TCER Working Paper Series

R&D in Clean Technology: A Project Choice Model with Learning

Koki Oikawa

April 2015

Working Paper E-93
http://tcer.or.jp/wp/pdf/e93.pdf

TOKYO CENTER FOR ECONOMIC RESEARCH
1-7-10-703 Iidabashi, Chiyoda-ku, Tokyo 102-0072, Japan

Abstract

In this study, we investigate the qualitative and quantitative effects of an R&D subsidy for clean technology and a Pigouvian tax on a dirty technology on environmental R&D when it is uncertain how long the research takes to complete. The model is formulated as an optimal stopping problem, in which the number of successes required to complete the R&D project is finite and which incorporates learning about the probability of success. We show that the optimal R&D subsidy with the consideration of learning is higher than that without it. We also find that an R&D subsidy performs better than a Pigouvian tax unless the government can induce suppliers to make cost reduction efforts even after the new technology successfully replaces the old one. Moreover, by a two-project model, we show that a uniform subsidy is better than a selective subsidy.

Koki Oikawa
TCER
and
Waseda University
School of Social Sciences
1-6-1 Nishiwaseda Shinjuku-ku, Tokyo
oikawa.koki@gmail.com

R&D in Clean Technology: A Project Choice Model with Learning☆

Koki Oikawa[a,*], Shunsuke Managi[b]

[a] *School of Social Sciences, Waseda University. 1-6-1 Nishiwaseda Shinjuku-ku, Tokyo, Japan 169-8050*

[b] *Graduate School of Environmental Studies, Tohoku University*

Abstract

In this study, we investigate the qualitative and quantitative effects of an R&D subsidy for clean technology and a Pigouvian tax on a dirty technology on environmental R&D when it is uncertain how long the research takes to complete. The model is formulated as an optimal stopping problem, in which the number of successes required to complete the R&D project is finite and which incorporates learning about the probability of success. We show that the optimal R&D subsidy with the consideration of learning is higher than that without it. We also find that an R&D subsidy performs better than a Pigouvian tax unless the government can induce suppliers to make cost reduction efforts even after the new technology successfully replaces the old one. Moreover, by a two-project model, we show that a uniform subsidy is better than a selective subsidy.

JEL classification codes: D83, O33, Q55, Q58

Keywords: Environmental technology, Learning, R&D subsidy, Pigouvian tax.

☆We thank Hideki Konish, Takeshi Ogawa, and participants in the seminars and conferences at Hitotsubashi University, GRIPS, Tohoku University, JEA meeting, and SEEPS meeting. This research was supported by the Research Institute of Science and Technology for Society (RISTEX), Japan Science and Technology Agency (JST).

*Corresponding author. E-mail: oikawa.koki@gmail.com, Tel/Fax: +81-3-5286-1441.

1. Introduction

Because tackling global environmental problems such as climate change imposes significant costs on society, authors such as Nordhaus (2007) suggest a gradual approach to emissions reduction. Given that Acemoglu et al. (2012) claim that an immediate switch toward R&D related to clean technologies is important, the knowledge gained through R&D should be useful for designing optimal environmental policy. In this study, we develop a model in which a research institute conducts R&D on a new clean technology that aims to reduce costs, although the pace of progress is unknown. Before the new technology is developed, producers use the old technology, which generates external costs such as pollution. The central authority wants to replace the old (dirty) technology with the new (clean) technology. Our main question is how the optimal R&D policy is affected by learning-by-doing in research. Note that unlike in the literature on learning-by-doing in the production sector, in which costs are reduced by increasing cumulative output (cf. Popp et al. (2010)), we assume that learning occurs in the R&D sector.

To investigate the effect of learning, we formulate an R&D project as a step-by-step process. The research institute updates its belief about the potential of the current research project and makes a "go" or "no-go" decision after each trial to pass the current research step. Both private and social returns to R&D are only obtained when the research institute completes all steps, which results in the new clean technology being put to practical use and replacing the old technology. Because of incomplete information about the research's potential and the externalities associated with the old technology, the ex post optimal allocation of R&D investment is rarely achieved. To eliminate this inefficiency, at least to some degree, we consider the effects of two policies: an R&D investment subsidy and a Pigouvian tax on the old technology.[1] We show that the optimal R&D subsidy is higher when learning

[1] Many studies have examined environmental tax and subsidy policy. In the long run, an important criterion for judging environmental policies is the extent to which they promote the new clean technology (Kneese and Schulze (1975)). Some evidence from environmental economics research suggests that market-based instruments such as taxes are more effective at inducing technological change than are command-and-control instruments because they

by a researcher on the quality of its research project is taken into account, compared with the case without such learning. A Pigouvian tax on the dirty technology can also stimulate R&D activity by raising the prices of goods produced by the old technology; this reduces the extent to which the new substitute technology must be developed to be commercially viable. However, we show that a Pigouvian tax raises the probability of a bad project being carried out relative to granting an R&D subsidy. In other words, learning to find a better project is less efficient under a Pigouvian tax. However, the comparison of efficiency in terms of welfare across policies depends on the context, because a Pigouvian tax tends to induce the early adoption of the new technology.

The model presented herein is an optimal stopping problem in which R&D is completed after a fixed number of successes in Bernoulli trials and there is Bayesian-style learning about the unknown probability of success. Our study is conceptually close to that of Cowan (1991), who analyzes a model of technology adoption with increasing returns in an armed-bandit setting. He shows that although the central authority can improve the technology adoption process by internalizing the positive externality from increasing returns, it is still possible for technology adoption to converge to an inferior option. The distinction between our study and that of Cowan (1991) is threefold. First, we introduce a fixed goal for the trials, which is the completion of an R&D project. Neither the research institute nor society obtains any return from their R&D investment until the project is completed and the new technology is commercialized. Therefore, the remaining steps, or the expected duration of the R&D project, crucially affects decision making. As the number of remaining steps to completion decreases, it becomes less important to infer the superior option in terms of the probability of success at each step. Second, by considering two or more projects, we do not incorporate into our model the multi-armed bandit problem. This approach is taken because there are no instant rewards when an arm (i.e., a project in this study) is

provide incentives to avoid tax (Newell et al. (1999)). Laffont and Tirole (1996) and Jung et al. (1996) examine how the choice of policy instruments affects innovation.

successfully pulled and the best arm provides not the best pulls in perpetuity but rather the shortest expected time to enjoying a fixed payoff. Third, we define indices to capture the efficiency of learning and the social welfare loss to evaluate the examined policies quantitatively. The former is defined as the probability that the socially desired option is chosen and the latter is defined as the gap between the expected social net benefit obtained under a policy and the first-best outcome. Moreover, the parameters used in each simulation are chosen based on the body of research findings on this topic and interviews with representatives of the Toyota Motor Corporation about R&D into electric vehicles, an important part of R&D into clean automobiles.

Our model is also related to the project choice models presented in the industrial organization literature. Many of these studies incorporate strategic relationships among innovators and/or venture capitalists (see, for example, Dasgupta and Maskin (1987) and, more recently, Cardon and Sasaki (1998), Gerlach et al. (2005), and Bergemann and Hege (2005)). Because the focus of these studies is not the success or failure of particular R&D investments, R&D processes are appropriately simplified, typically to one-shot processes. On the contrary, we investigate R&D processes in more detail by focusing on a single decision maker's problem.[2] No previous studies have thus far developed a systematic framework for analyzing R&D choice with environmental policies.

The remainder of this paper is organized as follows. In Section 2, we present a model with one R&D project and policy interventions. In Section 3, we introduce learning efficiency and social loss to evaluate the policy impacts by simulations. In Section 4, we extend the model to incorporate two projects. In Section 5, we provide concluding remarks and discuss.

2. A Model with One R&D Project

Suppose that a mature product generates external costs or pollution as a byproduct of production. Hence, there is an incentive, especially for the

[2] In Appendix D, we show that the impact of learning in one-step R&D is quantitatively small, based on our parameter setting.

government, to develop environmentally friendly alternatives. For example, electric vehicles or fuel-cell vehicles are new alternatives to petrol-powered vehicles. Thermal power versus solar power is also applicable to this framework.

Suppose that the initial unit cost of the new clean technology is c_0, which is assumed to exceed the cost level required for marketability. A research institute may conduct R&D into this technology at a fixed cost of 1 per period. Through R&D investment in period t, c_{t+1} falls to $(1-\alpha)\,c_t$ with a probability of q, where $\alpha \in (0,1)$.

We assume that q is unknown. Hence, this R&D process is a sequence of Bernoulli trials with an unknown probability of success and a known goal. Let c^* be the target cost level at which the new clean technology can replace the old dirty technology in the market. To complete the project, the institute needs M successes, such that

$$(1-\alpha)^{M-1}\,c_0 > c^* \geq (1-\alpha)^{M}\,c_0$$

or

$$M = \left\lceil \frac{\ln c_0 - \ln c^*}{-\ln(1-\alpha)} \right\rceil, \tag{1}$$

where $\lceil x \rceil \equiv \min\{n \in \mathbb{Z} | x \leq n\}$.

For simplicity, we assume that the researcher does not obtain any return until c_t reaches c^*, and the new technology immediately prevails when c^* is attained. After M successes, the research institute and society obtain V and X, respectively; we assume $X > V$. The social return X can be decomposed into the external costs emitted by the old technology plus the private return, V, minus the lost profits of firms using the old technology. We assume that the research institute is risk neutral.

2.1. The R&D Decision with Learning

2.1.1. Prior and Posterior Distributions of q

We assume that the initial prior about q is uniform, $U[0,1]$, which is equivalent to a beta distribution, $Be(q|1,1)$. It is known that, starting with a beta distribution, the posterior distribution is also a beta distribution depending on the numbers of successes and failures.

Let n be the total number of trials and m the remaining steps. The posterior density function is

$$f(q|n,m,M) = \frac{q^{M-m}(1-q)^{n-M+m}}{B(M-m+1, n-M+m+1)},$$

where B is the beta function. For future reference, note that

$$E(q|n,m) = \frac{M-m+1}{n+2}. \tag{2}$$

2.1.2. The R&D Decision Rule

A state (n,m) is *feasible* if $(n,m) \in \mathbb{Z}_+ \times \mathbb{Z}_+$, $n \geq M-m$, and $m \leq M$. The research institute conducts R&D for a given feasible state (n,m) to maximize the expected net benefit from the project; this constitutes an optimal stopping problem. Let the cost of R&D in each trial be set at 1. The value of the project, $v(n,m)$, satisfies

$$v(n,m) = \begin{cases} \max\left\{0, -1+\beta\begin{bmatrix} E(q|n,m)\, v(n+1,m-1) \\ +(1-E(q|n,m))\, v(n+1,m) \end{bmatrix}\right\} & \text{for } m \geq 1, \\ V \quad \text{for } m = 0, \end{cases} \tag{3}$$

where β is a discount factor. Hence, for $m \geq 1$, the institute conducts R&D if

$$E(q|n,m)\, v(n+1,m-1) + (1-E(q|n,m))\, v(n+1,m) > \frac{1}{\beta}.$$

Let $\hat{n}(m)$ be the stopping trial in the sense that the research project is terminated when the total number of research trials reaches $\hat{n}(m)$ if the remaining steps to the goal equal m.

Proposition 1. *There uniquely exists an $\hat{n}(m)$ for each $m \in \{1, 2, ..., M\}$ such that*

$$v(n,m) > 0 \quad \Leftrightarrow \quad n < \hat{n}(m),$$

where

$$\hat{n}(m) = \min\{n \geq M-m \,|\, \beta E(q|n,m)\, v(n+1,m-1) \leq 1\}.$$

Moreover, $\hat{n}(m)$ decreases with m.

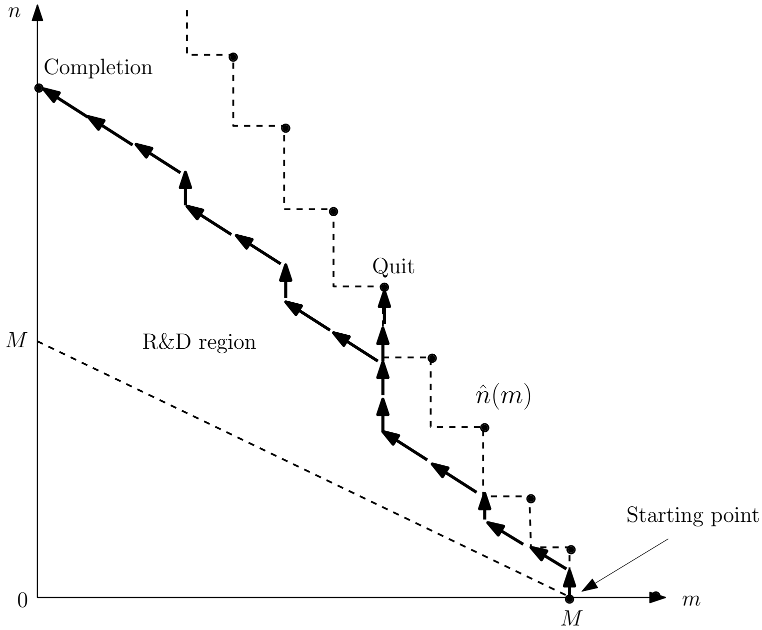


Figure 1: The R&D region and two sample paths. Starting at $(M, 0)$, both sample paths move to $(M, 1)$, which implies that the first research trial fails. However, because the path is still beneath $\hat{n}(M)$, the research institute continues R&D. In this example, the second trial is a "success" and thus it reaches $(M-1, 2)$. As long as the path is in the R&D region (the area below $\hat{n}(m)$), R&D continues. R&D stops if it goes off the R&D region ("quit") or reaches $m = 0$ ("completion"). The Paths never cross the line representing $m = M - n$ because $n < M - m$ is not feasible.

Intuitively, the researcher should stop the research project when it provides no expected profit even if the next trial ends in success. According to Proposition 1, the R&D decision is straightforward given the sequence of $\hat{n}(m)$ values. Appendix A provides a proof of Proposition 1 and shows how to compute $v(n, m)$, where $\hat{n}$ and v are decided sequentially.

Figure 1 shows the sample paths of the R&D project. The research institute continues the project if and only if the current pair (n, m) satisfies $n < \hat{n}(m)$. It terminates the project in two cases: one is "completion," when m reaches 0; the other is "quit," when n reaches $\hat{n}(m)$.

When the research institute pursues the social benefit of R&D, X, as if the government conducts the project, the above argument works by replacing V with X. This is the *second-best* case in the sense that the inefficiency stemming from the externality is resolved and the only inefficiency is derived from the incomplete information about the capability of the project.

2.2. The R&D Decision Without Learning

To evaluate the contribution of learning, we first check what happens if there is no learning. Even when the true probability of success is unknown, the solution without learning is simple. The expected value of the R&D project with m remaining steps, $v_{NL}(m)$, is

$$v_{NL}(m) = E\left(\beta^N V - \sum_{i=0}^{N-1} \beta^i \middle| m\right) \tag{4}$$

where N is the required number of periods until the completion of the R&D project. N is a random variable whose distribution is a negative binomial distribution with parameters m and q. We assume that the common belief about q follows a uniform distribution on $[0, 1]$.

(4) can be developed as

$$\begin{aligned} v_{NL}(m) &= \left(V + \frac{1}{1-\beta}\right) E\left(\beta^N \middle| m\right) - \frac{1}{1-\beta} \\ &= \left(V + \frac{1}{1-\beta}\right) \int_0^1 \left(\frac{\beta q}{1-\beta(1-q)}\right)^m dq - \frac{1}{1-\beta}, \end{aligned} \tag{5}$$

where we used the fact that

$$E\left(\beta^N | q, m\right) = \left(\frac{\beta q}{1-\beta(1-q)}\right)^m. \tag{6}$$

The researcher conducts R&D if and only if $v_{NL}(m) > 0$. Given that $v_{NL}(m)$ is strictly decreasing in m as long as $v_{NL}(m) > 0$, the researcher starts and continues the R&D project until completion if and only if $v_{NL}(M) > 0$.

2.3. The First-best Decision: Complete Information and the Pursuit of the Social Return of R&D

The second situation that should be compared with the decision-making process in Section 2.1 is the first-best case in which q is known and the return of R&D is set at $X > V$. For a given probability of success, $q \in [0, 1]$, the expected net social benefit of the R&D project is

$$E\left(\beta^N X - \sum_{i=0}^{N-1} \beta^i \middle| q, m\right) = \left(X + \frac{1}{1-\beta}\right) E\left(\beta^N | q, m\right) - \frac{1}{1-\beta}, \tag{7}$$

Thus, from (6), the institute should conduct the project if

$$\left(\frac{\beta q}{1-\beta(1-q)}\right)^m > \frac{1}{(1-\beta)X+1}.$$

Because the left-hand side of the above inequality is decreasing in m, the research institute starts and continues to carry out the project if and only if

$$q > \frac{1-\beta}{\beta} \frac{1}{\left[(1-\beta)X+1\right]^{\frac{1}{M}} - 1} \equiv \underline{q}. \tag{8}$$

We use $\underline{q}$ to define *learning efficiency* in Section 3, which indicates how likely the research institute is to choose the truly promising project. From (6)-(8), the first-best value of the R&D project for a given q is

$$v_{FB}(q) = \begin{cases} \left(X + \frac{1}{1-\beta}\right)\left(\frac{\beta q}{1-\beta(1-q)}\right)^M - \frac{1}{1-\beta} & \text{if } q > \underline{q}, \\ 0 \quad \text{otherwise.} \end{cases} \tag{9}$$

2.4. R&D Subsidy

In this subsection, we investigate how an R&D subsidy affects the R&D process. Let us denote $s \in [0, 1)$ as the R&D subsidy, with which the R&D cost in each trial becomes $1 - s$. Because a negative externality is generated by the old dirty technology, the government may have an incentive to promote R&D in the new clean technology. The government's decision depends on whether it considers learning in the research sector.

First, we consider the case with the learning process described in Section 2.1. Let us denote $v(n, m\,|V, s)$ as the value function when the return is V and the subsidy is s. Then, we obtain the following proposition and corollary.

Proposition 2. *Assume $V < X$ and $s \in [0,1)$.*

$$\begin{aligned} v\left(n, m \left| \frac{V}{1-s}, 0\right.\right) &= \frac{1}{1-s} v\left(n, m \left|V, s\right.\right), \\ \hat{n}\left(m \left| \frac{V}{1-s}, 0\right.\right) &= \hat{n}\left(m \left|V, s\right.\right). \end{aligned}$$

Corollary 1. *An R&D subsidy induces the second-best R&D decision by setting $s = s^* \equiv 1 - V/X$.*

The proof of Proposition 2 is in Appendix B. The proof of its corollary is straightforward. The R&D decision under a subsidy of s is identical to that under no subsidy if we replace the return from the completion of R&D with $\frac{V}{1-s}$. As a result, by setting $s^* = 1 - V/X$, the government can make the private R&D decision mimic the one that would be taken by a research institute pursuing social benefit.

Proposition 2 also implies that $\hat{n}\left(m \left|V, s_2\right.\right) \geq \hat{n}\left(m \left|V, s_1\right.\right)$ if $s_2 > s_1$ because $v\left(n, m \left|V, 0\right.\right)$ increases with V. Figure 2 shows the effect of the R&D subsidy for a given m. An increase in the R&D subsidy shifts v upwards and increases the stopping trial $\hat{n}$. As a result, the research institute is more likely to conduct and complete the R&D project under a positive subsidy.

Now, we compare the optimal subsidies with and without learning. Let s_{NL} be the optimal subsidy without learning. The private expected net benefit under subsidy s in the no-learning case is

$$v_{NL}^s(m|s) = \left(V + \frac{1-s}{1-\beta}\right) E\left(\beta^N \middle| m\right) - \frac{1-s}{1-\beta}. \tag{10}$$

Obviously, $v_{NL}^s(m|0) = v_{NL}(m)$.

If the government wants a private firm to conduct an R&D project without any consideration of learning about project quality, it is sufficient to make $v_{NL}^s(M|s) > 0$. Therefore, the optimal subsidy without learning is $s_{NL} = 0$ if $v_{NL}(M) > 0$. $s_{NL} > 0$ only when $v_{NL}(M) \leq 0$ and the project is socially worthwhile. Since s^* makes the private expected benefit meet the social expected benefit, $v_{NL}^s(M|s^*) > 0$ must hold if the project is promising from

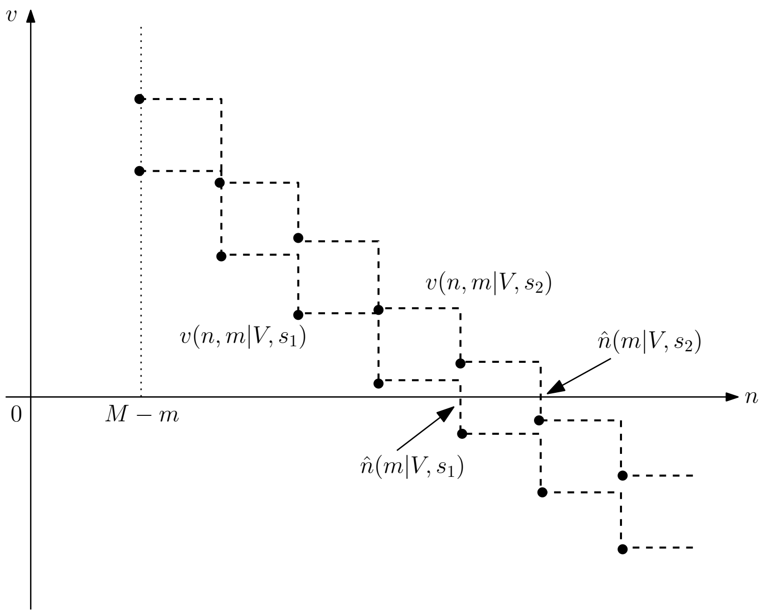


Figure 2: The effect of an R&D subsidy.

the social point of view. Therefore, there exists some $s_{NL} < s^*$ to satisfy $v^s_{NL}(M|s_{NL}) > 0$.[3] This argument leads to the next proposition.

Proposition 3. *The optimal subsidy when the learning process is taken into account, s^*, is greater than that without consideration of learning, s_{NL}.*

2.5. A Pigouvian Tax on the Dirty Technology

Suppose that the government imposes a Pigouvian tax on production based on the old technology. (A Pigouvian subsidy on the new technology has the equivalent effect.) Although the primary purpose of a Pigouvian tax is to reduce external costs, it influences on the R&D decision through its effect on the target cost level. Because the production cost of using the old

[3]From (10), the required subsidy to make $v^s_{NL}(m|s) > 0$ is that slightly higher than

$$1 - \frac{E(\beta^N|M)}{1 - E(\beta^N|M)}(1 - \beta)V.$$

technology is raised by the tax, the new technology can be commercialized despite its higher production cost if the impact of the tax is sufficiently large.

Let $(1+\tau)\,c^*$ be the new target cost after the Pigouvian tax is imposed ($\tau \geq 0$) and let g_τ be the change in the required number of successes before commercialization. From modification of (1),

$$g_\tau \equiv M - \left\lceil \frac{\ln c_0 - \ln (1+\tau)\, c^*}{-\ln (1-\alpha)} \right\rceil . \qquad (11)$$

The Pigouvian tax may bring the research institute closer to achieving its goal according to (11). Given τ, the value function (3) is redefined as

$$v\,(n, m|\tau) = \begin{cases} \max \left\{ 0, -1 + \beta \begin{bmatrix} E\,(q|n,m)\, v\,(n+1, m-1|\tau) \\ +\,(1 - E\,(q|n,m))\, v\,(n+1, m|\tau) \end{bmatrix} \right\} & \text{for } m \geq g_\tau + 1, \\ V \qquad \text{for } m = g_\tau. \end{cases}$$

As shown in Proposition 4, the Pigouvian tax may also induce private R&D, similar to the R&D subsidy policy.

Proposition 4. *For $\tau, \tau' \in [0, c_0/c^* - 1)$ and $\tau > \tau'$,*

$$\hat{n}\,(m|\tau) \geq \hat{n}\,(m|\tau') \qquad \forall m \in [g_\tau + 1, M]\,.$$

The proof is in Appendix C.

3. Simulations of Policy Impacts

3.1. Learning Efficiency

The literature on learning considers the probability that the learning process converges to the best choice as the process is repeated infinitely many times (e.g., Cowan (1991)). We similarly define learning efficiency, LE, in terms of the probability that the choice of project converges to the one that would be made under complete information and from the social point of view. Under complete information, the socially best strategy is that R&D continues to completion if and only if $q > \underline{q}$. Thus, LE under a policy of granting a subsidy of s for a given q is

$$LE\left(s|q\right)=\begin{cases}\Pr\left(\text{“completion”}|q,s\right) & \text{if } q>\underline{q},\\ \Pr\left(\text{“quit”}|q,s\right) & \text{if } q\leq\underline{q}.\end{cases}$$

To evaluate a policy (ex ante) in terms of learning efficiency, we take the expectations of $LE(s|q)$ over q, for which we use $q\sim U\left[0,1\right]$, the same as the initial prior.

$$LE\left(s\right)=\int_{0}^{\underline{q}}\Pr\left(\text{“quit”}|q,s\right)dq+\int_{\underline{q}}^{1}\Pr\left(\text{“completion”}|q,s\right)dq. \tag{12}$$

Clearly, if the research institute knows q and pursues the social benefit, as described in Section 2.3, $LE=1$ for any s. Deviation from complete information, the social viewpoint or both lowers learning efficiency to below unity. The minimum value of LE is attained when the private research institute does not update its belief, as in Section 2.2. In this no-learning case, learning efficiency, say LE_{NL}, is

$$LE_{NL}(s)=\begin{cases}1-\underline{q} & \text{if } v_{NL}\left(M|s\right)>0,\\ \underline{q} & \text{otherwise.}\end{cases}$$

The gap between LE and LE_{NL} indicates the degree to which the probability of making the right choice is improved by learning about the capability of the project.

3.2. Simulation

In this subsection, we investigate the policy effects by simulations. The parameters used in each simulation are chosen based on previous research and interviews with representatives of the Toyota Motor Corporation about R&D into electric vehicles.[4]

[4]Ito and Managi (2015) study the cost-effective strategy for reducing CO_2 emissions from the transport sector. They analyzes conditions of clean vehicles to be replaced for old vehicles (see also Managi (2012)).

3.2.1. Parameter Settings

Initial cost and target cost. We set c_0 at the initial unit cost of electric vehicles and the target cost level, c^*, at the current price of a typical small car. Then, $c_0 = 460$ and $c^* = 150$ (in ten thousand Japanese yen). Both capture the full cost of production. Note that only the ratio c_0/c^* matters in the decision-making process.

Cost reduction rate. The rate at which costs are reduced has been investigated in the literature on learning curves. The surveys by Dutton and Thomas (1984) and Popp et al. (2010) indicate a progress ratio or learning rate of about 20% in many industries; the progress ratio is defined as the percentage reduction in unit cost following a doubling of cumulative output. More specifically, if the cost curve is $n^{-b}c_0$, where n is the cumulative amount and b is a positive parameter, the progress ratio is defined as $1 - (2n)^{-b}/n^{-b} = 1 - 2^{-b}$. The engine of cost reduction in the main body of studies on learning-by-doing does not easily fit our R&D analysis. However, Söderholm and Sundqvist (2007) examine the learning curve with both cumulative outputs and R&D expenditure in the wind energy industry and find that 15% of the progress ratio comes from R&D expenditure (with 5% coming from typical learning-by-doing). Therefore, we assume that the progress ratio related to R&D is 15% and thus the corresponding b is about 0.23.

We set the cost reduction rate, α, at a value that is consistent with the standard progress ratio presented above. One problem with this approach is that our cost curve is not isoelastic such as $n^{-b}c_0$ but exponential on average for a given q: $(1-\alpha q)^n c_0$ from the binomial theorem. To bridge this gap, we search αq which equals the average of

$$1 - e^{-b\frac{\log n}{n}},$$

to have $n^{-b} \approx (1 - \alpha q)^n$ on average. We choose the coverage of n as 1 to 15 because the analysis in Söderholm and Sundqvist (2007) ranges up to 14 years, depending on the country. Then, we find αq being about 0.05. Since the ex-ante average of q is 0.5 under our assumption, we set $\alpha = 0.1$.

R&D reward and external costs. In the above model, we set the per-period cost of an R&D project as 1 because the decision problem, given by (3), is equivalent for multiplication by any positive constant. We keep this normalization in our numerical analyses.

The private reward for R&D, V, should be the expected discounted sum of profit flows (divided by the per-period cost of an R&D project). However, it is hard to choose a number that represents the expected sales volume of electric vehicles. Rather, we estimate the per-unit expected profits of electric vehicles and use this as the proxy for V.

We estimate the per-unit expected profits based on the interview as follows. First, we calculate the profit rate for a typical small car using the unit profit and manufacturing cost, which is about 30%. Then we apply this profit rate to the expected unit manufacturing cost of electric vehicles. Because the cost of batteries in electric vehicles cannot be determined exactly, we use two estimated values of V, $\{48.4, 77.8\}$ (in ten thousand Japanese yen), which represent low and high battery costs.

Social benefit, X, must consist of the negative externality derived from using the old technology. Because it is difficult to quantify this externality precisely, we simply set $X = 100$. Our results are robust to other values of X that exceed both values of V.

In summary, we use two parameter sets that share $c_0 = 460$, $c^* = 150$, $\alpha = 0.1$, and $X = 100$ in common, but have distinct V: $\{48.4, 77.8\}$. Under these parameter values, $M = 11$, $\underline{q} = 0.298$, and $v_{NL}(M) > 0$. Hence, the best policy when the government ignores learning in the private sector is $s = 0$ and $\tau = 0$. Further, LE in the case of no-learning is constant over s and τ.

3.2.2. R&D subsidy

An R&D subsidy increases the number of research trials and raises the completion rate. When the true value of q is below the social threshold $\underline{q}$, although the project is not worth the expense, the subsidy policy nevertheless inefficiently encourages researchers to pursue the R&D project. However, the quantitative significance of this source of inefficiency depends on the situation.

Figure 3 shows LE across R&D subsidies.[5] First, learning substantially increases the probability of making the right choice: even without a subsidy, LE surges by 12% and 17% for the lower and higher values of V, respectively, compared with the no-learning case. Second, LE is nonmonotonic and there is a maximum at some subsidy. Because an R&D subsidy stimulates research activity (i.e., $\hat{n}(m|s)$ is nondecreasing in s for any m), there is a surge in LE if, without any policy, the current project is highly likely to be terminated before completion even though the true q is high. This occurs primarily when V is low. On the contrary, LE declines when a subsidy facilitates project completion even though the true q is low. The impact of the R&D subsidy is determined by the net effect of this tradeoff. In the current simulation, there are two peaks of LE for each V. For each V, the LE-maximizing subsidies locate across the optimal subsidy, $1 - V/X$ as in Corollary 1.[6] The maximum values should be the same because the problem with V and s is identical to the problem with $\frac{V}{1-s}$ and $s = 0$ (see Proposition 2). The optimal R&D subsidy policy raises the probability of making the right choice by 7.1% under the lower V and 1.7% under the higher V.

To see the impact of learning from the alternative point of view, we define another efficiency measure that considers social loss relative to the first-best outcome, defined in (9). We define social loss, SL, as the average difference between the first-best value and the present-discounted value of the social net benefit of the R&D process computed from the simulation. Social loss

[5]Discrete jumps in LE come from discrete changes in the cutoff number of trials, $\hat{n}(m|s)$. Those changes vary across m for the same shift in s, while $\hat{n}$ is more frequently updated for larger m. The R&D decision with a large m tends to occur when the project is less promising. In such a case, the subsidy tends to inefficiently encourage incapable projects and reduce LE.

[6]More precisely, the subsidy of 0.430 and 0.576 achieves almost the same LE, 0.8916, under $V = 48.4$; and those of 0.090 and 0.312 achieves the same level of LE under $V = 77.8$.

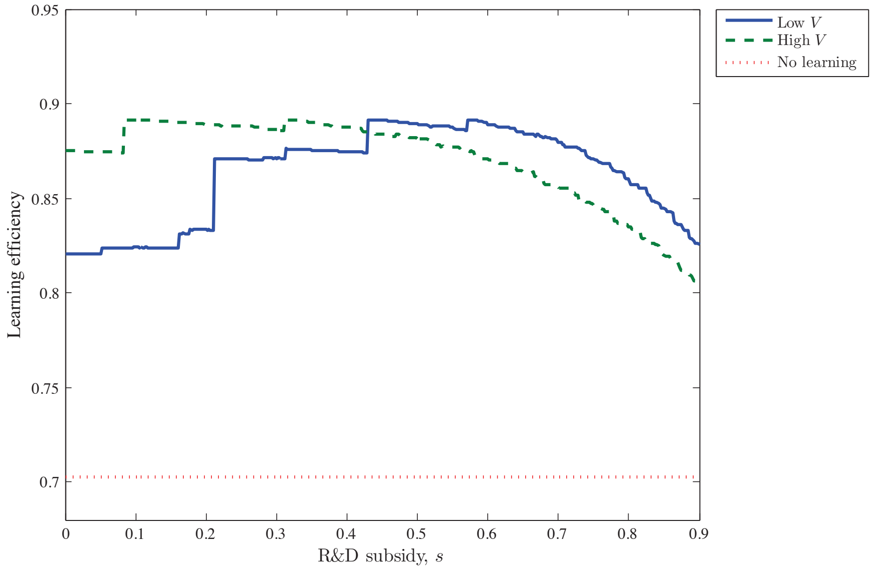


Figure 3: Learning efficiency with and without learning across R&D subsidies. The number of draws is 10^8 (we need this number of draws because there are two types of randomness: one is the true value of the probability of success q, which we assume to be uniform on $[0, 1]$; the other is the sample paths given q.)

under a policy of granting a subsidy of s for a given q, $SL(s|q)$, is[7]

$$SL(s|q) = v_{FB}(q) - \begin{cases} \beta^N X - \frac{1-\beta^N}{1-\beta} & \text{if completed,} \\ -\frac{1-\beta^N}{1-\beta} & \text{if quit after } N \text{ trials,} \end{cases} \tag{13}$$

where N is obtained from the simulation. We compute $SL(s)$ as the mean of $SL(s|q)$ over $q \sim U[0, 1]$.

When we ignore learning, social loss becomes

$$SL_{NL}(s) = \begin{cases} E[v_{FB}(q)] - \left[\left(X + \frac{1}{1-\beta}\right) \int_0^1 \left(\frac{\beta q}{1-\beta(1-q)}\right)^M dq - \frac{1}{1-\beta}\right] & \text{if } v^s_{NL}(M|s) > 0, \\ E[v_{FB}(q)] & \text{otherwise.} \end{cases}$$

Figure 4 depicts the simulation outcome. Because of incomplete infor-

[7] In the case of an R&D subsidy, the research institute incurs an R&D cost of only $1 - s$. However, because the government pays s for each trial, the social cost of each trial is 1.

mation, SL cannot reach zero even when the external cost is internalized by the optimal subsidy in Corollary 1. Compared with the no-learning case, learning reduces social loss by 16.0% under the lower V and 7.6% under the higher V without any subsidy. Further, compared with the case in which the research institute learns the actual capability of the ongoing project under $s = 0$, the optimal R&D subsidy reduce social loss by 10.5% under the lower V and 1.5% under the higher V. Thus, the impact of an R&D subsidy in terms of welfare is significant especially when the gap between the social and private benefits is large.

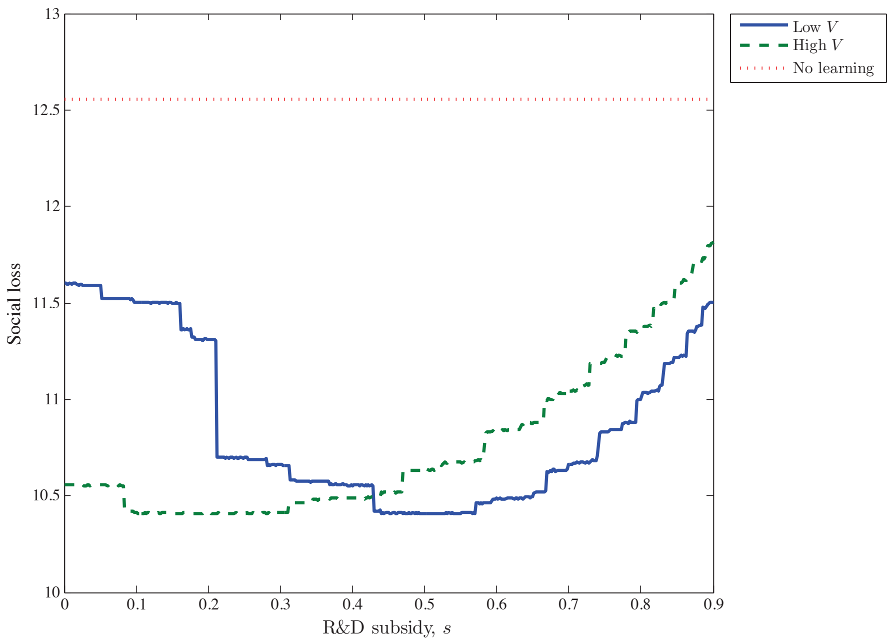


Figure 4: The effect of an R&D subsidy on social loss.

3.2.3. Pigouvian Tax on the Old Technology

Figure 5 shows the simulated result for LE under the imposition of a Pigouvian tax on the old technology. The dotted line indicates the maximum LE achieved under R&D subsidies. Although the Pigouvian tax stimulates R&D in the new technology, too much intervention reduces LE and, hence, some τ maximizes LE. However, it is impossible to exceed the maximum LE

achieved through an R&D subsidy. Because a Pigouvian tax increases the unit cost of using the old technology, fewer successes are required to replace the old technology with the new one. Then, even if the research institute estimates the true q to be low, it tends to be beneficial for the researcher to continue the project. Consider a simple example. Suppose that $V = 50$ and $m = 2$. Further, assume $\beta = 1$ for simplicity. To make the project promising, we need at least $50q^2 > 2$, or $q > 0.2$. Now, suppose that a Pigouvian tax reduces m to 1. Then, the minimum level of q required for the project to be worthwhile is only 0.02 ($50q > 1$). Such a Pigouvian tax lowers the threshold level of q 10-fold. In other words, a Pigouvian tax makes learning about project quality less important.

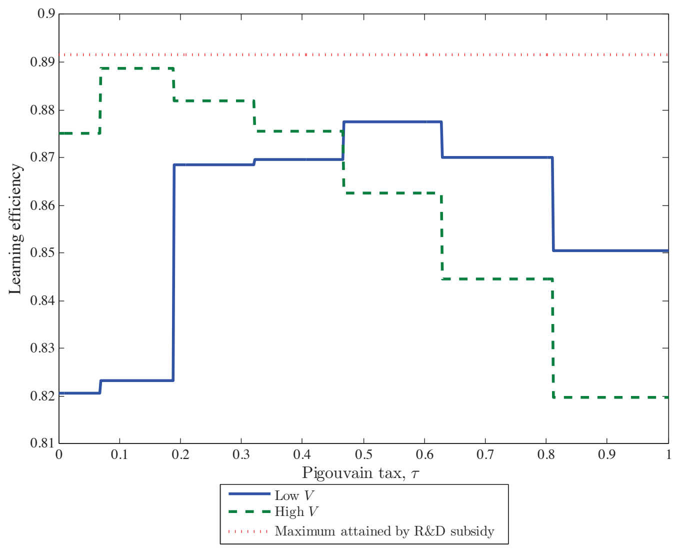


Figure 5: Learning efficiency under Pigouvian tax.

Social loss under a Pigouvian tax. We need to extend the model to examine the behavior of social loss over Pigouvian taxes because we only have the benefit of the early resolution of the negative externality in the above model. Here, we introduce the welfare cost of the Pigouvian tax and consider two cases.

The welfare cost stems from the loss in social surplus caused by a higher marginal cost because smaller cost-reduction efforts are required for the commercialization of the new technology under a Pigouvian tax. We first examine social loss with assuming that the higher marginal cost as a result of a Pigouvian tax remains in the market forever. Then, we examine how the result changes if the R&D project is continued even after the commercialization. We consider the second case to see the welfare cost from choosing a bad project. If a bad project happen to succeed due to a Pigouvian tax, it tends to stop before achieving the original target cost and a higher marginal cost remains in the market like in the first case.

To introduce the loss in social surplus, we consider a simple monopoly market. The marginal cost curve is horizontal at the cost level after the completion of the R&D process. If there is no Pigouvian tax, it is $(1-\alpha)^M c_0 \leq c^*$. However, under a Pigouvian tax, it is $(1-\alpha)^{M-g_\tau} c_0$ from (11). This gap in marginal costs generates the gap in social surplus. Assuming that the decrease in total demand from a higher price caused by this cost gap is sufficiently small, we can consider the loss in social surplus to be nearly demand times the cost gap. Since we approximate the social benefit, X, in per-unit term, it is natural to consider the social benefit after the completion of the R&D project under a Pigouvian tax as

$$X - \left(\frac{1}{(1-\alpha)^{g_\tau}} - 1 \right) (1-\alpha)^M c_0. \tag{14}$$

The upper panel of Figure 6 shows the social loss when this loss in social surplus is taken into account. SL is defined similar to (13), except that X is replaced with (14). The horizontal dotted line indicates the minimum social loss attained by the optimal R&D subsidy, s^*. Under the current set of parameters, the social loss from a Pigouvian tax is always higher than that under the optimal subsidy and is getting worse as τ increases.

This result holds only if the cost gap caused by a Pigouvian tax remains constant forever. Suppose that cost-reduction efforts are made even after the new technology dominates the market because, for example, the government may do away with the Pigouvian tax policy in the future. Then the researchers hired by the producer holding the new technology continue to conduct the same project, aiming to reach the original goal, c^*. Thus, the loss in social surplus, the second term in (14), decreases gradually as the project progresses. If the project is aborted as a result of learning, the loss in social surplus does not decline any more. If a project with a low q has been continued thanks to a Pigouvian tax, it will be aborted and end in an insufficient cost reduction with a relatively high probability. When this impact is significant, the result should be similar to that in the first experiment.[8]

The bottom panel in Figure 6 illustrates that the second experiment is different from the first one. A Pigouvian tax can lower the social loss below the minimum level attained by the R&D subsidy. Hence, under our parameter setting, the early resolution of the negative externality dominates if R&D is continued after the commercialization of the new technology. The loss from making a wrong choice does not significantly matter because projects are already filtered when it achieves the target cost level under a Pigouvian tax, at least for small τ. Hence, the superiority of R&D subsidy depends on whether the government can induce suppliers to make cost reduction efforts even after the new technology successfully replaces the old one. We present this advantage of Pigouvian taxes depends on the number of projects in Section 4.2.

4. A Model with Two R&D Projects

So far, we have assumed only one R&D project. In this section, we extend the above-presented model to incorporate two projects, labeled 1 and 2. Both projects create clean technologies, and the completed one replaces the old dirty technology. We assume that the unknown probabilities of success, q_1 and q_2, differ and are independent. Further, no cost is incurred to

[8] To run this experiment, we decompose X into a per-period term. We define $x(g)$ as the per-period social benefit when the number of remaining steps of cost reduction after the commercialization of the new technology is g, such as

$$x(g) = (1-\beta)\left[X - \left((1-\alpha)^{-g} - 1\right)(1-\alpha)^M c_0\right].$$

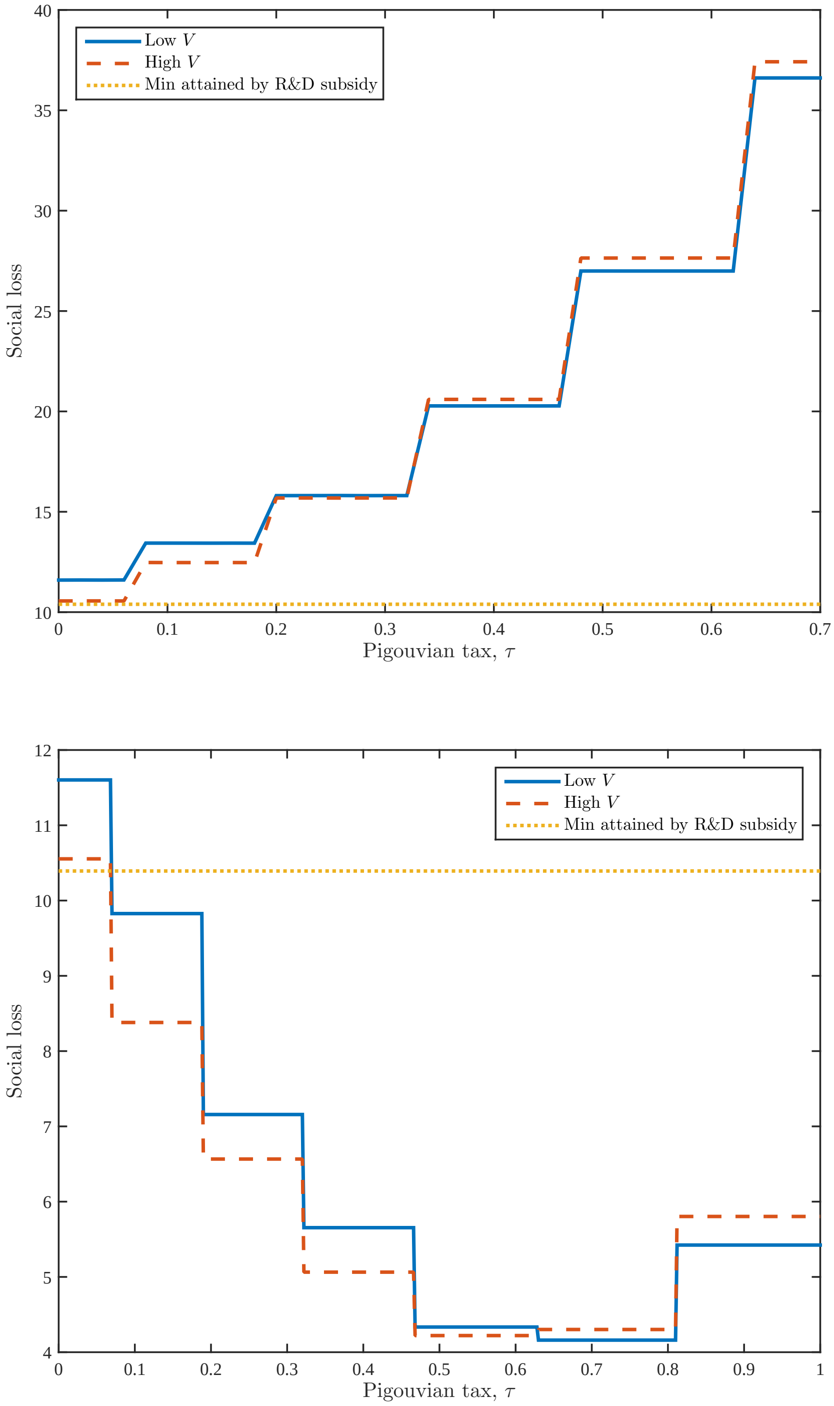


Figure 6: Social loss over Pigouvian taxes.

switch projects. Let c_{10} and c_{20} be the initial costs, and M_1 and M_2 be the corresponding steps required for commercialization.

4.1. The R&D Decision

This extension does not simply generate a two-armed bandit problem because the value of project 1 depends on the state of project 2 since the shortest expected time to stopping matters. Specifically, if one of the projects is completed, there is no reason to conduct the other: $v(n_1, m_1|n_2, 0) = 0$ even if $v(n_1, m_1|n_2, m_2) > 0$ for $m_2 > 0$.[9] Therefore, we cannot divide the two-project model into two independent one-project models.

The value function for state (n_1, m_1, n_2, m_2) is

$$v(n_1, m_1, n_2, m_2) = \begin{cases} \max\{v_1(n_1, m_1, n_2, m_2), v_2(n_1, m_1, n_2, m_2)\} & \text{if } \min\{m_1, m_2\} > 0, \\ V & \text{if } \min\{m_1, m_2\} = 0, \end{cases}$$

where

$$v_1(n_1, m_1, n_2, m_2) = \max\left\{0, -1 + \beta \left[\begin{array}{l} E(q_1|n_1, m_1)\, v(n_1+1, m_1-1, n_2, m_2) \\ +(1 - E(q_1|n_1, m_1))\, v(n_1+1, m_1, n_2, m_2) \end{array}\right]\right\},$$

and

$$v_2(n_1, m_1, n_2, m_2) = \max\left\{0, -1 + \beta \left[\begin{array}{l} E(q_2|n_2, m_2)\, v(n_1, m_1, n_2+1, m_2-1) \\ +(1 - E(q_2|n_2, m_2))\, v(n_1, m_1, n_2+1, m_2) \end{array}\right]\right\}.$$

Given the state (n_1, m_1, n_2, m_2), the research institute has three options: conduct no R&D, conduct project 1, or conduct project 2. $v_1(n_1, m_1, n_2, m_2)$ is the value of the state when the researcher undertakes project 1. $v_2(n_1, m_1, n_2, m_2)$ is its counterpart for project 2. The research institute stops R&D if $v(n_1, m_1, n_2, m_2) = 0$. Otherwise, it conducts project 1 (2) if $v_1(n_1, m_1, n_2, m_2) > (<) v_2(n_1, m_1, n_2, m_2)$.

[9] See Gittins et al. (2011) for details of multi-armed bandit problems.

We assume that when $v_1(n_1, m_1, n_2, m_2) = v_2(n_1, m_1, n_2, m_2) > 0$, one of the projects is randomly chosen.

We can compute these value functions backwardly. Leaving the details to Appendix E, here we sketch the algorithm briefly. Our calculation starts from state $(\hat{n}_1(1) - 1, 1, \hat{n}_2(1) - 1, 1)$, where $\hat{n}_i(1)$ is the stopping trial of project i when $m_i = 1$, defined in the one-project model (Proposition 1). In this state, one success in either project results in the final reward and one failure in project i results in project i becoming worthless and thereafter the problem reduces to the one-project model with project $j \neq i$. More clearly, v_1 (v_2 is analogous) at this state is

$$\begin{aligned} & v_1(\hat{n}_1(1) - 1, 1, \hat{n}_2(1) - 1, 1) \\ & = \max \left\{ 0, -1 + \beta \left[\begin{array}{l} E(q_1 | n_1, m_1) V \\ + (1 - E(q_1 | n_1, m_1)) \hat{v}_2(n_2, m_2) \end{array} \right] \right\}, \end{aligned}$$

where $\hat{v}_2$ indicates the value of project 2 in the one-project model, defined in (3). Then, we calculate the v values for all feasible states (n_1, m_1, n_2, m_2) by descending in n_i and ascending in m_i, sequentially, until coming at state $(0, M_1, 0, M_2)$.

4.2. R&D Subsidy and Pigouvian Tax

Learning Efficiency. Because the logic in the one-project model applies, both an R&D subsidy and a Pigouvian tax stimulate R&D in the new technology. The new feature of the two-project model is that the Pigouvian tax may change the project choice, whereas the R&D subsidy does not. To see this clearly, consider the following extreme example. Suppose that $m_2 < m_1$ but $v_1 > v_2$, which occurs when $E(q_2 | n_2, m_2)$ is sufficiently small relative to $E(q_1 | n_1, m_1)$. However, if the Pigouvian tax leads to $g_{2,\tau} \geq m_2$ ($g_{i,\tau}$ is defined by (11) for $M = M_i$), then project 1 is never chosen. Imposing a Pigouvian tax brings the goal closer, and its impact depends on the original distance from the goal. An R&D subsidy does not interfere in project selection because it increases the values of the states at the same rate, $1/(1-s)$.

Before moving to the simulation, we must redefine the efficiency measures for the two-project model. For simplicity, we consider the symmetric case,

$c_{10} = c_{20}$, so that $M_1 = M_2 = M$ and $\underline{q}_1 = \underline{q}_2 = \underline{q}$. (It is straightforward to extend to the case in which $c_{10} \neq c_{20}$.) Pointwise LE at (q_1, q_2) is

$$LE\left(s|q_1,q_2\right) = \begin{cases} \Pr\left(\text{“quit”}\right) & \text{if } \max\left\{q_1,q_2\right\} \leq \underline{q}_X, \\ \Pr\left(\text{“completion with 1”}\right) & \text{if } q_1 > \max\left\{\underline{q}, q_2\right\}, \\ \Pr\left(\text{“completion with 2”}\right) & \text{if } q_2 > \max\left\{\underline{q}, q_1\right\}. \end{cases}$$

Hence, the unconditional LE is

$$\begin{aligned} LE\left(s\right) &= \int_0^{\underline{q}} \left[\begin{array}{c} \int_0^{\underline{q}} \Pr\left(\text{“quit”}|q_1,q_2,s\right) dq_2 \\ + \int_{\underline{q}}^1 \Pr\left(\text{“completion with 2”}|q_1,q_2,s\right) dq_2 \end{array} \right] dq_1 \\ &+ \int_{\underline{q}}^1 \left[\begin{array}{c} \int_0^{q_1} \Pr\left(\text{“completion with 1”}|q_1,q_2,s,\tau\right) dq_2 \\ + \int_{q_1}^1 \Pr\left(\text{“completion with 2”}|q_1,q_2,s,\tau\right) dq_2 \end{array} \right] dq_1. \end{aligned}$$

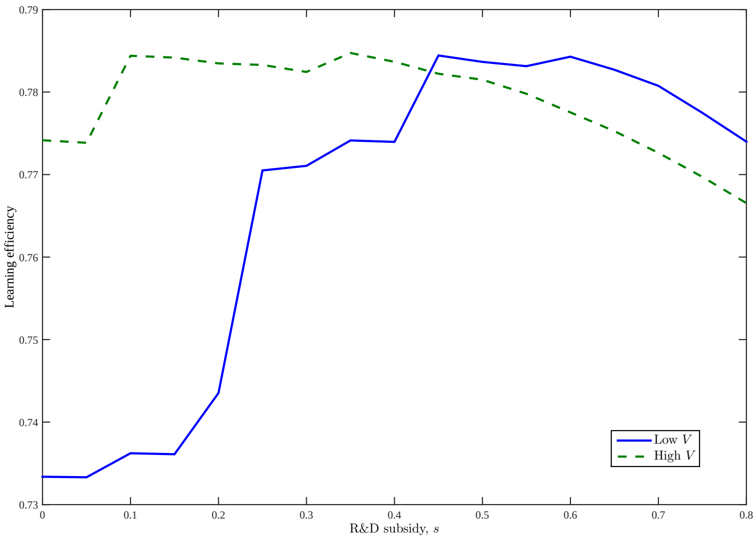
The top panel of Figure 7 illustrates the simulated impacts of R&D subsidies on LE for the two-project model. The two projects are identical except for their true values of q, and the parameters are the same as those used in the previous section. The best R&D subsidy in terms of learning efficiency is close to that shown in the one-project model. The bottom panel of Figure 7 shows LE under Pigouvian taxes, compared with LE attained by the R&D subsidy at best. The maximum LE under a Pigouvian tax is below that shown in the one-project model, which is consistent to the above argument.

Social Loss. Social loss in the two-project model is defined as follows. For each pair (q_1, q_2), the first-best outcome is

$$v_{FB}\left(q_1,q_2\right) = \begin{cases} \left(X + \frac{1}{1-\beta}\right)\left(\frac{\beta\bar{q}}{1-\beta(1-\bar{q})}\right)^M - \frac{1}{1-\beta} & \text{if } \bar{q} \equiv \max\{q_1,q_2\} > \underline{q}, \\ 0 & \text{otherwise.} \end{cases}$$

Since we assume q_i independently follows the standard uniform distribution, $\bar{q}$ follows a beta distribution, $Beta(2,1)$. Then, social loss with two projects is analogously defined as in the one-project model.

Figure 8 is the simulation result for SL over R&D subsidies. Social loss is globally smaller than that in the one-project model because the research institute has multiple options whose qualities are statistically independent.



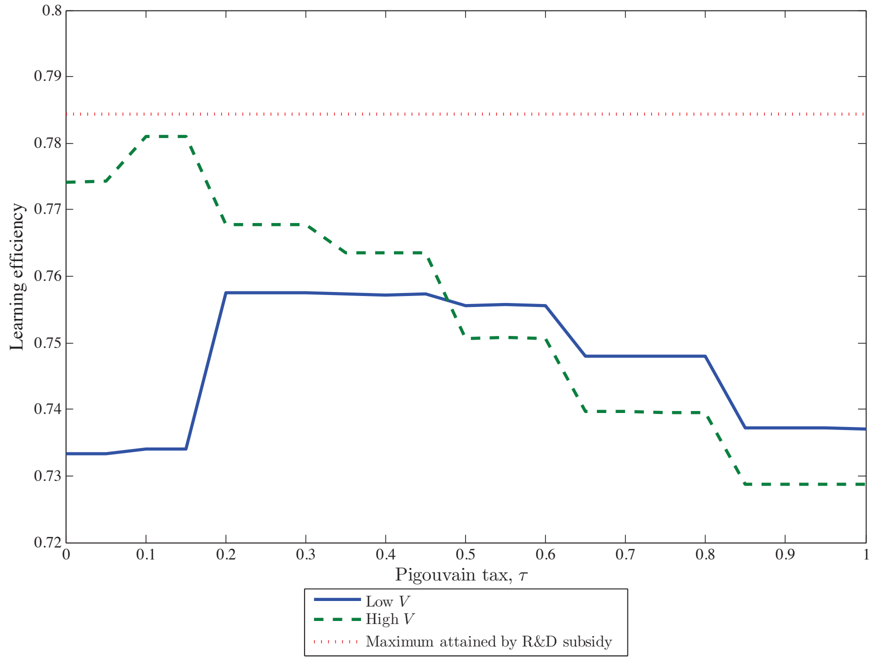


Figure 7: The effects on learning efficiency in the two-project model. The top panel is for R&D subsidies and the bottom panel is for Pigouvian taxes (the dotted horizontal line indicates the maximum level of LE attained by the R&D subsidy). The parameters are the same as those used in the one-project model and the two projects are identical except for their true values of q.

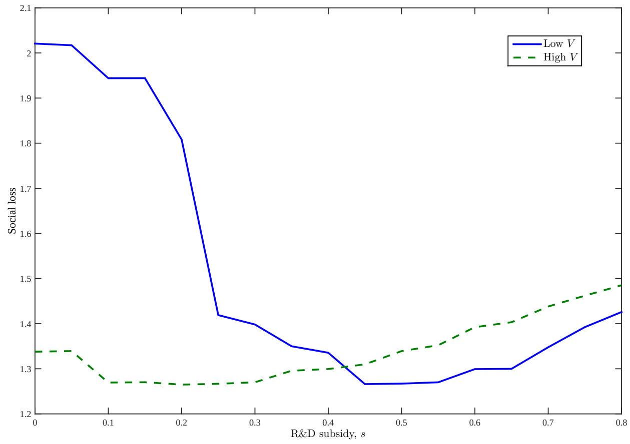


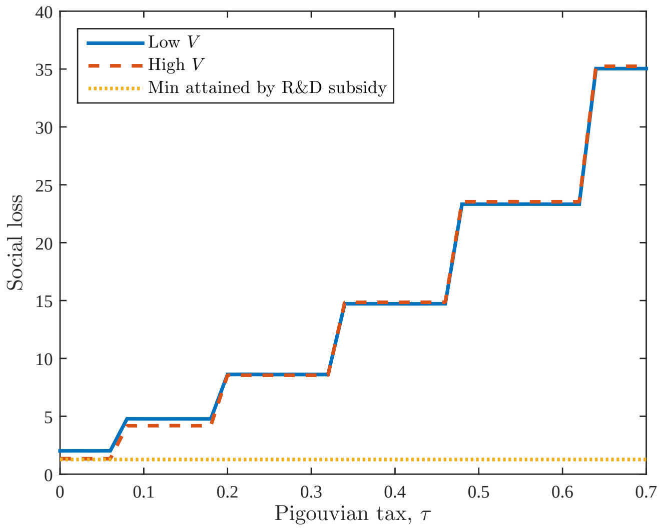
Figure 8: The effect of an R&D subsidy on social loss in the two-project model.

Other than the level difference, the behaviors of SL are similar to the basic model.

Figure 9 illustrates SL under Pigouvian taxes for the two cases, which are described in Section 3.2.3. The top panel shows SL when R&D stops at the commercialization of the new technology and a high marginal cost remains forever. The bottom panel is when R&D continues even after the commercialization, where we assume that the researcher does not switch projects after one of the two is commercialized. Relative to the one-project model, the advantage of the Pigouvian tax policy in the second case is not strong.

4.3. A Selective Subsidy

In this subsection, we present an application of the two-project model. So far, we have only considered an R&D subsidy policy that treats all projects equally, which we refer to as a *uniform subsidy*. In this case, the government does not interfere in the private research institute's project selection. In this subsection, we consider more active government intervention. That is, we consider a *selective subsidy*, under which the government offers preferential



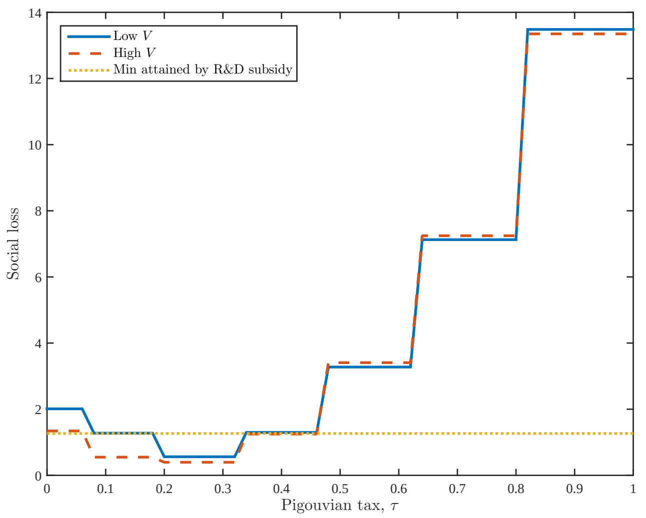


Figure 9: The effect of a Pigouvian tax on social loss in the two-project model. The top panel: R&D stops once the target is achieved. The bottom panel: R&D is continued even after commercialization but no switch over projects.

treatment to one of the two projects.[10]

Suppose that the government grants a subsidy of s to one of the projects at the beginning of period T. The target project is unchanged and the subsidization policy is maintained until the completion of the R&D project. Provided that the government has the same information as the research institute has, it should choose the target project based on $v\left(n_{1,T-1}, m_{1,T-1}, n_{2,T-1}, m_{2,T-1}\right)$.[11] Because the project with a higher value in period T incurs a lower private R&D cost thereafter, the selective subsidy policy tends to lock the research institute into its target project. This leads the institute to bias its information collection toward the project being conducted under the subsidy program, thus discouraging experimentation in the other research project. Therefore, LE is decreased by the selective subsidy relative to the level under a uniform subsidy program.

Figure 10 illustrates the deleterious effect of a selective subsidy with $T = 5$ and $V = 48.4$. The solid curve represents the relative LE under selective subsidies, defined as the ratio of LE under a selective subsidy to that under a uniform subsidy. The dashed curve represents the relative SL defined analogously. Although a selective subsidy is better than no policy, a large gap between the two policy programs exists, especially in terms of social loss. We can see that the gap in social loss is sizable if we set the subsidy at the optimal level under a uniform subsidy, which is indicated as the vertical line in the figure. This result is robust to other values of T and V. Therefore, it is better for the government not to choose a project as the subsidy target as long as the private research institute makes a rational decision.

Lerner (1999) suggests that one benefit of a selective scheme is the certification effect in the sense that government selection conveys information about the quality of a new firm to investors. However, the government must

[10] Cowan (1991) introduces a "selective subsidy" as a tool to be used by the government to determine private technology choice. Our setting is more moderate in the sense that the research institute still has options under the selective subsidy.

[11] We can also consider the case in which the government changes the target project based on v in each period. This policy is equivalent to granting a uniform subsidy because the government reduces the cost of the R&D project chosen by the private researcher.

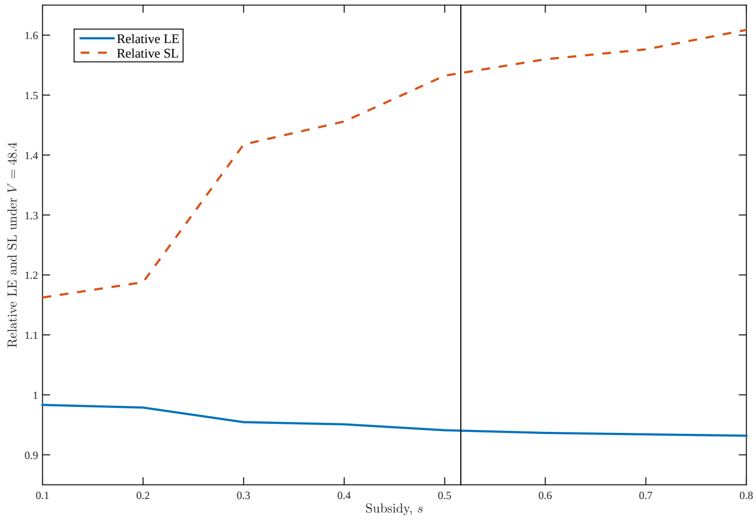


Figure 10: The learning efficiency and social loss of a selective subsidy, relative to those of a uniform subsidy. The vertical line indicates the optimal R&D subsidy. $V = 48.4$. (Similar for a higher V.)

know qualities of projects for the certification effect to be effective. In this subsection, we examined the impact of a selective subsidy when the government knows no more than the private sector does. In this case, granting a selective subsidy substantially lowers learning efficiency and heightens social loss, as shown in Figure 10.

5. Conclusion and Discussion

R&D subsidies for climate change and renewable energy are common in many countries, with government-sponsored research programs and grants available in the United States, the United Kingdom, Denmark, Ireland, Germany, Japan, and the Netherlands.

The decentralized economy underinvests in R&D in the absence of taxes and subsidies (Jones and Williams (2000)). In the energy sector, governmental funding plays an important role. For example, private sector investments in energy have decreased during the past decade, with current governmental funding accounting for 76% of all R&D expenditures in the US energy sector

(Nemet and Kammen (2007)). In the case of action on climate change, market failures in innovation, such as research spillovers, may justify research subsidies (Schneider and Goulder (1997)). Popp and Newell (2009) estimate social returns on environmental R&D and find that these returns are typically higher than social returns on other R&D investments.

Carraro and Siniscaico (1994) argue that research needs more of an environmental dimension in the long-term to tackle environmental problems. They show that high taxes are required to decrease CO_2 emissions in the absence of innovation. Similarly, Hart (2004) demonstrates the importance of environmentally oriented research in growth models.

In this study, we formulated an R&D project choice model in which the research institute does not know how long each project will take until the new technology is commercialized. By focusing on an environmental technology, we considered two sources of inefficiency: incomplete information and external costs. Our model is an optimal stopping model with a finite goal. By analyzing a simulation based on parameter values distilled mainly from information on R&D investment in electric vehicles, we showed that an R&D subsidy achieves the second-best R&D decision and performs better than a Pigouvian tax on the dirty technology under certain conditions. Moreover, the optimal R&D subsidy is different depending on whether the government takes into account the learning process in the research sector.

Analyses of the time paths of subsidies for competing technologies in a first-best economy (e.g., Kverndokk and Rosendahl (2007)) suggest that subsidies should be higher for technologies with higher spillovers. The authors of these studies assume that the regulator can select winners for the optimal subsidization. However, having taken account of learning in R&D activity, we showed that such selective subsidization is less effective than a uniform R&D subsidy. Likewise, having studied the role of technology subsidies in climate policy by using a simple dynamic equilibrium model Kverndokk and Rosendahl (2007) find that picking winners is complicated. Therefore, we suggest that governments should grant uniform subsidies rather than trying to choose winners.

Further research on our model could be used to analyze the correlations between projects, because trialing one project might help another. Incorporating an endogenous number of trials or R&D intensity in one period could also be a worthwhile extension. Although we focused on R&D activities and policies, broader approaches can be applied to environmental policy selection and to their effects on technological change (see Krysiak (2011) for a review). Considering learning in the context of R&D subsidies and understanding how these processes interact with other market-based policies represent useful directions for future research.

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012). The environment and directed technical change. *American Economic Review 102*(1), 131–66.

Bergemann, D. and U. Hege (2005). The financing of innovation: Learning and stopping. *RAND Journal of Economics 36*(4), 719–752.

Cardon, J. and D. Sasaki (1998). Preemptive search and R&D clustering. *RAND Journal of Economics 29*(2), 324–338.

Carraro, C. and D. Siniscaico (1994). Environmental policy reconsidered: The role of technological innovation. *European Economic Review 38*(3-4), 545–554.

Cowan, R. (1991). Tortoises and hares: Choice among technologies of unknown merit. *Economic Journal 101*(407), 801–14.

Dasgupta, P. and E. Maskin (1987). The simple economics of research portfolios. *Economic Journal 97*(387), 581–95.

Dutton, J. M. and A. Thomas (1984). Treating progress functions as a managerial opportunity. *The Academy of Management Review 9*(2), 235–247.

Gerlach, H. A., T. Rønde, and K. Stahl (2005). Project choice and risk in R&D. *Journal of Industrial Economics 53*(1), 53–81.

Gittins, J., K. Glazebrook, and R. Weber (2011). *Multi-armed Bandit Allocation Indices* (2nd ed.). Wiley.

Hart, R. (2004). Growth, environment and innovation–a model with production vintages and environmentally oriented research. *Journal of Environmental Economics and Management 48*(3), 1078–1098.

Ito, Y. and S. Managi (2015). The potential of alternative fuel vehicles: A cost-benefit analysis. *Research in Transportation Economics*. forthcoming.

Jones, C. I. and J. C. Williams (2000). Too much of a good thing? the economics of investment in R&D. *Journal of Economic Growth 5*(1), 65–85.

Jung, C., K. Krutilla, and R. Boyd (1996). Incentives for advanced pollution abatement technology at the industry level: An evaluation of policy alternatives. *Journal of Environmental Economics and Management 30*(1), 95–111.

Kneese, A. and C. Schulze (1975). *Pollution, Prices, and Public Policy.* Washington DC: Brookings.

Krysiak, F. C. (2011). Environmental regulation, technological diversity, and the dynamics of technological change. *Journal of Economic Dynamics and Control 35*(4), 528–544.

Kverndokk, S. and K. E. Rosendahl (2007). Climate policies and learning by doing: Impacts and timing of technology subsidies. *Resource and Energy Economics 29*(1), 58–82.

Laffont, J.-J. and J. Tirole (1996). Pollution permits and environmental innovation. *Journal of Public Economics 62*(1-2), 127–140.

Lerner, J. (1999). The government as venture capitalist: The long-run impact of the sbir program. *The Journal of Business 72*(3), 285–318.

Managi, S. (2012). Analysis of alternative fuel vehicles by disaggregated cost benefit. RIETI Discussion Paper Series, 12-E-035.

Nemet, G. F. and D. M. Kammen (2007). U.s. energy research and development: Declining investment, increasing need, and the feasibility of expansion. *Energy Policy 35*(1), 746–755.

Newell, R. G., A. B. Jaffe, and R. N. Stavins (1999). The induced innovation hypothesis and energy-saving technological change. *The Quarterly Journal of Economics 114*(3), 941–975.

Nordhaus, W. D. (2007). A review of the *stern review on the economics of climate change*. *Journal of Economic Literature 45*(3), 686–702.

Popp, D. and R. G. Newell (2009, October). Where does energy R&D come from? examining crowding out from environmentally-friendly r&d. NBER Working Papers 15423, National Bureau of Economic Research, Inc.

Popp, D., R. G. Newell, and A. B. Jaffe (2010). Energy, the environment, and technological change. In B. H. Hall and N. Rosenberg (Eds.), *Handbook of the Economics of Innovation*, Volume 2, Chapter 21, pp. 874–937. Elsevier.

Schneider, S. and L. Goulder (1997). Achieving low-cost emissions targets. *Nature 389*, 13–14.

Söderholm, P. and T. Sundqvist (2007). Empirical challenges in the use of learning curves for assessing the economic prospects of renewable energy technologies. *Renewable Energy 32*(15), 2559–2578.

Appendix A. Proof of Proposition 1 and Construction of the Value Function in the One-project Model

Proposition 1 is proved by using Lemmas 1–3. Note that the variables n and m below are all nonnegative integers.

Lemma 1. *The stopping trial for* $m = 1$ *is uniquely determined by*

$$\hat{n}(1) = \min\{n|\beta E(q|n,1)V \leq 1\}$$

and $v(n,1) > v(n+1,1)$ *for any* $n = M-1, ..., \hat{n}(1) - 1$.

Proof. First, we show $v(n,1) > 0$ if and only if $\beta E(q|n,1)V > 1$. If $\beta E(q|n,1)V > 1$, then $v(n,1) > 0$. To show the "only if" part, suppose that $\beta E(q|n,1)V \leq 1$ and $v(n,1) > 0$ for some n. Then,

$$v(n+1,1) = \frac{v(n,1) + 1 - \beta E(q|n,1)V}{\beta(1 - E(q|n,1))} > v(n,1).$$

Moreover, as $E(q|n,1)$ is decreasing in n, $\beta E(q|n+1,1)V \leq 1$. Hence, $v(n,1)$ is increasing in n and the research project is completed regardless of the number of trials that end in failure. However, in such a case, we can analytically compute $v(n,1)$ as the expected net benefit when the research institute commits to the project, as follows:

$$\begin{aligned} v(n,1) &= E\left(\beta^N V - \sum_{i=0}^{N-1}\beta^i \middle| n,1\right) \\ &= \left(V + \frac{1}{1-\beta}\right)\int_0^1 \frac{\beta q}{1-\beta(1-q)} F(dq|n,1) - \frac{1}{1-\beta}, \end{aligned} \tag{A.1}$$

where N is the number of trials before completion and $F(q|n,1)$ is the belief about q, which follows the beta distribution with parameters M and $n-M+2$. Because $F(q|n,1)$ first-order stochastically dominates $F(q|n+1,1)$ and $\frac{\beta q}{1-\beta(1-q)}$ is increasing in q, the expected net benefit (A.1) decreases with n (and becomes negative for sufficiently large n), which contradicts the above result that $v(n,1)$ is strictly increasing in n. Hence, $\beta E(q|n,1)V \leq 1$ implies $v(n,1) = 0$. Therefore,

$$\hat{n}(1) = \min\{n|\beta E(q|n,1)V \leq 1\}.$$

Clearly, $\hat{n}(1)$ is unique because $E(q|n,1)$ is strictly decreasing in n and converges to zero as $n \to \infty$.

Given $\hat{n}(1)$, we can show $v(n,1) > v(n+1,1)$ for $n = M-1, ..., \hat{n}(1)-1$ by induction. Note

$$\begin{aligned} v(\hat{n}(1)-2,1) &= -1 + \beta\{E(q|\hat{n}(1)-2,1)V + (1 - E(q|\hat{n}(1)-2,1))v(\hat{n}(1)-1,1)\} \\ &> -1 + \beta E(q|\hat{n}(1)-2,1)V > -1 + \beta E(q|\hat{n}(1)-1,1)V = v(\hat{n}(1)-1,1). \end{aligned}$$

Suppose that $v(n,1) > v(n+1,1)$ for some $n \in \{M-1, ..., \hat{n}(1)-1\}$. Then,

$$\begin{aligned} v(n-1,1) &= -1+\beta\{E(q|n-1,1)V+(1-E(q|n-1,1))v(n,1)\} \\ &> -1+\beta\{E(q|n,1)V+(1-E(q|n,1))v(n+1,1)\} = v(n,1). \end{aligned}$$

Hence, for any $n = M-1, ..., \hat{n}(1)-1$, $v(n,1) > v(n+1,1)$. ■

Lemma 2. *The stopping trial for $m \geq 2$ is uniquely determined by*

$$\hat{n}(m) = \min\{n|\beta E(q|n,m)v(n+1,m-1) \leq 1\}$$

and $v(n,m) > v(n+1,m)$ for any $n = M-1, ..., \hat{n}(m)-1$.

Proof. Let $m \geq 2$. Suppose that $\hat{n}(m-1)$ uniquely exists and $v(n,m-1) > v(n+1,m-1)$ for any $n \in \{M-1, ..., \hat{n}(m-1)-1\}$. If $\beta E(q|n,m)v(n+1,m-1) > 1$, then $v(n,m) > 0$ by definition. Suppose that there exists an $\bar{n}$ such that

$$\beta E(q|\bar{n},m)v(\bar{n}+1,m-1) \leq 1, \text{ and } v(\bar{n},m) > 0.$$

Then, $v(\bar{n}+1,m) > v(\bar{n},m)$. Given that $E(q|n,m)v(n+1,m-1)$ is decreasing in n, $\{v(n,m)\}_{n=\bar{n}}^{\infty}$ must be an increasing sequence for the given m. The research institute does not stop the project until it reaches $m-1$. The value obtained from one success should be less than $v(\bar{n},m-1)$ because $v(n,m-1)$ is decreasing in n. Thus, for $n \geq \bar{n}$,

$$0 < v(n,m) < \left(v(\bar{n},m-1)+\frac{1}{1-\beta}\right)\int_0^1 \frac{\beta q}{1-\beta(1-q)}F(dq|n,m) - \frac{1}{1-\beta}. \tag{A.2}$$

However, given that

$$\lim_{n\to\infty}\int_0^1 \frac{\beta q}{1-\beta(1-q)}F(dq|n,m) = 0,$$

the right-hand side in (A.2) becomes strictly negative for sufficiently large n. This is a contradiction. Hence, $v(n,m) > 0$ if and only if $\beta E(q|n,m)v(n+1,m-1) > 1$. Because $E(q|n,m)v(n+1,m-1)$ is decreasing in n and converges to 0, the stopping trial is uniquely determined by

$$\hat{n}(m) = \min\{n|\beta E(q|n,m)\, v(n+1, m-1) \leq 1\}.$$

Moreover, $v(n,m) > v(n+1,m)$ for any $n \in \{M-m, ..., \hat{n}(m)-1\}$ similar to the proof of Lemma 1.

Given that this argument holds for $m = 1$ according to Lemma 1, the proposition is proved by induction. ■

Lemma 3. $\hat{n}(m) < \hat{n}(m-1)$ *for any* $m = 2, 3, ..., M$.

Proof. Suppose that $\hat{n}(m) \geq \hat{n}(m-1)$ for some $m \in \{2, 3, ..., M\}$. Then,

$$0 < v(\hat{n}(m)-1, m) = -1+\beta \left[\begin{array}{l} E(q|\hat{n}(m)-1, m)\, v(\hat{n}(m), m-1) \\ +(1-E(q|\hat{n}(m)-1, m))\, v(\hat{n}(m), m) \end{array} \right] = -1$$

because $v(\hat{n}(m), m-1) \leq v(\hat{n}(m)-1, m-1) = 0$ and $v(\hat{n}(m), m) = 0$. This is a contradiction. ■

Construction of $v(n,m)$. From Lemma 1, the threshold number of trials, $\hat{n}(1)$, is determined by

$$\hat{n}(1) = \min\{n\,|\beta E(q|n,1)\, V \leq 1\} = \lceil \beta V M \rceil - 2. \tag{A.3}$$

If $\hat{n}(1) \leq M-1$, then the R&D project is never promising. Thus, we suppose that βV is sufficiently high to make $\hat{n}(1) > M-1$.

Starting from $v(\hat{n}(1), 1) = 0$, $v(n,1)$ for $n < \hat{n}(1)$ is computed by using

$$v(n,1) = \max\{0, -1+\beta[E(q|n,1)\, V + (1-E(q|n,1))\, v(n+1,1)]\}.$$

For $m \geq 2$, we again consider

$$\hat{n}(m) = \min\{n\,|\beta E(q|n,m)\, v(n+1, m-1) \leq 1\},$$

and define

$$v(n,m) = -1+\beta \left\{ \begin{array}{l} E(q|n,m)\, v(n+1, m-1) \\ +(1-E(q|n,m)\, v(n+1,m)) \end{array} \right\} \quad \text{for } n \leq \hat{n}(m)-1,$$

where $v\left(\hat{n}\left(m\right),m\right)=0$ by definition.

Appendix B. Proof of Proposition 2

For any feasible state (n,m), the value function must satisfy

$$v\left(n,m\left|V,s\right.\right) = \begin{cases} \max\left\{0,-\left(1-s\right)+\beta\left[\begin{array}{c} E\left(q|n,m\right)v\left(n+1,m-1\left|V,s\right.\right) \\ +\left(1-E\left(q|n,m\right)\right)v\left(n+1,m\left|V,s\right.\right) \end{array}\right]\right\} & \text{for } m\geq 1, \\ V \qquad \text{for } m=0. \end{cases}$$

Dividing each term by $1-s$ yields

$$\frac{v\left(n,m\left|V,s\right.\right)}{1-s} = \begin{cases} \max\left\{0,-1+\beta\left[\begin{array}{c} E\left(q|n,m\right)\frac{v\left(n+1,m-1\left|V,s\right.\right)}{1-s} \\ +\left(1-E\left(q|n,m\right)\right)\frac{v\left(n+1,m\left|V,s\right.\right)}{1-s} \end{array}\right]\right\} & \text{for } m\geq 1, \\ \frac{V}{1-s} \qquad \text{for } m=0. \end{cases}$$

This functional equation has the same structure as (3). Therefore,

$$v\left(n,m\left|\frac{V}{1-s},0\right.\right)=\frac{v\left(n,m\left|V,s\right.\right)}{1-s}.$$

This relationship ensures that the stopping trials $\left\{\hat{n}\left(m\right)\right\}_{m=1}^{M}$ are identical and R&D decisions are equivalent for situations between (V,s) and $\left(\frac{V}{1-s},0\right)$.

Appendix C. Proof of Proposition 4

First, note that $v\left(n,m|\tau\right)$ is decreasing in distance from the goal, $m-g_{\tau}$, because $v\left(n,m|\tau\right)$ has the same structure as $v\left(n,m\right)$, defined in (3), even though the terminal condition is modified. Now, $g_{\tau}\geq g_{\tau'}\geq 0$ by definition. Consider $m=g_{\tau}+h$ for $h=1,2,...,M-g_{\tau}$. Then,

$$0=v\left(\hat{n}\left(g_{\tau}+h|\tau\right),g_{\tau}+h|\tau\right)\geq v\left(\hat{n}\left(g_{\tau}+h|\tau\right),g_{\tau}+h|\tau'\right)$$

because $E\left(q|\hat{n}\left(g_{\tau}+h|\tau\right),g_{\tau}+h\right)$ are common and $\left(g_{\tau}+h\right)-g_{\tau'}\geq h$. Therefore, $\hat{n}\left(g_{\tau}+h|\tau\right)\geq\hat{n}\left(g_{\tau}+h|\tau'\right)$.

Appendix D. Comparing One-step and Many-step R&D Settings

In this study, we consider many-step R&D. By contrast, much of R&D economics research deals with one-step R&D. (For example, Bergemann and Hege (2005) consider the case of $M = 1$ with learning.) Although our analysis also applies to one-step R&D, the quantitative impact of learning is weak in such a one-step environment. Figure D.11(a) illustrates the LE-maximizing R&D subsidies for different values of M (with the other parameters unchanged). Figure D.11(b) shows how the maximum improvements in LE evolve along with M, normalized by those indices in the no-learning case. When only a few steps are required to finish the R&D project, the LE-maximizing subsidy is high, but the improvement in efficiency brought about by the subsidy policy is negligible. The degree of improvement increases gradually with M until $M = 11$ because learning opportunities increase with the total number of steps. The jump at $M = 12$ occurs because of a structural change. In the absence of policy and learning, the R&D project does not start when $M \geq 12$ under our current parameter settings.

In summary, we can ignore learning in small-step R&D. Learning matters only when there is a long way to go.

Appendix E. Construction of the Value Function in the Two-project Model

First, define $\hat{n}_i(m_i)$ for $i = 1, 2$, which is the stopping trial when the researcher focuses on project i. If $n_i \geq \hat{n}_i(m_i)$, the researcher does not conduct project i even when the outside value is 0, which is no more than the actual outside value because the other project may be worth conducting. Hence, the values of (n_1, m_1, n_2, m_2) equal the values of (n_j, m_j) in the one-project model if $n_i \geq \hat{n}_i(m_i)$ $(i \neq j)$. Clearly, $v(n_1, m_1, n_2, m_2) = 0$ if $n_i \geq \hat{n}_i(m_i)$ for all i and $v(n_1, m_1, n_2, m_2) = V$ if $m_i = 0$ for some i. In addition, $v(n_1, m_1, n_2, m_2)$ is not defined if $n_i < M_i - m_i$ for some i.

The above argument narrows the range of states whose values are defined. Now, we consider $n_i < \hat{n}_i(m_i)$.

We set $m_2 = 1$. The relevant n_2 should be between $M_2 - 1$ and $\hat{n}_2(1) - 1$. Let us set $n_2 = \hat{n}_2(1) - 1$. If $m_1 = 1$, then the relevant n_1 lies between $M_1 - 1$

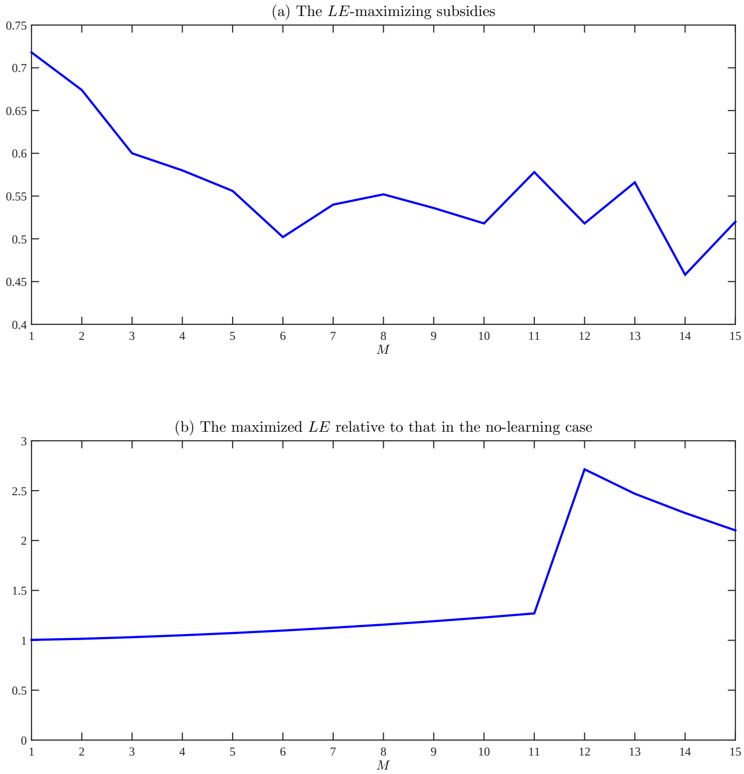


Figure D.11: Panel (a): LE-maximizing R&D subsidies for different values of M. Panel (b): the maximized LE relative to that in the no-learning case.

and $\hat{n}_1(1)-1$. Starting from $n_1=\hat{n}_1(1)-1$,

$$\begin{aligned}
& v_1\left(\hat{n}_1(1)-1,1,\hat{n}_2(1)-1,1\right) \\
= \quad & -1+\beta\left\{\begin{array}{l} E\left(q_1|\hat{n}_1(1)-1,1\right)V \\ +\left(1-E\left(q_1|\hat{n}_1(1)-1,1\right)\right)v\left(\hat{n}_1(1),1,\hat{n}_2(1)-1,1\right) \end{array}\right\} \\
= \quad & -1+\beta\left\{\begin{array}{l} E\left(q_1|\hat{n}_1(1)-1,1\right)V \\ +\left(1-E\left(q_1|\hat{n}_1(1)-1,1\right)\right)\hat{v}_2\left(\hat{n}_2(1)-1,1\right) \end{array}\right\}, \qquad \text{(E.1)}
\end{aligned}$$

where $\hat{v}_i(n_i,m_i)$ is the value of (n_i,m_i) in the one-project model. Because v_2 for the same state can be defined similarly, we obtain $v\left(\hat{n}_1(1)-1,1,\hat{n}_2(1)-1,1\right)$ as the maximum of these two values.

By having $v(\hat{n}_1(1)-1,1,\hat{n}_2(1)-1,1)$ as the terminal value, we repeat the calculation to obtain

$$v_1(n_1,m_1,n_2,m_2) = -1+\beta \left\{ \begin{array}{l} E(q_1|n_1,m_1)\, v(n_1+1,m_1-1,n_2,m_2) \\ +(1-E(q_1|n_1,m_1))\, v(n_1+1,m_1,n_2,m_2) \end{array} \right\},$$

$$v_2(n_1,m_1,n_2,m_2) = -1+\beta \left\{ \begin{array}{l} E(q_2|n_2,m_2)\, v(n_1,m_1,n_2+1,m_2-1) \\ +(1-E(q_1|n_1,m_1))\, v(n_1,m_1,n_2+1,m_2) \end{array} \right\},$$

$$v(n_1,m_1,n_2,m_2) = \max\{v_1(n_1,m_1,n_2,m_2), v_2(n_1,m_1,n_2,m_2)\},$$

over all the remaining states in the ascending order of m_i and descending order of n_i.