employed, however, a broad definition of textile inventions, including developments in dyeing (chemicals in Moser's categorisation) as well as textile machinery (manufacturing machinery in Moser). This figure of 44 per cent therefore masks wide differentiation in patenting rates of different types of textile invention. For this earlier period, they found results similar to Moser, with low patent rates for dyes and higher rates for textile machinery.[25]

Because they use a different data source from an earlier period, O'Brien et al.'s study provides strong corroborative evidence for Moser's results relating to differences in intersectoral propensities to patent. Moser attributes her results to variations in the underlying level of scientific knowledge in different technological sectors. Where the level of scientific understanding made reverse-engineering possible, secret working would be ineffective, forcing inventors to pursue other methods of commercial exploitation, namely patenting. Where, however, scientific knowledge was such that reverse-engineering was not practicable, secret working was preferred. Moser illustrates the phenomenon by analysing the patent rates of American chemical exhibits over four exhibitions between 1851 and 1893. The rates increase rapidly after the introduction of the periodic table in 1869 which, by categorising the known elements by their physical properties, made the reverse-engineering of chemicals into their constituent parts (the elements) much easier.[26]

The logic of this argument implies that the efficacy of the first appropriation method outlined by Cohen – the exploitation of complementary capabilities and lead time – would have been declining during the eighteenth and nineteenth centuries. This strategy relies on recouping the costs of invention during the time between when the invention is first commercialised by its progenitor and when competitors are able to market the invention competitively as well. During the early modern period, such a strategy might have been relatively effective. Most technical knowledge was 'implicit' and experience-based, limiting how such

[24] Trevor Griffiths, Philip Hunt and Patrick O'Brien, 'Inventive activity in the British textile industry, 1700–1800', *JEH*, 52 (1992), 885.

[25] It should, however, be noted that they do not attribute these differences to the practicability of reverse-engineering, but rather the level of involvement of sponsoring institutions, such as the Society of Arts, in different areas of development. Ibid., 888.

[26] In 1851, not a single American chemical exhibit was patented. At the 1893 World's Fair in Chicago, thirty-two years after the introduction of the periodic table, 16 per cent of US chemical exhibits were patented. Moser, Why don't inventors patent?, 29.

knowledge could be expressed and transmitted to competitors.[27] This might have given premodern inventors sufficient lead time to recoup the outlay of invention without patent protection. Over time, however, a growing body of scientific and technological knowledge was set down and codified, making the adoption of new technology easier and reducing the period of lead time in which monopolistic profits could be appropriated.[28]

Moser also argues that the provision of patent protection influences the direction of inventive activity between sectors where technology can be reverse-engineered and where it cannot. In a comparison of the sectoral distribution of exhibits between countries that awarded patents with those that did not, Moser found that countries without patent laws concentrated in sectors with strong alternative mechanisms for protecting intellectual property, and away from sectors that did require patents.[29] For example, in the two countries without patent systems in 1851, Switzerland and Denmark, scientific instruments constituted 27 per cent and 23 per cent of their exhibits respectively, whereas the median among the other exhibiting countries was 6 per cent.[30] A similar pattern occurs in countries with weakly enforced patents. For example, Bavaria had particularly ‘ill-enforced’ patents and after Switzerland and Denmark, it exhibited the highest proportion of scientific instruments.[31] In 1851, America probably possessed the most sophisticated patent system in the world, and the fact that a similar proportion of American and British exhibits were patented and that they had a similar sectoral distribution implies that the British patent system offered comparable access to effective patent protection. Moser only concerns herself with the ‘direction’ of inventive activity and is careful not to speculate on whether patents may have affected overall levels of inventive activity.

[27] Stephan Epstein, ‘Property rights to technical knowledge in pre-modern Europe, 1300–1800’, *AER*, 94 (2004), 382–83.

[28] Joel Mokyr estimates that the annual number of medical, scientific and technological titles published during the eighteenth century quintupled. Joel Mokyr, *The Enlightened Economy: An Economic History of Britain, 1700–1850* (London: Yale University Press, 2009), 46.

[29] Petra Moser, ‘How do patent laws influence innovation? Evidence from nineteenth century world’s fairs’, *AER*, 95 (2005), 1224.

[30] The ‘Scientific instruments’ sector includes technologies such as optical lenses, watch movements and barometers. These were technologies where methods of production often remained ‘intuitive’ and where the scale of production was often small. As such, scientific instruments lent themselves to secret working. Ibid.

[31] Ibid.

It is, however, worth mentioning that exhibits from sectors that were particularly reliant on patent protection, such as manufacturing machinery and engines, 'were extremely rare in the Swiss data'.[32] In contrast, countries with patent laws had many more exhibits in these sectors.

This research also relates to the second main argument employed to justify patent rights, specifically their role in the disclosure and commercialisation of new technology. If patenting is employed only in technological sectors, where secrecy is inherently difficult to maintain and hence, where technical information can disseminate quickly, awarding the inventor the right to exclude others will inevitably frustrate the flow of technology.[33] This argument, however, not only overlooks the role of patenting in encouraging the development of technology in the first place (without patenting, there may not be the technology developed to disseminate later), but it also overlooks the relative advantage of patenting over secret working. Returning to Moser's work on chemical exhibits in the nineteenth century, she has shown how just as the propensity to patent chemical inventions increased after 1869, there was a concomitant decline in the geographic concentration of chemical exhibits at American fairs. This decline cannot be attributed to any change in the location of production in the American chemical industry (which remained largely unchanged); instead, Moser posits that patents contributed to the geographic diffusion of technical information and inventive activities in the chemical industry.[34]

Neoclassical economic theory suggests that patents should play an important role in encouraging the development of new technology. Recent work, however, has shown that matters are not so straightforward. In particular, an important assumption underlying this theory is that imitation costs are negligible. Wherever this has been the case, the propensity to patent new inventions has been high, as was the case in steam engineering and manufacturing machinery during the Industrial Revolution. The empirical evidence, however, suggests that imitation costs vary significantly by industry and this is partly a function of the

32 The Swiss data refers to both the Great Exhibition in 1851 and the Centennial Exhibition in Philadelphia, 1876. Ibid., 1228.

33 Boldrin and Levine, *Against Intellectual Monopoly*, 166.

34 Moser's regressions show that a 1 per cent increase in the share of patented exhibits is associated with a 1.3 per cent decrease in geographic concentration. Petra Moser, 'Do patents weaken the localization of innovations? Evidence from World's Fairs', *JEH*, 71 (2011), 377–78.