Green Bandits - preliminary version *

Katinka Holtsmark†

February 23, 2016

Abstract

I consider a dynamic model of N countries making decisions over fossil energy consumption and green technology investments. The investment technology is modeled as a one-armed bandit: it is unknown to all countries whether investments may or may not yield positive returns. The potential for returns is perfectly correlated between countries, investments in each country generate information that is valuable in all countries. I find that a state-dependent investment subsidy is needed in addition to a Pigouvian tax, to implement the first-best investment and emission levels. This is the case even though there are no technological spillovers or hold-up problems in the model. The paper hence identifies a new channel creating underinvestment in green technology. Moreover, I show that the subsidy required to implement efficient investment levels is decreasing in the probability that the technology will succeed, over the range it is used. In other words, when the subsidy is positive, it is higher for less promising technologies.

Keywords: Bandits, Uncertainty, Investment, Climate change.
JEL classification: Q50, Q55, D83, F50.

*I want to thank Bård Harstad, Halvor Mehlum, Åsmund Sunde Valseth and Bjart Holtsmark for helpful comments.
†Dept. of Economics, University of Oslo, Norway, k.k.holtsmark@econ.uio.no

1 Introduction

A healthy climate is an international common good, suffering from over-exploitation due to lack of appropriate regulation. In economic theory, the answer to the *tragedy of the commons* [Hardin, 1968] is well-known: An appropriate price on exploitation of a common good can implement the welfare-maximizing allocation of resources [Pigou, 1920, Samuelson, 1954]. If consumers of fossil energy were forced to pay a price for their consumption in line with the damage it inflicts on others, efficient consumption levels would prevail.

Pricing of greenhouse gas emissions would reduce emissions both by reducing total energy consumption, and by spurring investment in green technologies. However, I show in this paper that fundamental uncertainty regarding the potential of different technologies, will lead to investment levels that are too low, even when carbon emissions are taxed according to Pigou [1920]. Furthermore, I show that an investment subsidy is needed, and that for technologies that should be subsidized, the subsidy needed to implement efficient investment levels is higher when the potential for a technology is more uncertain. In other words, among technologies that should be subsidized, the less promising should be subsidized more heavily. This is the case even though there are no technological spillovers, only the investor himself get access to new technological developments. Drawing on the literature on bandits and experimentation, the paper sheds light on a channel for underinvestment in green technology that so far has not received much attention.

As it seems unlikely that the world will reach an ambitious global agreement on emission reductions, unilateral actions and bottom-up approaches are increasingly important. And investment in renewable energy technologies are rapidly increasing in all parts of the world. According to United Nations Environment Programme [UNEP, 2015], worldwide investments in renewable energy reached $270 billion in 2014, following a 17 % increase from the year before. Important contributors to these investments were China ($83.3 billion), the US ($38.3 billion) and Europe ($57.5 billion). Investments are driven by cost reductions and increasing demand from new markets, but also by policy. It is therefore crucial to understand the different market failures related to green technology development, to direct resources to its most productive use. This paper sheds light on one such market failure.

I consider a dynamic model in continuous time of N countries each making decisions both over fossil energy consumption and green technology investments. Fossil energy consumption causes climate change, which inflicts damage on all countries. Investment in any country may result in innovations that yield access to a new abatement technology. Investment generates externalities in two different ways. Firstly, due to the climate externality, an innovation benefits all countries, through reducing the aggregate climate damage. However, the underlying rate of return from investments - the state of the world - is unknown, and this is the second source of externality in the model. Specifically, the investment technology is modeled as a one-armed bandit, and the bandit is perfectly correlated between countries: If the state of the world is good, investments will generate returns at random times, while if the state of the world is bad, there will never be any returns. Because the state of the world is the same in all countries, observing either positive returns or investments that are not paying off in other countries, provides each country with valuable information. The model set up is built on the model presented in Keller et al. [2005], and I add the climate externality to their framework.

Carbon capture and storage (CCS) is an example of a technology in development, where there is large uncertainty regarding the potential for using the technology large scale. CCS-facilities are already in place some places in the world, and several countries, including Norway, have spent substantial resources in attempts to develop full-scale abatement of emissions from power plants, using this type of technology. However, there is still uncertainty related especially to storage over time.It is this type of uncertainty facing investors that I consider in this paper.

I show that when there is this type of uncertainty related to green technology investments, underinvestment is not eliminated by a Pigouvian tax. A tax on consumption of fossil energy increases investments through making each country's own emissions more costly, and hence access to clean energy more valuable. However, in this setting, investments in a given country benefits others not only through decreasing expected fossil consumption in the investing country, but also by increasing expected future investments - and thereby decreasing expected future emissions - in *all* other countries. This externality - working through generation of information - is not internalized by a Pigouvian tax. An investment subsidy is therefore needed in order to implement the first-best investment levels. Because the incentive to invest is lower when the uncertainty is large, less promising technologies require a larger subsidy. For levels of uncertainty where the first-best investment levels are positive, the optimal subsidy is therefore decreasing in the probability that the technology will succeed.

The paper builds on Keller et al. [2005] , who consider the problem of pure information externalities, and show that in the symmetric Markov perfect equilibrium, there is underinvestment in experimentation compared to the first-best allocation. They also consider asymmetric equilibria, and characterize conditions under which the equilibrium information generation may be close to the efficient level. I build on their model, and show that including a payoff externality in the framework creates additional underinvestment, and find the state-dependent subsidy necessary in addition to a Pigouvian tax to implement the first-best allocation in the model.

The framework is also similar to that of Rothschild [1974]. However, Rothschild [1974] do not consider the effect of strategic interaction, which is the main focus of both Keller et al.

[2005], and this paper. For an overview of other early contributions to the literature on bandit models in economics, see Bergemann and Välimäki [2010].

Strategic interaction resulting in inefficient investment allocation is considered by Bolton and Harris [1999] and Bolton and Harris [2001], in a framework similiar to that of Rothschild [1974]. In both papers, they consider a problem similar to the one studied by Keller et al. [2005]. However, they model the technology as yielding payoffs with Brownian noise. Keller et al. [2005] show that using exponential bandits, as I also do in this paper, greatly simplifies the mathematical framework. Among other papers who also used exponential bandits, are both Bergemann and Hege [1998] and Bergemann and Hege [2005]. They consider investment in risky projects where the value is uncertain, and where investors can learn about this value over time from observing the development of the project. However, the strategic interaction they consider is connected mainly to problems of asymmetric information between the investor and an agent, and are hence of a different type than that considered both in Keller et al. [2005], and in this paper.

In environmental economics, the most well-known reason for under-provision of new technologies is *technological spillovers*. When others than the investor himself get access to new technological developments, a carbon price is no longer sufficient in order to implement efficient levels of consumption *and* investment. See for example Golombek and Hoel [2005]. When countries cooperate on emission reductions, there may also be underinvestment in green technologies for strategic reasons: the so-called *hold-up problem* arises if countries cut down on investments in order to avoid a weak bargaining position in future negotiations over emission reductions (see Buchholz and Konrad [1994] and Harstad [2015]). This paper contributes to this literature in environmental economics by applying the model framework from the bandit literature to the climate problem. By adding a payoff externality to the externality from information generation, the paper sheds light on a source of underinvestment in green technology that has not been in focus so far.

The paper is organized as follows: I will introduce the model set up in Section 2.1, and consider the cooperative solution to the problem in Section 2.2. The non-cooperative game, and the symmetric MPE is then discussed in Section 2.3, and I continue in Section 2.4 by showing how a combination of a tax on consumption and a subsidy on investment can implement the first-best allocation. Finally, some concluding remarks are given in Section 3.

2 The Model

2.1 The Model Setup

The setup of the model is close to that of Keller et al. [2005]. I consider $N > 1$ identical countries, each making decisions concerning both energy consumption and investment levels over time. Time $t \in [0, \infty)$ is continuous, and the discount rate is $r \in (0, 1)$.

The expected returns from investment depend on the state of the world, which is initially unknown to all countries. The underlying investment technology is modeled as a one-armed bandit with unknown type, where the countries invest in experimentation with the technology. The state of the world is the same for all countries, and there are two possible states, "good" or "bad". Countries do not know which state they are in when the game starts.

$\omega \in \{0, 1\}$ denotes the state; if the state of the world is good ($\omega = 1$), the investment will generate payoffs - or innovations - after exponentially distributed random times. The rate at which innovations are generated is proportional to the amount that is invested, for each country. Let I_{it} denote investment by country i at time t. Then, if $\omega = 1$, the conditional probability that country i will get an innovation from his risky investment during a time interval of length dt starting at time t is given by $Pr(\text{innovation in country } i \mid \omega = 1, I_{it}) = \lambda I_{it} dt$, with $\lambda > 0$.

If the state of the world is bad ($\omega = 0$), investments will never generate any innovations. Both investments and successful innovations are observed by all countries at all times. This means that if investment generates an innovation for one country at some point in time, all countries will know that the true state of the world is good. As long as no innovations are observed, the countries will use Bayesian updating on their beliefs about ω, and gradually become more pessimistic. At time $t = 0$, all countries have a common prior belief, p_0, denoting the subjective probability put on the state of the world being good. Because all countries share the same information at all times, the posterior - or current - belief will continue to be the same across countries. This current belief will be denoted by $p_t \in [0, 1]$.

Given the prior belief, the chosen investment levels, any observed innovations, p_t will develop over time. More precisely, if an innovation is observed, p_t will jump to 1, and stay there forever. As long as no innovation is observed, p_t will gradually decline. This decline in p_t over time can be described using Bayes' rule:

$$\begin{aligned} &\Pr{}_{t+dt}\left(\omega = 1 \mid \text{ no innovations, } \{I_{jt}\}_{j=1}^{N}, p_t\right) \\ &\quad = \frac{\Pr{}_{t+dt}\left(\text{no innovations, } \mid \omega = 1, \{I_{jt}\}_{j=1}^{N}, p_t\right) \Pr{}_{t+dt}\left(\omega = 1 \mid \{I_{jt}\}_{j=1}^{N}, p_t\right)}{\Pr{}_{t+dt}\left(\text{no innovations, } \mid \{I_{jt}\}_{j=1}^{N}, p_t\right)}. \end{aligned}$$

Within the time interval $[t, t+dt]$, the probablity of any country generating an innovation conditional on $\omega = 1$, is given by $\lambda \sum_j I_j dt$. In continuous time the probability that more than one country gets an innovation at the same moment in time can be disregarded. Hence, the probability of no innovations being observed over this time interval, conditional on $\omega = 1$ will be given by $1 - \lambda \sum_j I_{jt} dt$. The unconditional probability of no observed innovations is given by $p_t(1 - \lambda \sum_j I_{jt} dt) + (1 - p_t) = 1 - p_t \lambda \sum_j I_{jt} dt$. The conditional probability of the state of the world being good at time $t + dt$ can then be written as follows:

$$p_t + dp_t = \frac{\left(1 - \lambda \sum_j I_{jt} dt\right) p_t}{1 - p_t \lambda \sum_j I_{jt} dt} \in (0, 1),$$

where dp_t represents the change in the belief over the time interval $[t, t + dt]$, conditional on no country experiencing an innovation within this time interval.

In continuous time, the belief will then develop according to:

$$\begin{aligned} \dot{p}_t &= \lim_{dt \to 0} \frac{(p_t + dp_t) - p_t}{dt} \\ &= -\lambda p_t (1 - p_t) \sum_j I_{jt}, \end{aligned} \tag{1}$$

as in Keller et al. [2005]. In the following sections, the development of p_t over time will be represented by the expression in Equation (1).

As is intuitively clear, a higher λ will increase the speed at which p_t declines if no innovations are observed. Furthermore, the closer the belief at time t is to $1/2$, the faster will p_t adjust. Importantly, the speed of development of the belief of each country depends on aggregate investments. This creates the information externality: investments by one country generate information that is useful to all countries.

I consider Markov perfect equilibria (MPE) in this model, and the current belief p_t will be the (only) state variable. A country's strategy will be the investment decision at all t, given p_t.

Given the investment decisions, each country decides on energy consumption, in each time period. Energy consumption of country i is given by e_{it}, and consists of fossil energy. A successful innovation in country i gives this country access to an abatement technology that makes an amount $A > 0$ of their consumption "clean". The fossil consumption of country i creating climate damage is given by f_{it}, while the clean part is given by z_{it}, so that $e_{it} = f_{it} + z_{it}$. country i derives a flow utility from energy consumption according to the increasing and concave utility function $u(e_{it})$. As a normalization, fossil energy is available at no cost, but consumption is limited by a flow damage from climate change incurred by all countries, resulting from aggregate fossil energy consumption. This climate damage is assumed to be linear in the aggregate fossil energy consumption, and the constant marginal damage for each country is given by $D > 0$.

This access to clean energy following an innovation is long-lasting, but the abatement technology will be subject to depreciation. Specifically, from an innovation at time s in country i, the clean energy that is left at time $t > s$ is given by $e^{-\delta(t-s)}A$, where $\delta \in (0,1)$ is the rate of depreciation. And at any time t, when there might have been any number of innovations earlier, the amount of clean energy available in country i is given by:

$$z_{it} = \int_0^t e^{-\delta(t-s)} A_{is} \, ds,$$

where $A_{is} = A$ if there was an innovation in country i at time s, and 0 otherwise.

The total discounted value of an innovation at time s to country i is thus given by $D\hat{A}$, with:

$$\hat{A} \equiv \int_0^\infty e^{-(r+\delta)t} A \, dt = \frac{1}{r+\delta} A,$$

which also represents the benefit to each of the other countries from an innovation by country i because the reduced climate damage is a public good. [1]

The investment cost is linear, and given by cI_{it} for country i at time t, with $c > 0$. However, the investment is limited to $I_{it} \in [0,1], \forall i, t$. Because both the investment cost and the benefit from innovation is linear, each country will either want to invest fully ($I_{it} = 1$), not to invest ($I_{it} = 0$), or he will be indifferent between all possible investment levels. The current belief put on the state of the world being good, p_t, will determine the level of investment maximizing expected utility.

I assume throughout that the size of the parameters are such that if the state of the world is known to be good ($p_t = 1$), the (individually) utility maximizing investment level is always $I_{it} = 1$, while if the state of the world is known to be bad ($p_t = 0$), it is always optimal to choose $I_{it} = 0$. Furthermore, I assume that there exist strictly positive current beliefs such that the first-best investment levels are zero. [2]

[1] I will throughout assume that there is always positive demand for fossil energy from all countries, meaning that the investment decision is only affected by earlier innovations through the current belief p_t. This assumption boils down to assuming that $e^* > \frac{1}{\delta}A$, where e^* is defined by Equation (3).

[2] These assumptions boil down to $0 < c < \lambda D\hat{A}$.

The expected total utility of country i over the whole time horizon is given by:

$$E\left[\int_0^{\infty} e^{-rt}\left(u(f_{it}+z_{it}) - D\sum_j f_{jt} - cI_{it}\right)dt\right], \tag{2}$$

where the expectation is over z_{it} and I_{it} for country i, and f_{jt} for all countries j. At each point in time, every country makes two decisions: First, the investment levels are decided, and if there is an innovation, it is realized immediately. Then, fossil consumption is determined, depending on the realized z_{it}. The consumption level, f_{it}, will depend negatively on the number of earlier innovations in country i.

Because of the negative externality from fossil energy consumption, an innovation for country i will not only directly benefit this country, but will increase the utility of all other countries. Indeed, access to clean energy can be considered a public good.

2.2 The Cooperative Solution

In this section, I solve for the symmetric investment strategy that maximizes aggregate welfare, following the method of Keller et al. [2005]. First, it is necessary to first characterize the levels of fossil energy consumption that maximize aggregate welfare, given investments. Time subscripts are dropped from this point on, unless they are strictly needed.

I define aggregate welfare as the discounted sum of utilities over the whole time horizon across all countries. Within a given time interval, $[t, t+dt]$, the aggregate welfare is then given by:

$$\sum_i \left[u(e_i) - D\sum_j f_j - cI_i\right] dt.$$

This expression must be maximized with respect to $\{e_i, f_i\}_{i=1}^N$, given that $e_i = z_i + f_i$, and $z_{it} = \int_0^t e^{-\delta(t-s)} A_{is}\, ds$. The solution is given by the following N first-order conditions:

$$u'(e_i^*) = ND \qquad \forall i. \tag{3}$$

Energy consumption should be set so that the individual marginal utility equals the aggregate marginal damage, which defines the first-best consumption levels; $e_i^* = e^*$ and $f_i^*(z_i) = f^*(z_i) = e^* - z_i$ To simplify notation in what follows, define also $U^* \equiv u(e^*) - DNe^*$, and $I \equiv \sum_j I_j$.

In order to characterize the first-best aggregate investment levels when there has been no innovation so that $p < 1$, the following maximization problem must be solved: [3]

$$W(p) = \max_{I\in[0,N]} \left\{ N\left[U^* - c\frac{I}{N} + p\lambda ID\hat{A}\right] dt + e^{-rdt} E\big[W(p_{t+dt})|p, I\big] \right\},$$

where p_{t+dt} is the belief at the end of the time interval. The expectation in the last term can be written as follows:

$$E\big[W(p_{t+dt})|p, I\big] = p\lambda IdtW(1) + (1 - p\lambda Idt)W(p + dp),$$

where dp represents the change in the belief, p, given that no innovation is observed, and;

$$\begin{aligned} W(1) &= \int_0^\infty e^{-rt} N\left[U^* - c + \lambda ND\hat{A}\right] dt \\ &= \frac{1}{r} N\left[U^* - c + \lambda ND\hat{A}\right], \end{aligned}$$

because, by assumption, the welfare-maximizing investment level is $I = N$ when the state of the world is known to be good.

Using the expression given in Equation (1), we have $p+dp = p-I\lambda p(1-p)dt$, while $W(p+dp) = W(p) + dpW'(p)$ and $e^{-rdt} = 1 - rdt$ when the length of the time interval, dt, is sufficiently small.

The maximization problem can then be rewritten as follows:

$$\begin{aligned} W(p) = \max_{I\in[0,N]} &\left\{ N\left[U^* - c\frac{I}{N} + p\lambda ID\hat{A}\right] dt \right. \\ &\left. + (1 - rdt)\big[p\lambda IdtW(1) + (1 - p\lambda Idt)W(p) - (1 - p\lambda Idt)I\lambda p(1 - p)dtW'(p)\big] \right\}. \end{aligned}$$

Finally, subtracting $(1 - rdt)W(p)$ on both sides, simplifying and letting $dt \to 0$, gives:

$$\begin{aligned} rW(p) = NU^* + \max_{I\in[0,N]} I &\left\{ p\lambda\big[W(1) - W(p) - (1 - p)W'(p)\big] \right. \\ &\left. - \left[\frac{c}{N} - p\lambda D\hat{A}\right] N \right\}. \end{aligned}$$

The value function, $W(p)$, is now expressed in continuous time. This expression consists of three main parts: First, and independent of the investment level, is the flow utility U^* for each country. Next is the value investment, in two parts. The first part is the expected future benefit of increased investment; increased probability of an innovation, giving a change in

[3] Because the clean energy availability, $\{z_j\}_{j=1}^N$, is not payoff relevant, the only state variable of the problem is p.

the value function of $W(1) - W(p)$, and a larger change in the posterior, p, should there be no innovation from investments. Finally, the last term can be considered the net cost of investment in terms of reduced flow utility at time t. This net cost is the investment cost subtracted the direct gain from innovation: reduced climate damage.

Given the linearity of both investment costs and climate damage, aggregate welfare is maximized either by choosing $I = 0$ or choosing $I = N$, or any $I \in [0, N]$ solves the problem. In order to characterize the values of the current belief, p, for which each of these alternatives is true, I start by expressing the aggregate welfare in the two following cases:

- When p is such that the problem of the social planner is solved by $I = 0$, aggregate welfare is independent of the state, and is given by:

$$rW_0(p) = NU^*$$

- When p is such that the problem of the social planner is solved by $I = N$, aggregate welfare is given by the differential equation:

$$rW_N(p) = NU^* + N\Big(p\lambda\big[W(1) - W(p) - (1-p)W'(p)\big] - \left[\frac{c}{N} - p\lambda D\hat{A}\right] N\Big), \tag{4}$$

  which is solved by:

$$rW_N(p) = N\left[\lambda pND\hat{A} + U^* - c + rC^*(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{\lambda N}}\right], \tag{5}$$

  where C^* is a constant of integration. See Appendix A.1 for calculations.

Intuitively, the marginal value of investment is nondecreasing in the state p, since higher p means a higher probability that increased investment results in a successful innovation. It must therefore exist a threshold value for p such that the socially optimal investment level is equal to zero below this threshold, and equal to N above this threshold. The following result characterizes of the first-best investment allocation:

Let I^* denote the solution to the social planner's problem.

Lemma 1. *Given the number of countries, N, there exists a threshold value, $p^*(N)$, such that the welfare maximizing aggregate investment level is zero for all current beliefs below this* *value, and N for all current beliefs above this value:*

$$I^* = \begin{cases} 0 & if \quad p < p^*(N) \\ N & if \quad p > p^*(N) \end{cases}.$$

The threshold value $p^*(N)$ is given by:

$$p^*(N) = \frac{cr/\lambda}{(r+\lambda N)ND\hat{A} - Nc} > 0,$$

and the aggregate welfare is described by the following expression:

$$rW_N(p) = N\Bigg[\lambda N D\hat{A}p - c + U^* - \left(\lambda N D\hat{A}\ p^*(N) - c\right)\frac{1-p}{1-p^*(N)}\left(\frac{1-p}{p}\frac{p^*(N)}{1-p^*(N)}\right)^{\frac{r}{\lambda N}}\Bigg] \quad (6)$$

when $p \geq p^*(N)$, and by $rW_0(p) = NU^*$ when $p \leq p^*(N)$.

Proof. Given $I_i^* = 0$ if $p = 0$ and $I_i^* = 1$ if $P = 1$, and the marginal value of investment increasing in p, there must exist a threshold $p^*(N)$. Welfare must then be independent of the investment level at the threshold;

$$rW_N(p^*(N)) = rW_0(p^*(N)) \quad \text{(Value Matching)},$$

and the change in welfare when the belief increases at the threshold must be zero:

$$rW_N'(p^*(N)) = 0 \qquad \text{(Smooth Pasting)}.$$

These two conditions give the constant of integration, C^*, and the threshold, $p^*(N)$. See appendix A.2 for elaboration on this proof. □

This result is parallel to the result presented in Proposition 3.1 in Keller et al. [2005]. Note that for $p = p^*(N)$, any $I \in [0, N]$ solves the maximization problem of the social planner. If $N = 1$, the threshold for the current belief is given by:

$$p^*(1) = \frac{cr/\lambda}{(r+\lambda)D\hat{A} - c} \geq p^*(N).$$

For $N > 1$ the inequality is strict. The social planner will have a lower threshold above which he prefers full investment by all countries when the number of countries is higher, and as the number of countries approaches infinity, the threshold will approach zero. This is because investment by one country benefits (in expectation) all the other countries, both at the time of investment and in the future, while the investment cost is private. The gain from investment can be divided in two, both parts equally beneficial to all countries. First, increased investment gives higher probability of an innovation that results in lower climate damage. Second, higher investment generates more information on the state of the world.

2.3 The Non-Cooperative Solution

With the first-best investment allocation from the last section as a benchmark, I will now look at the non-cooperative game where each country chooses both investments and energy consumption over time, in order to maximize their own utility over the whole time horizon, again following the method used in Keller et al. [2005].

The expected utility of country i over the entire time horizon is given by Equation (2), and each country maximizes utility by choosing investments and energy consumption over time, given the state variable p. The development of the state variable - the current belief of the countries of the state of the world being good - is still given by Equation (1). I restrict my attention to symmetric MPE of the game. The main reason for focusing on symmetric equilibria is that introducing only a slight convexity in the investment cost function, would reduce exclude asymmetric equilibria in the model. [4]

First, consider the choice of energy consumption of country i given investments, within a time interval $[t, t+dt]$. country i solves the following maximization problem:

$$\max_{e_i, f_i} \left\{ \left[u(e_i) - D \sum_j f_j - cI_i \right] dt \right\}$$

subject to $e_i = z_i + f_i$ and $z_{it} = \int_0^t e^{-\delta(t-s)} A_{is}\ ds$, which gives the following first-order condition:

$$u'(e_i^E) = D,$$

where superscript E denotes the MPE. Each country chooses energy consumption such that his marginal utility equals his (private) marginal damage, which defines the equilibrium consumption levels, $e_i^E = e^E$ and $f_i^E(z_i) = f^E(z_i) = e_i^E - z_i\ \forall i$. Define $U^E \equiv u(e^E) - DNe^E$, to simplify notation in the following.

The standard market failure of a common bad is apparent here; each country chooses energy consumption without taking the damage inflicted on other countries into account, and the

[4] See Keller et al. [2005] for a characterization of asymmetric equilibria in a similar game, without the payoff externality introduced in this paper.

result is consumption above the first-best level; $e^E > e^*$. The utility derived from energy consumption, net of climate damages, is lower than in the first-best solution; $U^E < U^*$. However, note that the immediate benefit from an innovation for one country - lower energy consumption for this country and hence lowered damages for all countries - is the same as in the first-best benchmark:

The dynamic maximization problem of country i, when there has been no innovations so that $p < 1$, is given by:

$$V(p) = \max_{I_i \in [0,1]} \left\{ \left[U^E - cI_i + p\lambda I D\hat{A} \right] dt + e^{-rdt} E\big[V(p_{t+dt})|p, I\big] \right\}.$$

As for the social planner's problem in Section 2.2, the dynamic problem for a single country can be rewritten. First, the expectation can be expressed by:

$$E\big[V(p_{t+dt})|p, I\big] = p\lambda I dt V(1) + (1 - p\lambda I dt)V(p + dp).$$

Then, making use of the fact that for $dt \to 0$, e^{-rdt} can be replaced by $1 - rdt$ and $V(p + dp)$ by $V(p) + dpV'(p)$, while $dp = -\lambda p(1-p)Idt$, and simplifying, the maximization problem for country i in continuous time becomes:

$$\begin{aligned} rV(p) =& U^E + I_{-i}\lambda p\big[V(1) - V(p) - (1-p)V'(p) + D\hat{A}\big] \\ &+ \max_{I_i} I_i \left\{ \lambda p\big[V(1) - V(p) - (1-p)V'(p)\big] - \left[c - \lambda p D\hat{A}\right] \right\}, \end{aligned} \tag{7}$$

where $I_{-i} \equiv \sum_{j \neq i} I_j$. If an innovation is observed and the state variable jumps to 1, the value function for country i becomes:

$$\begin{aligned} V(1) =& \int_0^\infty e^{-rt} \left[U^E - c + \lambda N D\hat{A} \right] dt \\ =& \frac{1}{r} \left[U^E - c + \lambda N D\hat{A} \right] \end{aligned}$$

An increase in the investment of country i increases country i's utility both by increasing the probability of an innovation today, valued by $D\hat{A}$, and by increasing the amount of information generated about the state of the world. More information can either take the form of a jump in p to 1, or the form of a faster downwards adjustment in p. In both cases, the information is valuable. The individual maximization problem stated in the above deviates from the social planner's problem in that country i only takes his own value of an innovation into account, both in terms of reduced climate damage and information generation.

Let I_i^E, I_{-i}^E and I^E denote the MPE investment levels. Given the problem in Equation (7), it is clear that country i will choose investment levels as follows:

$$I_i^E \begin{cases} = 1 & \text{if } \lambda p\big[V(1) - V(p) - (1-p)V'(p)\big] > \Big[c - \lambda p D\hat{A}\Big] \\ = \text{ any } I_i \in [0,1] & \text{if } \lambda p\big[V(1) - V(p) - (1-p)V'(p)\big] = \Big[c - \lambda p D\hat{A}\Big] \\ = 0 & \text{if } \lambda p\big[V(1) - V(p) - (1-p)V'(p)\big] < \Big[c - \lambda p D\hat{A}\Big] \end{cases}, \tag{8}$$

If the gain from information generation exceeds the net cost of investment today, country i will invest, while if the gain falls short of the net cost, he will not. If the gain equals the cost, country i is indifferent between investing any amount, or not investing.

If country i does not invest ($I_i = 0$) it follows from Equation (7) that the value function, $V^0(p)$, is given by:

$$\begin{aligned} &I_{-i}\lambda p(1-p)V_0'(p) + (r + I_{-i}\lambda p)V_0(p) \\ &\qquad = \frac{1}{r} I_{-i}\lambda p \lambda N D\hat{A} + \frac{1}{r} I_{-i}\lambda p(U^E - c) + U^E + I_{-i}\lambda p D\hat{A}. \end{aligned}$$

For values of p such that no country invests we will have $I_{-i} = 0$, which gives:

$$V_0^E(p) = \frac{1}{r} U^E.$$

On the other hand, if i does invest ($I_i = 1$) the value function is given by the differential equation:

$$\begin{aligned} &I\lambda p(1-p)V_1'(p) + (r + I\lambda p)V_1(p) \\ &\qquad = \left(1 + \frac{1}{r} I\lambda p\right)(U^E - c) + \left(1 + \frac{1}{r}\lambda N\right)\lambda p I D\hat{A}, \end{aligned} \tag{9}$$

which is solved by:

$$V_1(p) = C_1(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{\lambda I}} + \frac{1}{r}(U^E - c) + \frac{1}{r} p\lambda N D\hat{A} - \frac{1}{I + r/\lambda} p(N - I)D\hat{A}, \tag{10}$$

where C_1 is a constant of integration (see Appendix B.1 for calculations).

Finally, if country i is indifferent between investing and not investing it follows from the equality in (8) that the value function, $\tilde{V}(p)$, is given by the differential equation:

$$\lambda p(1-p)\tilde{V}'(p) + \lambda p\tilde{V}(p) = \lambda p\frac{1}{r}(U^E - c) - c + \left(1 + \frac{1}{r}\lambda N\right)\lambda p D\hat{A}, \tag{11}$$

which is solved by:

$$\tilde{V}(p) = (1-p)\tilde{C} + (1-p)\ln\left(\frac{1-p}{p}\right)\frac{1}{\lambda}c + \frac{1}{r}\left(U^E - c\left(\frac{r}{\lambda}+1\right)\right) + \left(1+\frac{1}{r}\lambda N\right)D\hat{A}, \quad (12)$$

where $\tilde{C}$ is a constant of integration.

Moreover, it also follows from the dynamic problem as formulated in Equation (7), given the equality in (8) that for values of p such that country i is indifferent, we must have $r\tilde{V}(p) = U^E + I_{-i}c$, which can be rewritten to

$$I^E_{-i} = \frac{r\tilde{V}(p) - U^E}{c}. \quad (13)$$

For given investments by other countries, an increase in p will increase the net return to investment for country i. In a symmetric equilibrium, it must therefore be the case that I_i is a non-decreasing function of p. Given that $I^E = N$ for $p = 1$ and $I^E = 0$ for $p = 0$, the equilibrium investment choices must be characterized by the following: Each country i invests zero for low values of p and invest one for high values of p, while there exist an interval $[\underline{p}, \overline{p}]$, with $\underline{p} \leq \overline{p}$, over which country i is indifferent between investing and not investing. Equation (13) gives the relationship between total investments and the current belief over the range where countries are indifferent. Assuming that $\tilde{V}(p)$ is an increasing function (which can be verified by taking the derivative of Equation (14)), total investments must be increasing over the range $[\underline{p}, \overline{p}]$. The intuition here is that while the marginal gain from investments in country i (in expected terms) is increasing in p, it is decreasing in I_{-i}. Hence, as p increases over the interval, the increased marginal gain is offset by an increase in investments in other countries, leaving each country indifferent.

In the following, I will characterize the threshold probabilities $\underline{p}$ and $\overline{p}$, subsequently.

Proposition 1. *In the symmetric MPE of the non-cooperative game, all countries choose no investment, $I^E_i = 0$, for all $p < \underline{p}^E$, with:*

$$\underline{p}^E = \frac{cr/\lambda}{(r+N\lambda)\,D\hat{A} - c} \in \big(p^*(N), p^*(1)\big).$$

Hence, there is a welfare loss from underinvestment over the range $(p^*(N), \underline{p}^E)$ of the current belief.

Proof. Given that (by assumption) the optimal investment levels are such that $I^E_i = 1$ if $p = 1$ and, $I^E_i = 1$ if $p = 0$, and given that we consider a symmetric equilibrium, investments must be non-decreasing in p. Then there must exist a range $[\underline{p}, \overline{p}]$ with $0 \leq \underline{p} \leq \overline{p} \leq 1$, such that country i is indifferent between investing and not investing, with $I^E_i = I^E_{-i} = 0$ at $\underline{p}^E$.

Furthermore, we must have that:

$$\tilde{V}(\underline{p}^E) = V_0(\underline{p}^E) = \frac{1}{r}U^E \qquad \text{when} \quad I_{-i} = 0 \qquad\qquad \text{(Value Matching)},$$
$$\tilde{V}'(\underline{p}^E) = 0 \qquad\qquad \text{when} \quad I_{-i} = 0 \qquad\qquad \text{(Smooth Pasting)}.$$

The threshold, $\underline{p}^E$, (together with the constant of integration $\tilde{C}$) follows from these two conditions. See Appendix B.2 for a more elaborate version of this argument.

The welfare loss follows from the fact that investments are lower than their first-best levels for all values of p such that $p^*(N) < p \leq \underline{p}^E$.

Finally, it follows from the next Lemma that $\tilde{V}(p)$ is indeed non-decreasing in p (see Appendix B.2). □

Lemma 2. *For any p such that country i is indifferent between investing and not investing, the value function for any country i in the symmetric MPE is given by:*

$$\tilde{V}^E(p) = \frac{1}{r}U^E + \frac{1}{\lambda}c\left[\frac{1}{\underline{p}^E}\left(p - \underline{p}^E\right) - (1-p)\ln\left(\frac{1-\underline{p}^E}{\underline{p}^E}\cdot\frac{p}{1-p}\right)\right]. \tag{14}$$

For $p \leq \underline{p}^E$, the equilibrium value function is given by:

$$V_0^E(p) = \frac{1}{r}U^E.$$

Proof. The expression in Equation (14) follows from Equation (12) when Value Matching and Smooth Pasting are used to solve for $\underline{p}^E$ and $\tilde{C}$. For $p \leq \underline{p}^E$, the value follows directly from Equation (7), as $I_i = I_{-i} = 0$ in equilibrium. □

The next proposition shows how the allocation of investments is determined for higher levels of p:

Proposition 2. *In the symmetric MPE of the non-cooperative game, all countries choose $I_i^E = 1$ for all $p \geq \overline{p}^E$, with $\overline{p}^E > \underline{p}^E$ defined by:*

$$\frac{r}{\lambda}c\left[\frac{1}{\underline{p}^E}\left(\overline{p}^E - \underline{p}^E\right) - (1-\overline{p}^E)\ln\left(\frac{1-\underline{p}^E}{\underline{p}^E}\cdot\frac{\overline{p}^E}{1-\overline{p}^E}\right)\right] = (N-1)c. \tag{15}$$

For $p \in (\underline{p}^E, \overline{p}^E)$, investment is given by:

$$I_i^E = \frac{r\tilde{V}^E(p) - U^E}{(N-1)c} \in (0,1) \;\;, \;\; \forall i. \tag{16}$$

Hence, there is a welfare loss from underinvestment over the range $(\underline{p}^E, \overline{p}^E)$ *of the current belief.*

Proof. The investment level given by Equation (16) follows from symmetry, $I_{-i}^E = (N-1)I_i^E$, together with the indifference condition given by Equation (13). As $\tilde{V}^E(p)$ is increasing in p investments must increase in p over the interval where the countries are indifferent, and we must have:

$$1 = \frac{r\tilde{V}^E(\overline{p}^E) - U^E}{(N-1)c},$$

in $\overline{p}^E$which by inserting for $\tilde{V}^E(\overline{p}^E)$ gives the expression in Equation (15).

The welfare loss follows from the fact that investments are lower than their first-best levels over the interval: $I < N$ for p such that $\underline{p}^E < p < \overline{p}^E$. □

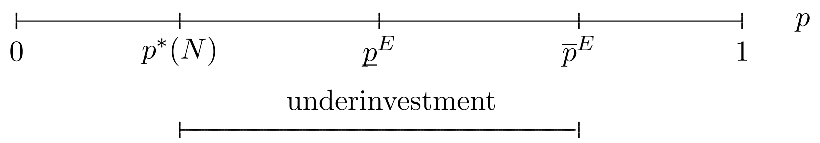
The results presented here in Proposition 2 are close to those presented in the first part of Propositon 5.1 in Keller et al. [2005], while the following lemma is comparable to the final part of the proposition.

Lemma 3. *For any* $p > \overline{p}^E$*, the value function for country* i *in the symmetric MPE,* $V_1^E(p)$*, is given by Equation* (10)*, where* $I = N$ *and* C^1 *solves* $\tilde{V}^E(\overline{p}^E) = V_1^E(\overline{p}^E)$*.*

Proof. This follows from the fact that at $\overline{p}^E$, both $\tilde{V}^E(\cdot)$ and $V_1^E(\cdot)$ represent the value function for country i (Value Matching). □

The complete allocation of investments in the symmetric MPE is now described, and Figure 1 summarizes the solution.

Figure 1: Threshold probabilities



The underinvestment arises because the value of investment to other countries - reduced climate damage in expectation and higher information generation - is not taken into account when investments are determined in country i. Furthermore, the externality from information generation is higher because the innovation is a public good. The information generation from investment in country i is beneficial to another country j not only because j can adjust his own investment, but also because it is beneficial to j that all $k \neq j$ will increase their investments following a jump in p resulting from i's investments.

2.4 Optimal regulation

In absence of the information externality inherent in this model, the efficient regulation of the climate externality is well-known, and described already by Pigou [1920]: A price on emissions should be introduced so that each country faces a cost of fossil energy consumption internalizing the damage their own consumption inflicts on other countries. Given a per unit tax τ on consumption of fossil energy, the static optimization problem of country i is given by:

$$\max_{e_i, f_i} \left\{ \left[u(e_i) - D \sum_j f_j - \tau f_i - cI_i \right] dt \right\}$$

subject to $e_i = z_i + f_i$ and $z_{it} = \int_0^t e^{-\delta(t-s)} A_{is}\, ds$. The solution to this problem is given by the following first-order condition:

$$u'(e_i^\tau) = D + \tau \quad \Rightarrow \quad e_i^\tau = e^\tau \text{ and } f_i^\tau(z_i) = f^\tau(z_i) = e^\tau - z_i \quad \forall i,$$

where superscript τ denotes MPE values given the tax. When the tax is in place, consumption reduction in country i is beneficial to i not only in terms of reduced climate damage, but also in terms of lower tax payments. This additional gain from consumption reduction is of course private: it accrues only to country i. In line with the reasoning of Pigou, the first-best consumption levels - given investments - are implemented when $\tau = \tau^* \equiv (N-1)D$.

For any tax, let $U^\tau \equiv u(e^\tau) - DNe^\tau - \tau e^\tau$. Then, when $p < 1$, the dynamic optimization problem of country i, given τ, is the following:

$$V^\tau(p) = \max_{I_i \in [0,1]} \left\{ \left[U^\tau - cI_i + p\lambda I_i \tau \hat{A} + p\lambda I D \hat{A} \right] dt + e^{-rdt} E\left[V^\tau(p_{t+dt})|p, I\right] \right\},$$

with

$$E\left[V^\tau(p_{t+dt})|p, I\right] = p\lambda I dt V^\tau(1) + (1 - p\lambda I dt) V^\tau(p + dp),$$
$$V^\tau(1) = \frac{1}{r} \left[U^\tau - c + \lambda \tau \hat{A} + \lambda N D \hat{A} \right].$$

Using the same method as in the basic model in Section 2.3, the continuous time version of the problem is given by:

$$\begin{aligned} rV^{\tau}(p) =& U^{\tau} + I_{-i}\lambda p\big[V^{\tau}(1) - V^{\tau}(p) - (1-p)V'^{\tau}(p) + D\hat{A}\big] \\ &+ \max_{I_i} I_i\Big\{\lambda p\big[V^{\tau}(1) - V^{\tau}(p) - (1-p)V'^{\tau}(p)\big] - \Big[c - \lambda p\tau\hat{A} - \lambda pD\hat{A}\Big]\Big\}. \end{aligned} \tag{17}$$

The value for country i of an innovation is higher with a higher tax, because an innovation will give lower fossil consumption and hence lower tax payments, now and in the future. In addition, the information generated from investment is more valuable when the tax is in place - to the country itself and to all other countries. The threshold probabilities, $\underline{p}^{\tau}$ below which there will be no investment, and $\overline{p}^{\tau}$ above which there will be full investment, can be calculated using exactly the same methods as in Section 2.3, and are presented in the following results. Let I_i^{τ} denote the MPE investment in country i given the tax τ.

Lemma 4. *In the symmetric MPE where countries pay a tax, $\tau > 0$, per unit of fossil energy consumption, all countries choose no investment, $I_i^{\tau} = 0$, when $p < \underline{p}^{\tau}$, and all countries choose full investment, $I_i^{\tau} = 1$, when $p > \overline{p}^{\tau}$, with:*

$$\underline{p}^{\tau} = \frac{cr/\lambda}{(r+\lambda)\,\tau\hat{A} + (r+N\lambda)\,D\hat{A} - c}, \tag{18}$$

and $\overline{p}^{\tau} > \underline{p}^{\tau}$ defined by:

$$\frac{r}{\lambda}c\left[\frac{1}{\underline{p}^{\tau}}\left(\overline{p}^{\tau} - \underline{p}^{\tau}\right) - \left(1-\overline{p}^{\tau}\right)\ln\left(\frac{1-\underline{p}^{\tau}}{\underline{p}^{\tau}}\cdot\frac{\overline{p}^{\tau}}{1-\overline{p}^{\tau}}\right)\right] = (N-1)(c - \lambda\overline{p}^{\tau}\tau\hat{A}). \tag{19}$$

For $p \in (\underline{p}^{\tau}, \overline{p}^{\tau})$, investment is given by:

$$I_i^{\tau} = \frac{r\tilde{V}^{\tau}(p) - U^{\tau}}{(N-1)(c - \lambda p\tau\hat{A})} \in (0,1)\ \ ,\ \ \forall i. \tag{20}$$

Proof. Inserting $I_i = 0$ and $I_i = 1$, respectively, in Equation (17) and solving the resulting differential equations define $V_0^{\tau}(p)$ and $V_1^{\tau}(p)$, exactly as in the basic model presented in Section 2.3. Furthermore, the following condition must hold if country i is indifferent between no investment and full investment:

$$\lambda p\big[V^{\tau}(1) - V^{\tau}(p) - (1-p)V^{\tau\prime}(p)\big] - \Big[c - \lambda p\tau\hat{A} - \lambda pD\hat{A}\Big], \tag{21}$$

and inserting this condition into (17) defines $\tilde{V}^{\tau}(p)$, and and the investment level given in

Equation (20). In $\underline{p}^\tau$ we must have that

$$\tilde{V}^\tau(\underline{p}^\tau) = V_0^\tau(\underline{p}^\tau) = \frac{1}{r}U^\tau \qquad \text{when} \quad I_{-i} = 0 \qquad\qquad \text{(Value Matching)},$$
$$V_0^{\tau\prime}(\underline{p}^\tau) = 0 \qquad\qquad \text{when} \quad I_{-i} = 0 \qquad\qquad \text{(Smooth Pasting)},$$

which gives $\underline{p}^\tau$ as expressed in Equation (18), together with an explicit solution for $\tilde{V}^\tau(p)$ (see Lemma 5).

Finally, inserting $I_i = 1$, together with the solution for $\tilde{V}^\tau(p)$ and $p = \overline{p}^\tau$ in Equation (20) gives Equation (19). □

Lemma 5. *The value function for country i in the symmetric MPE where the countries pay the tax, $\tau > 0$, per unit of fossil energy consumption is given by:*

$$V_0^\tau(p) = \frac{1}{r}U^\tau \qquad \textit{when } p < \underline{p}^\tau,$$
$$\tilde{V}^\tau(p) = \frac{1}{r}U^\tau + \frac{1}{\lambda}c\left[\frac{1}{\underline{p}^\tau}(p - \underline{p}^\tau) - (1-p)\ln\left(\frac{1-\underline{p}^\tau}{\underline{p}^\tau}\cdot\frac{p}{1-p}\right)\right] \qquad \textit{when } p \in [\underline{p}^\tau, \overline{p}^\tau],$$
$$V_1^\tau(p) = C_1^\tau(1-p)\left(\frac{1-p}{p}\right)^{\frac{r}{\lambda N}} + \frac{1}{r}(U^\tau - c) + \frac{1}{r}p\lambda(\tau\hat{A} + ND\hat{A}) \qquad \textit{when } p > \overline{p}^\tau,$$

where C_1^τ solves $V_1^\tau(\overline{p}^\tau) = \tilde{V}^\tau(\overline{p}^\tau)$.

Proof. The expressions for $V_0^\tau(p)$ and $\tilde{V}^\tau(p)$ follow directly the proof of Lemma 4. $V_1^\tau(p)$ follows from Equation (17) when $I_i = 1\ \forall i$. □

We can now consider the investment levels prevailing when the Pigouvian tax is in place, compared to the first-best investment levels. Let superscript τ^* denote values when the tax is set at $(N-1)D$.

Proposition 3. *The welfare loss from underinvestment is not eliminated by the Pigouvian tax, $\tau^* = (N-1)D$, on fossil energy consumption: there is still underinvestment for current beliefs $p \in (p^*(N), \overline{p}^{\tau^*})$.*

Proof. It follows from Lemma 4, when inserting τ^* for τ, that $p^*(N) < \underline{p}^{\tau^*} < \overline{p}^{\tau^*}$. □

The full value of investment in country i is not internalized by the Pigouvian tax. There are three ways in which other countries are positively affected by investments in country i, all in terms of expectation. Firstly, the investment increases the probability of an innovation which will lower consumption in country i today and in all future periods. This direct effect is internalized by the Pigouvian tax. Secondly, the investment generates information that is valuable to all countries. And finally, by generating information, country i increases the expected future investment levels in all countries, which benefits everyone. These last to effects are *not* internalized by the Pigouvian tax. Therefore, additional policy is needed.

Intuitively, an investment subsidy, dependent on on the state, p, could implement the first-best investment levels. Denote this subsidy $\psi(p)$.

Given the Pigouvian tax *and* such a subsidy scheme, the dynamic optimization problem of country i for $p < 1$ is given by

$$\begin{aligned} rV^{\tau^*\psi}(p) = U^{\tau^*} &+ I_{-i}\lambda p\big[V^{\tau^*\psi}(1) - V^{\tau^*\psi}(p) - (1-p)V'^{\tau^*\psi}(p) + D\hat{A}\big] \\ &+ \max_{I_i} I_i \Big\{ \lambda p\big[V^{\tau^*\psi}(1) - V^{\tau^*\psi}(p) - (1-p)V'^{\tau^*\psi}(p)\big] \\ &\qquad - \Big[c - \psi(p) - \lambda p \tau^* \hat{A} - \lambda p D\hat{A}\Big] \Big\}. \end{aligned} \tag{22}$$

The net cost of investment is now given by $c - \psi(p)$, and the value function given that a successful innovation is observed will potentially be affected by the subsidy:

$$V^{\tau^*\psi}(1) = \frac{1}{r}\Big[U^{\tau^*} - (c - \psi(1)) + \lambda\tau^*\hat{A} + \lambda N D\hat{A}\Big].$$

As before, inserting $I_i = 0$ and $I_i = 1$, respectively, in Equation (22) and solving the resulting differential equations define $V_0^{\tau^*\psi}(p)$ and $V_1^{\tau^*\psi}(p)$. Using Value Matching (as before) at $p = \underline{p}^{\tau^*}$: $V_0^{\tau^*\psi}(\underline{p}^{\tau^*}) = \frac{1}{r}U^{\tau^*}$ for $I_{-i} = 0$, gives:

$$V_0^{\tau^*\psi}(p) = \frac{1}{r}U^{\tau^*} + \frac{I}{I + r/\lambda} p \left(\frac{1}{r}\Big(\lambda N D\hat{A} + \lambda\tau^*\hat{A} - (c - \psi(1))\Big) + D\hat{A} \right) \tag{23}$$

A subsidy implementing the first-best investment level must ensure that the value function for country i given full investment, $V_1^{\tau^*\psi}(p)$, is at least as high as $V_0^{\tau^*\psi}(p)$, when $I_{-i} = N-1$, for all current beliefs $p \geq p^*(N)$. Define the optimal subsidy path as the lowest possible subsidy for every p that achieves this goal. The following proposition characterizes this optimal path of the subsidy.

Proposition 4. *The subsidy path, $\psi^*(p)$, needed per unit of investment in order to implement the first-best investment levels is given by:*

$$\begin{aligned} &\psi^*(p) = 0 && \textit{for } p \leq p^*(N) \textit{ and } p \geq \overline{p}^{\tau^*}, \\ &\psi^*(p) = c - \lambda p\tau^*\hat{A} - \frac{rV_0^{\tau^*\psi}(p) - U^{\tau^*}}{N-1} > 0 \quad \textit{with } I_{-i} = N-1, && \textit{for } p \in (p^*(N), \overline{p}^{\tau^*}). \end{aligned} \tag{24}$$

Over the range where the subsidy is positive, it is decreasing in the perceived probability, p, of a successful innovation: $\psi^{*\prime}(p) < 0$.

Proof. For $p \leq p^*(N)$ and $p \geq \overline{p}^{\tau^*}$ the result is trivial: when this is the case, the investment level in the MPE with $\tau = \tau^*$ and $\psi(p) = 0$ is equal to the first-best investment level.

For values of $p \in (p^*(N), \overline{p}^{\tau^*})$, we must have:

$$I_{-i} = \frac{r\tilde{V}^{\tau^*\psi}(p) - U^{\tau^*}}{c - \psi(p) - \lambda p \tau^* \hat{A}} \tag{25}$$

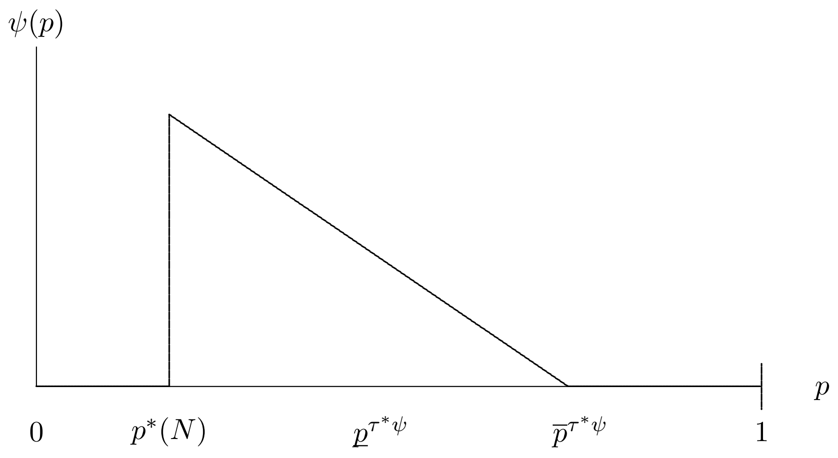
for country i to be indifferent (and hence willing to invest fully). When country i is indifferent, we also have:

$$V_0^{\tau^*\psi}(p) = \tilde{V}^{\tau^*\psi}(p) = V_1^{\tau^*\psi}(p).$$

We can then insert $V_0^{\tau^*\psi}(p)$ for $\tilde{V}^{\tau^*\psi}(p)$ together with $I_{-i} = N - 1$ into Equation (25). The expression in Equation (24) follows. And it follows directly from this expression that $\psi^{*\prime}(p) < 0$. □

The path of the optimal subsidy is illustrated in Figure 2. As is stated in the proposition, the subsidy is decreasing in the current belief, p, over the range where it is positive. In other words, the less promising the technology is within the range of current beliefs $p \in (p^*(N), \overline{p}^{\tau^*})$, the higher is the optimal subsidy. Or, if more than one technology of the type discussed in this paper is available, the optimal subsidy is highest for the least promising technologies among those that are subsidized.

Figure 2: The optimal subsidy



The intuition for this final result is simply that when the potential of a technology is low, a higher subsidy is necessary for the private value of investment to be sufficiently high.

3 Conclusion

Building on the literature on bandits and experimentation, this paper sheds light on a source of underinvstment in green technology that has not received much attention so far. The paper shows that a Pigouvian tax is not sufficient to implement first-best investment levels and that the path-dependent investment subsidy necessary is higher for technologies with lower potential, within the range where the subsidy is used.

These findings are important in light of the current situation with little international cooperation on climate policies, making unilateral policies and technological development central. Understanding what characterizes technologies where the underinvestment problems are most severe is of high value, especially taking into account the large amount of money that governments in many countries put into research and development of green technologies. The findings of this paper suggest that in addition to direct technological spillovers between firms and countries, uncertainty regarding the potential of new developments for different technologies is a source of underinvestment. Furthermore, the findings imply that the largest investment subsidies are needed for technologies where the uncertainty is relatively large.

This paper is still in progress, and there are several paths I consider interesting to follow within this framework. Firstly, it is not necessarily so that innovations are always observable in all countries, and the dynamics of the model might change if countries make inference based on investment levels without observing successful innovations in other countries. Rosenberg et al. [2007] consider private information on own payoffs in a model without payoff externalities, and I hope to be able to extend my model in this direction. Similarly, including the the possibility of imperfect correlation in successes, or the possibility to conceal information regarding own investment (see Heidhues et al. [2012]), may also give insight to interesting mechanisms.

It would also be useful to extend the model in several directions, in order to discuss investment in specific technologies. In Keller and Rady [2010] the model of Keller et al. [2005] is extended to a setting where there can be positive returns on investments in all states of the world, which for some technologies will be the case. This is thus an interesting extesion to do also in the model presented in this paper. The same is true for the development of the model made in Klein and Rady [2011], who consider bandits that are negatively correlated across countries. If the potential of a technology depends on physical characteristics of a country for example, there may exist technologies where information about successful innovations in one country is bad news in other countries.

Finally, related to the huge literature on climate change within economics, I also believe that including international negotiations in the model could provide new insights. The optimal design of a climate treaty should account for information externalities such as those considered in this paper, and for how they interact with the basic market failure of fossil energy consumption.

A The Cooperative Solution

A.1 Value function under aggregate welfare maximization

Equation (4) can be rewritten as follows:

$$\begin{aligned}&W'(p)N\lambda p(1-p) + W(p)(r + N\lambda p)\\&\qquad = N\left[p(\lambda H + \lambda NDA)\left(1 + N\frac{\lambda}{r}\right) + (U^* - c)\left(1 + Np\frac{\lambda}{r}\right)\right].\end{aligned}$$

This differential equation is solved by the following:

$$W(p) = e^{-\int f(p)dp}\left(C^* + \int e^{\int f(p)dp} g(p)dp\right),$$

where

$$\begin{aligned}f(p) &\equiv \frac{r + N\lambda p}{N\lambda p(1-p)}\\ g(p) &\equiv \frac{N}{N\lambda p(1-p)}\left[p(\lambda H + \lambda NDA)\left(1 + N\frac{\lambda}{r}\right) + (U^* - c)\left(1 + Np\frac{\lambda}{r}\right)\right].\end{aligned}$$

A.2 Proof of Lemma 1

The threshold $p^*(N)$ is calculated using the following two conditions:

- Welfare must be independent of the investment level *at* the threshold:

  $$rW^N(p^*(N)) = rW^0(p^*(N)) \quad \text{(Value Matching)}.$$

  This condition is intuitive: at the threshold, the social planner must be indifferent between choosing $I = 0$ and $I = N$, and then it cannot be the case that welfare is higher in one of the two cases.

- The change in welfare when the current belief increases *at* the threshold must be zero:

  $$rW^{N\prime}(p^*(N)) = 0 \qquad \text{(Smooth Pasting)}.$$

  This condition must hold because it cannot be the case that welfare is either decreasing or increasing in p when the social planner is indifferent between no investment and full investment: Choosing no investment will give the same innovation independently of p, and a decrease in p can hence never *reduce* welfare. On the other hand, a decrease in welfare following an increase in p is obviously not coherent with welfare maximizing investments.

Using the expression in Equation (5) gives the following derivative:

$$W^{N\prime}(p) = \left[\frac{1}{r}(\lambda H + \lambda NDA) - C^{*}\left(\frac{1-p}{p}\right)^{\frac{r}{\lambda N}}\left(1+\frac{r}{\lambda Np}\right)\right]N.$$

Smooth Pasting and Value Matching then gives the expression for the threshold provided in the Lemma. Solving for the constant C^{*}, and inserting this into Equation (5) gives the explicit solution for the value function (Equation (6)).

B The Non-Cooperative Solution

B.1 Value functions in the non-cooperative game

The differential equation in (9) is solved by:

$$V^{1}(p) = e^{-\int f^{1}(p)dp}\left(C^{1} + \int e^{\int f^{1}(p)dp} g^{1}(p)dp\right),$$

where

$$f^{1}(p) \equiv \frac{r + I\lambda p}{I\lambda p(1-p)}$$
$$g^{1}(p) \equiv \frac{1}{I\lambda p(1-p)}\left[\left(1+\frac{\lambda}{r}I\right)p(\lambda H + \lambda NDA) + \left(1+\frac{\lambda}{r}Ip\right)(U^{E}-c) - \lambda p(N-I)DA\right],$$

which gives Equation (10).

Equivalently, $\tilde{V}(p)$ in Equation (12) is solved for by using the following:

$$\tilde{V}(p) = e^{-\int \tilde{f}(p)dp}\left(\tilde{C} + \int e^{\int \tilde{f}(p)dp}\tilde{g}(p)dp\right),$$

where, from (11);

$$\tilde{f}(p) \equiv \frac{1}{1-p}$$
$$\tilde{g}(p) \equiv \frac{1}{\lambda p(1-p)}\left[\left(1+\frac{\lambda}{r}\right)p(\lambda H + \lambda NDA) - c + \frac{\lambda}{r}p(U^{E}-c) - \lambda p(N-1)DA\right].$$

which gives Equation (12).

B.2 Proof of Proposition 1

Value Matching and Smooth Pasting must hold at the threshold $\underline{p}^E$, and the intuition is similar to that for the cooperative solution:

- Firstly, at the threshold where country i is indifferent, it cannot be the case that his value function is higher in one of the two cases ($I_i = 1$ or $I_i = 0$) than in the other. Furthermore, the lowest possible value for p at which country i can be indifferent, must be the one where $I_{-i} = 0$, given Equation (13). Value Matching for $I_{-i} = 0$ at $\underline{p}^E$ follows.
- Secondly, it cannot be the case that the value function of country i is increasing or decreasing at the threshold, as long as there is zero investment by others. Since country i can immediately change from $I_i > 0$ to $I_i = 0$ should he play $I_i > 0$ and see a decrease in p, this cannot lower his value further, if it is the case that $I_i = 0$ is a utility maximizing choice in the first place. Furthermore, a decrease in the value function following an increase in p cannot be the case if there is no investment by others at the threshold ($I_{-i} = 0$ at $p = \underline{p}^E$), and country i's strategy is utility maximizing.

Equation (12) gives:

$$\tilde{V}'(p) = -\tilde{C} - \frac{c}{\lambda} \ln\left(\frac{1-p}{p}\right) - \frac{c}{\lambda p},$$

and the two conditions can then be used to solve for $\underline{p}^E$ and $\tilde{C}$.

Finally, taking the derivative of the equilibrium value function in Equation (14), gives:

$$\tilde{V}^{eq\prime}(p) = \frac{c}{\lambda}\left[\ln\left(\frac{1-\underline{p}^E}{1-p}\frac{p}{\underline{p}^E}\right) + \left(\frac{1}{\underline{p}^E} - \frac{1}{p}\right)\right],$$

which is ≥ 0 for $p \geq \underline{p}^E$.

References

Julien Beccherle and Jean Tirole. Regional initiatives and the cost of delaying binding climate change agreements. *Journal of Public Economics*, 95(11-12):1339–1348, 2011.

Dirk Bergemann and Ulrich Hege. Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance*, 22(6):703–735, 1998.

Dirk Bergemann and Ulrich Hege. The financing of innovation: learning and stopping. *RAND Journal of Economics*, 36(4):719–752, 2005.

Dirk Bergemann and Juuso Välimäki. Bandit problems. *The new Palgrave Dictionary of Economics*, (ed. by S.N.Durlauf and L.E.Blume):336–340, 2010.

Patrick Bolton and Christopher Harris. Strategic experimentation. *Econometrica*, 67(2): 349–374, 1999.

Patrick Bolton and Christopher Harris. Strategie experimentation: the undiscounted case. *Incentives, Organizations and Public Economics - Papers in Honour of Sir James Mirrlees*, (ed. by P.J.Hammond and G.D.Myles Oxford: Oxford University Press):53–68, 2001.

Wolfgang Buchholz and Kai A. Konrad. Global environmental problems and the strategic choice of technology. *Journal of Economics*, 60(3):299–321, 1994.

Rolf Golombek and Michael Hoel. Climate Policy under Technology Spillovers. *Environmental & Resource Economics*, 31(2):201–227, 2005.

Garrett Hardin. The tragedy of the commons. *Science*, 162(3859):1243–1248, 1968.

Bård Harstad. The Dynamics of Climate Agreements. *Journal of the European Economic Association*, forthcoming, 2015.

Paul Heidhues, Sven Rady, and Philipp Strack. Strategic experimentation with private payoffs. Technical report, SFB/TR 15 Discussion Paper No. 307, 2012.

Godfrey Keller and Sven Rady. Strategic experimentation with poisson bandits. *Theoretical Economics*, 5(2):275–311, 2010.

Godfrey Keller, Sven Rady, and Martin Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

Nicolas Klein and Sven Rady. Negatively correlated bandits. *The Review of Economic Studies*, 78(2):693–732, 2011.

A. C. Pigou. The economics of welfare. *London: Macmillan, 4th. edn.*, 1920.

Dinah Rosenberg, Eilon Solan, and Nicolas Vieille. Social learning in one-arm bandit problems. *Econometrica*, 75(6):1591–1611, 2007.

Michael Rothschild. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9(2):185–202, 1974.

Paul A. Samuelson. The pure theory of public expenditure. *The Review of Economics and Statistics*, 36(4):pp. 387–389, 1954.

United Nations Environment Programme UNEP. Global trends in renewable energy investment 2015. *Frankfurt School Collaborating Centre for Climate and Sustainable Energy Finance and Bloomberg New Energy Finance*, 2015.