Innovation, duplication, and the contract theory of patents

Vincenzo Denicolò and Luigi Alberto Franzoni[1]
Department of Economics, University of Bologna

Paper prepared for the
XVII International School of Economic Research
Intellectual Property, Innovation and Competition
Siena, 11-18 July 2004

Preliminary version

Abstract

We analyze optimal patent design when innovators can rely on secrecy to protect their innovations. Secrecy has no fixed term, but does not preclude accidental disclosure nor independent creation by others. We prove that under weak conditions it is socially preferable to increase patent life as much as is necessary to induce first inventors to patent.

JEL Nos: L00 (Industrial organization), O31, O34 (Intellectual property rights).

[1] Mailing address: Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy. E-mail: denicolo@economia.unibo.it; franzoni@economia.unibo.it.

1. INTRODUCTION

Surveys of US firms have repeatedly found that secrecy is more highly ranked than patents as a protection mechanism for both product and process innovations, and has increased in importance over the last decade.[2] Cohen, Nelson and Walsh, (2000) report that for process innovations, only 23% of all respondents consider patents as an effective appropriability mechanism as compared to 50% and 38% of respondents for secrecy and lead time, respectively.[3] For product innovations, patents are considered relatively more effective (41%), but still less effective than either secrecy (51%) or lead time (50%).

These findings suggest that economists may have overestimated the importance of the reward motive for granting intellectual property rights.[4] According to a popular critique of the patent system, this empirical evidence shows that the incentive effect of patents is small or even negligible, and that firms would have a substantial incentive to innovate even in the absence of intellectual property rights. However, secrecy or lead time provide a reward to innovators only inasmuch as they are a source of market power, and market power is exploited by pricing above marginal cost independently of its origin.

These arguments naturally raise the question as to whether promoting the diffusion

[2] See Levin et al. (1987) and Cohen, Nelson and Walsh (2000).

[3] Lead time differs from secrecy in that innovative technical knowledge is not intentionally hidden by the inventor, but is similar to secrecy in other respects. Both lead time and secrecy mean that the mere fact that the inventor practices her innovation may not disclose enough information to allow instant and costless replication by others.

[4] The *reward theory of patents* maintains that the function of the patent system is to remunerate innovators so as to encourage R&D effort. With strong backing from the US Constitution, the reward theory is by far the most prominent approach to the economic analysis of patents.

of innovative knowledge can be a distinct rationale for the patent system. Indeed, when innovations can be concealed, patents can be viewed as a "contract" between innovators and society whereby a temporary property right is granted in exchange for disclosure. This *contract* (or *exchange*) *theory of patents* is widely accepted and respected by lawyers and courts, but has long been discredited among economists. Critics of the contract theory argue that to have innovators patenting and disclose innovations, patent life must be at least as long as the expected duration of the secret. As pointed out by Oxford economist J.E.T. Rogers (1863) long ago,

> ...the bargain of the inventor with the public, is thoroughly one-sided. If it be his interest to keep his secret, he infallibly does so, not so much from the cause that a patent is expensive, as because it is his interest. [...] It is perfectly true, indeed it is insisted on by the advocates of the rights of invention, that nothing can compel him to disclose his discovery. Does he ever do so except on the ground that the profits of the monopoly would be more valuable than the profits of the secret? (Rogers (1863: 127)

In a similar vein, more recently Boldrin and Levine (2004: 129)) note that

> Granting a legal monopoly in exchange for revealing the "secret" of the innovation is one, apparently clean, way to make innovations more widely available in the long run. However, this argument has not been subject to much scrutiny by economists, and indeed, in the simplest case it fails. Suppose that each innovation can be kept secret for some period of time, with the actual length varying from innovation to innovation, and that the length of legal patent protection is 20 years. Then the innovator

will choose secrecy in those cases where it is possible to keep the secret for longer than 20 years, and choose patent protection in those cases where the secret can be kept only for less than 20 years. In this case, patent protection has a socially damaging effect. Secrets that can be kept for more than 20 years are still kept for the maximum length of time, while those that without patent would have been kept for a shorter time, are now maintained for at least 20 years.

This argument to the effect that disclosure of the innovation is not an independent motive for granting patents does indeed hold true in the "simplest case" in which innovators relying on secrecy face only the risk of inadvertent disclosure of the secret to the public.[5] But secrecy precludes neither inadvertent disclosure nor independent creation by others. The possibility that other inventors duplicate the original innovations alters the comparison between patents and secrets in a fundamental way. First, duplicative efforts exerted for the independent rediscovery of the innovation are largely wasteful. Second, the evolution of the post-innovation market structure will be different under patents and secrets. Typically, market structure evolves more discontinuously in the case of patent protection because of the "catastrophic" event of patent expiration. Thus, even if patent protection is strengthened as much that

[5] The argument has been formalized by Gallini (1992). Gallini models the innovator's choice whether to patent or keep the innovation secret, assuming that if the innovation is kept secret there is an exogenous probability of leakage, whereas if the innovation is patented there is free entry of imitators with a fixed imitation cost k, where k is a measure of patent breadth, and the patent expires after a finite duration. She proves that if the patent life that would be optimal if leakage was certain and patent breadth was infinite is lower than the patent life that makes the innovator indifferent between patenting and not, then the optimal patent life can be set equal to zero with no loss of generality (Proposition 2 (i)).

the innovator is indifferent between patenting and not, the deadweight loss associated with the two options is generally different. Indeed, we find that under a weak condition (i.e., the presence of a business stealing effect) it is socially desirable to strengthen patent protection as much as is necessary to induce inventors to patent, even if encouraging R&D efforts is not a concern.

Section 2 introduces the model, section 3 describes the equilibrium behavior of the players. Section 4 is devoted to the optimal patent life, and section 5 concludes.

2. THE MODEL

To disentangle the disclosure from the reward motive for granting patents, we assume that the innovation is already available. We concentrate instead on the costs and benefits of patents, on the assumption that the innovation can alternatively be protected by trade secrecy.

Consider a product innovation or, equivalently, a drastic cost-reducing innovation.[6] Let demand be given by $X(p)$, where p is price, X is output, and $X(.)$ is a strictly decreasing and differentiable function on $[0, \bar{p}]$ and is zero on $[\bar{p}, \infty)$. It follows that inverse demand, $p(X)$, is decreasing and differentiable on $[0, X(0)]$. Let the post-innovation marginal cost be constant at c. The innovator must decide whether to patent the innovation or keep it secret. By patenting, the innovator reaps monopoly profits π_m for the duration of the patent, T; when the patent expires, the innovator's profits are driven to zero.[7] Alternatively, the innovator can rely on secrecy. Here the

[6] An innovation is drastic if the innovator is unconstrained by outside competition and can therefore engage in monopoly pricing. Our results hold *a fortiori* with non-drastic innovations.

[7] Patents are assumed to be so broad that imitation of a patented innovation is impossible without infringing the patent. As such, the strength of patent protection is fully captured by the patent's

risk is accidental disclosure (a "leak") or successful duplication by the follower. We assume that leakage of the secret has the same effects as expiry of the patent, i.e., the innovation becomes public and profits are driven to zero. The random event of a leak occurs with an exogenous probability γ.[8] The innovator also loses his monopoly if the innovation is duplicated. We assume that if the innovation is not patented there is free entry by duplicators, upon payment of a fixed duplication cost k.[9]

Let s denote the number of active firms, other than the original innovator. We denote by $\pi(s)$ the individual flow of profit with $s+1$ active firms (i.e., the original innovator and s duplicators); $\pi(0) \equiv \pi_m$ is the monopoly profit. For a variety of reasons, the social returns from the innovation typically exceed the private returns. In our simplified framework, the social returns include any gain that accrues to consumers before the patent expires. Such gains depend on the equilibrium price and output. Let $S(s)$ and S_c denote the instantaneous social returns from the innovation with $s+1$ active firms and under perfect competition, respectively, and let $S_c - S(s) = D(s)$ denote the deadweight losses with $s+1$ active firms; $D_m = D(0)$ is the monopoly deadweight loss. We assume that individual profits and the deadweight loss decrease with the number of active firms. All future profits are discounted at the

life. Explicit analysis of the breadth-length trade-off (see Denicolò, 1996, and the literature cited therein) would complicate matters and add little to the issues that we focus on in this paper.

[8]Leakage may occur for a variety of reasons: hackers may violate protected files, new techniques enabling to reverse-engineer the innovation may become available, the government may expropriate the trade secret. The probability of a leak depends both on the technical difficulty of concealing the innovation and the strength of trade secret protection. Both are taken as exogenous in our analysis.

[9]We rule out the possibility that the innovator licenses potential duplicators, like in Maurer and Scotchmer (2002). Such licensing agreements are problematic when the innovation is not patented and there if free entry of potential duplicators.

common discount rate r. It is convenient to define

$$\textit{“normalized” patent length: } z \equiv \left(1 - e^{-rT}\right).$$

z equals the proportion of overall discounted monopoly profits, $\frac{\pi_m}{r}$, that accrues to the patentee. It is lower than one if patent life is finite. Indeed, z ranges from 0 to 1 as T goes from 0 to ∞.

To summarize, the timing of events is as follows. First, the innovator decides whether or not to patent. Second, if the innovator has not patented, uncertainty concerning leakage is resolved. If the innovation is kept secret by the inventor, and the secret does not leak out, duplicators decide whether or not to enter the market. Finally, profits are realized. All parameter values and actions are common knowledge.

3. BENCHMARK

Suppose that the duplication cost k is infinite, so that no duplication occurs at equilibrium. We compare the welfare loss associated with patents and secrecy, respectively. These losses are evaluated *ex post*, i.e., assuming that the innovation is already available.

To begin with, let us consider the innovator's problem. The innovator, who already owns the innovation, must decide whether or not to patent. If the innovator patents, she earns monopoly profits until the patent expires:

$$\int_0^T e^{-rt}\pi_m dt = z\frac{\pi_m}{r}.$$

If she does not patent, her payoff is $(1-\gamma)\frac{\pi_m}{r}$: the discounted value of a constant flow of a monopoly profit, multiplied by the probability that the secret does not leak out. Clearly, the innovator patents if and only if $z \geq (1-\gamma)$.[10]

Next consider social welfare. As compared to the first best *ex post* solution, in which the market is perfectly competitive and there are no deadweight losses, if the innovator chooses to patent social welfare is reduced by

$$WL_{\text{patents}} = \int_0^T D_m e^{-rt} dt = z\frac{D_m}{r}$$

as the monopoly deadweigth loss is borne for the duration of the patent. If instead the innovation is kept secret, society bears the monopoly deadweight loss forever, conditional on the secret not leaking out:

$$WL_{\text{secrets}} = (1-\gamma)\frac{D_m}{r}.$$

Consequently, in this benchmark case, if patent length is long enough as to induce the innovator to patent, i.e., $z \geq (1-\gamma)$, the welfare loss under patents exceeds the welfare loss under secrecy. This is the logic underlying the arguments of Rogers (1863) and Boldrin and Levine (2004).

4. DUPLICATION

Suppose now that the duplication costs k is small enough that some duplication occurs at equilibrium. With a finite duplication cost and free entry of duplicators, the equilibrium number of entrants, s^*, is the largest integer s such that $\frac{\pi(s)}{r} \geq k$. For simplicity, in what follows we shall treat s as a continuous variable. Let $\pi(s)$ be continuously differentiable with $\pi'(s) < 0$ as long as $\pi(s) > 0$ and $\pi'(s) = 0$ when

[10]To fix ideas, we assume that the innovator patents when she is indifferent between patenting and not, but our results are independent of this tie-breaking rule.

$\pi(s) = 0$. The free entry condition then becomes

$$\frac{\pi(s)}{r} = k \tag{1}$$

and determines s^* uniquely.

The innovator's expected profit, in case she relies on secrecy, is now

$$(1-\gamma)\frac{\pi(s^*)}{r} = (1-\gamma)k$$

because with probability γ the secret leaks out and the innovator gets nothing, and with the complementary probability $(1-\gamma)$ the secret does not leak out but s^* duplicator enter the market. Consequently, the innovator will prefer to patent if and only if

$$z \geq (1-\gamma)\frac{rk}{\pi_m}. \tag{2}$$

Ex post, it would be pointless to prolong patent life beyond the minimum level which induces the innovator to patent. At such a minimum patent life $\bar{z} = (1-\gamma)\frac{rk}{\pi_m}$, the welfare loss under patents is

$$WL_{\text{patents}} = (1-\gamma)k\frac{D_m}{\pi_m} \tag{3}$$

Compare this with the welfare loss under secrecy. This now comprises the expected deadweight loss and duplication costs:

$$WL_{\text{secrets}} = (1-\gamma)\left[\frac{D(s)}{r} + sk\right]. \tag{4}$$

Proposition 1. *If a business stealing effect is present, so that individual output falls as the number of active firms increases, and the patent life is just long enough as to induce the innovator to patent, patents are better than secrecy.*

Proof. Comparing (3) and (4) we have that $WL_{\text{patents}} \leq WL_{\text{secrets}}$ if and only if

$$k\frac{D_m}{\pi_m} \leq \frac{D(s)}{r} + sk$$

Using (1), this condition reduces to

$$\frac{D_m}{\pi_m} \leq \frac{D(s)}{\pi(s)} + s. \tag{5}$$

To prove that (5) holds, define $H(s) \equiv \frac{D(s)}{\pi(s)} + s$ and note that the left hand side of (5) is $H(0)$. Therefore, it suffices to show that $H'(s) \geq 0$. Noting that $\pi(s) = (p-c)\frac{X}{s+1}$ and $D'(s) = -(p-c)X'(s)$, differentiating and rearranging terms we get

$$H'(s) = \frac{(p-c)^2 \frac{X}{s+1}\left[\frac{X}{s+1} - X'(s)\right] - D(s)\pi'(s)}{[\pi(s)]^2}$$

But the business stealing effect means that an increase in the number of firms reduces individual output. This implies that the term inside square brackets is positive, and so $H'(s) > 0$. ■

Figure 1 provides some intuition for why society prefers disclosure to secrecy. Note first of all that the goal of the policymaker is to minimize the sum of the expected deadweight losses and duplication costs per unit of the innovator's profit. The reason is that the innovator can choose between patents and secrecy and she will choose the more profitable alternative – patenting must be at least as profitable as secrecy. Panel *a* displays monopoly profits and monopoly deadweight losses – the left hand side of (5). Panel *b* displays the innovator's profits and the sum of deadweight losses and duplication costs in the presence of duplication – the right hand side of (5). With duplication, output is higher and price is lower, but the innovator's profit is only a fraction of the rents associated with pricing above marginal cost; the remaining fraction is dissipated in the duplication process.

5. CONCLUDING REMARKS

It is an historical fact that patents grew as alternatives to trade secrets (see, for instance, David 1993).[11] Many of the first patents (*privilegi*) were granted to people who had not invented the technology at hand, but just ferreted it out from foreign guilds.[12] In this perspective, patents were instrumental to the diffusion of jealously held technology.

In this paper we have argued that patents may serve a socially valuable function as alternative to trade secrets. To disentangle the disclosure motive from the traditional reward motive, we have assumed that the innovation is the outcome of "serendipity," so that stimulating R&D effort is not a concern. When duplication is not possible, and an innovator that relies on secrecy faces only the risk of a leakage, disclosure of the innovation is not an independent motive for granting patents. However, secrecy precludes neither inadvertent disclosure nor independent creation by others. When the possibility of duplication is taken into account, we have shown that it is socially preferable to increase patent life as much as is necessary to induce first inventors to patent.

[11] In England, the life of patents had long been 14 years long, that is twice the term of service of an apprentice.

[12] Although the Statute of Monopolies of 1624 stated that patents could only be granted to the "true and first inventor" of new manufactures, initially inventors could just be importers of foreign skills and know-how: "whether they learned by travel or by study, it is the same thing," declared an English court in 1693 (Machlup 1968).

The model that we have developed in this paper is very simple and it can be extended in a number of ways. First, one could assume that there is only one potential duplicator rather than free entry by duplicators. With restricted entry in the duplication activity, the presence of business stealing no longer suffices to guarantee that patents are socially preferable to secrecy, and stronger conditions are needed (see Denicolò and Franzoni 2004). One reason why entry by duplicators may be restricted is that successful duplicators can be eligible to patent protection. Which forms of protection can a successful duplicator adopt if the first inventor opts for secrecy is, indeed, a matter of policy. Notable differences are observable from nation to nation and over time. For instance, under the British 1956 Patent Act second inventors were not entitled to valid patents, but currently second inventors can patent in most European countries and the US.[13] Differences persist, however, regarding whether the first inventor is allowed to continue to practice an innovation patented by others. In Europe, being first inventor is a defense against infringement (in the legal jargon, first inventors are granted *prior user rights*). In the US, by contrast, second inventors can exclude anybody else, including first inventors, from the innovation. Denicolò and

[13] The old British rule was changed in 1977 to harmonize the British with the European patent law. In the US, under 35 U.S.C. § 102(g), a second-inventor can claim a valid patent if the innovation was "abandoned, suppressed or concealed" by the first inventor. The second-inventor's patent is therefore valid provided that secret use is interpreted as a form of "concealment." This interpretation has been put forward in *Gore v. Garlock* (721 F.2d 1540, 1983), where the court held that the prior user's secret use of a process to create a product (PTFE filament) did not invalidate the patent, despite the fact that the product had been commercially exploited. As explained by the court: "Early disclosure is a linchpin of the patent system. As between a prior inventor and a later inventor who promptly files a patent application ... the law favors the latter." Former decisions distinguished between two cases of public use: "hidden" and "non-informing;" only the former would count as a form of concealment. See Harriel (1996).

Franzoni (2004) compare these policy options on welfare grounds and show that in a fully optimized system prior user rights are never optimal. However, with a predetermined patent life, introducing prior user rights in the American patent system would increase the incentives to innovate at the cost of greater duplicative efforts, with an ambiguous overall welfare effect.

Second, the model could be extended to account for heterogeneous innovations. Indeed, one of Boldrin and Levine's arguments is that innovators will patent only those innovations that are most difficult to conceal. What kind of contract is this, where the innovator keeps the best innovations for himself and gives the worse (i.e., those which can hardly be kept secret) to the state? Denicolò and Franzoni (2003) assume that the probability of leakage varies across innovations, and show that the optimal patent life is positive. The reason is that for marginal innovations (i.e., those for which the innovator is just indifferent between patenting and not), patents are socially better than secrecy for the same reason as discussed above. However, patents also entail a cost, as they cannot be tailored to each individual innovation and so some infra-marginal innovations will end up being over-protected. For these infra-marginal innovations, in the absence of a reward effect increasing patent protection is socially costly. Therefore, an increase in the patent life has two effects on social welfare: the positive effect is that it induces disclosure of marginal innovations, which is always socially desirable; the negative effect is that it increases the deadweight loss associated with infra-marginal innovations that would have been disclosed anyway. But with a zero patent life nobody patents and so there are no infra-marginal innovations; as a consequence, only the former, positive effect is at work which means that the optimal patent life is necessarily positive.

Third, we have assumed that duplication costs are socially wasteful. The fact that duplication costs should be included into the calculation of the welfare loss is questionable.[14] As long as duplicators do not replicate the innovation exactly, duplication may have positive social value. For example, me-too drugs sometimes are better tolerated by, or are more effective for, certain patients. This extension can be important to explain why secrecy is also protected to some extent, and is left for future research.

[14] Boldrin and Levine (2004: 2xx) argue that

> one of the most frequently abused arguments supporting IP, and patents in particular [is] that patents, by forcing the disclosure of the innovative secret, avoid the socially wasteful "rediscovery" of the same idea by future imitators. This argument relies either on the existence of some negative external effect, whose nature is obscure to us, or on the assumption that pure or disembodied "ideas" have economic and productive value, which is patently false.

We find it difficult to deny that duplication costs can be real, however: for example, once it has been proved that a new drug has no toxic effects and the information is put in the public domain, society can spare the cost of replicating safety tests.

REFERENCES

Boldrin M. and D. Levine, 2004. Rent-seeking and innovation. *Journal of Monetary Economics* 51, 127-160

Cohen W., Nelson, R., Walsh J., 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). NBER Working Paper W7552.

David, P., 1993. Intellectual property institutions and the Panda's thumb: patents, copyrights, and trade secrets in economic theory and history. in: Wallerstein, M., Mogee. M.E., Schoen R. (Eds.), *Global dimensions of intellectual propertyrights in science and technology.* National Academy Press, Washington D.C.

Denicolò, V. 1996. Patent races and optimal patent breadth and length. *Journal of Industrial Economics* 44, 249-265.

Denicolò, V., Franzoni L.A., 2003. The contract theory of patents. *International Review of Law and Economics* 23, 365-380

Denicolò, V., Franzoni L.A. , 2004. Patent, secrets, and the first inventor defense. *Journal of Economics and Management and Strategy* 13, 517-538.

Gallini, N. 1992. Patent length and breadth with costly imitation, *RAND Journal of Economics* 23, 52-63.

Harriel, K. 1996. Prior User Rights in a first-to-invent patent system: Why not? *IDEA The Journal of Technology* 3, 543-568.

Levin, Klevorick, Nelson andWinter, 1987. Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity* 3, 783-820.

Machlup, F., 1968. Patents, in *International Encyclopedia of the Social Sciences*. Collier-MacMillan, London.

Machlup, F., Penrose, E., 1950. The patent controversy in the Nineteenth century. *The Journal of Economic History* 10/1, 1-29.

Maurer S. and S. Scotchmer, 2002. The independent invention defense in intellectual property. *Economica* 69, 535-547.

Rogers, J.E.T., 1863. On the rational and working of the patent laws. *Journal of the Statistical Society of London* 26, 125-.

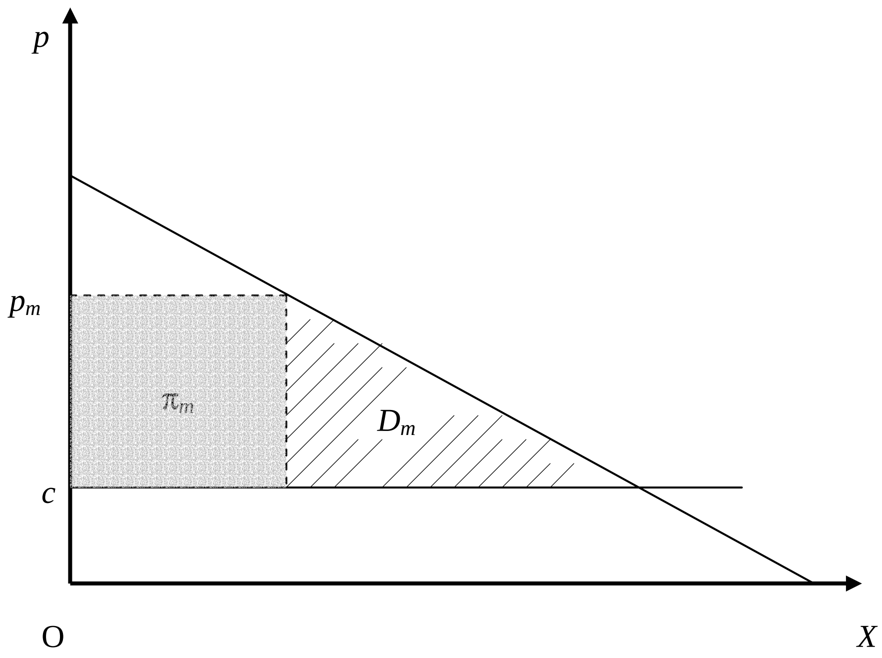

Figure 1*a*

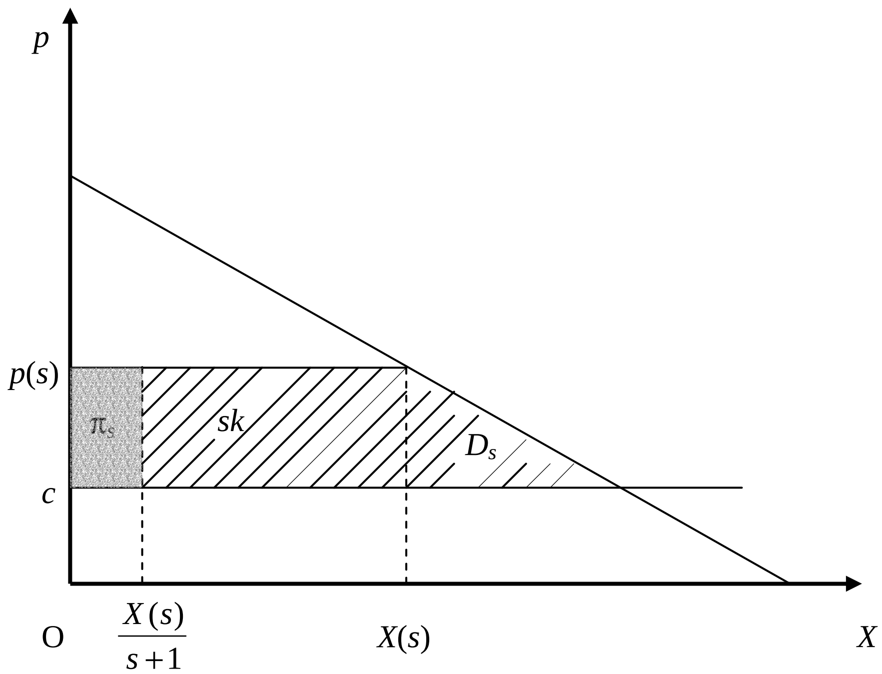

Figure 1*b*