Incentives in R&D markets: Analysing the relation between the elasticities of patent values and the optimality of patent policy

Leandro M. Meller

April 2023

Abstract

The objective of this article is to analyse the relation between the conditions for the optimality of patent policy and the elasticities of the social and private values of innovation. It is also an objective to study the influence of the coexistence of independent innovations and the cumulative nature of innovation in such conditions. Three models have been proposed: the first one Takalo (2001) represents a scenario with an isolated innovation, the second one features two independent innovations, and the third model features two linked innovations, one of which is a "research tool" for the development and the introduction of the other one. Sufficient conditions for the optimality of patents with maximum breadth or infinite length were presented.

1 Introduction

It is a commonly accepted fact that there is a connection between patents, understood as temporary rights on inventive ideas, and efficiency. In their seminal contributions, Arrow (1962) showed that competitive markets tend to underprovide innovation and Wright (1983) proved that, under certain conditions, patents constitute the best solution to this problem. However, patents also tend to generate deadweight loss (see Nordhaus, 1969, and related works), and even an excess of inventive activity (Reinganum, 1989). These inefficiencies can be minimised by assigning adequate values to the dimensions of the patent system. This possibility allowed the development of a new branch of economic literature, which focuses on optimal patent policy.

According to the literature (see Scotchmer, 2004, p. 97), patent systems have three basic dimensions: *length*, *breadth*, and *patentability requirements*. Patent length is the duration of patent protection, patent breadth is the degree of protection patents provide during their life, and patentability requirements define the boundary between patentable and non-patentable ideas. The reader should be warned that "patent breadth" has many different meanings in patent literature: it has been associated with a fixed imitation cost (Gallini, 1992), the maximum price the owner of the patent is allowed to charge on the patented product (Gilbert & Shapiro, 1990), or the instantaneous profit of the innovator, among other definitions (see Bergin, 2022).

This paper focuses on the first two dimensions of patent policy: patent length and breadth. The main objective is to identify, in a partial equilibrium model of a closed economy, conditions under which market efficiency is maximised with a given combination of patent length and breadth, focusing on combinations with maximum patent breadth or infinite patent length. It is also an objective to explore the effects of the cumulative nature of patentable knowledge on such conditions. In order to reach that objective, a model was built assuming that there are only two possible innovations, both of which are patentable. Furthermore, it was also built an extension where one of these innovations is a "research tool" that facilitates the discovery or development of the other one.

These issues are relevant to economic theory and applied economics. The optimisation of patent length and breadth is a "non trivial problem" (Denicolò & Zanchettin, 2022, footnote on p. 4), whose solution might have important consequences for the development and diffusion of new remedies and vaccines against the COVID-19 and other diseases (Volpert & Riepe, 2021). In this context, it should be taken into account that the relation between patent dimensions and technological trajectories is affected by the cumulative nature of knowledge in healthcare industry (Khazabi & van Quyen, 2017). A sign of such cumulativeness in this industry is the existence of prominent "research tools", such as the polymerase chain reaction, the basic technology of bioengineering in the Cohen-Boyer patent, and cell lines (see Scotchmer, 2004, p. 142).

The rest of the paper is structured as follows. Next section reviews literature on optimal (welfare-maximising) patent length and breadth, focusing on the differences between stand-alone and cumulative innovation settings. Section 3 presents Takalo's model, which analyses optimal patent policy mix in a single-innovation setting. That model constitutes the base for the models in Sections 4 and 5. The first of those models represents a setting with two independent innovations, whereas the second one describes a scenario with a "research tool" and a second generation innovation. Conditions for the optimality of maximum patent breadth and infinite patent length constitute the main input for the conclusions in Section 6.

2 Literature review

There is a well-established tradition in economic literature that stresses the lack of incentives to innovate in competitive markets and the importance of patents as a possible solution to this problem. Arrow (1962) provided one of the first formal analyses of inefficiencies in markets for inventions. However, the first author to analyse optimal patent policy in such a context was Nordhaus (1967, 1969), who discussed with Scherer (1972) about the problem of finding the optimal level of patent length. Nordhaus (1967) also wondered what the optimal "breadth" (p. 22) of a patent is, but he did not manage to find a determinate answer except for a situation where specific assumptions hold. Perhaps Tandon (1982) was the first author to draw meaningful conclusions from a model where patent length and breadth are regarded as policy tools to encourage innovation (for instance, the author in question found that there might be a trade-off between extending length and raising breadth to improve welfare).

Models of optimal patent policy can be divided into several groups. Concretely, Rockett (2010) recognised three sorts: (a) *single-innovation*, (b) *cumulative innovation*, and (c) *complementary innovations* models. Single-innovation models assume that inventive ideas are unrelated, and usually focus on a single patentable idea. Cumulative innovation models, instead, represent one of the following two cases (Hall & Harhoff, 2012): an innovation may lead to one or more second-generation innovations (the "research tool" case), or there may be a sequence of inventive ideas where each innovation constitutes an improvement on the previous ones (the "quality ladder" case). Finally, complementary innovations models of patent policy represent situations where several patentable ideas are the inputs for a single second-generation innovation.

The first theoretical contributions that explored the problem of finding the welfare-maximising combination of patent length and breadth, which focused on single-innovation models, led to mixed conclusions. For instance, infinite length is optimal in Tandon (1982), Gilbert and Shapiro (1990), and Galasso et al. (2016). However, finite length and maximum breadth maximise welfare in Gallini (1992), Chen et al. (2019), and Eswaran and Gallini (2019). An additional source of complexity comes from the fact that there are models where both infinite length patents and maximum breadth ones might be optimal (Bagchi & Roy, 2012; Klemperer, 1990) or any length-breadth combination maximises welfare (Nordhaus, 1972). These results are summarised in Table 1.

Despite this seeming inconsistency, there have been two successful attempts to summarise these conclusions. According to Denicolò (1996), if the static social welfare and the relative incentive to innovate are concave in patent breadth, with one of them being strictly so, infinite patent length is optimal. On the other hand, if those functions are convex in breadth, with one of them being strictly convex, maximum breadth maximises welfare. Takalo (2001) provided a result which is even more general than the previous one. This author inferred sufficient conditions for the optimality of infinite length and maximum breadth from the elasticities of private and social values of innovation to changes in patent length and breadth (see next section of the present paper).

Models of patent policy in a "research tool" setting also led to diverse conclusions. For instance, Green and Scotchmer (1995) analysed patent policy in a scenario where a finite patent length is long enough to compensate both the creator of the "research tool" and the developer of the second-generation innovation. After introducing additional assumptions, they showed that finite length and maximum breadth are optimal in this context. However, Matutes et al. (1996) built another model of patent policy in a "research tool" setting and reached the opposite conclusion: infinite length and less-than-maximum breadth maximise welfare. Meng (2017) analysed optimal patent policy from an evolutionary perspective instead, and found that both sorts of solutions might be optimal. Table 2 summarises these results.

Table 1

Optimal patent length and breadth in selected stand-alone innovation models

Article	(A) Could infinite length be optimal?	(B) Could maximum breadth be optimal?
Nordhaus (1972)	✓	✓
Tandon (1982)	✓	
Gilbert and Shapiro (1990)	✓	
Klemperer (1990)	✓	✓
Gallini (1992)		✓
Bagchi and Roy (2012)	✓	✓
Galasso et al. (2016)	✓	
Chen et al. (2019)		✓
Eswaran and Gallini (2019)		✓

Table 2

Optimal patent length and breadth in selected "research tool" innovation models

Article	(A) Could infinite length be optimal?	(B) Could maximum breadth be optimal?
Green and Scotchmer (1995)		✓
Matutes et al. (1996)	✓	
Meng (2017)	✓	✓

This paper contributes to that branch of the literature. Concretely, it presents a model to show that several conclusions on the optimality of infinite length and maximum breadth in a "research tool" setting can be inferred from the elasticities of the social and private values of the innovations to changes in patent length and breadth, as it was done in Takalo (2001) for the case of a stand-alone innovation. The model in question assumes that there is a "research tool" and only one second-generation innovation, and ignores any conceivable interaction with other industries in the economic system. For this reason, it could be thought of as a generalisation of partial equilibrium models. Welfare analysis, which constitutes the most important part of this paper, is focused on efficiency. Inequality analysis, albeit relevant, is beyond the scope of this paper.

Research on optimal patent policy in "quality ladder" and complementary innovations models, although it has also been carried out (see Hall & Harhoff, 2012; Rockett, 2010), is not analysed in the present paper. However, it is reasonable to think that using Takalo's approach to analyse such settings would also constitute an enlightening contribution. For that reason, it is expected to provide here, as an additional contribution, a basis for future research on optimal patent length and breadth in that sort of models.

3 Takalo's model

The model in Takalo (2001) can be understood as a two-stage game with two players: a potential innovator and a regulator. In the first stage, the regulator chooses patent policy, which consists of a value $T \in [0, \infty)$ for patent length and a level $w \in [0, 1]$ of patent breadth. Patentability requirements, which define the boundaries between patentable and non-patentable ideas, constitute another dimension of patent policy. However, this paper does not focus on them. Instead, it was assumed that any innovation is patentable, and that the regulator cannot modify this rule. This paper also assumes "full commitment", which means that the regulator, once it has chosen its patent policy, is not able to change its decision. As a consequence of this assumption, dynamic inconsistencies that would arise if the regulator were allowed to modify its patent policy at any moment were discarded. In the second stage, the potential innovator chooses a level $e \in [0, 1]$ of "innovative effort" or "probability of success". In other words, it is assumed that there is a one-to-one relationship between the innovator's effort and its probability of getting a new product to introduce into the market. Another key assumption in the model is that such product is patentable.

Assuming the agents (i.e., the government and the innovator) are rational, the model will be solved using backward induction. However, before solving the model, it is necessary to define the incentives of the players. Since the potential innovator is the last player who moves, it will be appropriate to analyse its behaviour first. Let $C(\alpha)$ be a function giving the Research and Development (R&D) investment required to achieve success probability α, and assume that this "cost function" is $C(\alpha) = R \cdot \alpha^2/2$, where $R > 0$ is an exogenous parameter. In addition, let $P(T, w)$ be the present value of the benefits the innovator will get in case of success when patent length and breadth are T and w, respectively. Such value can also be interpreted as the market price of the patent that this innovator might get if it develops its own idea. Since the innovator is risk-neutral, its objective function will be $P(T, w) \cdot \alpha - R \cdot \alpha^2/2$. It follows then that the innovator will aim to solve the following problem.

$$\begin{aligned} \max_{\alpha} \quad & P(T, w) \cdot \alpha - R \cdot \alpha^2/2 \\ \text{s.t.} \quad & \alpha \leq 1, \\ & 0 \leq \alpha \end{aligned} \tag{1}$$

The optimal solution, $\alpha^*(T, w)$, is given by equation (2), which represents the innovator's optimal reaction to any given length-breadth combination. This expression reveals that the optimal choice of innovative effort is non-decreasing in patent value and non-increasing in the scale parameter of the initial R&D, which seems to be a reasonable result.

$$\alpha^*(T, w) = \begin{cases} 0 & \text{if} \quad P(T, w) < 0 \\ P(T, w)/R & \text{if} \quad 0 \leq P(T, w) < R \\ 1 & \text{if} \quad R \leq P(T, w) \end{cases} \tag{2}$$

The regulator, who is in charge of setting patent length and breadth, knows the innovator's reaction functions and uses that information to maximise a social welfare function. Let the function $S(T, w)$ represent the market surplus —i.e., the discounted sum of innovator's profit and consumer surplus— when the innovation arises. Then, $\alpha \cdot S(T, w)$ is the expected social value of the innovation provided the level of effort chosen by the potential patentee. Since the government will take the innovator's reaction into consideration before setting patent policy, it would make sense to replace α with $\alpha^*(T, w)$. As a result, $\alpha^*(T, w) \cdot S(T, w) - R \cdot \alpha^2/2$ must be the difference between the expected social value of the innovation and its cost. Assume this is the function the government aims to maximise, except that the government wants the innovator to choose an effort level $\alpha = 1$, so that the following constraint should be imposed: $\alpha^*(T, w) = 1$. Then, expression (3) shows the problem of the regulator.

$$\begin{aligned} \max_{T, w} \quad & S(T, w) - R/2 \\ \text{s.t.} \quad & R \leq P(T, w), \\ & w \leq 1, \\ & 0 \leq T, \\ & 0 \leq w \end{aligned} \tag{3}$$

At this point, it is useful to introduce several definitions. First, let $T(w) : [0, 1] \to [0, \infty)$ be a function returning, for a given level of patent breadth, the minimum value for patent length that provides an incentive to innovate with success probability $\alpha = 1$, i.e.

$$T(w) = \min\{T : R < P(T, w)\} \tag{4}$$

Secondly, provided that $\partial P(T, w)/\partial T \geq 0$ and $\partial P(T, w)/\partial w \geq 0$, the minimum levels of patent length and breadth that satisfy the restriction can be defined as follows.

$$\underline{T} = T(1) \tag{5}$$

$$\underline{w} = \hat{w} : \lim_{w \to \hat{w}} T(w) = \infty \tag{6}$$

Finally, the length and breadth elasticities of the social value of the only possible innovation are defined by expressions (7) and (8) respectively. On the other hand, the elasticities of the private value of the innovation —which is equal to the market value of the patent, assuming frictionless markets— are defined by (9) and (10).

$$\varepsilon_{ST} = \frac{\partial S}{\partial T}\frac{T}{S} \tag{7}$$

$$\varepsilon_{Sw} = \frac{\partial S}{\partial w}\frac{w}{S} \tag{8}$$

$$\varepsilon_{PT} = \frac{\partial P}{\partial T}\frac{T}{P} \tag{9}$$

$$\varepsilon_{Pw} = \frac{\partial P}{\partial w}\frac{w}{P} \tag{10}$$

The optimisation problem in expression (3) and the definitions presented in the previous paragraphs are the building blocks of Takalo's (2001) main result, which is summarised in Propositions 1 and 2. More concretely, these three propositions state sufficient conditions for the optimality of patents with maximum breadth (Proposition 1) or infinite length (Proposition 2).

Proposition 1 (Takalo, 2001)**.** *Patents with minimum length ($T = \underline{T}$) and maximum patent breadth ($w = 1$) are optimal if, for any pair $(T, w) \in [0, \infty) \times [0, 1]$, the following condition holds.*

$$\frac{\varepsilon_{Sw}}{\varepsilon_{ST}} < \frac{\varepsilon_{Pw}}{\varepsilon_{PT}} \tag{11}$$

Proof. See Appendix A □

Proposition 2 (Takalo, 2001)**.** *Patents with infinite length ($T \to \infty$) and minimum breadth ($w = \underline{w}$) are optimal if, for any pair $(T, w) \in [0, \infty) \times [0, 1]$, the following condition holds.*

$$\frac{\varepsilon_{Sw}}{\varepsilon_{ST}} > \frac{\varepsilon_{Pw}}{\varepsilon_{PT}} \tag{12}$$

Proof. See Appendix B □

The next sections will show to what extent these results could be extended to the analysis of a setting with two innovations. First, a case where there is no connection between these innovations will be analysed. Then, a setting where access to one of these innovations (the "research tool") is a necessary condition for the development of the other one (the second-generation innovation) will be analysed. This will make it possible to analyse the consequences of the cumulative nature of innovation to the expression of optimality conditions in terms length and breadth elasticities of innovation values.

4 Two independent innovations

This part of the article extends the analysis in the previous section to a setting with two isolated innovations. Unlike the only-one-innovation setting, now there are three players: a regulator and two potential innovators. Despite this change, the game still has two stages. As before, in the first stage the government sets patent policy, which is defined by a patent policy vector $(T, w) \in [0, \infty) \times [0, 1]$. Then, in the second stage, the innovators play simultaneously. Each innovator has to choose a level of innovative effort $e_i \in [0, 1]$, where e_i is the effort level of the i-th innovator and its probability of success. It is assumed that each innovation is patentable. Besides, patent length and breadth are uniform, which means that both patents must have the same duration, and there is a one-to-one correlation between the breadth of the first patent and the breadth of the second one. In other words, it is impossible to modify the breadth of the first patent without modifying the breadth of the second one. This happens, for instance, if patent breadth is set by court rulings and legislation that affect every patent.

As in the previous section, the game will be solved using backward induction. This means that the behaviour of the potential innovators will be analysed first. To keep things simple, it is assumed that the effort and the outcome of each innovator have no effect on the value of the patent the other innovator is trying to get. Then, the i-th innovator invests $C(\alpha_i) = R_i \cdot \alpha_i^2/2$ to get a patent whose expected private value is $\alpha_i \cdot P_i(T, w)$. There is no relationship between $P_i(\cdot)$ and α_{-i}) because it is supposed that the innovators serve different, unrelated markets.

There is a little more to say about the shape of $P_1(\cdot)$ and $P_2(\cdot)$. First, it is assumed that both functions are differentiable on T and w, and that their second partial derivatives exist and are continuous. Secondly, this paper only deals with situations where the innovators' streams of revenue are higher during the life of the patents than after their expiration. As a consequence of this assumption, the more the patents last, the more valuable they become, which formally means that $\partial P_1/\partial T > 0$ and $\partial P_2/\partial T > 0$. Furthermore, because it is assumed that the innovations serve different, unrelated markets, a raise in patent breadth would only bring higher profits for the patentees at the consumers' expense. Formally, this implies that $\partial P_1/\partial w > 0$ and $\partial P_2/\partial w > 0$.

Since the i-th innovator wants to maximise its expected benefit, its problem is given by the following expression.

$$\begin{aligned} \underset{\alpha_i}{\text{maximise}} \quad & P_i(T, w) \cdot \alpha_i - R_i \cdot \alpha_i^2/2 \\ \text{subject to} \quad & \alpha_i \leq 1, \\ & 0 \leq \alpha_i \end{aligned} \tag{13}$$

Then, the optimal response of each innovator to patent policy is represented by expression (14), which shows, as expression (2), that the profit-maximising level of innovative effort is positively related to the private value of the patent the innovator is pursuing.

$$\alpha_i^*(\alpha_{-i}, T, w) = \begin{cases} 0 & \text{if} \quad P_i(T, w) < 0 \\ P_i(T, w)/R_i & \text{if} \quad 0 \leq P_i(T, w) < R_i \\ 1 & \text{if} \quad R_i \leq P_i(T, w) \end{cases} \tag{14}$$

Since it has been assumed that the effort level of each innovator has no effect on the reward to the other innovator, the reaction of the i-th innovator, $\alpha_i^*(\alpha_{-i}, T, w)$, will not depend on the choice of the other one, α_{-i}. Formally, this means that $\partial \alpha_i^*/\partial \alpha_{-i} = 0$. This situation leads to an equilibrium like that depicted in Figure 1, where each innovator has chosen its level of innovative effort independently.

Several assumptions were imposed to focus on situations where patent systems are necessary to encourage both innovations because otherwise, optimising the patent system would not be a relevant issue. More concretely, the analysis was limited to settings where the innovators would not invest if there is no patent system at all ($T = 0$ or $w = 0$). Formally, this is achieved assuming that

$$P_1(0, w) < R_1 \quad \text{and} \quad P_2(0, w) < R_2 \quad \forall \quad w \in [0, 1] \tag{15}$$

$$P_1(T, 0) < R_1 \quad \text{and} \quad P_2(T, 0) < R_2 \quad \forall \quad T \in [0, \infty) \tag{16}$$

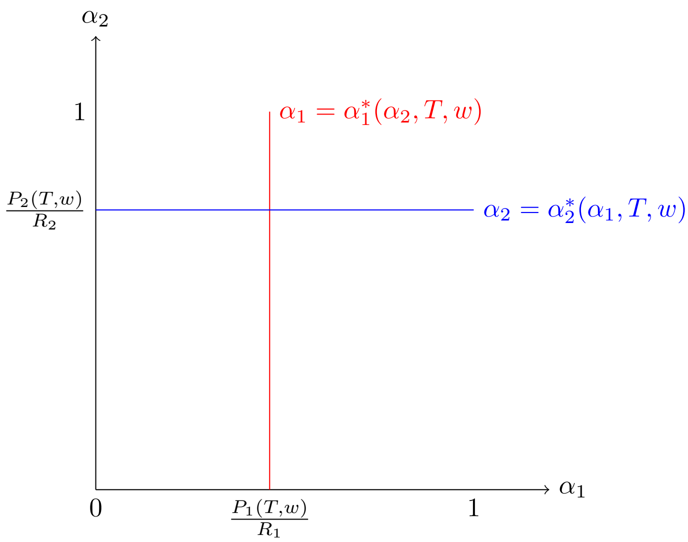
Furthermore, it was assumed that there exists a combination of patent length and breadth such that each initial investment is compensated by the expected private value of the corresponding patent.

$$\exists (T, w) \in (0, \infty) \times (0, 1] : P_1(T, w) \geq R_1 \text{ and } P_2(T, w) \geq R_2 \tag{17}$$

After defining the equilibrium for the second stage of the game, $\alpha_i^*(T, w)$, $i \in 1, 2$, it is time to analyse the regulator's choice. Social welfare would be zero if no potential innovator decided to introduce its idea into the market ($\alpha_1 =$

Figure 1

Equilibrium of the second stage of the game when the innovations are independent (assuming $0 < P_i(T,w) < R_i$ for $i = 1,2$)



$0, \alpha_2 = 0)$. In addition, let $S_{10}(T,w)$ be social welfare —i.e., the discounted market surplus— if only the first innovation arises. Similarly, let the function $S_{01}(T,w)$ represent the actualised social surplus if only the second one is introduced into the market. To be consistent with the assumption that there is no relation between the innovations, assume that social welfare if both innovations arise is

$$S_{11}(T,w) = S_{10}(T,w) + S_{01}(T,w) \tag{18}$$

Certain things about functions $S_{10}(\cdot)$ and $S_{01}(\cdot)$ should be taken into consideration. Since marshallian surplus must be lower during the life of the patent than after its expiration (due to social deadweight loss caused by monopoly power), the more a patent lasts, the lower discounted welfare must be. Formally, this means that $\partial S_{10}/\partial T < 0$ and $\partial S_{01}/\partial T < 0$. Furthermore, wider patents imply more monopoly power, higher social deadweight loss and lower discounted welfare. Formally, this means that $\partial S_{10}/\partial w < 0$ and $\partial S_{10}/\partial w < 0$. As an additional assumption, in this paper functions $S_{10}(\cdot)$ and $S_{01}(\cdot)$ are supposed to be double-differentiable.

Other assumptions in this paper are critical as well, because they should be taken into account in the proofs of the main Propostions in this paper. For instance, negative externalities were excluded from the analysis to focus on the distortions that are expected to be corrected by the patent system, i.e. those arising from the "public good" nature of knowledge. The reader who is looking for an analysis of optimal patent length and breadth with negative externalities should be referred to Eswaran and Gallini (2019). What is more, the social value of an innovation in this model is greater than its private value –i.e., the market value of the corresponding patent supposing there are no frictions.

$$S_{10}(T,w) > P_1(T,w) \tag{19}$$

$$S_{01}(T,w) > P_2(T,w) \tag{20}$$

Then, assuming that effort outcomes are independent, the expected social value of innovation in this model can be defined as $\alpha_1(1-\alpha_2) \cdot S_{10}(T,w) + \alpha_2(1-\alpha_1) \cdot S_{01}(T,w) + \alpha_1\alpha_2 \cdot S_{11}(T,w) = \alpha_1 \cdot S_{10}(T,w) + \alpha_2 \cdot S_{01}(T,w)$. However, since success probabilities α_1 and α_2 are the result of the innovators' reactions to patent policy, which can be inferred from expression (14), it would be more reasonable to write $\alpha_1^*(T,w)$ and $\alpha_2^*(T,w)$ instead. Furthermore, since the cost of getting such probabilities is $\sum_{i\in\{1,2\}} R_i \cdot \alpha_i^*(T,w)^2/2$, the value added by innovation in this model should be written as $\alpha_1^*(T,w)\cdot S_{10}(T,w)+\alpha_2^*(T,w)\cdot S_{01}(T,w)-\sum_{i\in\{1,2\}} R_i\cdot\alpha_i^*(T,w)^2/2$. Assuming that the regulator wants both innovations to happen (i.e., it wants that $\alpha_i^*(T,w)=1$, $i\in\{1,2\}$), the problem of the regulator is given by the following expression.

$$\begin{aligned} \underset{T,w}{\text{maximise}} \quad & S_{10}(T,w)+S_{01}(T,w)-\sum_{i\in\{1,2\}} R_i/2 \\ \text{subject to} \quad & R_i \le P_i(T,w), \qquad i\in 1,2 \\ & w\le 1, \\ & 0\le T, \\ & 0\le w \end{aligned} \tag{21}$$

Before presenting the main results for this model, it would be useful to introduce a series of definitions. Firstly, the length and breadth elasticities of the social value of the innovations are defined by expressions (22) to (25) respectively. On the other hand, the elasticities of the private value of the innovations —which are equal to the market value of the patents, assuming frictionless markets— are defined by (26) to (29).

$$\varepsilon_{S_{10}T}=\frac{\partial S_{10}}{\partial T}\frac{T}{S_{10}} \tag{22}$$

$$\varepsilon_{S_{10}w}=\frac{\partial S_{10}}{\partial w}\frac{w}{S_{10}} \tag{23}$$

$$\varepsilon_{S_{01}T}=\frac{\partial S_{01}}{\partial T}\frac{T}{S_{10}} \tag{24}$$

$$\varepsilon_{S_{01}w}=\frac{\partial S_{01}}{\partial w}\frac{w}{S_{10}} \tag{25}$$

$$\varepsilon_{P_1T}=\frac{\partial P_1}{\partial T}\frac{T}{P_1} \tag{26}$$

$$\varepsilon_{P_1w}=\frac{\partial P_1}{\partial w}\frac{w}{P_1} \tag{27}$$

$$\varepsilon_{P_2T}=\frac{\partial P_2}{\partial T}\frac{T}{P_2} \tag{28}$$

$$\varepsilon_{P_2w}=\frac{\partial P_2}{\partial w}\frac{w}{P_2} \tag{29}$$

These definitions facilitate the expression of the following propositions, which present sufficient conditions for the optimality of patents with maximum breadth, infinite length or any length-breadth combination that satisfies the constraints given by the problem in a setting with two independent innovations. The fact that those conditions are expressed in terms of the length and breadth elasticities of patent values makes it possible to compare them with Takalo's (2001) conclusions regarding optimal patent policy.

The first proposition concerns sufficient conditions for the optimality of maximum-patent breadth.

Proposition 3. *In a model with two independent, patentable innovations, there exists a maximum-breadth patent policy that locally maximises welfare if*

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} \le \min\left\{\frac{\varepsilon_{P_1w}}{\varepsilon_{P_1T}},\frac{\varepsilon_{P_2w}}{\varepsilon_{P_2T}}\right\},\ \textit{or} \tag{30}$$

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} < \max\left\{\frac{\varepsilon_{P_1w}}{\varepsilon_{P_1T}},\frac{\varepsilon_{P_2w}}{\varepsilon_{P_2T}}\right\}\ \textit{and}\quad \exists\ \ T_i\in\mathbb{R}_+ : R_i=P_i(T_i,1)\ \wedge\ R_{-i}\le P_{-i}(T_i,1),\quad i=\underset{j\in\{1,2\}}{\textit{arg max}}\ \frac{\varepsilon_{P_jw}}{\varepsilon_{P_jT}} \tag{31}$$

Proof. See Appendix C.

□

This Proposition have two important corollaries. To begin with, define $T_i(w)$ as the minimum patent length that, given patent breadth $w\in[0,1]$, encourages the i-th innovator to innovate with success probability $\alpha_i=1$.

$$T_i(w)=\min\{T : R_i\le P_i(T,w)\}, \qquad i=1,2 \tag{32}$$

Also define $\underline{T}_i$ as the lower bound of this function. Since $\partial P_i(T,w)/\partial T > 0$ and $\partial P_i(T,w)/\partial w > 0$, it is true that $\partial T_i/\partial w < 0$. Combining this expression with the fact that the domain of the function is $(\underline{w}_i, 1]$, with $\underline{w}_i = \hat{w} \in [0,1] : \lim_{w \to \hat{w}^+} T_i(w) = \infty$, it follows that $\underline{T}_i$ can also be defined as follows.

$$\underline{T}_i = T_i(1) \tag{33}$$

Now consider a situation in which $\underline{T}_1(1) \geq \underline{T}_2(1)$. In this setting, if patent breadth is set at its maximum, any level of patent length satisfying the restriction of the first innovator also satisfies the second restriction. This means that the constraint given by the incentives for the second innovator can be ignored when evaluating a solution with $w = 1$. Moreover, if the social welfare function and the restriction given by the first restriction are such that welfare for a policy mix with $w = 1$ is higher than for any other combination within that constraint, then the second restriction can also be ignored for any combination with $w < 1$. This leads to the first corollary: Assuming that $\underline{T}_1(1) \geq \underline{T}_2(1)$, maximum patent breadth ($w = 1$) is optimal if, for any pair $(T, w) \in [0, \infty) \times [0, 1]$, the following condition holds:

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} < \frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}} \tag{34}$$

The proof of this corollary consists of two steps. In the first one, show that, if the constraint $R_2 \leq P_2(T, w)$ is ignored, then the only solution that fulfils the Karush-Kuhn-Tucker optimality conditions is $(T = \underline{T}_1, w = 1)$. In the second step, prove that such solution also satisfies the ignored constraint if $\underline{T}_1(1) \geq \underline{T}_2(1)$.

A situation in which $\underline{T}_1(1) \leq \underline{T}_2(1)$ can be analysed using an analogous reasoning. If the social welfare function and the restriction given by the incentives of the second innovator are such that a patent policy with $w = 1$ maximises social welfare subject to the constraint of the first innovator, patent breadth in this setting should be set at its maximum. Then, assuming that $\underline{T}_1(1) \leq \underline{T}_2(1)$, maximum patent breadth ($w = 1$) is optimal if, for any pair $(T, w) \in [0, \infty) \times [0, 1]$, the following condition holds:

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} < \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}} \tag{35}$$

After having demonstrated the first corollary, the proof of the second one is trivial. First prove that, if the constraint $R_1 \leq P_1(T, w)$ is ignored, then the only solution that fulfils the Karush-Kuhn-Tucker optimality conditions is $(T = \underline{T}_1, w = 1)$. In a second step, show that such solution also satisfies the ignored constraint if $\underline{T}_1(1) \leq \underline{T}_2(1)$.

It is interesting to observe that conditions (34) and (35) are similar. Concretely, they both compare the derivatives of social and private values of innovations with respect to patent length and breadth. The only difference is that condition (34) focus on the private value of the first innovation, whereas condition (35) focus on the private value of the second one.

In a similar manner, sufficient conditions for an infinite-patent-length policy can be inferred. However, things are a little more complicated in this case. Formally, there is no such thing as a solution with infinite length, so this part of the exposition will only show that, under the proposed conditions, the regulator wants to extend patent duration as much as possible. Besides, it will be shown that optimal patent breadth in such a context would be "minimum" provided the constraints of the problem. Proposition 4 presents those conditions in a concise manner.

Proposition 4. *In a model with two independent, patentable innovations, an infinite-length patent policy would be the best choice if*

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} \geq \max\left\{\frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}}, \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}}\right\}, \ \textit{or} \tag{36}$$

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} > \min\left\{\frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}}, \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}}\right\} \ \textit{and} \ \ \exists \ \ w_i \in [0,1] : R_i = \lim_{T \to \infty} P_i(T, w_i) \ \wedge \ R_{-i} \leq \lim_{T \to \infty} P_{-i}(T, w_i), \quad i = \textit{arg min}_{j \in \{1,2\}} \frac{\varepsilon_{P_j w}}{\varepsilon_{P_j T}} \tag{37}$$

Proof. See Appendix D. □

Analogously to the maximum-breadth analysis, the conditions can be organised into three propositions, depending on the values of $\underline{w}_1$ and $\underline{w}_2$. Remember that

$$\underline{w}_i = \hat{w} : \lim_{w \to \hat{w}^+} T_i(w) = \infty \tag{38}$$

Now consider a situation in which $\underline{w}_1 \geq \underline{w}_2$. From the definitions of limit it follows that there exists a value for patent duration $\hat{T}$ such that, any patent length $T > \hat{T}$ satisfying the first restriction of the problem also satisfies the second one. This implies that the constraint given by the incentives for the second innovator can be ignored as $T \to \infty$. Then, assuming that $\underline{w}_1 \geq \underline{w}_2$, infinite patent length ($T \to \infty$) would be optimal if, for any pair $(T, w) \in [0, \infty) \times [0, 1]$, the following condition holds.

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} > \frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}} \tag{39}$$

The proof of this corollary is straightforward. If condition (39) holds, for any length-breadth combination (T_0, w_0) such that $R \leq P_1(T_0, w_0)$, there must exist another policy mix (T_1, w_1) such that $R \leq P_1(T_1, w_1)$, $T_0 < T_1$, and $S_{10}(T_0, w_0) + S_{01}(T_0, w_0) < S_{10}(T_1, w_1) + S_{01}(T_1, w_1)$. Furthermore, if $\underline{w}_1 \geq \underline{w}_2$ and $\hat{T} < T_0$, then $R \leq P_2(T_0, w_0)$ and $R \leq P_2(T_1, w_1)$. In other words, if the condition in the corollary is true, then for every solution with long patents that satisfies both restrictions, there exists another solution that also satisfies them and makes the regulator reach a higher level of social welfare.

Situations in which $\underline{w}_1(\infty) \leq \underline{w}_2(\infty)$ can be analysed using an analogous reasoning. More concretely, if the social welfare function and the restriction given by the second restriction are such that a patent policy with $T \to \infty$ maximises social welfare subject to that constraint, and the first restriction is not active at such point, then patent length in this setting should be set at its maximum. In other words, assuming that $\underline{w}_1 \leq \underline{w}_2$, infinite patent breadth ($T \to \infty$) would be optimal if, for any pair $(T, w) \in [0, \infty) \times [0, 1]$, the following condition holds.

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} > \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}} \tag{40}$$

As before, conditions in these corollaries are similar, because they compare the derivatives of the social and private values of the innovations. But condition (39) focus on the private value of the first innovation, whereas condition (40) focus on the private value of the second one.

A comparison between the Propositions in the previous section and the Propositions in this one leads to several conclusions. First, in both the stand-alone case and the scenario with two independent innovations it was possible to infer sufficient conditions for the optimality of maximum patent breadth and infinite length from the elasticities of the innovation values. Second, the following parallelism can be traced between the Propositions: if the stand-alone condition for the optimality of a policy is true for both innovations, then that policy is also optimal in the case with two independent innovations. Third, alternative sets of conditions can be formulated if the private value of each innovation under the evaluated policy is known. Concretely, if the stand-alone condition for the optimality of maximum breadth (infinite length) holds for the less valuable innovation –i.e., the innovation which needs a higher value for patent length (breadth) to become profitable–, then it can be concluded that such policy maximises welfare.

It would be interesting to figure out if the same conclusions could be extended to a cumulative innovation setting. For instance, it can be analysed what would happen if one of these innovations was a “research tool” that should be used in the development of the other one. The next section presents deals with that question.

5 Two linked innovations

The previous models omitted issues arising from the cumulative nature of innovation, i.e. the fact that new ideas are usually the result of building on the contributions made by other inventors. This section makes an attempt to fill this gap, concretely by extending the previous section approach to study a "research tool" scenario, which constitutes just one of those settings where the cumulative nature of knowledge has consequences on the optimal patent policy. As before, the model in this section is a two-stage game where the players are a regulator and two potential innovators, who will be identified as the "first" innovator and the "second" one.

According to Scotchmer (1991), who introduced the idea that the cumulative nature of innovation must have implications for the optimality of patent policy, the fact that the first innovation is a "research tool" for the second one might mean one of the following things: (a) that the second innovation could not be invented if the first innovation was not revealed, (b) that the cost of developing the second innovation would be lower if the first innovation was disclosed, or (c) that the development of the second innovation would demand less time if the first innovation was introduced in the market.

This work deals with the first case only. In other words, supposing that each innovative idea constitutes "private knowledge", it is assumed that one of them –the idea of the first innovator, or the "first innovation"– is a "research tool" for the development of the other one –the idea of the second innovator, or the "second innovation". This means that the idea of the second innovator cannot be introduced in the market if the idea of the first one has not been revealed, because the first idea is the only thing that makes the second one possible. Such interdependency constitutes the essence of "research tool" innovation.

As in the previous section, the model assumes that R&D is costly for both potential innovators, but R&D costs are independent. This means that cost functions are the same as in the model with independent innovations. However, the innovators' independence disappears when profits are taken into consideration. The second innovator earns nothing if the first innovation has not arisen, but might collect some revenue if the first innovator has disclosed his idea. On the other hand, the first innovator might prefer a world without the second innovation or a setting where such idea has been revealed. The first situation might arise if the second innovation is a competing product or an improvement that erodes the first innovator's profit, and the potential innovators cannot sign a licence agreement or are not allowed to do so. The second situation might occur when the first innovator is allowed to collect some revenue from the second one by means of a licence agreement.

However, in this paper it is assumed that licence agreements are illegal. In this context, suppose that $P_1(T, w)$ is the present value of the benefits that the first innovator will get in case of success when patent length and breadth are T and w, respectively. Observe that $P_1(\cdot)$ does not depend on the second innovator's outcome. This assumption reflects the fact that the innovators' markets are not related and the first innovator cannot earn a revenue from licence agreements. Allowing for further abuse of notation, assume that $P_2(T, w)$ is the value of the benefits that the second innovator will get in case of success. In other words, this function returns the discounted revenue of the second innovator assuming that the first innovator has been successful. This explains why the first innovator's outcome cannot be seen as an argument of this function.

Taking the assumptions in the previous paragraphs into consideration, the problem of the first innovator must still be given by expression (13), whereas the problem of the second must be represented by (41).

$$\begin{aligned} \underset{\alpha_2}{\text{maximise}} \quad & P_2(T, w) \cdot \alpha_1 \cdot \alpha_2 - R_2 \cdot \alpha_2^2/2 \\ \text{subject to} \quad & \alpha_2 \leq 1, \\ & 0 \leq \alpha_2 \end{aligned} \tag{41}$$

The optimal decision for the first innovator is still given by (14), but the reaction of the second innovator must be (42), which shows that its optimal level of effort depends on the decision of the first innovator.

$$\alpha_2^*(\alpha_1, T, w) = \begin{cases} 0 & \text{if} \quad P_1(T, w) \cdot P_2(T, w) < 0 \\ \alpha_1 \cdot \frac{P_2(T, w)}{R_2} & \text{if} \quad 0 \leq P_1(T, w) \cdot P_2(T, w) < R_1 \cdot R_2 \\ 1 & \text{if} \quad R_1 \cdot R_2 \leq P_1(T, w) \cdot P_2(T, w) \end{cases} \tag{42}$$

Formally, the "equilibrium functions" for the first innovator and the second one are (43) and (44), respectively.

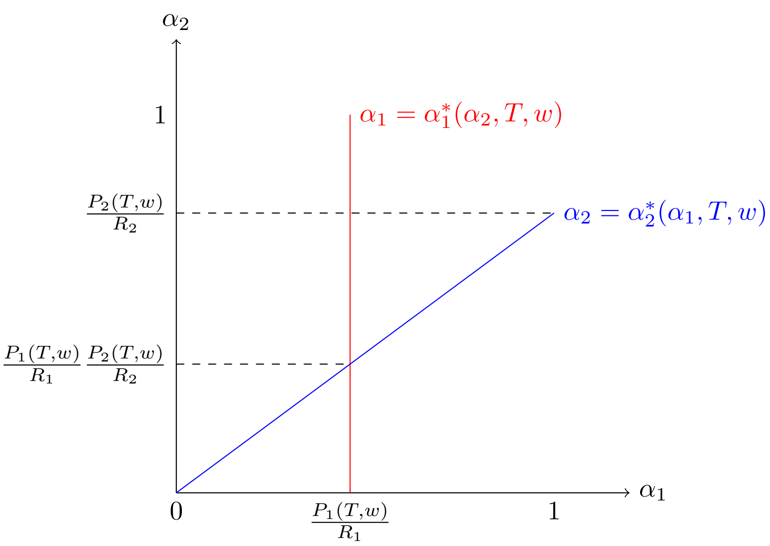
$$\alpha_1^*(T,w) = \begin{cases} 0 & \text{if} \quad P_1(T,w) < 0 \\ P_1(T,w)/R_1 & \text{if} \quad 0 \leq P_1(T,w) < R_1 \\ 1 & \text{if} \quad R_1 \leq P_1(T,w) \end{cases} \tag{43}$$

$$\alpha_2^*(T,w) = \begin{cases} 0 & \text{if} \quad P_1(T,w)P_2(T,w) < 0 \\ [P_1(T,w)/R_1][P_2(T,w)/R_2] & \text{if} \quad 0 \leq P_1(T,w) < R_1 \text{ and } 0 \leq P_1(T,w)P_2(T,w) < R_1R_2 \\ P_2(T,w)/R_2 & \text{if} \quad R_1 \leq P_1(T,w) \text{ and } 0 \leq P_2(T,w) < R_2 \\ 1 & \text{if} \quad R_1R_2 \leq P_1(T,w)P_2(T,w) \text{ and } R_2 \leq P_2(T,w) \end{cases} \tag{44}$$

The equilibrium of this subgame (assuming $0 < P_1(T,w) < R_1$ and $0 < P_1(T,w)P_2(T,w) < R_1R_2$) is shown in Figure 2. A comparison with Figure 1 reveals that, profit and cost functions held constant, the cumulative nature of knowledge has a negative effect on the optimal effort of the second innovator. Of course, the situation would be different if the innovators could coordinate or sign a license agreement, but the analysis of such a scenario is beyond the scope of this paper.

Figure 2

Equilibrium of the second stage of the game when the innovations are linked
(assuming $0 < P_1(T,w) < R_1$ and $0 < P_1(T,w)P_2(T,w) < R_1R_2$)



The main question is how patent length and breadth should be in such a context. This requires analysing the regulator's choice. As in the previous section, the regulator is in charge of patent length (T) and breadth (w). The main differences in relation to section 4 concern the objective function of the regulator. Firstly, it must be assumed that

$$S_{01}(T,w) = 0 \tag{45}$$

to represent the fact that it is impossible for the second innovator to introduce something valuable into the market –or that its innovative effort would be sterile– if the first innovator has not been successful. Secondly, assumption (18) must be replaced by

$$S_{11}(T,w) > S_{10}(T,w) \tag{46}$$

The last expression means that introducing the second innovation will produce a rise in social welfare in comparison with the situation where only the first innovation arises.

The rest of the assumptions in section 4 hold. Combining them with (45), (46), and the equilibrium functions $\alpha_1^*(T,w)$ and $\alpha_2^*(T,w)$ results in the following representation of the regulator's objective function: $\alpha_1^*(T,w)(1-\alpha_2^*(T,w))S_{10}(T,w) + \alpha_1^*(T,w)\alpha_2^*(T,w)S_{11}(T,w) - \sum_{i\in\{1,2\}} R_i \cdot \alpha_i^*(T,w)^2/2$. In addition, if the restriction that $\alpha_i^*(T,w) = 1$, $i \in \{1,2\}$, is taken into consideration, the regulator's problem is

$$\begin{aligned}
\underset{T,w}{\text{maximise}} \quad & S_{11}(T,w) - \sum_{i\in\{1,2\}} R_i/2 \\
\text{subject to} \quad & R_1 \leq P_1(T,w), \\
& R_1 R_2 \leq P_1(T,w)P_2(T,w), \\
& R_2 \leq P_2(T,w), \\
& w \leq 1, \\
& 0 \leq T, \\
& 0 \leq w
\end{aligned} \tag{47}$$

This problem resembles (21). In fact, the condition that $R_1R_2 \leq P_1(T,w)P_2(T,w)$ in (47) is redundant, and $S_{10}(T,w) + S_{01}(T,w)$ in (21) could be replaced by $S_{11}(T,w)$. What is more, Proofs in Appendixes C and D can also be used to analyse optimal patent policy in the "research tool" setting (just replace $\partial S_{10}/\partial w + \partial S_{01}/\partial w$ and $\partial S_{10}/\partial T + \partial S_{01}/\partial T$ by $\partial S_{11}/\partial w$ and $\partial S_{11}/\partial T$ in the proofs in question). As a consequence, Propositions 3 and 4 are also valid in this context.

The explanation of this behaviour is straightforward. The main difference between the independent innovations scenario and the "research tool" setting consisted of a lower effort for the second innovator in the model with linked innovations. Since the restrictions of the problems imposed specific levels of effort on the innovators, such difference was eliminated.

Having said that, it would be interesting to infer under which conditions the regulator would really want to impose the proposed levels of innovative effort ($\alpha_1 = \alpha_2 = 1$). This step consists of limiting the analysis to situations where both innovations are socially valuable. To reach that goal, an alternative consists of assuming that *ex-post* social surplus ($S_{11}(T,w) - R_1/2 - R_2/2$) is positive for any length-breadth combination.

$$S_{11}(T,w) > R_1/2 + R_2/2 \quad \forall \quad (T,w) \in [0,\infty) \times [0,1] \tag{48}$$

Secondly, to avoid reducing the analysis of the model to the study of a single-innovation setting, it should be assumed that, for any length-breadth combination, the second innovation generates a positive surplus in comparison with the situation where only the first innovative idea is introduced into the market. Concretely, the social surplus generated for the second innovation, given patent length $T \in [0,\infty)$ and breadth $w \in [0,1]$, is $S_2(T,w) - S_1(T,w) - R_2/2$, and the assumption in question implies

$$S_{11}(T,w) > S_{10}(T,w) + R_2 \quad \forall \quad (T,w) \in [0,\infty) \times [0,1] \tag{49}$$

There is an additional supposition about $S_1(\cdot)$ and $S_2(\cdot)$ that should be imposed. Concretely, an interesting situation arises if the following condition holds.

$$\lim_{T\to\infty} S_{11}(T,1) > S_{10}(0,0) + R_2/2 \tag{50}$$

This ensures, in a "research tool" setting, that using the patent system to incentivise both innovations is always better than using it to incentivise only the first one. This means that it could be assumed that $\alpha_1 = 1$ and $\alpha_2 = 1$ without affecting the optimal solution to the problem. Then, solving the problem of the regulator without imposing restrictions on the desired effort levels is equivalent to solving (47), because the optimal solutions to these problems coincide.

The independent innovations model requires a different set of conditions. First, it should be assumed that the social surplus generated by each innovation is positive.

$$S_{10} > R_1/2 \tag{51}$$

$$S_{01} > R_2/2 \tag{52}$$

Secondly, since in this model it is possible to get the second innovation without introducing the first one, it should be assumed that it is always better to encourage both innovations than incentivising just one of them. As a consequence, in addition to (50), the following assumption must be imposed.

$$\lim_{T\to\infty} S_{11}(T,1) > S_{01}(0,0) + R_1/2 \tag{53}$$

These assumptions are enough to reach the conclusion, in a model with two independent innovations, that solving the problem of the regulator without imposing restrictions on the desired effort levels is equivalent to solving (21).

6 Conclusion

This article aimed at finding a relation between the conditions for the optimality of specific patent policies and the elasticities of the (private and social) values of innovations. It was also an objective to analyse the effect of the coexistence of independent innovations and the cumulativeness of knowledge on that relation. To reach such objectives, three models were presented and analysed. The first one was Takalo's (2001) model of optimal patent policy for an isolated innovation. The second one was an extension of such model that was built to analyse a setting with two independent innovations. The third model featured two innovations as well, but they were linked: firtst one was a "research tool", or a necessary condition, for the development and the introduction of the second one. In all of them it was analysed the optimal patent policy under the restriction that both innovators should be encouraged to invest in R&D to ensure the emergence of both innovators.

The results show that it is possible to formulate optimality conditions for a policy with maximum patent breadth, or infinite patent length, departing from the elasticities of the private and social value of innovations to changes in patent length and breadth. In the first model, comparing the elasticity ratio for the social value of the innovation (lets call it "the social ratio") with the ratio for its private value ("the private ratio") is enough to infer if maximum breadth or infinite length, or neither of them, was optimal. In the models with two innovations, comparing the ratios makes it possible to get a conclusion when the social ratio is greater or lower than both private ratios (i.e., the ratio of the first patent and the ratio of the second one). However, when the social ratio lied between the private ratios, it was also possible to make a conclusion depending on which patent was less profitable. Another interesting result was the fact that optimality conditions were the same for both sorts of models with two innovations.

The last result seems to suggest that the cumulative nature of knowledge has no effect on the optimality of patent policy. Nothing could be further from the truth. That result arose because it was imposed on the optimisation problems the restriction that innovators must be encouraged to choose the maximum level of innovative effort. Exploring what happens in an unrestricted version of the problem is a task that should be addressed in future work. Furthermore, it would be interesting to explore what would happen with the conditions if the regulator was allowed to set different values of length and breadth for different innovations. It would also be enlightening to explore more complex models, for instance, those featuring alternative models of innovation ("quality ladder", complementary innovations, etc.). Other interesting extensions should include patent races, open economies, asymmetric information, or general equilibrium effects. The new models would be more appropriate to analyse patent policy in a real world situation and would be more likely to have a significant, positive impact in the allocation of resources in R&D markets.

References

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In R. Nelson (Ed.), *The rate and direction of inventive activity: Economic and social factors* (pp. 609–626). Princeton University Press. http://www.nber.org/chapters/c2144

Bagchi, A., & Roy, A. (2012). Endogenous research and development and intellectual property laws in developed and emerging economies. *Southern Economic Journal*, *78*(3), 895–930. https://doi.org/10.4284/0038-4038-78.3.895

Bergin, J. (2022). Patent licensing, non-practising entities, and investment in R&D. *The Journal of Industrial Economics*, *70*(2), 396–462. https://doi.org/10.1111/joie.12289

Chen, J., Liu, J., Long, Y., & Luo, J. (2019). Medicine pricing, optimal patent length and social welfare. *Revista de Cercetare si Interventie Sociala*, *67*, 205–222. https://doi.org/10.33788/rcis.67.13

Denicolò, V. (1996). Patent races and optimal patent breadth and length. *Journal of Industrial Economics*, *44*(3), 249–266. https://doi.org/10.2307/2950496

Denicolò, V., & Zanchettin, P. (2022). Patent protection for complex technologies. *International Journal of Industrial Organization*, *81*, 1–20. https://doi.org/10.1016/j.ijindorg.2021.102811

Eswaran, M., & Gallini, N. (2019). Can competition and patent policies avert the antibiotic crisis? *Canadian Public Policy*, *45*(1), 74–92. https://doi.org/10.3138/cpp.2018-021

Galasso, A., Mitchell, M., & Virag, G. (2016). Market outcomes and dynamic patent buyouts. *International Journal of Industrial Organization*, *48*, 207–243. https://doi.org/10.1016/j.ijindorg.2016.06.007

Gallini, N. (1992). Patent policy and costly imitation. *The RAND Journal of Economics*, *23*(1), 52–63. https://doi.org/10.2307/2555432

Gilbert, R., & Shapiro, C. (1990). Optimal patent length and breadth. *The RAND Journal of Economics*, *21*(1), 106–112. https://doi.org/10.2307/2555497

Green, J., & Scotchmer, S. (1995). On the division of profit in sequential innovation. *The RAND Journal of Economics*, *26*(1), 20–33. https://doi.org/10.2307/2556033

Hall, B., & Harhoff, D. (2012). Recent research on the economics of patents. *Annual Review of Economics*, *4*, 541–565. https://doi.org/10.1146/annurev-economics-080511-111008

Khazabi, M., & van Quyen, N. (2017). The search for new drugs: A theory of R&D in the pharmaceutical industry. *Journal of Economic Studies*, *44*(5), 690–726. https://doi.org/10.1108/JES-01-2016-0002

Klemperer, P. (1990). How broad should the scope of patent protection be? *The RAND Journal of Economics*, *21*(1), 113–130. https://doi.org/10.2307/2555498

Matutes, C., Regibeau, P., & Rockett, K. (1996). Optimal patent design and the diffusion of innovations. *The RAND Journal of Economics*, *27*(1), 60–83. https://doi.org/10.2307/2555792

Meng, S. (2017). Optimal patent protection: A macroeconomic perspective. *Asian Journal of Law and Economics*, *8*(2), 1–16. https://doi.org/10.1515/ajle-2016-0021

Nordhaus, W. (1967). *The optimal life of a patent* (Discussion Paper No. 241). Cowles Foundation for Research in Economics at Yale University. New Haven, CT. https://elischolar.library.yale.edu/cowles-discussion-paper-s%20eries/474

Nordhaus, W. (1969). *Invention, growth and welfare: A theoretical treatment of technological change*. The MIT Press.

Nordhaus, W. (1972). The optimum life of a patent: Reply. *The American Economic Revew*, *62*(3), 428–423. http://www.jstor.org/stable/1803389

Reinganum, J. (1989). The timing of innovation: Research, development and diffusion. In R. Schmalensee & R. Willig (Eds.), *Handbook of industrial organization* (pp. 849–908, Vol. 1). North-Holland.

Rockett, K. (2010). Property rights and invention. In B. H. Hall & N. Rosenberg (Eds.), *Handbook of the economics of innovation* (pp. 315–380, Vol. 1). North-Holland. https://doi.org/10.1016/S0169-7218(10)01007-5

Scherer, F. (1972). Nordhaus' theory of optimal patent life: A geometric reinterpretation. *American Economic Review*, *62*(3), 422–427. http://www.jstor.org/stable/1803388

Scotchmer, S. (1991). Standing on the shoulders of giants: Cumulative research and the patent law. *Journal of Economic Perspectives*, *5*(1), 29–41. https://doi.org/10.1257/jep.5.1.29

Scotchmer, S. (2004). *Innovation and incentives*. The MIT Press.

Takalo, T. (2001). On the optimal patent policy. *Finnish Economic Papers*, *14*(1), 33–40.

Tandon, P. (1982). Optimal patents with compulsory licensing. *Journal of Political Economy*, *90*(3), 470–486. https://doi.org/10.1086/261070

Volpert, T., & Riepe, M. (2021). Patentdimensionen und die entwicklung und herstellung von corona-impfstoffen. *Wirtschaftsdienst*, *101*(5), 387–393. https://doi.org/10.1007/s10273-021-2923-4

Wright, B. (1983). The economics of invention incentives: Patents, prizes and research contracts. *American Economic Review*, *73*(4), 691–707. https://www.jstor.org/stable/1816567

A Proof of Proposition 1

The Lagrangian for the problem in (3) is represented by the following expression (for simplicity, arguments have been omitted).

$$L = S - \frac{R}{2} + \lambda(P - R) + \lambda_w(1 - w) \tag{54}$$

The Karush-Kuhn-Tucker (first-order) conditions for a local optimum are

$$\begin{aligned}
L_T &= \frac{\partial S}{\partial T} + \lambda \frac{\partial P}{\partial T} \leq 0 & (55)\\
T \cdot L_T &= T \cdot \left[\frac{\partial S}{\partial T} + \lambda \frac{\partial P}{\partial T} \right] = 0 & (56)\\
L_w &= \frac{\partial S}{\partial w} + \lambda \frac{\partial P}{\partial w} - \lambda_w \leq 0 & (57)\\
w \cdot L_w &= w \cdot \left[\frac{\partial S}{\partial w} + \lambda \frac{\partial P}{\partial w} - \lambda_w \right] = 0 & (58)\\
L_\lambda &= P - R \geq 0 & (59)\\
\lambda \cdot L_\lambda &= \lambda \cdot (P - R) = 0 & (60)\\
L_{\lambda_w} &= 1 - w \geq 0 & (61)\\
\lambda_w \cdot L_{\lambda_w} &= \lambda_w \cdot (1 - w) = 0 & (62)\\
\lambda &\geq 0 & (63)\\
\lambda_w &\geq 0 & (64)
\end{aligned}$$

It was assumed that $P(0, w) < R$, and $\{(T, w) \in [0, \infty) \times [0, 1] : R \leq P(T, w)\} \neq \emptyset$. This implies that, in the optimum, $T > 0$ and the following expressions must be true.

$$\begin{aligned}
\frac{\partial S}{\partial T} + \lambda \frac{\partial P}{\partial T} &= 0 & (65)\\
\frac{\partial S}{\partial T} &= -\lambda \frac{\partial P}{\partial T} & (66)\\
-\frac{\frac{\partial S}{\partial T}}{\frac{\partial P}{\partial T}} &= \lambda & (67)
\end{aligned}$$

Since we are looking for sufficient conditions for the optimality of $w = 1$, it should be assumed that $w > 0$. Then, it follows that

$$\begin{aligned}
\frac{\partial S}{\partial w} + \lambda \frac{\partial P}{\partial w} - \lambda_w &= 0 & (68)\\
\frac{\partial S}{\partial w} + \lambda \frac{\partial P}{\partial w} &= \lambda_w & (69)
\end{aligned}$$

The next steps will show sufficient conditions for $\lambda_w > 0$, since in that situation the optimal w must be the maximum ($w = 1$). To see why, look at condition (62). If the objective function and the restrictions are such that in any feasible point the only λ_w that fulfils them is a positive value, then $(1 - w)$ in that point must be equal to zero. In other words, it must be true that $w = 1$ in a point that satisfies the optimality conditions.

Concretely, according to those conditions, it must be true that $\lambda_w > 0$ if

$$\lambda_w = \frac{\partial S}{\partial w} + \lambda \frac{\partial P}{\partial w} > 0 \tag{70}$$

Combining (67) and (70), it follows that

$$\lambda_w = \frac{\partial S}{\partial w} + \lambda \frac{\partial P}{\partial w} > 0 \tag{71}$$

$$\frac{\partial S}{\partial w} - \frac{\frac{\partial S}{\partial T}}{\frac{\partial P}{\partial T}} \frac{\partial P}{\partial w} > 0 \tag{72}$$

$$\frac{\partial S}{\partial w} > \frac{\frac{\partial S}{\partial T}}{\frac{\partial P}{\partial T}} \frac{\partial P}{\partial w} \tag{73}$$

$$\frac{\frac{\partial S}{\partial w}}{\frac{\partial S}{\partial T}} < \frac{\frac{\partial P}{\partial w}}{\frac{\partial P}{\partial T}} \tag{74}$$

$$\frac{\varepsilon_{Sw}}{\varepsilon_{ST}} < \frac{\varepsilon_{Pw}}{\varepsilon_{PT}} \tag{75}$$

In sum, if $\frac{\varepsilon_{Sw}}{\varepsilon_{ST}} < \frac{\varepsilon_{Pw}}{\varepsilon_{PT}}$ and there is a point which fulfils the Karush-Kuhn-Tucker conditions for an optimal solution, then it must be true that $\lambda_w > 0$ in that point. By (62), it must also be true that $w = 1$ in such optimum. This completes the proof.

B Proof of Proposition 2

Consider a restricted version of the problem in (3):

$$\begin{aligned} \underset{T,w}{\text{maximise}} \quad & S(T,w) - R/2 \\ \text{subject to} \quad & R \leq P(T,w), \\ & T \leq \overline{T}, \\ & w \leq 1, \\ & 0 \leq T, \\ & 0 \leq w \end{aligned} \tag{76}$$

The Lagrangian for that problem is represented by

$$L = S - \frac{R}{2} + \lambda(P - R) + \lambda_T(\overline{T} - T) + \lambda_w(1 - w) \tag{77}$$

Expressions (57) to (64), which can be found in Appendix A, and (78) to (82) provide the Karush-Kuhn-Tucker (first-order) conditions for a locally optimal solution to such problem.

$$L_T = \frac{\partial S}{\partial T} + \lambda\frac{\partial P}{\partial T} - \lambda_T \leq 0 \tag{78}$$

$$T \cdot L_T = T \cdot \left[\frac{\partial S}{\partial T} + \lambda\frac{\partial P}{\partial T} - \lambda_T\right] = 0 \tag{79}$$

$$L_{\lambda_T} = \overline{T} - T \geq 0 \tag{80}$$

$$\lambda_T \cdot L_{\lambda_T} = \lambda_T \cdot (\overline{T} - T) = 0 \tag{81}$$

$$\lambda_T \geq 0 \tag{82}$$

Since it was assumed that $P(T,0) < R$, and $\{(T,w) \in [0,\infty) \times [0,1] : R \leq P(T,w)\} \neq \emptyset$, $w > 0$ in the optimum. As a consequence, expressions (68) and (69), which can be found in Appendix A, still hold. Combining them with (64) leads to

$$\lambda_w = \frac{\partial S}{\partial w} + \lambda\frac{\partial P}{\partial w} \geq 0 \tag{83}$$

$$\frac{\partial S}{\partial w} \geq -\lambda\frac{\partial P}{\partial w} \tag{84}$$

$$-\frac{\frac{\partial S}{\partial w}}{\frac{\partial P}{\partial w}} \leq \lambda \tag{85}$$

Furthermore, since we are looking for sufficient conditions for the desirability of $T \to \infty$, it should be assumed that $T > 0$. Then, it follows that

$$\frac{\partial S}{\partial T} + \lambda\frac{\partial P}{\partial T} - \lambda_T = 0 \tag{86}$$

$$\frac{\partial S}{\partial T} + \lambda\frac{\partial P}{\partial T} = \lambda_T \tag{87}$$

The next steps will show sufficient conditions for $\lambda_T > 0$, because in that situation the optimal T must be $\overline{T}$. To see why, look at condition (81). If the objective function and the restrictions are such that in any feasible point the only

λ_T that fulfils them is a positive value, then $(\overline{T} - T)$ in that point must be equal to zero. In other words, it must be true that $T = \overline{T}$ at any point satisfying the optimality conditions. Then, as $\overline{T} \to \infty$, problem (76) approaches problem (3), and the optimal patent length approaches infinity.

Concretely, an analysis of (87) reveals a sufficient condition for $\lambda_T > 0$.

$$\lambda_T = \frac{\partial S}{\partial T} + \lambda \frac{\partial P}{\partial T} > 0 \tag{88}$$

After combining (85) and (87), a more restrictive, albeit enlightening, condition is obtained.

$$\lambda_T \geq \frac{\partial S}{\partial T} - \frac{\frac{\partial S}{\partial w}}{\frac{\partial P}{\partial w}} \frac{\partial P}{\partial T} > 0 \tag{89}$$

$$\frac{\partial S}{\partial T} > \frac{\frac{\partial S}{\partial w}}{\frac{\partial P}{\partial w}} \frac{\partial P}{\partial T} \tag{90}$$

$$\frac{\frac{\partial P}{\partial w}}{\frac{\partial P}{\partial T}} < \frac{\frac{\partial S}{\partial w}}{\frac{\partial S}{\partial T}} \tag{91}$$

$$\frac{\varepsilon_{Pw}}{\varepsilon_{PT}} < \frac{\varepsilon_{Sw}}{\varepsilon_{ST}} \tag{92}$$

To sum up, if $\frac{\varepsilon_{Pw}}{\varepsilon_{PT}} < \frac{\varepsilon_{Sw}}{\varepsilon_{ST}}$ and there is an optimal solution to the problem in (76), then $\lambda_T > 0$ at such point. By (81), this implies that $T = \overline{T}$ in the optimum. Furthermore, as $\overline{T} \to \infty$, problem (76) approaches (3) and the optimal patent lenght approaches infinity as well $(T \to \infty)$. This completes the proof.

C Proof of Proposition 3

The Lagrangian for the problem in (21) is represented by the following expression (for simplicity, arguments have been omitted).

$$L = S_{10} + S_{01} - \frac{R_1}{2} - \frac{R_2}{2} + \lambda_1(P_1 - R_1) + \lambda_2(P_2 - R_2) + \lambda_w(1 - w) \tag{93}$$

The Karush-Kuhn-Tucker (first-order) conditions for a local optimum are

$$L_T = \frac{\partial S_{10}}{\partial T} + \frac{\partial S01}{\partial T} + \lambda_1 \frac{\partial P_1}{\partial T} + \lambda_2 \frac{\partial P_2}{\partial T} \leq 0 \tag{94}$$

$$T \cdot L_T = T \cdot \left[\frac{\partial S_{10}}{\partial T} + \frac{\partial S01}{\partial T} + \lambda_1 \frac{\partial P_1}{\partial T} + \lambda_2 \frac{\partial P_2}{\partial T} \right] = 0 \tag{95}$$

$$L_w = \frac{\partial S_{10}}{\partial w} + \frac{\partial S01}{\partial w} + \lambda_1 \frac{\partial P_1}{\partial w} + \lambda_2 \frac{\partial P_2}{\partial w} - \lambda_w \leq 0 \tag{96}$$

$$w \cdot L_w = w \cdot \left[\frac{\partial S_{10}}{\partial w} + \frac{\partial S01}{\partial w} + \lambda_1 \frac{\partial P_1}{\partial w} + \lambda_2 \frac{\partial P_2}{\partial w} - \lambda_w \right] = 0 \tag{97}$$

$$L_{\lambda_1} = P_1 - R_1 \geq 0 \tag{98}$$

$$\lambda_1 \cdot L_{\lambda_1} = \lambda_1 \cdot (P_1 - R_1) = 0 \tag{99}$$

$$L_{\lambda_2} = P_2 - R_2 \geq 0 \tag{100}$$

$$\lambda_2 \cdot L_{\lambda_2} = \lambda_2 \cdot (P_2 - R_2) = 0 \tag{101}$$

$$L_{\lambda_w} = 1 - w \geq 0 \tag{102}$$

$$\lambda_w \cdot L_{\lambda_w} = \lambda_w \cdot (1 - w) = 0 \tag{103}$$

$$\lambda_1 \geq 0 \tag{104}$$

$$\lambda_2 \geq 0 \tag{105}$$

$$\lambda_w \geq 0 \tag{106}$$

It was assumed that $P_1(0, w) < R_1$, $P_2(0, w) < R_2$, and $\{(T, w) \in [0, \infty) \times [0, 1] : R_1 \leq P_1(T, w) \wedge R_2 \leq P_2(T, w)\} \neq \emptyset$. This implies that, in the optimum, $T > 0$ and the following expression must be true.

$$\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} + \lambda_1 \frac{\partial P_1}{\partial T} + \lambda_2 \frac{\partial P_2}{\partial T} = 0 \tag{107}$$

Let i be the index of the patent with the highest elasticity ratio: $i = j \in \{1, 2\} : \frac{\varepsilon_{P_j w}}{\varepsilon_{P_j T}} = \max \left\{ \frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}}, \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}} \right\}$. Similarly, let $-i$ be the index which represents the patent with the lowest ratio: $-i = j \in \{1, 2\} : \frac{\varepsilon_{P_j w}}{\varepsilon_{P_j T}} = \min \left\{ \frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}}, \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}} \right\}$. (If the ratios are equal, any patent can be indexed as i or $-i$). From (107), it follows that

$$\lambda_i = -\frac{\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} + \lambda_{-i} \frac{\partial P_{-i}}{\partial T}}{\frac{\partial P_i}{\partial T}} \tag{108}$$

Since we are looking for sufficient conditions for the optimality of $w = 1$, it should be assumed that $w > 0$. Then, it follows that

$$\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} + \lambda_1 \frac{\partial P_1}{\partial w} + \lambda_2 \frac{\partial P_2}{\partial w} - \lambda_w = 0 \tag{109}$$

$$\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} + \lambda_1 \frac{\partial P_1}{\partial w} + \lambda_2 \frac{\partial P_2}{\partial w} = \lambda_w \tag{110}$$

The next steps will show sufficient conditions for $\lambda_w > 0$, since in that situation the optimal w must be the maximum ($w = 1$). To see why, look at condition (103). If the objective function and the restrictions are such that in any feasible point the only λ_w that fulfils them is a positive value, then $(1 - w)$ in that point must be equal to zero. In other words, it must be true that $w = 1$ in a point that satisfies the optimality conditions.

Concretely, according to those conditions, it must be true that $\lambda_w > 0$ if

$$\lambda_w = \frac{\partial S_{10}}{\partial w} + \frac{\partial S01}{\partial w} + \lambda_1 \frac{\partial P_1}{\partial w} + \lambda_2 \frac{\partial P_2}{\partial w} > 0 \tag{111}$$

Combining (108) and (111), it follows that

$$\lambda_w = \frac{\partial S_{10}}{\partial w} + \frac{\partial S01}{\partial w} + \lambda_i \frac{\partial P_i}{\partial w} + \lambda_{-i} \frac{\partial P_{-i}}{\partial w} > 0 \tag{112}$$

$$\frac{\partial S_{10}}{\partial w} + \frac{\partial S01}{\partial w} - \frac{\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} + \lambda_{-i}\frac{\partial P_{-i}}{\partial T}}{\frac{\partial P_i}{\partial T}} \frac{\partial P_i}{\partial w} + \lambda_{-i} \frac{\partial P_{-i}}{\partial w} > 0 \tag{113}$$

$$\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} + \lambda_{-i}\frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} - \lambda_{-i}\frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w} > 0 \tag{114}$$

$$\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} > \lambda_{-i}\left(\frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w} - \frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T}\right) \tag{115}$$

One of the following two things must happen. First, the elasticity ratios might be equal. Secondly, they might be different. The first situation would have the following implications:

$$\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} = \frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}} \tag{116}$$

$$\varepsilon_{P_{-i}T} \cdot \varepsilon_{P_i w} = \varepsilon_{P_{-i}w} \cdot \varepsilon_{P_i T} \tag{117}$$

$$\frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w} = \frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T} \tag{118}$$

$$\frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w} - \frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T} = 0 \tag{119}$$

Combining the previous expression with (115), it follows that a sufficient condition for $\lambda_w > 0$ and, as a consequence, the (local) optimality of $w = 1$ is

$$\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} > 0 \tag{120}$$

$$\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} > \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} \tag{121}$$

$$\frac{\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}}{\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}} < \frac{\frac{\partial P_i}{\partial w}}{\frac{\partial P_i}{\partial T}} \tag{122}$$

$$\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} < \frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} \tag{123}$$

(Notice that the last substitution is a consequence of the fact that it has been assumed that $S_{10} + S_{01} = S_{11}$). This leads to the following result: *a policy with $w = 1$ is locally optimal if the following condition holds:*

$$\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} < \frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}} = \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}} \tag{124}$$

On the other hand, if the ratios are different, the analysis should be extended. First, notice that in such case

$$\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} > \frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}} \tag{125}$$

$$\varepsilon_{P_{-i} T} \cdot \varepsilon_{P_i w} > \varepsilon_{P_{-i} w} \cdot \varepsilon_{P_i T} \tag{126}$$

$$\frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w} > \frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T} \tag{127}$$

$$\frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w} - \frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T} > 0 \tag{128}$$

Combining the last expression with (115), it follows that $\lambda_w > 0$ if there exists $\lambda_{-i} \geq 0$ such that

$$\frac{\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w}}{\frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w} - \frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T}} > \lambda_{-i} \tag{129}$$

A necessary condition for the existence of such value is $\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} > 0$, which leads to (123) again. Having said that, it must be noticed that another condition might be necessary. The need to introduce it depends on $\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}}$ and $\frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}}$. To see this, observe that, in (129), turning i into $-i$ and vice-versa leads to

$$\frac{\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_{-i}}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_{-i}}{\partial w}}{\frac{\partial P_i}{\partial T}\frac{\partial P_{-i}}{\partial w} - \frac{\partial P_i}{\partial w}\frac{\partial P_{-i}}{\partial T}} < \lambda_i \tag{130}$$

(Notice that the sign has been inverted because $\frac{\partial P_i}{\partial T}\frac{\partial P_{-i}}{\partial w} - \frac{\partial P_i}{\partial w}\frac{\partial P_{-i}}{\partial T} < 0$). According to this expression, to ensure there is a local optimum in a solution where $w = 1$ there must exist $\lambda_i \geq 0$ that satisfies condition (130).

This is not a problem if $\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} \leq \frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}}$, as can be inferred from the following expressions.

$$\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} \leq \frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}} \tag{131}$$

$$\varepsilon_{S_{11} w} \cdot \varepsilon_{P_{-i} T} \geq \varepsilon_{S_{11} T} \cdot \varepsilon_{P_{-i} w} \tag{132}$$

$$\varepsilon_{S_{11} w} \cdot \varepsilon_{P_{-i} T} - \varepsilon_{S_{11} T} \cdot \varepsilon_{P_{-i} w} \geq 0 \tag{133}$$

$$\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_{-i}}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_{-i}}{\partial w} \geq 0 \tag{134}$$

$$\frac{\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_{-i}}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_{-i}}{\partial w}}{\frac{\partial P_i}{\partial T}\frac{\partial P_{-i}}{\partial w} - \frac{\partial P_i}{\partial w}\frac{\partial P_{-i}}{\partial T}} \leq 0 \tag{135}$$

The lower bound is negative or equal to zero, which does not interfere with the condition that $\lambda_i \geq 0$. In conclusion, if $\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} < \frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}}$ and $\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} \leq \frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}}$, then $\lambda_w > 0$. This leads to the following result: *a policy with* $w = 1$ *is optimal if the following condition holds:*

$$\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} \leq \min\left\{\frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}}, \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}}\right\} \tag{136}$$

Notice that condition (136) is a generalisation of (124). For this reason, the former constitutes the first part of Proposition 3, whereas the latter appears only implicitly.

Now consider what happens when $\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} > \frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}}$, since this assumption has several implications which should be taken into account.

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} > \frac{\varepsilon_{P_{-i}w}}{\varepsilon_{P_{-i}T}} \tag{137}$$

$$\varepsilon_{S_{11}w} \cdot \varepsilon_{P_{-i}T} < \varepsilon_{S_{11}T} \cdot \varepsilon_{P_{-i}w} \tag{138}$$

$$\varepsilon_{S_{11}w} \cdot \varepsilon_{P_{-i}T} - \varepsilon_{S_{11}T} \cdot \varepsilon_{P_{-i}w} < 0 \tag{139}$$

$$\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_{-i}}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_{-i}}{\partial w} < 0 \tag{140}$$

$$\frac{\left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_{-i}}{\partial T} - \left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_{-i}}{\partial w}}{\frac{\partial P_i}{\partial T}\frac{\partial P_{-i}}{\partial w} - \frac{\partial P_i}{\partial w}\frac{\partial P_{-i}}{\partial T}} > 0 \tag{141}$$

According to the last expression, the lower bound for λ_i must be greater than zero. This means that $\lambda_i > 0$. By conditions (99) and (101), this would only be possible if there is a point $(T_i, 1)$ such that $R_i = P_i(T_i, 1)$. That point will be a feasible one if $R_{-i} \leq P_{-i}(T_i, 1)$. Assuming that such point exists, if $\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} < \frac{\varepsilon_{P_iw}}{\varepsilon_{P_iT}}$, and $\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} > \frac{\varepsilon_{P_{-i}w}}{\varepsilon_{P_{-i}T}}$, then $\lambda_w > 0$ in the optimal solution to the problem. This leads to the last part of Proposition 3: *there is a locally optimal solution with $w = 1$ if*

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} < \max\left\{\frac{\varepsilon_{P_1w}}{\varepsilon_{P_1T}}, \frac{\varepsilon_{P_2w}}{\varepsilon_{P_2T}}\right\} \tag{142}$$

and there exists a point $(T_i, 1)$ such that $R_i = P_i(T_i, 1)$ and $R_{-i} \leq P_{-i}(T_i, 1)$. This completes the proof.

D Proof of Proposition 4

Consider a restricted version of the problem in (21):

$$\begin{aligned}
\underset{T,w}{\text{maximise}} \quad & S_{10}(T,w) + S_{01}(T,w) - \sum_{i\in\{1,2\}} R_i/2 \\
\text{subject to} \quad & R_i \leq P_i(T,w), \qquad i \in 1,2 \\
& T \leq \overline{T}, \\
& w \leq 1, \\
& 0 \leq T, \\
& 0 \leq w
\end{aligned} \tag{143}$$

The Lagrangian for that problem is represented by

$$L = S_{10} + S_{01} - \frac{R_1}{2} - \frac{R_2}{2} + \lambda_1(P_1 - R_1) + \lambda_2(P_2 - R_2) + \lambda_T(\overline{T} - T) + \lambda_w(1 - w) \tag{144}$$

Expressions (96) to (106), which can be found in Appendix C, and (145) to (149) provide the Karush-Kuhn-Tucker (first-order) conditions for a locally optimal solution to such problem.

$$L_T = \frac{\partial S_{10}}{\partial T} + \frac{\partial S01}{\partial T} + \lambda_1 \frac{\partial P_1}{\partial T} + \lambda_2 \frac{\partial P_2}{\partial T} - \lambda_T \leq 0 \tag{145}$$

$$T \cdot L_T = T \cdot \left[\frac{\partial S_{10}}{\partial T} + \frac{\partial S01}{\partial T} + \lambda_1 \frac{\partial P_1}{\partial T} + \lambda_2 \frac{\partial P_2}{\partial T} - \lambda_T \right] = 0 \tag{146}$$

$$L_{\lambda_T} = \overline{T} - T \geq 0 \tag{147}$$

$$\lambda_T \cdot L_{\lambda_T} = \lambda_T \cdot (\overline{T} - T) = 0 \tag{148}$$

$$\lambda_T \geq 0 \tag{149}$$

Since it was assumed that $P_1(T,0) < R_1$, $P_2(T,0) < R_2$, and $\{(T,w) \in [0,\infty) \times [0,1] : R_1 \leq P_1(T,w) \wedge R_2 \leq P_2(T,w)\} \neq \emptyset$, $w > 0$ in the optimum. As a consequence, expressions (109) and (110), which can be found in Appendix C, still hold. Combining them with (106) and letting λ_i be the multiplier of the patent with the highest elasticity ratio, as in Appendix C, lead to

$$\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} + \lambda_1 \frac{\partial P_1}{\partial w} + \lambda_2 \frac{\partial P_2}{\partial w} \geq 0 \tag{150}$$

$$\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} + \lambda_{-i} \frac{\partial P_{-i}}{\partial w} + \lambda_i \frac{\partial P_i}{\partial w} \geq 0 \tag{151}$$

$$\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} + \lambda_{-i} \frac{\partial P_{-i}}{\partial w} \geq -\lambda_i \frac{\partial P_i}{\partial w} \tag{152}$$

$$-\frac{\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} + \lambda_{-i} \frac{\partial P_{-i}}{\partial w}}{\frac{\partial P_i}{\partial w}} \leq \lambda_i \tag{153}$$

Furthermore, since we are looking for sufficient conditions for the desirability of $T \to \infty$, it should be assumed that $T > 0$. Then, it follows that

$$\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} + \lambda_1 \frac{\partial P_1}{\partial T} + \lambda_2 \frac{\partial P_2}{\partial T} - \lambda_T = 0 \tag{154}$$

$$\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} + \lambda_1 \frac{\partial P_1}{\partial T} + \lambda_2 \frac{\partial P_2}{\partial T} = \lambda_T \tag{155}$$

The next steps will show sufficient conditions for $\lambda_T > 0$, because in that situation the optimal T must be $\overline{T}$. To see why, look at condition (148). If the objective function and the restrictions are such that in any feasible point the only λ_T that fulfils them is a positive value, then $(\overline{T} - T)$ in that point must be equal to zero. In other words, it must be true that $T = \overline{T}$ in a point that satisfies the optimality conditions. Then, as $\overline{T} \to \infty$, problem (143) approaches problem (21), and the optimal patent length approaches infinity.

Concretely, an analysis of (153) reveals a sufficient condition for $\lambda_T > 0$.

$$\lambda_T = \frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} + \lambda_1 \frac{\partial P_1}{\partial T} + \lambda_2 \frac{\partial P_2}{\partial T} > 0 \tag{156}$$

$$\lambda_T = \frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} + \lambda_{-i} \frac{\partial P_{-i}}{\partial T} + \lambda_i \frac{\partial P_i}{\partial T} > 0 \tag{157}$$

After combining (153) and (155), a more restrictive, albeit workable, condition is obtained.

$$\lambda_T \geq \frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} + \lambda_{-i} \frac{\partial P_{-i}}{\partial T} - \frac{\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} + \lambda_{-i} \frac{\partial P_{-i}}{\partial w}}{\frac{\partial P_i}{\partial w}} \frac{\partial P_i}{\partial T} > 0 \tag{158}$$

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} \right) \frac{\partial P_i}{\partial w} - \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} \right) \frac{\partial P_i}{\partial T} > \lambda_{-i} \left(\frac{\partial P_{-i}}{\partial w} \frac{\partial P_i}{\partial T} - \frac{\partial P_{-i}}{\partial T} \frac{\partial P_i}{\partial w} \right) \tag{159}$$

If elasticity ratios $\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}}$ and $\frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}}$ are equal, the following expressions must be true:

$$\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} = \frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}} \tag{160}$$

$$\varepsilon_{P_{-i} T} \cdot \varepsilon_{P_i w} = \varepsilon_{P_{-i} w} \cdot \varepsilon_{P_i T} \tag{161}$$

$$\frac{\partial P_{-i}}{\partial T} \frac{\partial P_i}{\partial w} = \frac{\partial P_{-i}}{\partial w} \frac{\partial P_i}{\partial T} \tag{162}$$

$$0 = \frac{\partial P_{-i}}{\partial w} \frac{\partial P_i}{\partial T} - \frac{\partial P_{-i}}{\partial T} \frac{\partial P_i}{\partial w} \tag{163}$$

Combining the previous expression with (159), it follows that a sufficient condition for $\lambda_T > 0$ and, as a consequence, the desirability of $T \to \infty$ is

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} \right) \frac{\partial P_i}{\partial w} - \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} \right) \frac{\partial P_i}{\partial T} > 0 \tag{164}$$

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T} \right) \frac{\partial P_i}{\partial w} > \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w} \right) \frac{\partial P_i}{\partial T} \tag{165}$$

$$\frac{\frac{\partial P_i}{\partial w}}{\frac{\partial P_i}{\partial T}} < \frac{\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}}{\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}} \tag{166}$$

$$\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} < \frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} \tag{167}$$

(The last substitution is a consequence of assuming $S_{10} + S_{01} = S_{11}$). Then, it can be said that extending patent life as much as possible ($T = \overline{T}$) would be the best policy if $\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} > \frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}} = \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}}$. Furthermore, analysing what happens as $\overline{T} \to \infty$ makes it possible to say that *patents should last forever ($T = \infty$) if that condition holds:*

$$\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} > \frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}} = \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}} \tag{168}$$

On the other hand, if ratios $\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}}$ and $\frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}}$ are different,

$$\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} > \frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}} \tag{169}$$

$$\varepsilon_{P_{-i} T} \cdot \varepsilon_{P_i w} > \varepsilon_{P_{-i} w} \cdot \varepsilon_{P_i T} \tag{170}$$

$$\frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w} > \frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T} \tag{171}$$

$$0 > \frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T} - \frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w} \tag{172}$$

Combining the last expression with (159), it follows that $\lambda_T > 0$ if there exists $\lambda_{-i} \geq 0$ such that

$$\frac{\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} - \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T}}{\left(\frac{\partial P_{-i}}{\partial w}\frac{\partial P_i}{\partial T} - \frac{\partial P_{-i}}{\partial T}\frac{\partial P_i}{\partial w}\right)} < \lambda_{-i} \tag{173}$$

If the left-hand side of that expression is less than zero, then such value for λ_{-i} exists. Provided that the denominator of (173) is less than zero, the lower bound for λ_{-i} will be negative if the following inequality holds.

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} - \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} > 0 \tag{174}$$

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} > \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} \tag{175}$$

$$\frac{\frac{\partial P_i}{\partial w}}{\frac{\partial P_i}{\partial T}} < \frac{\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}}{\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}} \tag{176}$$

$$\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} < \frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} \tag{177}$$

This means that $\lambda_T > 0$ if $\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} < \frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}}$, where $\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} > \frac{\varepsilon_{P_{-i} w}}{\varepsilon_{P_{-i} T}}$. In other words, patents should last as much as possible ($T = \overline{T}$) if $\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} > \max\left\{\frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}}, \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}}\right\}$. The previous finding implies, as $\overline{T} \to \infty$, that *patents should last forever ($T \to \infty$) if the following condition holds:*

$$\frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} > \max\left\{\frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}}, \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}}\right\} \tag{178}$$

Notice that condition (178) is a generalisation of (168). For this reason, the former constitutes the first part of Proposition 3, whereas the latter appears only implicitly.

Finally, if the left-hand side of (173) is greater than zero, an additional condition is required to ensure that $\lambda_T > 0$ in the optimal solution. First, see that the denominator of (173) is less than zero, so λ_{-i} must be positive if

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} - \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} \leq 0 \tag{179}$$

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_i}{\partial w} \leq \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_i}{\partial T} \tag{180}$$

$$\frac{\frac{\partial P_i}{\partial w}}{\frac{\partial P_i}{\partial T}} \geq \frac{\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}}{\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}} \tag{181}$$

$$\frac{\varepsilon_{P_i w}}{\varepsilon_{P_i T}} \geq \frac{\varepsilon_{S_{11} w}}{\varepsilon_{S_{11} T}} \tag{182}$$

Assuming that such condition holds, by conditions (99) and (101) no solution with $\lambda_T > 0$ would fulfil all the Karush-Kuhn-Tucker conditions unless there is a feasible point $(\overline{T}, w_{-i})$ such that $R_{-i} = P_{-i}(\overline{T}, w_{-i})$. But that point will be a feasible one only if $R_i \leq P_i(\overline{T}, w_{-i})$.

What happens with the relation between $\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}}$ and $\frac{\varepsilon_{P_{-i}w}}{\varepsilon_{P_{-i}T}}$? First remember that, as a consequence of (172), $\frac{\partial P_i}{\partial w}\frac{\partial P_{-i}}{\partial T} - \frac{\partial P_i}{\partial T}\frac{\partial P_{-i}}{\partial w} > 0$. Then, (151) and (157) imply

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_{-i}}{\partial w} - \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_{-i}}{\partial T} > \lambda_i \left(\frac{\partial P_i}{\partial w}\frac{\partial P_{-i}}{\partial T} - \frac{\partial P_i}{\partial T}\frac{\partial P_{-i}}{\partial w}\right) \tag{183}$$

$$\frac{\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_{-i}}{\partial w} - \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_{-i}}{\partial T}}{\frac{\partial P_i}{\partial w}\frac{\partial P_{-i}}{\partial T} - \frac{\partial P_i}{\partial T}\frac{\partial P_{-i}}{\partial w}} > \lambda_i \tag{184}$$

Since $\lambda_i \geq 0$, the left hand-side of (184) must be positive. Provided that $\frac{\partial P_i}{\partial w}\frac{\partial P_{-i}}{\partial T} - \frac{\partial P_i}{\partial T}\frac{\partial P_{-i}}{\partial w} > 0$, the left-hand side of (184) will be positive if

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_{-i}}{\partial w} - \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_{-i}}{\partial T} > 0 \tag{185}$$

$$\left(\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}\right)\frac{\partial P_{-i}}{\partial w} > \left(\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}\right)\frac{\partial P_{-i}}{\partial T} \tag{186}$$

$$\frac{\frac{\partial P_{-i}}{\partial w}}{\frac{\partial P_{-i}}{\partial T}} < \frac{\frac{\partial S_{10}}{\partial w} + \frac{\partial S_{01}}{\partial w}}{\frac{\partial S_{10}}{\partial T} + \frac{\partial S_{01}}{\partial T}} \tag{187}$$

$$\frac{\varepsilon_{P_{-i}w}}{\varepsilon_{P_{-i}T}} < \frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} \tag{188}$$

This means that patents should last as much as possible ($T = \overline{T}$) if $\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} > \min\left\{\frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}}, \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}}\right\}$ and there exists a point $(\overline{T}, w_{-i})$ such that $R_{-i} = P_{-i}(\overline{T}, w_{-i})$ and $R_i \leq P_i(\overline{T}, w_{-i})$. Analysing what happens as $\overline{T} \to \infty$ leads to the last part of Proposition 3: *patents should last forever ($T \to \infty$) if*

$$\frac{\varepsilon_{S_{11}w}}{\varepsilon_{S_{11}T}} > \min\left\{\frac{\varepsilon_{P_1 w}}{\varepsilon_{P_1 T}}, \frac{\varepsilon_{P_2 w}}{\varepsilon_{P_2 T}}\right\} \tag{189}$$

and there exists $w_{-i} \in [0,1]$ such that $R_{-i} = \lim_{T\to\infty} P_{-i}(T, w_{-i})$ and $R_i \leq \lim_{T\to\infty} P_i(T, w_{-i})$. This completes the proof.