How Are Patent Decisions Affected by Environmental Regulation?

Pak-Sing Choi, Ana Espínola-Arredondo, Félix Muñoz-García, and Eugenio Diaz-Farina

March 20, 2024

Abstract

This paper examines the setting of patent lengths in polluting industries. In the absence of environmental regulation, the patent office faces a well-known tradeoff: a longer patent yields a welfare benefit from inducing more R&D investment, but generates a welfare loss from allowing a longer monopoly during the patent period. When environmental policy is present, we show that the welfare loss is emphasized (ameliorated) when the environmental agency is less (more) flexible than the patent office, thus inducing shorter (longer) patents. We also consider green innovations, showing that environmental regulation becomes less stringent, and patent decisions approach those in the absence of regulation.

Keywords: Patent length, emission fees, polluting industries, environmental regulation, environmental damages, green innovation

JEL classification: H23, O32, Q53, Q55
1 Introduction

The optimal setting of patents has been extensively examined since Nordhaus (1969) and Takalo (2001). These studies consider the tradeoff between the social loss from market monopolization during the patent period and the social benefit of inducing more investment in R&D, lowering production costs for the innovator during the patent and for all firms in the industry after the patent expires. More recently, a growing number of scholars have proposed the shortening of patent lengths, or their complete elimination for certain goods; see Boldrin and Levine (2013). These studies, however, consider that firms do not generate environmental damages and, as a consequence, assume that firms do not face environmental regulation. Many patentees compete in polluting industries, such as the Saudi Arabian Oil Company (holder of over 1,030 patents in 2023), Basf (333 patents), Walmart (266 patents), the Haier Group (253 patents), Exxon Mobil (238 patents), or Nippon Steel Corporation (211 patents); among others. In this paper, we show that patent lengths are shortened by the presence of environmental policy, thus providing a new argument to reconsider patent decisions. This result is robust to changes in the Environmental Protection Agency’s (EPA) ability to revise emission fees in each period, or the green features of the innovation.

For comparison purposes, we consider a setting similar to Takalo’s (2001), where in the first stage, the patent office (PO) sets a patent length; in the second stage, the innovator responds investing in R&D to reduce its marginal production cost; and, in the third stage, firms compete. During the patent period, the innovator uses its technology to operate as a monopolist, and once the patent expires, the technology is publicly available.

As a benchmark, we first examine an industry where the EPA is absent, where our results confirm Nordhaus’ tradeoff. We then study how environmental policy affects this tradeoff. For completeness, we allow for this policy to occur in three different stages: in the first, second, or third stage.

When emission fees are set in the first stage, before patent and investment decisions are made, our model represents regulatory settings where environmental agencies cannot easily revise future policies. While many countries adjust emission fees according to inflation or pollutants intensity, few update their policies according to changes in the market structure (e.g., whether the industry becomes more competitive, if large R&D investments are made, or if cost-reducing innovations occur), thus closely fitting our model. Examples include the EPA permit annual fees in the US (Title V Operating Permits), initially set at $32/ton in 1996, and only adjusted for inflation every year in September. Similarly, the sanctions for firms exceeding CO2 emissions in the EU-ETS, originally set at 100 Euro/ton in 2003, have remained unchanged, only being adjusted for inflation since 2013; and a similar argument applies to the criminal provisions of the US Clean Air Act (42

---

1 In many countries, a utility patent is granted to a product, process, machine, manufacture, chemical compound, or asexually reproducible plant for 20 years, while a design patent varies from 15 years in the US (The United States Patent and Trademark Office, 2017) and China (Zhou, 2021) or 20 years in Japan (Japan Patent Office, n.d.) to a maximum of 25 years in Europe (European Union Intellectual Property Office, n.d.).
U.S.C. 7413) which have not been revised since its enactment in 1990. Likewise, Finland carbon
tax law was not amended for 14 years, between 1997 and 2011, although it has been more recently
adjusted in 2018.\footnote{Globally, carbon tax revenues have been relatively constant from 2006 to 2016, as reported by World Bank (2023).}

In this context, we show that the EPA’s presence induces higher costs (net of taxes), less R&D
investments, and a lower social return of the innovation, which ultimately induces the PO to respond
with shorter patents than when the EPA is absent. This “patent differential” grows when pollution
is more severe and the innovation does not bring large cost-reduction effects, indicating that it is
in these contexts when the PO should be more vigilant about environmental policy being active.

If, instead, the EPA acts in the second stage, after the PO has set patent lengths, our model
still captures settings where the EPA cannot easily revise future environmental policies. Yet, the
EPA is, comparatively, more flexible than the PO in this setting. We adopt a standard approach
in the industrial organization literature, assuming less flexibility to strategies chosen close to the
first stage of the game. In this context, we demonstrate that the emission fee imposes the same
cost increase in all scenarios: before the patent expires, after it expires, and when no innovation
occurs. Therefore, the welfare loss that the PO considers when deciding to prolong the patentee’s
legal monopoly for one more period is unaffected by environmental regulation. Welfare gains from
a longer patent are, however, larger, since firms invest less with than without regulation, yielding
a larger marginal effect from these lower investments. As a result, the PO has incentives to set
longer patents than in the absence of environmental policy.

A similar argument applies when the EPA becomes more flexible, acting in the third stage after
patents and investments are made.\footnote{The sanctions in the US Clean Water Act (33 U.S.C. 1319), for instance, have been frequently revised, with their last modifications happening in 2018 and 2019.} In this case, the environmental regulator observes whether the
patent is still in force and whether the innovation was successful, thus being able to set emission
fees that induce the socially optimal output in every scenario. Thus, aggregate output coincides
before and after the patent expires, eliminating the welfare loss arising from allowing for a longer
monopoly (i.e., “breaking” the Nordhaus’ tradeoff), and inducing the PO to set longer patents than
under no regulation.

Overall, our results suggest that, as the EPA’s ability to revise emission fees increases, this
regulatory agency induces aggregate output that is closer to the first best in each industry setting
(i.e., before the patent expires, after it expires, and if there is no innovation). This ameliorates
the welfare loss from extending the patent for more periods (that is, allowing a legal monopoly),
ultimately reducing the Nordhaus’ tradeoff. As a consequence, the PO has incentives to set longer
patents than when the EPA is absent. In contrast, when this agency cannot easily revise environ-
mental policy, exhibiting more rigidity than the PO, our results indicate that this policy emphasizes
the welfare loss from allowing for longer patents, strengthening the Nordhaus’ tradeoff, ultimately
inducing the PO to set shorter patents than in the absence of regulation.

Finally, we test how our findings are affected when innovations not only reduce firms’ production
costs but also lower pollution intensity. Because the EPA anticipates less pollution and sets less stringent emission fees, firms’ investment decisions and the PO’s patent policy become more similar to the setting without regulation, thus generating smaller changes in their behavior. In summary, our results indicate that POs can essentially ignore the presence of environmental regulation when innovations are extremely green, but must consider this regulation when setting patent lengths otherwise.

**Related Literature.** Our model follows the seminal work by Nordhaus (1969) about optimal patent lengths, which identified the tradeoff between the marginal static loss of allowing a monopoly market that induces the innovator to recover its R&D costs, and the marginal dynamic gain that the innovation produces on society once the patent expires. More specifically, we use Takalo’s (2001) game-theoretic analysis of these two effects. This literature extended along different dimensions, such as: (1) the optimal scope (or breadth) of patents, as in Merges and Nelson (1990), Gilbert and Shapiro (1990), Klemperer (1990), and Gallini (1992); (2) licensing of the patent to other firms, as in Denicolò (1996) and Gallini and Scotchmer (2002); (3) the protection that the patent gives to the innovator against other products that could be infringing the patent (often referred to as the “height” of the novelty requirement), as in van Dijk (1996), La Manna (1992), and O’Donoghue et al. (1998); or (4) the development of complementary innovations, as in Heller and Eisenberg (1998) and Shapiro (2001). For a survey of the literature, see Eckert and Langinier (2014) to show the effect of agency interactions.

Our paper examines how optimal patents are affected when the PO sets patents in settings where the EPA is absent (no environmental regulation) or present (emission fees seeking to curb pollution). One of the few articles exploring the connection between the EPA and the PO’s decisions is Gerlagh et al. (2014), which also considers patents in clean energy R&D, but assumes that firms invest in abatement R&D to reduce their pollution intensity. In contrast, we study an innovation that reduces the firms’ production costs, thus being more similar to the standard patent literature described above, helping us identify how the EPA’s presence affects patent policy.

Langinier and Chaudhuri (2020) analyze the effects of knowledge appropriability on patentability requirements and emission fees in the presence of environmentally-conscious consumers. The authors show that green R&D is maximized when (exogenous) emission fees are neither too high to make the investment unprofitable nor too low to yield insufficient abatement. However, they consider fixed patent lengths, which we endogenize.

The remainder of this article is organized as follows. Section 2 describes the model. As a benchmark, section 3 identifies equilibrium behavior when the EPA is absent and section 4 examines how our results are affected when the EPA is present. Section 5 discusses our main results and their policy implications.

---

4 According to Globaldata.com, in 2022, for instance, there were over 43,800 new green patents in the automotive sector, more than 5,800 in the chemical industry, over 6,800 in the industrial goods and machinery industry, and 2,900 in agriculture and forestry.

5 For an introduction, see the presentation in Belleflamme and Peitz (2015, p.541-542).
2 Model

Consider an innovator with R&D cost function \( C(x) = \frac{1}{2}\gamma x^2 \), where \( \gamma \) denotes its R&D efficiency, that is, a lower \( \gamma \) represents a greater efficiency. As in Takalo (2001) and Belleflamme and Peitz (2015), we assume that \( \gamma \) is sufficiently large to yield \( x \in [0, 1] \) in equilibrium, thus allowing for \( x \) to be interpreted as the innovator’s probability of success. The inverse demand function is \( p(Q) = 1 - Q \), where \( Q \geq 0 \) denotes aggregate output.

Without the innovation, every firm faces marginal cost of production \( c \), where \( 0 \leq c \leq 1 \), and we assume that they interact in a perfectly competitive market, yielding zero profit, \( \pi^{NI} = 0 \), where superscript \( NI \) denotes “no innovation.” When the innovator is successful, its marginal cost decreases to \( c - \alpha \), where \( 0 \leq \alpha \leq c \) denotes the cost-reduction effect of the innovation. When \( \alpha = 0 \), the innovation is inconsequential but when \( \alpha = c \), the firm’s marginal cost is reduced to zero.

The innovator receives a patent during \( T \geq 0 \) periods, which lets the firm use the technology that decreases its marginal cost of production to \( c - \alpha \), earning monopoly profit \( \pi^P \), where superscript \( P \) represents a “patent” period. Once the patent expires, the innovation becomes public, allowing other firms to enjoy this technology as well, and every firm earns competitive profits \( \pi^C = 0 \), where superscript \( C \) denotes “competition.”

In the absence of patents, aggregate output occurs at the point where the demand \( p(Q) \) and the marginal cost \( c \) intersect with one another, that is, \( Q_0 \) solves \( p(Q_0) = c \), or \( Q_0 = 1 - c > 0 \).

We consider the following time structure of the game:

1. In the first stage, the EPA sets an emission fee, \( \tau \).
2. In the second stage, the patent office (PO) observes \( \tau \), and responds choosing a patent length \( T \).
3. In the third stage, the innovator observes the fee \( \tau \) and patent length \( T \), and responds with its R&D investment, \( x(T) \).
4. In the fourth stage, firms compete in output.

This time structure indicates that emission fees are administratively difficult to revise after the PO sets a patent length or after firms invest in R&D. Hence, it characterizes the policy constraints that several environmental regulatory agencies face, where fees have been rarely revised in the last decades, as reported in the above examples. (We separately examine these regulatory regimes in sections 4.2-4.3.)

The PO’s welfare function is \( W = CS + PS \), which includes consumer and producer surplus, since POs in most countries ignore pollution when setting patent lengths. In contrast, the EPA’s welfare function is

\[
W_{EPA} = CS + PS + Tax - Env, \tag{1}
\]
where $\text{Tax} = \tau Q$ denotes total tax collection, which guarantees that the emission fees are revenue neutral; and $\text{Env} = dQ^2$ denotes environmental damages, which are increasing and convex in aggregate output, and parameter $d$ satisfies $d > 1/2$ to guarantee positive emission fees in all settings.

We first analyze equilibrium behavior in a setting where pollution is not addressed by the EPA, becoming a three-stage game that we solve by backward induction.

3 Equilibrium Analysis when the EPA is absent

3.1 Third stage, output decisions

After the patent expires (in periods $t \geq T$), every firm $i$ enjoys production cost $c - \alpha$. In this context, firms interact in a perfectly competitive market, with aggregate output $Q^C = 1 - (c - \alpha)$, which solves $p(Q^C) = c - \alpha$, yielding zero profits for every firm, that is, $\pi^C = 0$.

Before the patent expires ($t < T$), the innovator enjoys production cost $c - \alpha$, choosing a monopoly output $q^m = \frac{1-(c-\alpha)}{2}$ that yields monopoly price $p^m = \frac{1+(c-\alpha)}{2}$. We define the cost-reduction effect of the innovation, as captured by $\alpha$, to be “non-radical” as in Belleflamme and Peitz (2015), if the monopoly price of the innovator during the patent exceeds its rivals’ common marginal cost, $p^m = \frac{1+(c-\alpha)}{2} > c$, or $\alpha < 1 - c$. For the innovator to be the only seller, it sets a price marginally below its rivals’ common marginal cost, $c$, that is, $p^P = c - \varepsilon$, where $\varepsilon \to 0$. As a result, patent output $q^P$ solves $1 - q^P = c$, or $q^P = Q_0 = 1 - c$.

In this context, the innovator earns monopoly profits in every patent period, as follows,

$$\pi^P = [p^P - (c - \alpha)] q^P = \alpha (1 - c)$$

which are positive during the patent but zero afterwards, that is, $\pi^P > \pi^C = 0$. Patent profits increase with the cost-reduction effect of innovation $\alpha$ but decrease in the initial cost $c$.

3.2 Second stage, R&D investment

Anticipating output decisions in the last stage, the innovator chooses its R&D investment, $x$, to solve

$$\max_{x \geq 0} x \Pi_0 (T) + (1 - x) \int_0^{+\infty} e^{-rt} \pi^N I dt - \frac{1}{2} \gamma x^2$$

where the first two terms represent the innovator’s expected return from his investment, which is successful with probability $x$ earning $\Pi_0 (T)$, but unsuccessful with probability $(1 - x)$ earning $\int_0^{+\infty} e^{-rt} \pi^N I dt - \frac{1}{2} \gamma x^2$. If, instead, the innovation was radical, meaning that $p^m < c$, the innovator would be able to set the monopoly price $p^m$, not needing the legal protection of the patent to become a monopolist. In that setting, the PO’s role would become irrelevant and patent lengths would be inconsequential.
profits $\pi^{NJ}$ in every period, and

$$\Pi_0(T) = \int_0^T e^{-rt}\pi_P dt + \int_T^{+\infty} e^{-rt}\pi_C dt = \frac{\alpha(1-c)(1-e^{-rT})}{r}$$

where $r$ denotes the discount factor satisfying $0 \leq r \leq 1$, and the second equality considers $\pi_C = 0$ and $\pi_P = \alpha(1-c)$. Note that $\Pi_0(T)$ is increasing in the patent length $T$, that is, $\Pi'_0(T) = \pi_P e^{-rT} \geq 0$, meaning that the innovator can earn monopoly profits $\pi_P$ during more periods.

**Lemma 1.** The innovator’s equilibrium investment in R&D is $x_0(T) = \frac{\Pi_0(T)}{\gamma} = \frac{\alpha(1-c)(1-e^{-rT})}{\gamma r}$, which is increasing and concave in $T$, increasing in $\alpha$, and decreasing in $c$ and $\gamma$.

As expected, the innovator has stronger incentives to invest in R&D with longer patents, thus providing him with a longer monopoly, although these incentives increase at a decreasing rate. In contrast, investment decreases when net costs increase (higher $c - \alpha$) or it becomes more costly (higher $\gamma$).

### 3.3 First stage, patent length

The PO anticipates the firms’ decisions in the subsequent stages of the game, choosing the patent length, $T$, that solves the following problem,

$$\max_{T \geq 0} x_0(T) \left( \int_0^T e^{-rt}W_0 P dt + \int_T^{+\infty} e^{-rt}W_0 C dt \right) + \left[ 1 - x_0(T) \right] \int_0^{+\infty} e^{-rt}W_0 N dt - \frac{1}{2} \gamma [x_0(T)]^2$$

where $x_0(T)$ originates from Lemma 1, and the first (second) term denotes the welfare with (without) innovation. In addition, $W_0 P = \frac{(1-c)^2}{2} + \alpha (1-c)$ denotes welfare in each patent period, $W_0 C = \frac{1-(c-\alpha)^2}{2}$ measures welfare after the patent expires, and $W_0 N = \frac{(1-c)^2}{2}$ represents welfare without the innovation. Simplifying, the above problem becomes

$$\max_{T \geq 0} x_0(T) S_0(T) + \int_0^{+\infty} e^{-rt}W_0 N dt - \frac{1}{2} \gamma [x_0(T)]^2$$

(3)

where $S_0(T)$ represents the social return of the innovation, relative to that without the innovation, defined as

$$S_0(T) = \left( \int_0^T e^{-rt}W_0 P dt + \int_T^{+\infty} e^{-rt}W_0 C dt \right) - \int_0^{+\infty} e^{-rt}W_0 N dt.$$

While profits satisfy $\pi_P > \pi_C = \pi_N = 0$, the welfare ranking is $W_0 N < W_0 P < W_0 C$ for all parameter values, indicating that firms prefer longer patents than the PO. Next, differentiating the
PO’s problem in expression (3) with respect to \( T \), yields
\[
\frac{\partial x_0(T)}{\partial T} S_0(T) = x_0(T) \left( \frac{\partial x_0(T)}{\partial T} - \frac{\partial S_0(T)}{\partial T} \right). \tag{4}
\]

The left-hand side of expression (4) measures the marginal dynamic gain, \( MDG_0(T) \), from extending the patent for one more period. Intuitively, the firm increases its R&D intensity \( x_0(T) \), which increases the expected social welfare. The right-hand side, in contrast, captures the marginal static loss, \( MSL_0(T) \), from a longer patent, in the form of a larger R&D cost (first term), and a lower consumer surplus due to extending the monopoly for more periods (second term). The next lemma examines how these effects are separately affected by a longer patent.

**Lemma 2.** \( MSL_0(T) \) increases in \( T \) at a decreasing rate when \( T < \frac{\log 2}{r} \), decreases in \( T \) at an increasing rate when \( \frac{\log 2}{r} \leq T < \frac{\log 4}{r} \), and decreases in \( T \) at a decreasing rate otherwise. In contrast, \( MDG_0(T) \) decreases in \( T \) at a decreasing rate, and more significantly than \( MSL_0(T) \) does if and only if \( \frac{\log 2}{r} < T < T_0 \), where cutoff \( T_0 = \frac{1}{r} \log \left[ \frac{4(1-(c/\alpha))}{\alpha} \right] \).

Figure 1 illustrates the results in Lemma 2, showing that \( MDG_0(T) \) is unambiguously decreasing in \( T \), while \( MSL_0(T) \) originates at zero, first increases in \( T \) until reaching a maximum at \( \frac{\log 2}{r} \), and then decreases in \( T \).\(^7\) Furthermore, \( MDG_0(T) \) decreases faster than \( MSL_0(T) \) for all \( T < T_0 \), but slower afterwards. The optimal patent length, \( T_0 \), then occurs at the point where \( MDG_0(T) = MSL_0(T) \), which we identify in Proposition 1.

![Figure 1. Patent length when the EPA is absent.](image-url)

\(^7\)For illustration purposes, the figure considers \( \alpha = 1/4, c = 2/3, \) and \( r = 1/10 \). Other parameter values yield similar results and can be provided by the authors upon request.
Lemma 3. \( MSL_0(T) \) and \( MDG_0(T) \) exhibit the following comparative statics:

1. Cost-reduction effect, \( \alpha \): Both \( MDG_0(T) \) and \( MSL_0(T) \) increase in \( \alpha \), but \( MSL_0(T) \) increases more substantially than \( MDG_0(T) \) does if and only if \( T > T_0 \).

2. Initial cost, \( c \): Both \( MDG_0(T) \) and \( MSL_0(T) \) decrease in \( c \), but \( MDG_0(T) \) decreases more significantly than \( MSL_0(T) \) does if and only if \( T < b_T \),

\[
\begin{align*}
T_0 &\equiv \frac{1}{r} \log \left[ \frac{2[3\alpha+2(1-c)]}{3\alpha} \right], \quad \hat{T}_0 \equiv \frac{1}{r} \log \left[ \frac{2[\alpha+2(1-c)]}{\alpha} \right], \quad \text{and cutoffs satisfy } \frac{\log 2}{r} < T_0 < \hat{T}_0 < T_0.
\end{align*}
\]

Thus, a larger cost-reducing effect (higher \( \alpha \)) produces a more significant pivot in \( MSL_0(T) \) than in \( MDG_0(T) \) if the patent is long enough (\( T > T_0 \), which is confirmed by Proposition 1 below), implying that the optimal patent becomes shorter. The opposite argument applies when initial cost, \( c \), increases.

**Proposition 1.** When the EPA is absent, the optimal patent length is \( T_0 = \frac{1}{r} \log \left[ \frac{2[1-(c-\alpha)]}{\alpha} \right] \), where \( T_0 \in (T_0, \hat{T}_0) \). In addition, \( T_0 \) decreases in both \( \alpha \) and \( c \). Finally, the innovator invests \( x_0 = \frac{\alpha(1-c)[\alpha+2(1-c)]}{2r[1-(c-\alpha)]} \), which increases in the cost-reduction effect \( \alpha \) but decreases in cost \( c \).

Therefore, when firms benefit from larger cost-reduction effects (higher \( \alpha \)), they experience stronger incentives to invest in R&D. Anticipating these incentives, the PO can set shorter patents. The above proposition also shows that, when firms’ initial cost is low or the cost-reduction effect of R&D is high, the net cost, \( c - \alpha \), decreases, providing firms with more incentives to invest.

### 4 Equilibrium Analysis when the EPA is present

In this section, we examine how our above equilibrium results are affected when firms face environmental regulation. We first study the case in which the EPA sets emission fees in the first stage (subsection 4.1), when the EPA chooses fees in the second stage (subsection 4.2), and finally when it acts in the third stage (subsection 4.3).

**4.1 The EPA acts in the first stage**

In this context, equilibrium behavior in stages 2-4 is analogous to that in stages 1-3 in the model without environmental regulation, but increasing firms’ costs in all scenarios: from \( c \) to \( c + \tau \) in the absence of innovation; and from \( c - \alpha \) to \( c - \alpha + \tau \) when innovation takes place.

In the third stage, every firm chooses its R&D investment, \( x \), to solve a problem analogous to (2), yielding an equilibrium investment of \( x(T, \tau) = \frac{\alpha(1-c)(1-e^{-\tau T})}{\gamma} \), which is decreasing in the emission fee \( \tau \). In addition, this investment satisfies \( 0 < x(T, \tau) < x_0(T) \) and \( 0 < x'(T, \tau) < x'_0(T) \), implying that firms invest less with than without regulation and that, when facing a longer patent, they increase their investments less significantly when facing regulation than otherwise.
The PO’s problem is analogous to (3), yielding similar expressions of $MDG(\tau)$ and $MSL(\tau)$, which are now a function of fee $\tau$. Figure 2 depicts these curves: first, evaluated at $\tau = 0$ which coincide with $MDG_0$ and $MSL_0$ in figure 1, denoted as $MDG(0)$ and $MSL(0)$; second, evaluated at $\tau = 0.1$, where both curves shift downwards relative to $\tau = 0$; and, finally, evaluated at $\tau = 0.2$, further shifting both curves downwards. As shown above, regulation lowers investment incentives, decreasing both $MDG(\tau)$ and $MSL(\tau)$. As illustrated by the figure, the downward shift in $MDG(\tau)$ dominates that in $MSL(\tau)$, implying that the crossing point between $MDG(\tau)$ and $MSL(\tau)$ moves leftward. Therefore, the equilibrium patent length decreases in the stringency of the emission fee, $\tau$; as confirmed by the patent $T$ that solves the PO’s problem, which is

$$T_{EPA} = \frac{1}{r} \log \left[ \frac{2[1 - (c + \alpha + \tau)]}{\alpha} \right].$$

(5)

Hence, the PO sets shorter patents when firms’ initial costs are higher ($c$ or $\tau$ are higher) because the patent gives rise to a smaller cost-reducing effect (for a given value of $\alpha$), ultimately reducing the social benefit of the patent.

The EPA anticipates the welfare in each case, $W^P_{EPA}$, $W^C_{EPA}$, and $W^N_{EPA}$, as defined in equation (1), setting fee $\tau$ to solve

$$\max_{T \geq 0} x(T, \tau) \left( \int_0^T e^{-rt}W^P_{EPA}dt + \int_T^{+\infty} e^{-rt}W^C_{EPA}dt \right)$$

$$+ [1 - x(T, \tau)] \int_0^{+\infty} e^{-rt}W^N_{EPA}dt - \frac{\tau}{2} [x(T, \tau)]^2.$$

(6)

Differentiating with respect to $\tau$ in problem (6) yields intractable results, and we rely on nu-
Numerical simulations. Figures 3a-3d depict emission fee $\tau^*$ as a function of pollution severity, $d$, in the horizontal axis; and evaluated at different values of $c$ (figure 3a), $\alpha$ (figure 3b), $\gamma$ (figure 3c), and $r$ (figure 3d). As a benchmark, we consider $\gamma = 1.5$, $\alpha = 0.25$, $c = 0.67$, and $r = 0.1$. Emission fees are more stringent as pollution becomes more damaging (higher $d$).

When production is more costly (either because $c$ increases or $\alpha$ decreases), the EPA anticipates less output and pollution, setting a less stringent emission fee. A similar argument applies when the cost of investing in R&D, $\gamma$, increases, where the EPA anticipates less innovation, higher expected production costs and, as a consequence, less pollution. Finally, an increase in discounting (higher $r$) induces less stringent fees. This occurs because the EPA assigns less importance to future periods, after the patent expires and all firms benefit from lower costs, caring more about present welfare under the patent. Since the patentee generates fewer emissions, the EPA can relax emission fees as a result.

Table I reports patent lengths at similar parameter values as in figures 3a-3d. The benchmark (top row) considers the same parameter values as in figure 3, and $d = 1.5$. The table indicates that patents become shorter when the innovation is more cost-reducing (higher $\alpha$). Intuitively, this
indicates the presence of a direct effect of $\alpha$ in $T_{EPA}$ since, as shown above, $T_{EPA}$ decreases in $\alpha$; but also an indirect effect, since $\alpha$ produces an increase in emission fee $\tau^*$, which in turn decreases $T$. Both effects, then, move in the same direction, ultimately reducing the patent length; as shown in Table II.

In contrast, more costly production (higher $c$) produces a negative direct effect on $T_{EPA}$; but induces a less stringent fee $\tau^*$ and, thus, a longer patent (positive indirect effect). Our results, however, identify that the negative direct effect dominates, yielding a shorter patent. This argument also applies when future periods are more heavily discounted (higher $r$), since the patent length becomes shorter (negative direct effect) but the emission fee is less stringent thus increasing the patent length (positive indirect effect). Overall, Table I shows that the direct effect dominates, shortening the patent length.

<table>
<thead>
<tr>
<th></th>
<th>Regulation</th>
<th>No regulation</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau^*$</td>
<td>$x_{EPA}$</td>
<td>$\Delta T$</td>
</tr>
<tr>
<td>Benchmark</td>
<td>0.2555</td>
<td>9.6409</td>
<td>0.0802</td>
</tr>
<tr>
<td>Higher $d$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d = 2$</td>
<td>0.2746</td>
<td>9.0403</td>
<td>0.0582</td>
</tr>
<tr>
<td>$d = 2.5$</td>
<td>0.2869</td>
<td>8.6328</td>
<td>0.0447</td>
</tr>
<tr>
<td>$d = 3$</td>
<td>0.2956</td>
<td>8.3378</td>
<td>0.0356</td>
</tr>
<tr>
<td>Higher $\gamma$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma = 3$</td>
<td>0.2528</td>
<td>9.7225</td>
<td>0.0417</td>
</tr>
<tr>
<td>$\gamma = 5$</td>
<td>0.2517</td>
<td>8.7563</td>
<td>0.0524</td>
</tr>
<tr>
<td>$\gamma = 7$</td>
<td>0.2512</td>
<td>9.7709</td>
<td>0.0163</td>
</tr>
<tr>
<td>Higher $c$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c = 0.4$</td>
<td>0.4531</td>
<td>11.5529</td>
<td>0.1677</td>
</tr>
<tr>
<td>$c = 0.5$</td>
<td>0.3792</td>
<td>10.8738</td>
<td>0.1335</td>
</tr>
<tr>
<td>$c = 0.6$</td>
<td>0.3051</td>
<td>10.1504</td>
<td>0.1009</td>
</tr>
<tr>
<td>Higher $\alpha$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha = 0.28$</td>
<td>0.2579</td>
<td>9.3161</td>
<td>0.0853</td>
</tr>
<tr>
<td>$\alpha = 0.30$</td>
<td>0.2598</td>
<td>9.1239</td>
<td>0.0880</td>
</tr>
<tr>
<td>$\alpha = 0.32$</td>
<td>0.2619</td>
<td>8.9473</td>
<td>0.0901</td>
</tr>
<tr>
<td>Higher $r$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r = 0.15$</td>
<td>0.2538</td>
<td>6.4632</td>
<td>0.0549</td>
</tr>
<tr>
<td>$r = 0.20$</td>
<td>0.2528</td>
<td>4.8612</td>
<td>0.0157</td>
</tr>
<tr>
<td>$r = 0.25$</td>
<td>0.2523</td>
<td>3.8957</td>
<td>0.0336</td>
</tr>
</tbody>
</table>

Table I. Equilibrium outcomes with and without regulation - The EPA acts first.

Changes in parameters $d$ and $\gamma$, however, only give rise to an indirect effect on $T_{EPA}$ by affecting the stringency of the emission fee, since these parameters are not arguments of $T_{EPA}$. To see this point, note that more severe pollution (higher $d$) entails a more stringent emission fee, $\tau^*$, which lowers the patent length, $T_{EPA}$. While the PO ignores pollution in its decisions (i.e., $T_{EPA}$ is unaffected by $d$), this agency anticipates that a more stringent emission fee increases firms’ net production cost, thus making innovation less valuable for society, ultimately leading to shorter
patents in equilibrium. The opposite argument applies when R&D becomes more costly (higher $\gamma$) since $T_{EPA}$ is not directly affected by $\gamma$, but it is indirectly via a less stringent emission fee (recall that the EPA anticipates less investment in R&D, followed by less output and pollution). Overall, a lower fee $\tau$ yields a longer patent. Table II summarizes the direct and indirect effects from changes in each parameter, as well as their overall effect on $T_{EPA}$.

<table>
<thead>
<tr>
<th></th>
<th>Direct effect</th>
<th>Indirect effect</th>
<th>Overall effect on $T_{EPA}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher $\alpha$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
</tr>
<tr>
<td>Higher $c$</td>
<td>$-$</td>
<td>$+$</td>
<td>$-$</td>
</tr>
<tr>
<td>Higher $r$</td>
<td>$-$</td>
<td>$+$</td>
<td>$-$</td>
</tr>
<tr>
<td>Higher $d$</td>
<td>None</td>
<td>$-$</td>
<td>$-$</td>
</tr>
<tr>
<td>Higher $\gamma$</td>
<td>None</td>
<td>$+$</td>
<td>$+$</td>
</tr>
</tbody>
</table>

Table II. Direct and indirect effects on $T_{EPA}$.

**Regulation vs. No regulation.** The last two columns in Table I examine the effect of regulation on the patent length, $\Delta T \equiv T_0 - T_{EPA}$, and on equilibrium investment, $\Delta x \equiv x_0 - x_{EPA}$. Both $\Delta T$ and $\Delta x$ are unambiguously positive, meaning that the presence of environmental policy induces less investment and shorter patents, as discussed above. As expected, this effect is particularly intense when pollution becomes more severe (higher $d$), R&D is less costly (lower $\gamma$), production is less costly (lower $c$), and the innovation does not bring large cost-reduction effects (lower $\alpha$). It is in these contexts that a PO should be more aware of the EPA’s policies, anticipating firms’ responses to regulation, thus shortening patent lengths more significantly.

### 4.2 The EPA acts in the second stage

In this section, we still consider the role of environmental regulation, but assume that the EPA acts in the second stage. This entails that, in the first stage, the PO sets the patent length $T$; in the second stage, the EPA responds setting the emission fee $\tau$; in the third stage, firms choose their investment in R&D, $x$; and in the fourth stage, they compete after the patent expires.

In this context, equilibrium behavior in stages 3-4 is analogous to those in section 4.1. In the second stage, the EPA anticipates per-period welfare before the patent expires, $W^{P}_{EPA}$, after the patent expires, $W^{C}_{EPA}$, and in the absence of innovation, $W^{N}_{EPA}$; as defined in (1). (For compactness, these expressions are presented in Appendix 1.) Therefore, the EPA sets the emission fee $\tau$ that solves

$$
\max_{\tau \geq 0} \ W = \ x(T, \tau) \left( \int_0^T W^{P}_{EPA}dt + \int_T^{+\infty} W^{C}_{EPA}dt \right) + [1 - x(T, \tau)] \int_0^{+\infty} W^{N}_{EPA}dt - \frac{\gamma}{2} [x(T, \tau)]^2
$$

(7)
which, as shown in Appendix 1, yields emission fee \( \tau(T) \) that can be presented as the following linear combination

\[
\tau(T) = \theta \tau^P + (1 - \theta) \tau^C
\]

where \( \tau^P = \frac{2d(1-c)-\alpha}{1+2d} \) denotes the fee in the monopoly market that emerges in each patent period and \( \tau^C = \frac{2d(1-c)+2d\alpha}{1+2d} \) represents the fee in the perfectly competitive market after the patent expires, where \( \tau^C > \tau^P \) indicates that pollution is larger after the patent expires, requiring a more stringent fee.

For compactness, weight \( \theta \in [0, 1] \) is presented in Appendix 1, and figure 4 depicts this weight as a function of the patent length \( T \). When \( T \) is relatively short, the EPA anticipates a short monopoly but a long period of perfect competition with low costs, yielding an overall increase in pollution. In this context, it seeks a more stringent fee, thus assigning a lower weight on \( \tau^P \), i.e., \( \theta \) decreases in \( T \). However, when patents become longer, the EPA anticipates a monopoly during more periods, and thus less pollution. This agency sets, then, a less stringent fee, assigning a higher weight on \( \tau^P \), i.e., \( \theta \) increases in \( T \).

Figure 4. Weight \( \theta \) on fee \( \tau^P \).

Figure 4 also shows that weight \( \theta \) increases in the severity of pollution (higher \( d \)), thus assigning more weight to the monopoly fee \( \tau^P \). However, this result does not imply that overall tax burden decreases when \( d \) increases, since both \( \tau^P \) and \( \tau^C \) are increasing in \( d \). As shown in Table A.1 in Appendix 1, the overall fee \( \tau(T) \) is also increasing in \( d \). This table reports comparative statics of \( \theta \), \( \tau^P \), \( \tau^C \), and \( \tau(T) \) with respect to \( \gamma \), \( c \), \( \alpha \), and \( r \).

In the first stage, the PO anticipates emission fee \( \tau(T) \), which induces investment level \( x(T, \tau(T)) \), and yields welfare levels \( W^P = \frac{(1-c-\tau(T))^2}{2} + \alpha (1 - c - \tau(T)) \), \( W^C = \frac{[1-(c-\alpha+\tau(T))]^2}{2} \), and \( W^N = \frac{[1-(c+\tau(T))]^2}{2} \). Unlike the EPA, the PO does not directly consider environmental damages, but these welfare levels are still indirectly affected by pollution severity, \( d \), since they are a function of the emission fee \( \tau(T) \). The PO, then, chooses the patent length \( T \) to solve a problem analogous to (3),

\[\text{Figure 4 also shows that weight } \theta \text{ increases in the severity of pollution (higher } d \text{), thus assigning more weight to the monopoly fee } \tau^P. \text{ However, this result does not imply that overall tax burden decreases when } d \text{ increases, since both } \tau^P \text{ and } \tau^C \text{ are increasing in } d. \text{ As shown in Table A.1 in Appendix 1, the overall fee } \tau(T) \text{ is also increasing in } d. \text{ This table reports comparative statics of } \theta, \tau^P, \tau^C, \text{ and } \tau(T) \text{ with respect to } \gamma, c, \alpha, \text{ and } r. \text{ In the first stage, the PO anticipates emission fee } \tau(T), \text{ which induces investment level } x(T, \tau(T)), \text{ and yields welfare levels } W^P = \frac{(1-c-\tau(T))^2}{2} + \alpha (1 - c - \tau(T)), \text{ } W^C = \frac{[1-(c-\alpha+\tau(T))]^2}{2}, \text{ and } W^N = \frac{[1-(c+\tau(T))]^2}{2}. \text{ Unlike the EPA, the PO does not directly consider environmental damages, but these welfare levels are still indirectly affected by pollution severity, } d, \text{ since they are a function of the emission fee } \tau(T). \text{ The PO, then, chooses the patent length } T \text{ to solve a problem analogous to (3),}\]
that is,

\[
\max_{T \geq 0} W = x(T, \tau(T)) \left( \int_0^T e^{-rt} W^P dt + \int_T^{+\infty} e^{-rt} W^C dt \right) \\
+ [1 - x(T, \tau(T))] \int_0^{+\infty} e^{-rt} W^N dt - \frac{1}{2} \gamma [x(T, \tau(T))]^2
\]  

(9)

Before solving (9), let us compare the PO’s problem with and without regulation. In the absence of environmental regulation, the PO anticipates that a longer patent gives rise to a monopoly for one more period, yielding a welfare loss of \(WL_0 \equiv W_C^0 - W_P^0 = \frac{\alpha^2}{T}\); and with regulation this welfare loss remains unaffected at \(WL \equiv W_C - W_P = \frac{\alpha^2}{T}\). Intuitively, production costs are increased by the same fee \(\tau\) before and after the patent expires, thus yielding the same output change.

While the marginal static loss from a longer patent is composed of two terms: (i) a longer monopoly, and (ii) higher investment costs; our results show that (i) is unchanged as a result of regulation but (ii) decreases. This occurs because investments are lower than without regulation \(x_1(T) < x_0(T)\) and, due to the convexity of the cost of R&D function, marginal costs are also lower. In contrast, the marginal dynamic gain is larger since R&D gains are concave. Overall, both effects induce the PO to set longer patents.

Differentiating (9) with respect to \(T\) yields a highly non-linear first-order condition, but the marginal dynamic gain from a longer patent exceeds its marginal static loss for all values of \(T\), as suggested above, and depicted in figure 5 (this figure considers the same parameter values as the benchmark of Table I). Appendix 1 shows that this ranking is robust to different parameter combinations.

![Figure 5. MDG and MSL with EPA acting second.](image)

4.3 The EPA acts in the third stage

Consider that the EPA has the ability to revise emission fees after patent lengths and R&D investments are chosen. For compactness, Appendix 2 identifies equilibrium investments, emission fees,
and welfare in this context, while here we focus on examining the PO’s decision.

In the current setting, the EPA sets emission fee \( \tau^P \) before the patent expires, \( \tau^C \) after the patent expires, and \( \tau^N = \frac{2d(1-c)}{(1+2d)} \) when no innovation occurs, which satisfy \( \tau^C > \tau^P > \tau^N \), to account for the market monopolization during patent periods. Therefore, these fees induce the same output before and after the patent expires, yielding a PO’s welfare of \( W^P = W^C = \frac{(1+4d)(1-c)^2}{2(1+2d)} \); whereas when innovation does not occur this welfare becomes \( W^N = \frac{(1+4d)(1-c)^2}{2(1+2d)} \).

As a consequence, the social return of innovation, relative to no innovation, becomes \( S_1 = \frac{1}{r} e^{rT(1-c)(1+2d)} - \frac{1}{2} \gamma [x_1(T)]^2 \), which is unaffected by the patent length, \( T \); as opposed to \( S_0(T) \). Hence, the PO’s problem simplifies to

\[
\max_{T \geq 0} x_1(T)S_1 - \int_0^{+\infty} e^{-rt}W^N dt - \frac{1}{2} \gamma [x_1(T)]^2
\]

where \( \int_0^{+\infty} e^{-rt}W^N dt = \frac{(1-c)^2(1+4d)}{2r(1+2d)^2} \) is also constant in \( T \). This maximization problem yields first-order condition \( S_1 - \gamma x_1(T) \geq 0 \), simplifying to \( e^{-rT} \geq \frac{(1-c)(1+6d)+2da}{2(1-c)(1+2d)} \). This condition holds with strict inequality for all parameter values, entailing that \( S_1 > \gamma x_1(T) \). Therefore, the marginal dynamic gain from a longer patent exceeds its marginal static loss, as the following proposition shows.

**Proposition 2.** Both \( MDG_1(T) \) and \( MSL_1(T) \) are unambiguously decreasing in \( d \); \( MDG_1(T) \) decreases in \( T \), \( MSL_1(T) \) decreases in \( T \) if and only if \( T > \frac{\log 2}{r} \); and \( MDG_1(T) > MSL_1(T) \) for all \( T \).

The Nordhaus’ trade-off, which emerges under no regulation, breaks down when the EPA sets policy in the third stage. This occurs because the PO anticipates that a longer patent will not alter output, since it coincides before and after the patent expires.

### 4.4 Allowing for green innovations

Consider that investments not only reduce production costs but also environmental damages (a “green” innovations). In particular, environmental damage decreases from \( d \) to \( d - \lambda \), where \( d \geq \lambda \geq 0 \); but remains unaffected if the innovation is not successful. When \( \lambda = 0 \), the innovation does not reduce pollution intensity, as in our main model; whereas when \( \lambda = d \), the innovation makes output completely clean. (For compactness, we study the effect of green innovations in the setting of section 4.1.)

In the fourth stage, equilibrium output levels are unaffected by the decrease in parameter \( d \). In the third stage, the innovator chooses investment level, \( x(T, \tau) = \frac{\alpha(1-c-\tau)(1-c-e^{-\tau T})}{\gamma^2} \), which coincides with that in section 4.1.

In the second stage, the PO chooses the optimal patent length similarly as in subsection 4.1, and is unaffected by the environmental damages or green innovations (i.e., it is not a function of \( d \) or \( d - \lambda \)), that is, \( T_{EPA} = \frac{1}{r} \log \left[ \frac{2(1-(c-\alpha+\tau))}{\alpha} \right] \).
Finally, in the first stage, the EPA anticipates patent length $T_{EPA}$ and chooses emission fee $\tau$ to solve

$$\max_{\tau \geq 0} \ x(T, \tau) \left( \int_0^T e^{-\tau t} W_{EPA}^P(\lambda) dt + \int_T^{+\infty} e^{-\tau t} W_{EPA}^C(\lambda) dt \right) + [1 - x(T, \tau)] \int_0^{+\infty} e^{-\tau t} W_{EPA}^N dt - \frac{\gamma}{2} [x(T, \tau)]^2$$

(6')

which is analogous to (6), but with welfare levels being a function of the green innovation parameter, $\lambda$.

Differentiating with respect to $\tau$ yields a highly nonlinear first-order condition, not allowing for an explicit solution for $\tau(\lambda)$. Table III numerically evaluates the emission fee, patent length, and investment in equilibrium, at the same parameters as in Table I. For comparison purposes, the top row considers $\lambda = 0$, obtaining the same results as in Table I; while lower rows allow more intense green innovations (higher $\lambda$).

As expected, when innovations are greener (higher $\lambda$), the EPA anticipates less pollution, and thus sets a less stringent emission fee $\tau(\lambda)$. Facing a lower fee, the PO responds with a longer patent $T_{EPA}$ in the second stage. Intuitively, the PO anticipates firms will face lower costs (net of taxes), implying that the cost-reduction effect of the patent is more intense. This makes the innovation more socially attractive and the PO sets longer patents. Finally, firms respond to longer patents and less stringent fees investing more in R&D. Therefore, while firms ignore the pollution-reduction effects of their innovation, they benefit from this innovation in the form of lower taxes and longer patents, ultimately providing them with incentives to invest more in R&D than when innovations do not reduce pollution.

<table>
<thead>
<tr>
<th>Regulation</th>
<th>No regulation</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau(\lambda)$</td>
<td>$T_{EPA}$</td>
<td>$x_{EPA}$</td>
</tr>
<tr>
<td>Benchmark, $\lambda = 0$</td>
<td>0.2555</td>
<td>9.6409</td>
</tr>
<tr>
<td>$\lambda = 0.05$</td>
<td>0.2547</td>
<td>9.6667</td>
</tr>
<tr>
<td>$\lambda = 0.10$</td>
<td>0.2538</td>
<td>9.6929</td>
</tr>
<tr>
<td>$\lambda = 0.15$</td>
<td>0.2529</td>
<td>9.7195</td>
</tr>
<tr>
<td>$\lambda = 0.20$</td>
<td>0.2520</td>
<td>9.7467</td>
</tr>
<tr>
<td>$\lambda = 0.25$</td>
<td>0.2511</td>
<td>9.7743</td>
</tr>
<tr>
<td>$\lambda = 0.30$</td>
<td>0.2502</td>
<td>9.8026</td>
</tr>
</tbody>
</table>

Table III. Equilibrium outcomes with green innovations.

The previous-to-last column in Table III evaluates the patent decrease due to environmental

\footnote{Welfare $W_{EPA}^N$ is unaffected by $\lambda$ since the innovation does not occur in this setting, thus not bringing pollution-reduction effects. Welfare $W_{EPA}^P(\lambda)$ and $W_{EPA}^C(\lambda)$ are both unambiguously increasing in $\lambda$.}
regulation, $\Delta T \equiv T_0 - T_{EPA}$, which is positive for all parameter values, but decreases in $\lambda$. As the innovation becomes greener, environmental policy is less necessary (lower emission fees), entailing that the setting with regulation approaches that without fees, yielding more similar patent lengths. Similarly, the last column reports the investment reduction as a result of regulation, $\Delta x \equiv x_0 - x_{EPA}$, which is also positive for all parameters and decreases in $\lambda$. In this case, firms face less stringent fees, and patents approach those without regulation, ultimately inducing similar R&D investments with and without environmental policy.

5 Discussion

Our results help understand how environmental regulation affects patent policy. The presence of emission fees induces firms to invest less in R&D, which reduce the expected social return of the innovation, thus inducing the PO to set shorter patents.

A negative direct effect arises both with and without regulation, with patents shortening as firms’ initial costs increase. With regulation, a positive indirect effect emerges, since emission fees become less stringent as costs increase, ultimately leading the PO to set longer patents. Overall, the negative direct effect dominates, yielding shorter patents. When the innovation becomes more cost-effective, however, the negative direct effect still arises, but the emission fee becomes more stringent, thus inducing a negative indirect effect, which reinforces the direct effect, implying an unambiguous patent reduction. A more severe pollution or less costly R&D do not affect patent lengths under no regulation, but they do with regulation. In particular, these changes induce a more stringent emission fee, thus producing a negative indirect effect on patents, lowering them.

When the EPA acts in the second stage, firms’ costs are uniformly affected by regulation in all scenarios (before and after the patent expires, and in the absence of innovation). This implies that the PO’s decision to allow for longer patents, thus letting the monopolist persist for more periods, does not affect the severity of the welfare loss from the patent, relative to that without environmental regulation. Nonetheless, environmental policy lowers investment costs, thus reducing the overall welfare loss from longer patents, so the PO can set longer patents to stimulate investments.

When the EPA acts in the third stage, it can induce the same output level in all three scenarios, yielding the same welfare. In this setting, a longer patent does not affect social welfare at all, implying that the traditional welfare loss from a longer monopoly does not arise in the presence of regulation. Since, in addition, firms’ R&D investment is lower with than without regulation, investment costs are lower, entailing that patent lengths are longer with than without regulation.

Overall, patents are shorter with than without environmental policy when the EPA is relatively less flexible than the PO, but become longer otherwise. Hence, considering the EPA’s administrative ability to revise emission fees is important, since this flexibility can produce different effects in the PO’s patent decisions. Environmental agencies often face bureaucratic hurdles, not allowing for rapid adjustments of emission fees based on industry conditions, thus suggesting that POs should set shorter patents with than without regulation. Ignoring the presence of environmental policy...
when designing patent lengths will induce unnecessarily long monopolies, ultimately reducing social welfare.

Finally, we show that our results are robust to green innovations, when R&D investments not only reduce firms’ production costs but also lower their pollution intensity. In particular, we demonstrate that green innovations help the EPA lower the stringency of emission fees, which lets firms increase their R&D investment, ultimately allowing the PO to set longer patents. In other words, as innovations become greener, environmental policy becomes less necessary, and patents approach those under no regulation. In these contexts, considering environmental regulation before setting patent lengths is less critical, but when innovations do not bring large reductions in pollution intensity, this consideration becomes more important for the PO.

Further research. Our model can be extended along several dimensions. First, we assumed that the PO can perfectly observe the severity of environmental damages that the EPA considers (parameter $d$). However, one could assume that the PO does not accurately observe this severity, potentially leading to further distortions in the patent policy. Second, the cost-reducing effect of the innovation may be a function of the firm’s investment (i.e., $\alpha$ being increasing in $x$). Finally, one could consider changes in the PO’s strategic plan, which require this agency to include environmental considerations in their assessments.

6 Appendix

6.1 Appendix 1 - EPA acts in the second stage

Second stage. In this stage, the EPA anticipates that per-period welfare when the patent is in force is given by

$$W_{EPA}^P = CS^P + PS^P + Tax^P - Env^P$$
$$= \frac{(1 - c - \tau)^2}{2} + \alpha (1 - c - \tau) + \tau (1 - c - \tau) - d (1 - c - \tau)^2$$
$$= \frac{(1 - c - \tau) [2 (\alpha + \tau) + (1 - 2d) (1 - c - \tau)]}{2}$$

Per-period welfare in every period after the patent expires is

$$W_{EPA}^C = CS^C + PS^C + Tax^C - Env^C$$
$$= \frac{[1 - (c - \alpha + \tau)]^2}{2} + 0 + \tau [1 - (c - \alpha + \tau)] - d [1 - (c - \alpha + \tau)]^2$$
$$= \frac{[1 - (c - \alpha + \tau)] [2\tau + (1 - 2d) [1 - (c - \alpha + \tau)]]}{2}$$
and welfare in every period without the innovation is

\[ W_{EPA}^N = CS^N + PS^N + Tax^N - Env^N \]

\[ = \frac{[1 - (c + \tau)]^2}{2} + 0 + \tau [1 - (c + \tau)] - d [1 - (c + \tau)]^2 \]

\[ = \frac{[1 - (c + \tau)] [2\tau + (1 - 2d) [1 - (c + \tau)]]}{2} \]

We can now use \( W_{EPA}^P, W_{EPA}^C \), and \( W_{EPA}^N \) to find the present discounted value of the innovation for the EPA

\[ W = x \left( \int_0^T W_{EPA}^P dt + \int_T^{+\infty} W_{EPA}^C dt \right) + (1 - x) \int_0^{+\infty} W_{EPA}^N dt - \frac{\gamma}{2} x^2 \]

Differentiating with respect to fee \( \tau \), yields

\[ \frac{\partial W_{EPA}}{\partial \tau} = \frac{\alpha^2 (1 - e^{-rT}) \left\{ [\alpha (2d - 1) - 2\tau (1 + 4d) + 8d (1 - c)] e^{-rT} - 2 (1 - c - \tau) \right\}}{2\gamma r^2} \]

\[ + \frac{2d (1 - c) - \tau (1 + 2d)}{r} \]

Setting the above first-order condition equal to zero, we find the equilibrium emission fee

\[ \tau(T) = \frac{2 (2d\gamma r - \alpha^2) (1 - c) + \alpha^2 e^{-rT} [\alpha (2d - 1) + 2 (1 - c) (1 + 4d)] - \alpha^2 e^{-2rT} [\alpha (2d - 1) + 8d (1 - c)]}{2\gamma r (2d + 1) - 2\alpha^2 (1 - e^{-rT}) [1 - (4d + 1) e^{-rT}]} \]

This fee can be alternatively expressed as a linear combination

\[ \tau(T) = \theta \tau^P + (1 - \theta) \tau^C \]

where \( \tau^P = \frac{2d(1-c)-\alpha}{1+2d} \) and \( \tau^C = \frac{2d[1-(c-\alpha)]}{1+2d} \), and weight \( \theta \) is

\[ \theta = \frac{2 \left[ \alpha (1 - c) - 2\alpha^2 d + 2d\gamma r (2d + 1) \right] + \alpha (2d + 1) e^{-rT} \left[ \alpha (1 + 6d) - 2 (1 - c) \right]}{2 (2d + 1) (\gamma r (2d + 1) - \alpha^2 (1 - e^{-rT}) [1 - (4d + 1) e^{-rT}])} \]

\[ - \frac{-\alpha [\alpha (1 + 4d + 12d^2) - 4d (1 - c)] e^{-2rT}}{2 (2d + 1) (\gamma r (2d + 1) - \alpha^2 (1 - e^{-rT}) [1 - (4d + 1) e^{-rT}])}. \]

(For a description on how to find fees \( \tau^P \) and \( \tau^C \), and their comparative statics, see Lemma A1 in Appendix 2.)

When \( T \to 0 \), weight \( \theta \) converges to \( \theta \to \frac{2d(1-c)}{1+2d} \), whereas when \( T \to +\infty \) this weight approaches \( \theta \to \frac{2d[(1+2d)\gamma-\alpha^2]+\alpha(1-c)}{(1+2d)((1+2d)\gamma-\alpha^2)} \). While \( \theta > 0 \) and \( \theta < 1 \) yield large intractable roots, weight \( \theta \) satisfies \( \theta \in [0,1] \) for a large set of parameter values, as reported in table A1. (The benchmark at the top row considers \( T = 10 \) and the same parameter values as table I in the main body of the paper.)
Table A2 considers similar parameter values.) For completeness, the table also reports fees $\tau^P$, $\tau^C$, and their weighted sum, $\tau(T) = \theta \tau^P + (1 - \theta) \tau^C$.

<table>
<thead>
<tr>
<th></th>
<th>$\theta$</th>
<th>$\tau^P$</th>
<th>$\tau^C$</th>
<th>$\tau(T)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td>0.7259</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2560</td>
</tr>
<tr>
<td>Higher $d$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d = 2$</td>
<td>0.7682</td>
<td>0.2167</td>
<td>0.4667</td>
<td>0.2746</td>
</tr>
<tr>
<td>$d = 2.5$</td>
<td>0.7976</td>
<td>0.2361</td>
<td>0.4861</td>
<td>0.2867</td>
</tr>
<tr>
<td>$d = 3$</td>
<td>0.8193</td>
<td>0.2500</td>
<td>0.5000</td>
<td>0.2952</td>
</tr>
<tr>
<td>Higher $\gamma$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma = 3$</td>
<td>0.7374</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2531</td>
</tr>
<tr>
<td>$\gamma = 5$</td>
<td>0.7423</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2519</td>
</tr>
<tr>
<td>$\gamma = 7$</td>
<td>0.7444</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2514</td>
</tr>
<tr>
<td>Lower $c$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c = 0.6$</td>
<td>0.7256</td>
<td>0.2375</td>
<td>0.4875</td>
<td>0.3061</td>
</tr>
<tr>
<td>$c = 0.5$</td>
<td>0.7245</td>
<td>0.3125</td>
<td>0.5625</td>
<td>0.3813</td>
</tr>
<tr>
<td>$c = 0.4$</td>
<td>0.7243</td>
<td>0.3875</td>
<td>0.6375</td>
<td>0.4564</td>
</tr>
<tr>
<td>Higher $\alpha$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha = 0.28$</td>
<td>0.7209</td>
<td>0.1800</td>
<td>0.4600</td>
<td>0.2582</td>
</tr>
<tr>
<td>$\alpha = 0.30$</td>
<td>0.7173</td>
<td>0.1750</td>
<td>0.4750</td>
<td>0.2598</td>
</tr>
<tr>
<td>$\alpha = 0.32$</td>
<td>0.7136</td>
<td>0.1700</td>
<td>0.4900</td>
<td>0.2617</td>
</tr>
<tr>
<td>Higher $r$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r = 0.15$</td>
<td>0.7441</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2515</td>
</tr>
<tr>
<td>$r = 0.20$</td>
<td>0.7529</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2493</td>
</tr>
<tr>
<td>$r = 0.25$</td>
<td>0.7566</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2484</td>
</tr>
<tr>
<td>Higher $T$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T = 15$</td>
<td>0.7413</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2522</td>
</tr>
<tr>
<td>$T = 20$</td>
<td>0.7557</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2486</td>
</tr>
<tr>
<td>$T = 25$</td>
<td>0.7669</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2458</td>
</tr>
<tr>
<td>$T = 30$</td>
<td>0.7747</td>
<td>0.1875</td>
<td>0.4375</td>
<td>0.2438</td>
</tr>
</tbody>
</table>

Table A1. Weight $\theta$ and fees $\tau^P$, $\tau^C$, and $\tau(T)$.

**First stage.** The PO solves

$$
\max_{T \geq 0} W = x(T, \tau(T)) \left( \int_0^T e^{-rt} W^P dt + \int_T^{+\infty} e^{-rt} W^C dt \right)
+ [1 - x(T, \tau(T))] \int_0^{+\infty} e^{-rt} W^N dt - \frac{1}{2} \gamma [x(T, \tau(T))]^2
$$

where $W^P = \frac{(1-c-\tau)^2}{2} + \alpha (1-c-\tau)$, $W^C = \frac{[1-(c+\alpha+\tau)]^2}{2}$, and $W^N = \frac{[1-(c+\tau)]^2}{2}$. This problem simplifies to

$$
\max_{T \geq 0} W = x(T, \tau(T)) S_1(\tau(T), T) + \int_0^{+\infty} e^{-rt} W^N dt - \frac{1}{2} \gamma [x(T, \tau(T))]^2
$$

21
where
\[
S_1(\tau(T), T) = \int_0^T e^{-rt}WP \, dt + \int_T^{+\infty} e^{-rt}W^C \, dt - \int_0^{+\infty} e^{-rt}W^N \, dt
\]
denotes the social return from the innovation. Inserting \( \tau(T) \), yields
\[
S_1(T) = \frac{[2\alpha^2(1-c) - \alpha^2(2d-1)](1 - e^{-rT})e^{-rT} + 2(1-c)\gamma r}{8\gamma r^2 \{\gamma r(2d+1) - \alpha^2(1-e^{-rT})[1 - (1+4d)e^{-rT}]\}^2} \times B
\]
where
\[
B = 2(1-c)\gamma^2 r^2 + 2\alpha^4e^{-rT}(1-c)(1-e^{-rT})^2(1 + e^{-rT}) + 2\alpha^2(1-c)\gamma r(1 + e^{-rT} - 2e^{-2rT})
- \alpha^5 e^{-rT}(2d+1)(1 - 3e^{-rT})(1-e^{-rT})^2 + \alpha^3 \gamma r(2d+3)e^{-rT}(1-e^{-rT}).
\]

Solving the PO’s problem above, the next table evaluates the marginal dynamic gain (\( MDG \))
and marginal static loss (\( MSL \)) from a longer patent, showing that \( MDG > MSL \) for a large
set of parameter combinations. Other numerical simulations can be provided by the authors upon
request.

<table>
<thead>
<tr>
<th>( \tau^{EPA} )</th>
<th>( MDG )</th>
<th>( MSL )</th>
<th>( MDG - MSL )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td>0.2560</td>
<td>0.0017</td>
<td>0.00014</td>
</tr>
<tr>
<td>Higher ( d )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( d = 2 )</td>
<td>0.2746</td>
<td>0.011</td>
<td>0.00005</td>
</tr>
<tr>
<td>( d = 2.5 )</td>
<td>0.2867</td>
<td>0.0008</td>
<td>0.00004</td>
</tr>
<tr>
<td>( d = 3 )</td>
<td>0.2952</td>
<td>0.0006</td>
<td>-0.00001</td>
</tr>
<tr>
<td>Higher ( \gamma )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma = 3 )</td>
<td>0.2531</td>
<td>0.0008</td>
<td>0.00003</td>
</tr>
<tr>
<td>( \gamma = 5 )</td>
<td>0.2519</td>
<td>0.0005</td>
<td>0.00001</td>
</tr>
<tr>
<td>( \gamma = 7 )</td>
<td>0.2514</td>
<td>0.0004</td>
<td>0.00005</td>
</tr>
<tr>
<td>Lower ( c )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c = 0.6 )</td>
<td>0.3061</td>
<td>0.0023</td>
<td>0.00023</td>
</tr>
<tr>
<td>( c = 0.5 )</td>
<td>0.3813</td>
<td>0.0034</td>
<td>0.00043</td>
</tr>
<tr>
<td>( c = 0.4 )</td>
<td>0.4564</td>
<td>0.0047</td>
<td>0.00068</td>
</tr>
<tr>
<td>Higher ( \alpha )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha = 0.28 )</td>
<td>0.2582</td>
<td>0.0022</td>
<td>0.00018</td>
</tr>
<tr>
<td>( \alpha = 0.30 )</td>
<td>0.2598</td>
<td>0.0026</td>
<td>0.00022</td>
</tr>
<tr>
<td>( \alpha = 0.32 )</td>
<td>0.5165</td>
<td>0.0059</td>
<td>0.00092</td>
</tr>
<tr>
<td>Higher ( r )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r = 0.15 )</td>
<td>0.2515</td>
<td>0.0007</td>
<td>-0.00003</td>
</tr>
<tr>
<td>( r = 0.20 )</td>
<td>0.2493</td>
<td>0.0003</td>
<td>-0.00005</td>
</tr>
<tr>
<td>( r = 0.25 )</td>
<td>0.2484</td>
<td>0.0001</td>
<td>-0.00004</td>
</tr>
<tr>
<td>Higher ( T )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T = 15 )</td>
<td>0.2519</td>
<td>0.0011</td>
<td>0.000004</td>
</tr>
<tr>
<td>( T = 20 )</td>
<td>0.2486</td>
<td>0.0007</td>
<td>-0.000027</td>
</tr>
<tr>
<td>( T = 25 )</td>
<td>0.2458</td>
<td>0.0005</td>
<td>-0.000027</td>
</tr>
<tr>
<td>( T = 30 )</td>
<td>0.2438</td>
<td>0.0003</td>
<td>-0.000021</td>
</tr>
</tbody>
</table>
6.2 Appendix 2 - EPA acts in the third stage

**Fourth stage, output decisions.** Before the patent expires \((t < T)\), the innovator enjoys lower production cost, \(c - \alpha + \tau\), thanks to the non-radical innovation, generating monopoly profits of

\[
\pi_1^P = (c + \tau)Q_1 - (c - \alpha + \tau)Q_1 \\
= \alpha Q_1
\]

which are positive since \(Q_1 \geq 0\). Hence, profits are positive during the patent but zero afterwards, \(\pi_1^P \geq \pi_1^C = 0\).

**Third stage, emission fees.** The EPA anticipates the firms’ decisions in the subsequent stage and chooses the emission fee, \(\tau\), which induces the socially optimal output that solves (1), as the following lemma identifies.

**Lemma A1.** Emission fees are:

1. \(\tau^P = \frac{2d(1-c)-\alpha}{1+2d}\) during the patent period, which is positive for all parameter values.
2. \(\tau^C = \frac{2d(1-c-\alpha)}{1+2d}\) after the patent expires, which is positive for all parameter values, increasing in \(d\) but decreasing in \(c\).
3. \(\tau^N = \frac{2d(1-c)}{1+2d}\) if no innovation occurs, which is positive for all parameter values, increasing in \(d\) but decreasing in \(c\).

In addition, emission fees satisfy \(\tau^C > \tau^N > \tau^P\) for all parameter values, and induce a socially optimal output \(q^{SO} = \frac{1-(c-\alpha)}{1+2d}\) before and after the patent, and \(q^{SO} = \frac{1-c}{1+2d}\) without innovation. Output is higher without than with regulation for all parameter values.

**Proof of Lemma A1.** Given the innovator’s profits of \(\pi(q) = (1-q)q - [(c - \alpha) + \tau]q\), the EPA solves:

\[
\max_{q \geq 0} \frac{1}{2} q^2 + [1 - q - (c - \alpha) - \tau]q + \tau q + \frac{q}{T} - dq^2 \\
= \left[1 - \frac{q}{2} - (c - \alpha)\right]q - dq^2
\]

Differentiating with respect to \(q\), and assuming interior solutions, we find \([1-q-(c-\alpha)]-2dq=0\). Rearranging and solving for \(q\), the socially optimal output becomes \(q_1^{SO} = \frac{1-(c-\alpha)}{1+2d}\). Setting the socially optimal output equal to the monopolist’s equilibrium output when the patent is in force, that is, \(q^P = 1 - (c + \tau) = q^{SO}\), we find the emission fee,

\[
\tau^P = \frac{2d(1-c)-\alpha}{1+2d}
\]
which is positive for all $d > \tilde{d} \equiv \frac{\alpha}{2(1-c)}$. However, cutoff $\tilde{d}$ lies above $\frac{1}{2}$ since $\alpha < 1 - c$ by definition, implying that the initial condition $d > \frac{1}{2}$ is more demanding than $d > \tilde{d}$. Emission fee $\tau^P$ is increasing in $d$ because $\frac{\partial \tau^P}{\partial d} = \frac{2(1-c+\alpha)}{(1+2d)^2} > 0$.

Setting the socially optimal output equal to the industry’s aggregate output after the patent expires, that is, $Q^C = 1 - ((c - \alpha) + \tau) = q^{SO}$, we obtain the following emission fee,

$$\tau^C = \frac{2d [1 - (c - \alpha)]}{1 + 2d}$$

which is unambiguously positive and increasing in $d$ since $\frac{\partial \tau^C}{\partial d} = \frac{2(1-c+\alpha)}{(1+2d)^2} > 0$.

In the case of no innovation, the socially optimal output becomes $q^{SO} = \frac{1-c}{1+2d}$. Equating this to the industry’s aggregate output, $Q^N = 1 - (c + \tau)$, we find the following emission fee,

$$\tau^N = \frac{2d (1 - c)}{1 + 2d}$$

which is decreasing in the initial cost $c$ but increasing in environmental damages $d$ because $\frac{\partial \tau^N}{\partial d} = \frac{2(1-c)}{(1+2d)^2} > 0$.

Comparing these three fees, we obtain that $\tau^C > \tau^N > \tau^P$ since

$$\frac{2d [1 - (c - \alpha)]}{1 + 2d} > \frac{2d (1 - c)}{1 + 2d} > \frac{2d (1 - c) - \alpha}{1 + 2d}$$

simplifies to $2d [1 - (c - \alpha)] > 2d (1 - c) > 2d (1 - c) - \alpha$, and further to $2d \alpha > 0 > -\alpha$ that holds for all parameter values.

Finally, before the patent expires, $q^{SO} = \frac{1-c}{1+2d} < 1 - c = Q^P$ holds if and only if $d > \tilde{d} \equiv \frac{\alpha}{2(1-c)}$, which holds since $d > \frac{1}{2}$ by definition. After the patent expires, $q^{SO} = \frac{1-c}{1+2d} < 1 - (c - \alpha)$ unambiguously holds; and, similarly, if there is no innovation, $q^{SO} = \frac{1-c}{1+2d} < 1 - c$, which also holds for all parameters. Q.E.D.

After the patent expires, firms face a more stringent fee, $\tau^C > \tau^P$, to induce them to produce the same output level. When the innovation does not take place, the emission fee $\tau^N$ is more (less) stringent than when the patent is in force (expired), $\tau^N > \tau^P$ ($\tau^N < \tau^C$), since the market is perfectly competitive but no firms can benefit from the cost-reduction technology.

Second stage, R&D investment. Using the fees in Lemma A1, we find the firm’s profits during the patent period,

$$\pi^P = (1 - q^{SO} - (c - \alpha) - \tau^P)q^{SO} = \frac{\alpha [1 - (c - \alpha)]}{1 + 2d},$$

while after the patent expires, or in the absence of innovation, every firm earns $\pi^C = \pi^N = 0$ in a
perfectly competitive market\textsuperscript{10} yielding the return from a successful innovation of

\[
\Pi_1(T) = \int_0^T e^{-rt} \pi P(\tau P) dt + \int_T^{+\infty} e^{-rt} \pi C(\tau C) dt
\]

\[
= \frac{\alpha [1 - (c - \alpha)] (1 - e^{-rT})}{r (1 + 2d)}
\]

which is monotonically increasing in the cost-reduction effect \(\alpha\) and the patent length, \(T\), at a rate of \(e^{-rT} \pi P\); but decreasing in the initial cost \(c\) and environmental damages, \(d\).

In this context, the innovator chooses its R&D investment, \(x\), to solve

\[
\max_{x \geq 0} x \Pi_1(T) + (1 - x) \int_0^{+\infty} e^{-rt} \pi^{NI} dt - \frac{1}{2} \gamma x^2
\]

Differentiating the above expression with respect to \(x\), yields

\[
x(T) = \frac{\Pi_1(T)}{\gamma}
\]

which falls below that when the EPA is absent, \(x(T) < x_0(T)\), if and only if \(\Pi_1(T) < \Pi_0(T)\), as the following lemma shows.

**Lemma A2.** The innovator’s R&D investment is \(x(T) = \frac{\alpha [1 - (c - \alpha)] (1 - e^{-rT})}{\gamma r (1 + 2d)}\), which is decreasing in \(\gamma, d, c,\) and \(r\), but increasing in \(\alpha\). Moreover, \(0 < x(T) < x_0(T)\) and \(0 < x'(T) < x_0'(T)\).

**Proof of Lemma A2.** Differentiating the firm’s profit function with respect to \(x\), its optimal investment becomes

\[
x(T) = \frac{\Pi_1(T)}{\gamma}
\]

\[
= \frac{\alpha [1 - (c - \alpha)] (1 - e^{-rT})}{\gamma r (1 + 2d)}
\]

which lies below that when the EPA is absent, \(x_0(T) = \frac{\alpha (1-c)(1-e^{-rT})}{\gamma T}\), if and only if \(d > \tilde{d}\), which holds since \(d > \frac{1}{2}\) by definition; and it is decreasing in \(\gamma\) and \(r\) because

\[
\frac{\partial x(T)}{\partial r} = \frac{\alpha [1 - (c - \alpha)] [1 + rT - e^{rT}]}{\gamma r^2 (1 + 2d) e^{rT}}
\]

which is negative since \(1 + rT < e^{rT}\) for all \(r, T \geq 0\).

In addition, differentiating \(x(T)\) with respect to \(T\) yields \(\frac{\partial x(T)}{\partial T} = \frac{\alpha (1-c) e^{-rT}}{\gamma (1 + 2d)}\), which satisfies

\textsuperscript{10}Inserting fee \(\tau C\) from Lemma A1 into the firm’s profits after the patent expires, we obtain \(\pi C = (1 - q^{SO} - c + \alpha - \tau C) q^{SO} = 0\). Similarly, substituting emission fee \(\tau N\) from Lemma 4 into the firm’s profits in the absence of innovation, yields \(\pi N = (1 - q^{SO} - c - \tau N) q^{SO} = 0\).
$0 < x'(T) < x'_0(T)$ if and only if $d > \hat{d}$. Next, $x(T)$ increases in $\alpha$ since
\[
\frac{\partial x(T)}{\partial \alpha} = \frac{[1 - (c - 2\alpha)] (1 - e^{-rT})}{\gamma r (1 + 2d)} > 0
\]
Finally, $x(T)$ decreases in $c$ because $\frac{\partial x(T)}{\partial c} = -\frac{\alpha (1 - e^{-rT})}{\gamma r (1 + 2d)} < 0$. Q.E.D.

The emission fee, then, induces less R&D investments for a given patent length, $x(T) < x_0(T)$; and reduces the incentives to invest in R&D when the patent becomes longer, $x'(T) < x'_0(T)$. Similarly, the innovator has less incentives to invest in R&D, for a given patent length, when the investment becomes less efficient (higher $\gamma$) or when its discount factor increases (higher $r$).

**First stage, patent length.** The PO anticipates the firms’ decisions in the subsequent stages of the game and chooses the patent length, $T$, to solve the following welfare maximization problem,
\[
\max_{T \geq 0} x(T) S_1(T) - \frac{1}{2} \gamma [x(T)]^2
\]
where $S_1(T)$ is the social return of innovation, relative to that without innovation,
\[
S_1(T) = \left( \int_0^T e^{-rt} W^P dt + \int_T^{\infty} e^{-rt} W^C dt \right) - \int_0^{\infty} e^{-rt} W^N dt
\]
and $W^P$ ($W^C$) denotes the social welfare in each patent period (after the patent expires), and $W^N$ stands for the social welfare in the absence of the innovation.

While profits during and after the patent satisfy $\pi^P > \pi^C = 0$, welfare ranking satisfies $W^P = W^C = W^{SO}$ since the same socially optimal output, $q^{SO}$, is induced before or after the patent expires. In particular, this welfare level is
\[
W^{SO} = \frac{(q^{SO})^2}{2} + (1 - q^{SO} - c + \alpha) q^{SO} = \frac{(1 + 4d) [1 - (c - \alpha)^2]}{2 (1 + 2d)^2}
\]
which increases in $\alpha$ but decreases in $c$ and $d$; and simplifies to $W^{SO} = \frac{[1 - (c - \alpha)]^2}{2}$ when damages are absent, $d = 0$. Recall that the PO does not consider environmental damages, but $d$ enters its welfare evaluation because the PO anticipates the emission fee that the EPA sets in the next stage.

Similarly, welfare in the absence of innovation is evaluated at $q_0^{SO}$, yielding $W^N = \frac{(1 + 4d) (1 - c)^2}{2(1 + 2d)^2}$, which decreases in $c$ and $d$, and collapses to $\frac{(1 - c)^2}{2}$ when $d = 0$. In addition, $W^{SO} - W^N = \frac{\alpha(1 + 4d)[2(1 - c) + \alpha]}{2(1 + 2d)^2}$, which is positive for all $\alpha > 0$, increasing in $\alpha$, but decreasing in $d$. 

26
Inserting $W_P$, $W_C$, and $W_N$ into expression (A1), we obtain

$$S_1 = \left( \int_0^T e^{-rt} W_P \, dt + \int_T^{+\infty} e^{-rt} W_C \, dt \right) - \int_0^{+\infty} e^{-rt} W_N \, dt$$

$$= \frac{\alpha (1 + 4d) [\alpha + 2 (1 - c)]}{2r (1 + 2d)^2}$$

which is not a function of $T$ since welfare levels coincide before and after the patent expires; and simplifies to $S_1 = \frac{\alpha [\alpha + 2 (1 - c)]}{2r}$ when $d = 0$. Finally, note that $S_1$ increases in $\alpha$, but decreases in $c$ and $d$, asymptotically converging to 0 when $d \to +\infty$.

### 6.3 Proof of Lemma 1

It is straightforward to show that $x_0'(T) = \frac{\alpha (1-c) e^{-rT}}{\gamma} > 0$ and $x_0''(T) = -\frac{\alpha (1-c) e^{-rT}}{\gamma^2} < 0$, with

$$\frac{\partial x_0(T)}{\partial \alpha} = \frac{\alpha (1-c)(1-e^{-rT})}{\gamma^2} > 0$$

and

$$\frac{\partial x_0(T)}{\partial \gamma} = -\frac{\alpha (1-c)(1-e^{-rT})}{\gamma^2} < 0.$$

### 6.4 Proof of Lemma 2

The social return on innovation is the sum of consumer and producer surplus, above that of no innovation, as follows,

$$S_0(T) = \left( \int_0^T e^{-rt} W_P^0 \, dt + \int_T^{+\infty} e^{-rt} W_C^0 \, dt \right) - \int_0^{+\infty} e^{-rt} W_N^0 \, dt$$

$$= \frac{\alpha [\alpha + 2 (1 - c) e^{rT}] e^{-rT}}{2r}$$

which satisfies $\frac{\partial S_0(T)}{\partial T} = -\frac{\alpha^2 e^{-rT}}{2}$. Inserting into expression (3), we find that the $MDG_0(T)$ is

$$MDG_0(T) = \frac{\partial x_0(T)}{\partial T} S_0(T)$$

$$= \frac{\alpha^2 (1-c) [\alpha + 2 (1 - c) e^{rT}] e^{-2rT}}{2\gamma^2}$$

which is monotonically decreasing in $T$ because

$$\frac{\partial MDG_0(T)}{\partial T} = -\frac{\alpha^2 (1-c) [\alpha + (1 - c) e^{rT}] e^{-2rT}}{\gamma} < 0$$

with its second-order condition being

$$\frac{\partial^2 MDG_0(T)}{\partial T^2} = \frac{\alpha^2 (1-c) [2\alpha + (1 - c) e^{rT}] e^{-2rT}}{\gamma}$$

which is unambiguously positive.
In this context, the $MSL_0(T)$ becomes

$$MSL_0(T) = x_0(T) \left( \gamma \frac{\partial x_0(T)}{\partial T} - \frac{\partial S_0(T)}{\partial T} \right)$$

$$= \frac{\alpha^2 (1 - c) [\alpha + 2 (1 - c)] e^{-rT} (1 - e^{-rT})}{2 \gamma r}$$

that is monotonically decreasing in $T$ if and only if $T > \frac{\log 2}{r}$, yielding

$$\frac{\partial MSL_0(T)}{\partial T} = \frac{\alpha^2 (1 - c) [\alpha + 2 (1 - c)] e^{-2rT} (2 - e^{rT})}{2 \gamma} < 0$$

and at a decreasing rate if and only if $T > \frac{\log 4}{r}$, under which

$$\frac{\partial^2 MSL_0(T)}{\partial T^2} = \frac{\alpha^2 r (1 - c) [\alpha + 2 (1 - c)] e^{-2rT} (e^{rT} - 4)}{2 \gamma} > 0.$$ 

The $MDG_0(T)$ decreases more significantly in $T$ than the $MSL_0(T)$ does if and only if $MDG'_0(T) < MSL'_0(T)$, where

$$\frac{- \alpha^2 (1 - c) [\alpha + (1 - c) e^{rT}] e^{-2rT}}{\gamma} < \frac{\alpha^2 (1 - c) [\alpha + 2 (1 - c)] e^{-2rT} (2 - e^{rT})}{2 \gamma}$$

reduces to $2 [\alpha + (1 - c) e^{rT}] > [\alpha + 2 (1 - c)] (e^{rT} - 2)$, which yields

$$T < T_0 = \frac{1}{r} \log \left[ \frac{4 (1 - (c - \alpha))}{\alpha} \right]$$

where cutoff $T_0$ satisfies $\frac{\log 2}{r} < T_0 < \frac{\log 4}{r}$.

### 6.5 Proof of Lemma 3

**Comparative statics with respect to $\alpha$.** The $MDG_0(T)$ increases with $\alpha$, since

$$\frac{\partial MDG_0(T)}{\partial \alpha} = \frac{\alpha (1 - c) [3\alpha + 4 (1 - c) e^{rT}] e^{-2rT}}{2 \gamma r} > 0$$

and so does the $MSL_0(T)$ given that

$$\frac{\partial MSL_0(T)}{\partial \alpha} = \frac{\alpha (1 - c) [3\alpha + 4 (1 - c)] e^{-rT} (1 - e^{-rT})}{2 \gamma r} > 0.$$ 

Comparing these derivatives, we find that the $MSL_0(T)$ increases more significantly than the
MDG₀(T) in α if and only if $\frac{\partial MSL₀(T)}{\partial \alpha} > \frac{\partial MDG₀(T)}{\partial \alpha}$, that is,

$$\frac{\alpha (1 - c) [3\alpha + 4 (1 - c)] e^{-rT} (1 - e^{-rT})}{2\gamma r} > \frac{\alpha (1 - c) [3\alpha + 4 (1 - c) e^{rT}] e^{-2rT}}{2\gamma r}$$

which simplifies to $[3\alpha + 4 (1 - c)] (1 - e^{-rT}) > [3\alpha + 4 (1 - c) e^{rT}] e^{-rT}$, and solving for T, yields

$$T > T₀ = \frac{1}{r} \log \left[ \frac{2 [3\alpha + 2 (1 - c)]}{3\alpha} \right]$$

where $T₀ > \frac{\log 2}{r}$ simplifies to $4 (1 - c) > 0$.

Comparative statics with respect to c. The MDG₀(T) decreases with c, since

$$\frac{\partial MDG₀(T)}{\partial c} = -\frac{\alpha^2 [\alpha + 4 (1 - c) e^{rT}] e^{-2rT}}{2\gamma r} < 0$$

and so does the MSL₀(T) because

$$\frac{\partial MSL₀(T)}{\partial c} = -\frac{\alpha^2 [\alpha + 4 (1 - c) e^{rT}] e^{-rT} (1 - e^{-rT})}{2\gamma r} < 0.$$

Comparing these derivatives, we obtain that the MDG₀(T) decreases more significantly than the MSL₀(T) in c if and only if $\frac{\partial MDG₀(T)}{\partial c} < \frac{\partial MSL₀(T)}{\partial c}$, or

$$-\frac{\alpha^2 [\alpha + 4 (1 - c) e^{rT}] e^{-2rT}}{2\gamma r} < -\frac{\alpha^2 [\alpha + 4 (1 - c) e^{rT}] e^{-rT} (1 - e^{-rT})}{2\gamma r}$$

which simplifies to $e^{-rT} [\alpha + 4 (1 - c) e^{rT}] > [\alpha + 4 (1 - c)] (1 - e^{-rT})$, which becomes

$$T < \tilde{T}_0 = \frac{1}{r} \log \left[ \frac{2 [\alpha + 2 (1 - c)]}{\alpha} \right].$$

Cutoffs $\tilde{T}_0$ and $T₀$ satisfy $\tilde{T}_0 > T₀$ since $4 (1 - c) > 0$, while cutoffs $T₀$ and $\tilde{T}_0$ satisfy $T₀ > \tilde{T}_0$ since $\alpha > 0$ by definition.

6.6 Proof of Proposition 1

Setting $MDG₀(T) = MSL₀(T)$, yields

$$\frac{\alpha^2 (1 - c) [\alpha + 2 (1 - c) e^{rT}] e^{-2rT}}{2\gamma r} = \frac{\alpha^2 (1 - c) [\alpha + 2 (1 - c)] e^{-rT} (1 - e^{-rT})}{2\gamma r}.$$
Rearranging, we obtain
\[
[\alpha + 2 (1 - c) e^{rT}] e^{-rT} = [\alpha + 2 (1 - c)](1 - e^{-rT}).
\]
Solving for \(T\), we find the optimal patent length, \(T_0\), as follows,
\[
T_0 = \frac{1}{r} \log \left[ \frac{2[1 - (c - \alpha)]}{\alpha} \right].
\]
It is straightforward to verify that \(T_0\) satisfies \(T_0 > T_0\) which is rearranged to yield \(1 - c > 0\) that holds. Similarly, \(T_0\) satisfies \(T_0 < \hat{T}_0\), as this inequality holds by definition.
Differentiating \(T_0\) with respect to \(\alpha\) and \(c\), we obtain that
\[
\frac{\partial T_0}{\partial \alpha} = -\frac{1 - c}{r \alpha [1 - (c - \alpha)]} < 0 \quad \text{and} \quad \frac{\partial T_0}{\partial c} = -\frac{1}{r [1 - (c - \alpha)]} < 0.
\]
Equilibrium investment. Inserting \(T_0 = \frac{1}{r} \log \left[ \frac{2[1 - (c - \alpha)]}{\alpha} \right]\) into \(x_0(T)\), equilibrium investment becomes
\[
x_0 = \frac{\alpha (1 - c) \left( 1 - \frac{\alpha}{2[1 - (c - \alpha)]} \right)}{\gamma r}
= \frac{\alpha (1 - c) \left[ \alpha + 2 (1 - c) \right]}{2 \gamma r [1 - (c - \alpha)]}
\]
which satisfies
\[
\frac{\partial x_0}{\partial \alpha} = \frac{(1 - c) \left[ \alpha^2 + 2 (1 - c) (1 - (c - \alpha)) \right]}{2 \gamma r [1 - (c - \alpha)]^2} > 0
\]
\[
\frac{\partial x_0}{\partial c} = -\frac{\alpha \left[ \alpha^2 + 4 \alpha (1 - c) + 2 (1 - c)^2 \right]}{2 \gamma r [1 - (c - \alpha)]^2} < 0.
\]
6.7 Proof of Proposition 2
Using \(x_1(T)\) and \(S_1(T)\) from Appendix 1, the \(MDG_1(T)\) becomes
\[
MDG_1(T) = \frac{\partial x_1(T)}{\partial T} S_1(T)
= \frac{\alpha^2 [1 - (c - \alpha)] [\alpha + 2 (1 - c)] (1 + 4d) e^{-rT}}{2 \gamma r (1 + 2d)^3}.
\]
Similarly, the $MSL_1 (T)$ is
\[
MSL_1 (T) = x_1 (T) \left( \gamma \frac{\partial x_1 (T)}{\partial T} - \frac{\partial S_1 (T)}{\partial T} \right) = \frac{\alpha^2 [1 - (c - \alpha)]^2 e^{-rT} (1 - e^{-rT})}{\gamma r (1 + 2d)^2}.
\]

$MDG_1 (T)$ is monotonically decreasing in $T$ because
\[
\frac{\partial MDG_1 (T)}{\partial T} = -\frac{\alpha^2 [1 - (c - \alpha)] [\alpha + 2 (1 - c)] (1 + 4d) e^{-rT}}{2\gamma (1 + 2d)^3}
\]

However, $MSL_1 (T)$ is monotonically decreasing in $T$ if and only if $T > \frac{\log 2}{r}$, since
\[
\frac{\partial MSL_1 (T)}{\partial T} = \frac{\alpha^2 [1 - (c - \alpha)]^2 e^{-2rT} (2 - e^{-rT})}{\gamma r (1 + 2d)^2}
\]
is negative when $2 - e^{-rT} < 0$, which simplifies to $T > \frac{\log 2}{r}$. In addition,
\[
\frac{\partial MDG_1 (T)}{\partial d} = -\frac{\alpha^2 [2(1 - c) + \alpha] (1 - c + \alpha) (1 + 8d) e^{-rT}}{r\gamma (1 + 2d)^4} < 0
\]
and
\[
\frac{\partial MSL_1 (T)}{\partial d} = -\frac{4\alpha^2 [1 - (c - \alpha)]^2 e^{-2rT} (e^{rT} - 1)}{\gamma r (1 + 2d)^3} < 0.
\]

Setting $MDG_1 (T) > MSL_1 (T)$, we find
\[
\frac{\alpha^2 [1 - (c - \alpha)] [\alpha + 2 (1 - c)] (1 + 4d) e^{-rT}}{2\gamma r (1 + 2d)^3} > \frac{\alpha^2 [1 - (c - \alpha)]^2 e^{-rT} (1 - e^{-rT})}{\gamma r (1 + 2d)^2}
\]
which simplifies to
\[
[\alpha + 2 (1 - c)] (1 + 4d) > 2 [1 - (c - \alpha)] (1 + 2d) (1 - e^{-rT})
\]
and further rearranges to $e^{-rT} > \frac{2(1-c)(1-2d)+\alpha}{2[1-(c-\alpha)][1+2d]}$. The denominator of this ratio is unambiguously positive, while the numerator is unambiguously negative since $d > \frac{1}{2}$ and $\alpha < 1 - c$ by definition. Hence, $MDG_1 (T) > MSL_1 (T)$ holds for all admissible parameters.

References


