Project Mechanisms and Technology Diffusion in Climate Policy

March 15, 2010

Abstract

The paper deals with the diffusion of GHG mitigation technologies in developing countries. We develop a model where an abatement technology is progressively adopted by firms and we use it to compare the Clean Development Mechanism (CDM) with a standard Cap and Trade scheme (C&T). In the presence of learning spillovers, we show that the CDM yields a higher social welfare than C&T if the first adopter receives CDM credits whereas the followers do not.

The analysis leads us to suggest two CDM design improvements: relaxing the additionality constraint for projects which generate significant learning externalities, and allowing collective CDM projects which could internalize learning externalities.

Keywords : climate policy, technology diffusion, Kyoto Protocol, Clean Development Mechanism, emissions trading

JEL code: H87, O33, Q55, Q54

1 Introduction

Due to economic growth, developing countries are expected to overtake industrialized countries as the leading source of greenhouse gases (GHG) in the medium or long term. The transfer and diffusion of climate-friendly technologies in these economies is seen as a key means for solving the climate change problem. Accordingly, technology issues are included in both the United Nations Framework Convention on Climate Change (UNFCCC) and its Kyoto Protocol. The Asia-Pacific Partnership on Clean Development and Climate initiated by the U.S. administration in 2005 also places a very strong emphasis on the development and sharing of more efficient energy technologies.

The Clean Development Mechanism (CDM) is considered by many as an important tool to stimulate technology transfer and diffusion. It is an arrangement under the Kyoto Protocol allowing industrialized countries with a greenhouse gas reduction commitment (the so-called Annex 1 countries) or firms located in these countries to invest in emission reducing projects in countries that have not made such commitments (the Annex 2 countries). These projects, usually carried out in developing countries, provide a cheaper alternative to costly emission reductions in industrialized countries. They also assist host countries in achieving sustainable developement. The CDM can finally contribute to international technology transfer by financing projects using technologies not available in the host countries and to technology diffusion within the countries.[1] Such transfers have gradually gained in importance in policy debates; they are in particular at the core of ongoing talks on the Post-Kyoto regime.

In this paper we develop a model to study whether emissions trading can yield the socially optimal path of technology diffusion[2]. We focus on the CDM,

[1] It is worth noting that the CDM does not have an explicit technology transfer and diffusion mandate under the Kyoto Protocol. But the CDM is clearly linked to the technological issue in the policy debate on climate change, and in particular, in post-Kyoto talks.

[2] The model implicitly relies on a relatively narrow definition of technology transfer. In

whose specificity lies in the additionality requirement, which is that firms may implement a CDM project only if emissions credits make the project profitable. In order to investigate the impact of additionality, we compare the CDM with the traditional Cap and Trade program (C&T) in which any abatement - whether privately profitable or not - makes emissions credits available. We then characterize two ways to improve the efficiency of the CDM: relaxing the additionality conditions and allowing collective CDMs.

The model describes n firms located in a host country which initially operate with an old technology. These firms could adopt a cleaner technology simultaneously or sequentially. The first adoption is usually the international technology transfer, since the technology was not previously available in the country. Subsequent adoptions correspond to the diffusion within the host country.

Adoption entails a fixed cost which endogenously decreases once the technology has been introduced in the host country. In reality, that decrease may occur because observing the outcome of the first adoption of the new technology reduces uncertainty about technology benefits for subsequent adopters (called "followers"), or the first company to adopt - the "leader" - accumulates learning-by-doing skills which diffuse through various channels (e.g. labor market) to potential adopters.

These learning spillovers generate two types of inefficiency. The first is the standard under-provision problem. The propensity for a firm to take the lead in adopting new technology is low, since firms who may consider the step fail to take into account positive externalities, thereby hindering technology transfer. The second inefficiency is a coordination problem that results from the dynamic character of the diffusion process. All firms would prefer to follow so that they can enjoy a reduced adoption cost. But following requires that one firm take the lead. This is a dynamic version of a "chicken game" where both firms derive a

particular, it is not designed to analyze issues such as human capacity building.

positive benefit from adoption but have conflicting views on who should go first. One possible outcome is that, although the first adoption may be (privately and socially) profitable, it is delayed.

We show that a standard C&T scheme does not implement the first best diffusion path. One reason is that each adopter receives the same number of credits whatever its adoption rank. The Clean Development Mechanism differs from Cap & Trade program, in that the credit price signal is not uniform across all firms - it is zero for non-additional projects. Not surprisingly, the welfare impact of the additionality requirement is clearly detrimental when adoption by the leader is not additional, since technology diffusion is too slow in the absence of credits. However, the CDM unambiguously improves welfare when the leader receives credits while the others do not. This is so because it reduces the followers' advantage, thereby mitigating the coordination problem.

This leads us to suggest two improvements in the design of the CDM: first, relaxing the additionality requirement imposed on the leader, by granting credits for the first adoption even though it would have been profitable, and second, allowing collective CDM projects. The former may be viewed as the classical Pigovian subsidy while the second may be viewed as the classical Coasean solution to the externality problem.

These results provide new insights on the link between CDM and technology diffusion, leading to operational policy lessons. A growing literature provides evidence of the effect of the CDM on technology transfers across and within countries (See Schneider et al. 2008 for a recent review). Based on a data set of about 600 CDM projects, Dechezlepretre et al. (2008) show for instance that about 44% of projects exhibit a transfer–a ratio that increases with project size and varies across sectors. Other papers focus on the barriers to technology transfers through the CDM. The large transaction costs entailed by the CDM are often cited as a significant barrier (Michaelowa et al., 2003). These costs are largely due to the adverse selection issues pertaining to the choice of projects and the costly screening and monitoring mechanisms they require (Millock, 2002; see also Liski and Virrankoski (2004) for a theoretical analysis of the consequences arising from transaction costs.

The problem of adverse selection has spurred many debates on additionality (see, for instance, Greiner & Michaleowa, 2003). A severe enforcement of additionality is considered a safeguard to avoid crediting projects that would have been carried out anyway. By taking into account the dynamics of technology diffusion, our analysis yields an argument that goes in the opposite direction. Concerns about CDM administration costs have also been the main argument in favor of collective (sectoral) CDM. We stress the fact that they can also organize technology diffusion when it involves learning externalities.

Apart from the specific policy literature on CDM, an important strand of theoretical literature has developed on environmental innovation and policy instruments (Jaffe & Stavins, 1995; Laffont & Tirole, 1996; Requate, 1998; Montero, 2002; Fischer, Parry, & Pizer, 2003). Like our approach, such studies show how environmental policy instruments can solve two externalities: the environmental externality and the externality associated with knowledge. There are, however, two major differences. First, the scope is broader: most papers compare different policy instruments. Second, except for Jaffe and Stavins (1995) and Milliman and Prince (1989), they pay little attention to technology diffusion and ignore learning spillovers, which are central in our own analysis.[3] An exception is a theoretical analysis by Golombek and Hoel (2005) who analyze the interplay between technology spillovers and environmental externalities in climate treaties. But their focus is broader than ours: they compare an interna-

[3] Blackman (1999) surveys the general economic literature on technology diffusion in order to derive lessons for climate policy.

tional agreement with a common tax and an agreement with tradable quotas. They do not compare different emissions trading schemes.

Our paper is also related to the literature on technology diffusion in industrial organization (see Hoppe, 2002, for a good survey). This literature aims to explain why new technologies diffuse only progressively. In most papers the timing of adoption depends on a trade-off between adoption costs that are exogenously decreasing with time and the competitive advantage of adopting a technology early (Reinganum, 1981; Fudenberg and Tirole, 1985). We depart from this pattern by endogenizing the decrease of the adoption cost and by undertaking a normative analysis of the optimal path of technology diffusion.

The paper is organized as follows. Section 2 presents a model of technology adoption by n firms, and characterizes the socially optimal technology diffusion path. Section 3 characterizes diffusion patterns under a C&T scheme. We investigate the CDM in Section 4 and compare it with C&T in Section 5. We analyze relaxed additionality and collective CDM projects in Section 6. Section 7 concludes.

2 Model and social optimum

We use a simple continuous time model that describes the adoption of a GHG mitigation technology by n symmetric firms under emissions trading.

2.1 Firms' payoffs

At the beginning of the game, firm i derives a market profit π° per time period. When the firm adopts the abatement technology, this profit changes. Let π denote the profit flow after adoption. Adoption also reduces GHG emissions. Without loss of generality, we assume that firms emit one unit per period before adoption and zero afterwards. The fact that emission and abatement is normalized to unity do not alter any result. Under these assumptions, firm i 's profit function is:

$$\pi(e_i) = \begin{cases} \pi^\circ \text{ if } e_i = 1 & \text{(pre-adoption)} \\ \pi \text{ if } e_i = 0 & \text{(post-adoption)} \end{cases}$$

where e_i is the firm's level of emissions[4]. Importantly, adoption can increase the profit ($\pi > \pi^\circ$) or decrease it ($\pi \leq \pi^\circ$). For ease of presentation, we maintain throughout that $\pi^\circ = 0$.

Note that technology adoption by a given firm does not affect others' profits. We assume that firms operate in a perfectly competitive market where a change in the production cost of one firm has negligible impacts on other firms' level of output and profit. With this assumption, we rule out strategic market issues that are potentially associated with technology adoption. It therefore greatly simplifies our analysis and allows us to focus sharply on the issue of technology diffusion.[5]

Adopting the new technology entails a fixed cost. To capture the learning spillovers following the introduction of the technology into the host country, we make the assumption that the adoption cost starts decreasing endogenously after the first adoption. Formally, c is the cost for the first adopter while a subsequent adopter bears $ce^{-\lambda d}$ where d is the time passed since the first adoption. When $\lambda > 0$, there is an incentive for the followers to delay adoption in order to benefit from the leader's experience. When $\lambda = 0$, the adoption has no positive externality.

The technology is competitively supplied at a uniform price that we normalize to zero, meaning there is no extra cost for the adopters. What we have in mind are generic technologies which are competitively supplied. Empirical

[4]The emissions variable e_i is discrete because we assume that technology adoption is a binary choice.

[5]Introducing imperfect competition would induce a cumbersome discussion about the potential of CDM to reduce market power in the product market, whereas dealing with imperfect competition is not the prime goal of a CDM.

studies like that of Dechezlepretre et al. (2008) suggest that real-world CDM projects do not rely on advanced proprietary technologies.

We now express the net present private profits. Let T denote the date of the first adoption and v^L the payoff discounted at time T of the first adopter ignoring credit sales/purchases. We have:

$$v^L = -c + \int_0^\infty \pi e^{-rt} dt = \frac{\pi}{r} - c \tag{1}$$

where r is a discount factor per time period which reflects the cost of waiting ($r > 0$). Turning next to followers, they derive zero market profit ($\pi^\circ = 0$) before adoption (between T and $T + d$). After adoption, they derive the market profit π. Their net present payoff excluding credit sales/purchases at time T is thus:

$$v^F(d) \equiv -ce^{-(r+\lambda)d} + \int_d^\infty \pi e^{-rt} dt = e^{-rd}\left(\frac{\pi}{r} - ce^{-\lambda d}\right) \tag{2}$$

2.2 Emissions trading

Emissions generate a constant marginal damage δ. At $t = 0$, there exists a climate regime with an international Cap and Trade scheme to mitigate the environmental damage. The credit market is competitive and the market price is equal to δ. That is, we assume that the trading scheme is designed to efficiently internalize the environmental externality, but ignores the learning externality problem. The model studies how the n firms can be included into this existing scheme. We consider two scenarios:

- The n firms are directly integrated in the international Cap and Trade scheme. More specifically, each firm initially receives a quantity of credits corresponding to its pre-adoption emissions[6]. These credits can be sold at price δ on the market.[7] Hence, the firm derives a benefit δ per time period after adoption.

- Each firm can implement a CDM project whereby it initially receives a number of credits proportional to to its pre-adoption emissions, *but only if adoption is not profitable without credits.* In CDM terminology, the abatement project should be additional.[8] Of course, these credits can also been sold at price δ.

The main difference between the two scenarios is that under C&T all adopters face the same price signal δ, whereas the CDM yields a price signal to additional projects only. Analyzing this difference is a key goal of the paper.

2.3 Timing

We consider a dynamic game in continuous time where the n firms decide whether and when they adopt the abatement technology. In doing so, they take into account the other firms' adoption decisions. The game has two stages:

- The first stage determines the date T of the first adoption.

- The second stage starts at time T and concerns the $n-1$ firms that did not adopt in the first stage. More specifically the follower indexed as i selects the adoption time $T+d_i$.

Firms act strategically, which substantially influences the results. Significantly, this does not mean that the n firms operate in the same oligopolistic

[6] Other initial allocation rules would exactly lead to the same results as outcomes are driven by relative payoffs and welfare changes.

[7] We implicitly assume here that the whole amount of emissions allowances is adjusted to maintain the market price δ after the inclusion of n additional firms into the scheme.

[8] Under this scenario, the amount of allowances is not adjusted as the additionality requirement preserves the environmental integrity of the international climate regime.

product market. In our game, the firms interact with other firms that could generate positive spillovers from which they would benefit. To do so, firms must be similar from a technological point of view, but they are at the same time necessarily competitors. That they operate on the same labor market, for instance, is much more relevant, as one spillover channel is labor mobility. In fact, what we assume is that the space containing the spillovers is sufficiently small for inducing strategic decisions by the potential adopters.

2.4 The socially optimal path of adoption

We now derive what should be the welfare-maximizing adoption path. The welfare function is the sum of the discounted firms' profits and environmental benefits. Therefore we need to find T^* and $d_2^*, .., d_n^*$ which maximize

$$\begin{aligned} W(T, d_2, .., d_n) &= \int_T^{\infty} \left(v^L + \delta\right) e^{-rt} dt + \sum_{i=2}^{n} \int_{T+d_i}^{\infty} \left(v^F(d_i) + \delta\right) e^{-rt} dt \\ &= \left(\frac{e^{-rT}}{r}\right) \left[v^L + \delta/r + \sum_{i=2}^{n} \left(v^F(d_i) + \delta/r\right) e^{-rd_i}\right] \quad (3) \end{aligned}$$

where firm 1 is arbitrarily the first firm that adopts at time T while the $n-1$ others follow with a delay d_i.

Our social welfare function is highly restrictive. We ignore the impact of new technologies on consumer surpluses through the product market. We also ignore the impact of diffusion on the incentives to innovate. In fact, our welfare analysis is entirely focused on diffusion.

When considering (3), it is obvious that the optimal date of first adoption, T^*, is either 0, if the term in brackets is positive, or ∞, otherwise (diffusion should not occur). In the case it is positive, the optimal delays are all the same as followers are symmetric: $d_i^* = d_j^* = d^*$ for any $i \neq 1$ and $j \neq 1$. Accordingly,

d^* is the solution of:

$$\max_d v^F(d) + \delta e^{-rd}/r \tag{4}$$

Substituting (2) and solving for d yields:

$$d^* = \begin{cases} \frac{1}{\lambda}\ln \frac{c}{\pi+\delta}(r+\lambda) & \text{if } c > \frac{\pi+\delta}{r+\lambda} \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Equation (5) essentially says that the higher the cost of adoption and the faster its decrease over time, the less likely followers are to adopt immediately. Should they decide to wait ($d^* > 0$), the delay d^* decreases with the profit flow π and the environmental benefit δ.

Obviously, d^* is meaningful only if welfare is positive at equilibrium. Thats is, if:

$$v^L + \delta/r + (n-1)\left[v^F(d^*) + \delta e^{-rd^*}/r\right] \geq 0 \tag{6}$$

which simplifies as

$$c \leq \beta\frac{\pi+\delta}{r} \ \text{ with } \ \beta \equiv \frac{1+(n-1)\,e^{-rd^*}}{1+(n-1)\,e^{-(r+\lambda)d^*}} \geq 1. \tag{7}$$

We gather these findings in a first lemma:

Lemma 1 *The socially optimal diffusion path is the following:*

1. *If* $c \leq \frac{\pi+\delta}{r+\lambda}$*, all firms should adopt simultaneously at* $T^* = 0$.
2. *If* $\frac{\pi+\delta}{r+\lambda} < c \leq \beta\frac{\pi+\delta}{r}$ *with* $\beta = \frac{1+(n-1)e^{-rd^*}}{1+(n-1)e^{-(r+\lambda)d^*}}$*, a first adoption should occur at* $T^* = 0$ *and the* $n-1$ *following adoptions at* $\widehat{T} + d^*$ *where* $d^* = \frac{1}{\lambda}\ln\left(\frac{c}{\pi+\delta}(r+\lambda)\right) > 0$.
3. *If* $c > \beta\frac{\pi+\delta}{r}$*, no adoption should take place.*

We will maintain throughout the rest of the paper that

Assumption: $c > \frac{\pi+\delta}{r+\lambda} \Leftrightarrow d^* > 0$.

This means that we exclude the case where there are no learning externalities in the social optimum as all firms adopt simultaneously at $T^* = 0$. This allows us to focus on the role of learning in the diffusion of technologies.

3 Diffusion under the Cap and Trade scheme

We now start the analysis of the impact of emissions trading. In this section we study first the simplest scheme, a C&T one that yields the same price signal to all adopters. More specifically, we explore whether this policy approach implements the social optimum. Given the existence of a positive externality, the answer is expectedly negative as the early adopters neglect learning benefits. However, we will see that the positive externality induces two types of inefficiency in our dynamic setting: the traditional under-provision problem and a coordination problem that leads to socially detrimental delays in adoption.

3.1 The second stage

Reasoning backwards, we consider first how followers react once one firm has adopted the technology. Under C&T, the followers derive the benefit δ per time period after adoption by selling credits when emissions fall to zero and the initial allocation of credits amounts to pre-adoption emissions. Hence, the welfare maximization program (4) and the followers' profit maximization program are identical. The equilibrium delay is thus d^*. This is not surprising. As followers' decisions entail zero learning externality, the decentralized outcome is socially optimal.

3.2 The first stage

Moving backward, we consider next the first adoption. We randomize the adoption decision at each time period $[t, t+dt)$ in order to obtain equilibria in mixed strategies. As we will see, this setup allows us to determine endogenous delays of adoption.

Let $x_i dt$ denote the probability that firm $i = 1, .., n$ adopts the technology between t and $t + dt$, provided that the technology has not been adopted yet at time t. Using these notations, a pure strategy consists of a probability which is either $x_i dt = 1$ or $x_i dt = 0$. That is, firm i adopts (or does not) in the short time interval $[t, t+dt]$. A mixed strategy is $0 < x_i dt < 1$.

Firm i's expected payoff at any time t is given by the following Bellman equation:

$$\begin{aligned} V_i(t) = \left[v^L + \delta/r\right] x_i dt + (1 - x_i dt) \left[1 - \prod_{k \neq i} (1 - x_k dt)\right] \left[v^F(d^*) + \delta e^{-rd^*}/r\right] \\ + \left[\prod_{k=1}^{n} (1 - x_k dt)\right] e^{-rdt} V_i(t + dt) \end{aligned} \tag{8}$$

In this expression, the first term $\left[v^L + \delta/r\right] x_i dt$ is the payoff of firm i if it adopts the technology times the probability of adoption $x_i dt$. The second term

$$(1 - x_i dt) \left[1 - \prod_{k \neq i} (1 - x_k dt)\right] \left[v^F(d^*) + \delta e^{-rd^*}/r\right]$$

is the expected payoff if firm i does not adopt in the time interval, which occurs with a probability $(1 - x_i dt)$, and if at least one firm $k \neq i$ adopts in the same period, which occurs with a probability $1 - \prod_{k \neq i} (1 - x_k(t))$ -. Finally,

$$\left[\prod_{k=1}^{n} (1 - x_k dt)\right] e^{-rdt} V_i(t + dt)$$

is the payoff when nobody adopts between t and $t + dt$. In this case, firm i

derives V_i in the next period which is discounted.

In the Appendix we solve the game for equilibria in pure and mixed strategies. This leads to a first proposition.

Proposition 1 *Depending on payoffs, we observe different equilibria:*

1. *If* $v^L + \delta/r < 0$*, or equivalently if* $c > \frac{\pi+\delta}{r}$*, then no firm ever adopts the technology.*

2. *If* $v^L + \delta/r \geq 0$*, or equivalently if* $c \leq \frac{\pi+\delta}{r}$*, there are:*

   (a) *n equilibria in pure strategies, whereby one firm adopts at* $T^* = 0$ *and the others follow at* $T^* + d^*$*.*

   (b) *one symmetric equilibrium in mixed strategies, in which each firm* $i = 1, .., n$ *adopts with a probability*

$$\hat{x} = \frac{rv^L + \delta}{[n-1]\left[v^F(d^*) - v^L - (\delta/r)\left(1 - e^{-rd^*}\right)\right]}$$

so that the expected delay until the first adoption is:

$$\hat{T} = \frac{n-1}{n} \frac{v^F(d^*) - v^L - (\delta/r)\left(1 - e^{-rd^*}\right)}{rv^L + \delta} \tag{9}$$

Proof. See Appendix. ■

This proposition is the first key result of the paper. The intuition underlying Case 1 is obvious. No firm ever adopts because adopting first is not profitable ($v^L + \delta/r < 0$). This is so either because the adoption cost c is high or because the gross adoption benefit $\frac{\pi+\delta}{r}$ is too low.

The most interesting possibility is Case 2, where we have multiple equilibria. In this case, the adoption cost is sufficiently low for making adoptions profitable. But followers prefer delaying adoption so as to derive learning benefits.

In this situation, we have $v^L + \delta/r < v^F(d^*) + \delta e^{-rd^*}/r$, meaning that the incentive to preempt is weaker than the incentive to follow. This is the dynamic version of the "chicken game," where all firms are willing to adopt but have conflicting views on who should go first. As is usual in chicken games, a coordination problem results, leading to multiple Nash equilibria.

The economic interpretation of the equilibria in pure strategies, as shown in 2a, is problematic because all firms have an incentive to acquire the free benefits attendant on another firm's decision to adopt first, so that no single firm wishes to adopt first. In that case, we can reasonably expect strategic delays in the first adoption. This corresponds to the symmetric equilibrium in mixed strategies, shown in 2b, where the expected date of the first adoption $E(T^*)$ is strictly positive. Given (9), the larger the gap between the leader's total payoff $v^L+\delta/r$ and the followers' total payoff $v^F(d^*)+\delta e^{-rd^*}/r$, the longer the delay before the first adoption. In the rest of the paper, we remain focused on this equilibrium.

3.3 Welfare properties

We are now able to investigate the welfare properties of the C&T regime. To begin with, we recall that the followers' decision is optimal as it does not generate any externalities of adoption. With regard to the leader in adoption of technology, Proposition 2 tells us that the first adoption will take place iff

$$v^L + \delta/r \geq 0 \iff c \leq \frac{\pi + \delta}{r} \tag{10}$$

Not surprisingly, the comparison of (10) with the optimality condition (7) shows that the credit price δ is not sufficiently high to induce socially optimal decisions by the leader as $\beta \geq 1$. This is the standard result - that positive externalities lead to too few adoptions.

A second inefficiency, interestingly, exists. Proposition 2 predicts an equilibrium in mixed strategies which involves a delay in the first adoption even while the optimal date is $\widehat{T} = 0$.

We summarize these findings in our second Proposition.

Proposition 2 *A C&T scheme does not implement the first best outcome when $c \leq \beta \frac{\pi+\delta}{r}$. More precisely,*

1. *When $c \leq \frac{\pi+\delta}{r}$, the first adoption is delayed while the optimal adoption date is $T^* = 0$.*

2. *When $\frac{\pi+\delta}{r} < c \leq \beta \frac{\pi+\delta}{r}$, the first adoption should take place at $T^* = 0$ but that never occurs.*

In summary, the social inefficiency exclusively concerns the leader: Diffusion starts either too late or reaches an impass. By contrast, the followers make efficient decisions.

4 Diffusion path under the CDM

4.1 Additionality

Contrary to a C&T system, the benefit of the CDM is conditional to an additionality requirement. Since the additionality requirement means that a firm that adopts a new technology get credits only if technology adoption is not profitable without the credits. Hence adoption by a leader is additional if $v_L \leq 0$, or equivalently, if $\pi/r \leq c$.

Similarly, a second adoption after a delay d is additional if $v_F(d) \leq 0$, or $\pi e^{\lambda d}/r \leq c$. As $e^{\lambda d} > 1$, additionality obviously subsists if followers adopt before a threshold delay $d^{\max}$ defined by $c \equiv \pi e^{\lambda d^{\max}}/r$.

To summarize:

Lemma 2 *Adoption is additional for the leader if $\pi/r \leq c$. It is additional for a follower if $d < d^{\max}$ with $d^{\max} = \frac{1}{\lambda} \ln \left(\frac{rc}{\pi} \right)$*

Diffusion can then be conveniently analyzed separately, when the first adoption is not additional ($\pi/r \geq c$) and when it is additional ($\pi/r < c$) .

4.2 The first adoption is not additional ($\pi/r \geq c$)

This case is extremely simple. When the initial adoption is profitable without credits, the same is obviously true for subsequent ones. This means that all firms face the same price signal as under C&T, *except that the price is zero*. Accordingly, we just need to substitute δ by 0 in Proposition 1 to derive the equilibrium diffusion paths. The fact that $\pi/r \geq c$ rules out the first case in the proposition so that we end up with:

Lemma 3 *If adoption cannot be additional ($c \leq \pi/r$), each firm adopts with the same probability*

$$\tilde{x}dt = \frac{rv^L}{(n-1)\left[v^F(\widetilde{d}) - v^L\right]}dt.$$

Once a firm has adopted, the others follow after the delay $d^{\circ} \equiv \frac{1}{\lambda} \ln \frac{c}{\pi} (r + \lambda)$.

4.3 The first adoption is additional ($\pi/r < c$)

Reasoning backwards, we identify first the equilibrium delay $\widetilde{d}$, assuming that a firm has taken the lead, which requires $c < (\pi + \delta)/r$. Under CDM, followers have two options. They either decide to get CDM credits by choosing a delay $\widetilde{d} < d^{\max}$, or they prefer to give up the credits by choosing a longer delay which reduces the adoption cost. We consider these two strategies in turn.

In the first case, we know from (5) that the delay is d^* with the credits. However, this is compatible with additionality only if $d^* \leq d^{\max}$. Hence, obtaining credits requires the delay:

$$\widetilde{d} = \min\left\{d^*, d^{\max}\right\}$$

A simple calculation shows that $d^* < d^{\max}$ is equivalent to

$$\frac{\pi + \delta}{r + \lambda} > \frac{\pi}{r} \Leftrightarrow \frac{\pi}{r} < \frac{\delta}{\lambda} \tag{11}$$

The second strategy is that the firm decides not to implement a CDM project. Lemma 2 says that this leads to a delay $\widetilde{d} = d^{\circ}$. This strategy is clearly less profitable than the first one when (11) is met: As the additionality constraint is not binding, the firms always prefer a CDM project, since they can select the optimal delay d^* and get credits. Things are more ambiguous when the condition is not met (e.g., if $\frac{\pi}{r} \geq \frac{\delta}{\lambda}$). In fact, the followers will make their decision by comparing their payoff with credits, $v_F\left(d^{\max}\right) + \frac{\delta}{r} e^{-r d^{\max}}$, and without, $v_F\left(d^{\circ}\right)$. Appropriate substitutions yield that they prefer the CDM if

$$\frac{\pi}{r} < \frac{\delta}{\lambda}\left(\frac{r + \lambda}{r}\right)^{\frac{r+\lambda}{\lambda}}$$

This is very intuitive: Followers opt for the CDM when the credit price is high and/or post-adoption market profit is low.

We summarize the whole analysis in the following:

Lemma 4 *In the case where the first adoption is additional ($c > \pi/r$), and assuming that a leader has adopted the technology, which requires $c < (\pi + \delta)/r$,* *the followers select the delay* $\tilde{d} > 0$ *given by:*

$$\tilde{d} = \begin{cases} d^* \ \text{if } \frac{\pi}{r} < \frac{\delta}{\lambda} \\ d^{\max} \ \text{if } \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} > \frac{\pi}{r} \geq \frac{\delta}{\lambda} \\ d^{\circ} \ \text{if } \frac{\pi}{r} \geq \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} \end{cases}$$

Note that $d^{\max} < d^* < d^{\circ}$: As compared to the socially optimal delay d^*, additionality can either induce too slow or too fast diffusion depending on the size of π/r. This suggests a complex welfare effect of additionality. We return to this point later.

We complete the analysis with stage 1. We already know that no firm ever adopts if $c \geq (\pi + \delta)/r$. When $c < (\pi + \delta)/r$, the coordination problem arises. Exploiting similarities with Proposition 1 and the results of Lemma 4, we easily obtain:

Proposition 3 *In the case where adoptions can be additional ($c > \pi/r$), no firm ever adopts if $c > (\pi + \delta)/r$). Otherwise, each firm adopts with the per-time period probability*

$$\tilde{x}dt = \frac{r(v^L + \frac{\delta}{r})}{(n-1)\left[u^F(\tilde{d}) - v^L - \frac{\delta}{r}\right]}dt$$

where $\tilde{d}$ *is defined in Lemma 3 and*

$$u^F(\tilde{d}) = \begin{cases} v^F(\tilde{d}) + \frac{\delta}{r}e^{-r\tilde{d}} \ \text{if } \frac{\pi}{r} < \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} \ \textit{(followers get credits)} \\ v^F(\tilde{d})) \ \text{if } \frac{\pi}{r} \geq \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} \ \textit{(followers do not get credits)} \end{cases}$$

5 Welfare comparison

We are now able to compare the welfare properties of C&T and CDM. To begin with, social welfare is obviously zero if no firm ever adopts the technology. This occurs with C&T and CDM under the same condition where $c \geq \frac{\pi+\delta}{r}$, implying that both schemes are welfare equivalent in this case.

When diffusion occurs under both schemes ($c < \frac{\pi+\delta}{r}$), we have seen that a first firm adopts the technology with a probability xdt per time period. The social welfare function in equilibrium can thus be written as

$$W = \int_0^\infty nxe^{-nxt} \left[v^L + \frac{\delta}{r} + (n-1)(v^F(d) + \frac{\delta}{r}e^{-rd}) \right] dt$$

where xdt is the equilibrium values of the per-time period probability of adoption and d, the equilibrium delay after the first adoption. This expression simplifies as follows

$$W(x,d) = \frac{nx}{r+nx} \left[v^L + (n-1)v^F(d) + (1+(n-1)e^{-rd})(\delta/r) \right] \qquad (12)$$

We will use this expression to compute the equilibrium welfare in the different cases.

5.1 Cap & Trade

By substituting x^* and d^* in (12), we obtain a very simple expression:

$$W_{C\&T} = n(v^L + \frac{\delta}{r}) \qquad (13)$$

Since $v^L + \frac{\delta}{r} < v^F(d) + \frac{\delta}{r}e^{-rd}$, this is obviously less than the first best level which would be

$$W^* = v^L + \frac{\delta}{r} + (n-1) \left[v^F(d^*) + \frac{\delta}{r}e^{-rd^*} \right]$$

In fact, welfare under C&T would be the same if all firms were to adopt immediately and simultaneously. Therefore the benefit of the delay between first and second adoption, which amounts to the difference between $v^F(d^*) + \frac{\delta}{r}e^{-rd}$ and $v^L + \delta/r$, is entirely dissipated by the delay before the first adoption. In other words, the learning benefits and the coordination cost cancel each other out. The intuition is that the higher the learning benefit, the lower the incentives to take the lead, and thus the longer the delay before the first adoption.

5.2 CDM

Depending on the value of parameters, there are three feasible scenarios: 1) none of the firms receive credits; 2) the leader receives credits, but followers do not; or 3) all firms receives credits. The first path is described by Lemma 2, the two last paths by Proposition 3. We consider these cases in turn.

Case 1: None of the firms receive credits ($\pi/r \geq c$)

We substitute $\tilde{x}$ and $\tilde{d}$ from Lemma 3 in (12) leading to

$$W_{CDM} = nv^L \left[1 + \frac{1 + (n-1)e^{-rd^\circ}}{v^L + (n-1)v^F(d^\circ(0))} \frac{s}{r}\right] \tag{14}$$

A straightforward calculation then shows that this welfare level is lower than (13). This is not surprising. The main reason is that followers having zero credits wait too long to adopt, while their response is optimal under C&T ($\hat{d} = d^*$). In addition to this, the second source of inefficiency - the delay before the first adoption - is not significantly affected by the absence of credits, as this loss relative to C&T concerns both leaders and followers.

Case 2: All firms receive credits ($\frac{\pi}{r} < c < \frac{\pi+\delta}{r}$ and $\frac{\pi}{r} < \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}}$)

In this case, all firms get credits and $\tilde{d} = \min\{d^*, d^{\max}\}$. If $\tilde{d} = d^*$, we obviously have $W_{CDM} = W_{C\&T}$. But the same is also true when $\tilde{d} = d^{\max}$, since plugging $\tilde{x}$ and $d^{\max}$ in (12) yields $W_{CDM} = n(v^L + \delta/r)$.

The latter result is counter-intuitive, as followers adopt too early ($d^{\max} < d^*$) to meet the additionality requirement under the CDM. This, however, does not reduce welfare. We can understand why by looking at (13). This equation says that the (optimal) learning benefit is entirely dissipated by the losses due to the initial delay under C&T. The same mechanism works when the followers do not wait for the optimal amount of time under the CDM. This distortion is compensated by a lower initial delay.

Case 3: Only the leader receives credits ($\frac{\pi}{r} < c < \frac{\pi+\delta}{r}$ and $\frac{\pi}{r} \geq \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}}$)

In this last case, the leader and the followers respectively derive a private benefit $v^L + \delta/r$ and $v^F(d^\circ)$. Substituting these payoffs in $\tilde{x}$ and then $\tilde{x}$ and d° in (12) leads to

$$W_{CDM} = n(v^L + \delta/r)\left[1 + \frac{(n-1)e^{-rd^\circ}}{v^L + \delta/r + (n-1)v^F(d^\circ)}\frac{\delta}{r}\right], \qquad (15)$$

a social welfare results that is obviously higher than under the C&T scheme.

This is a key result of the paper. As in the previous case, the followers distort their decision as they have no credits, but the leader now gets credits implying that the gap between payoffs is reduced. Hence diffusion starts earlier. We show here that the latter effect outweighs the former.

This result is quite counter-intuitive. Recall that the original problem is the existence of positive learning externalities generated by the first adopter whereas followers make efficient decisions if they face the appropriate price signal δ. The standard policy solution is thus to subsidize the leader. This is not at all what we do here: The leader derives the same benefit as under C&T, but the CDM deprives the followers of any carbon subsidy. We have shown that this solution partly mitigates the externality problem.

The Proposition summarizes the findings:

Proposition 4 *There are two cases where CDM & C&T are not welfare equivalent:*

1. *When adoption by the leader is not additional under CDM, so that no firms receive any credits, C&T dominates CDM .*

2. *The opposite is true when the first adoption is additional and subsequent ones are not, that is, when CDM credits are only granted to the first adopter.*

6 Improving CDM design

We complete the analysis by exploring options which could increase the ability of the CDM to tackle learning externalities. We consider two options: 1) relaxing the additionality constraint for projects generating learning externalities and 2) bundling individual projects in a single CDM project.

6.1 Relaxing the additionality requirement

In Proposition 4 the CDM outperforms a C&T scheme in the case where adopting first is additional whereas following is not. The superiority of the CDM is due to the fact that adoption incentives are differentiated. This suggests that, even if not additional, granting CDM credits to the leading firm could be an interesting option. In this subsection, we explore whether suppressing the additionality requirement for leading firms would improve welfare compared to the current CDM rule.

Answering the question is not straightforward. On the one hand, this speeds up diffusion by reducing the delay before the first adoption as taking the lead becomes more profitable. On the other hand, this damages the environmental integrity of the whole scheme by increasing GHG emissions. The reason is that the leader sells its credits to firms which would have abated emissions otherwise. Let us now compare the two scenarios.

We analyze here the case where the leader's adoption is not additional, meaning that $c \leq \pi/r$. Under the standard CDM, the diffusion outcome is described in Lemma 1 and the corresponding welfare at equilibrium is given by (14). To characterize diffusion under the new rule, we just adapt Lemma 1 by adding credit sales to the leader's payoff to obtain its per-time period probability of adoption:

$$\tilde{x}'dt = \frac{r\left(v^L + \frac{\delta}{r}\right)}{(n-1)\left[v^F(d^\circ) - v^L - \frac{\delta}{r}\right]}dt \tag{16}$$

Unsurprisingly, this probability is higher than $\tilde{x}dt$ given in Lemma 2[9]. Then, social welfare function is given by:

$$W^{'}_{CDM} = \frac{n\tilde{x}'}{r + n\tilde{x}'}\left[v^L + (n-1)v^F(d) + (n-1)e^{-rd})(\delta/r)\right]$$

This functional form is very similar to (12). The only difference is that we omit the term δ/r in the left-hand term in brackets as adoption by the leader no longer avoids the environmental damage δ per-time period. Substituting (16)

[9] Simple calculations show that the differences per-time period probability with and without credits is equal to $\frac{\delta}{n-1}v^F(d^\circ)$.

and d° in this expression yields a revised level of welfare in equilibrium:

$$W_{CDM}^{'} = \frac{n\left(v^L + \frac{\delta}{r}\right)}{(n-1)v^F(d^{\circ}) + \left(v^L + \frac{\delta}{r}\right)}\left[v^L + (n-1)v^F(d^{\circ}) + (n-1)e^{-rd^{\circ}})(\delta/r)\right]$$

In appendix we compare the difference between $W_{CDM}^{'}$ and W_{CDM} (described by 14) and we establish the following:

Proposition 5 *A modified CDM which removes the additionality requirement for the leader's project improves welfare when the number of firms, n, and the learning parameter, λ, are sufficiently high and/or when marginal environmental damage δ is sufficiently low.*

This proposition is very intuitive. When the learning externality is high because there are many followers (n is large) or because adoption cost decreases quickly after the first adoption (λ is high), relaxing additionality improves welfare because it triggers the learning process earlier. When the environmental damage–as reflected by the parameter δ–is large, relaxing additionality is more costly as it increases emissions substantially as compared to the current CDM rule.

As the drawback of the revised rule is the hot air generated by non additional projects that are registered under the CDM, we have an interesting corollary.

Corollary 1 *Relaxing additionality always improves social welfare if the number of permits allocated in the international Cap and Trade scheme is reduced ex ante by a quantity equal to the quantity of hot air generated by non additional CDM projects.*

6.2 A sectoral mechanism

A mechanism grouping individual adoptions into a single CDM project is another option for improving the efficiency of the CDM. In post-Kyoto talks, intense discussions revolve around the potential for so-called sectoral mechanisms to gather into a single project the firms of a given sector or of a specific geographical area (Baron & Ellis, 2006). This solution is primarily viewed as way of reducing project administrative costs[10]. The Coase theorem suggests that grouping firms together might also be a possible way to internalize learning externalities and to coordinate adoptions.

The collective nature of the project does not change the way additionality is defined: Carbon credits would be generated after each individual adoption provided this adoption is not profitable in itself, as in the current CDM scheme.[11] Two types of mechanisms seem compatible with this additionality rule:

- **Type 1**: Mechanisms which would only authorize the merger of subprojects that are all additional. In our setting this limits the use of a sectoral mechanism to the case where $\frac{\pi}{r} < c < \frac{\pi+\delta}{r}$ and $\frac{\pi}{r} < \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}}$ (Case 2 in subsection 5.2).

- **Type 2**: Mechanisms which could include additional and non-additional subprojects. But, of course, the non-additional projects would not generate any credits. This allows using a sectoral mechanism in Cases 2 and 3.

We now show that both schemes improve welfare but the latter type is socially preferable.

Proposition 6 *When all individual adoptions are additional, both sectoral mech-*

[10] Already at its 21st meeting in 2005, the UNFCCC Board already agreed on general principles for bundling CDM projects. A bundle brings together several small-scale CDM activities to form a single CDM project.

[11] In practice, assessing the additionality of each individual sub-projects could be less strict than in the current scheme simply because additionality would be assessed globally.

anisms implements the first best social optimum. That is, a leader adopts at $T^* = 0$ and the others at $T^* + d^*$.

When only the leader's adoption is additional, a sectoral mechanism of type 2 is the only feasible option. The mechanism does not implement the first best outcome but it improves welfare. The leader adopts at $T^* = 0$ and the others at $T^* + d^\circ$.

Proof. Straightforward. As all adopters participate, they select the diffusion path which maximizes the sum of profits and credit sales. In the case where all projects generate credits, the objective function coincides with the social welfare function (3). In the case where only the leader gets credits, the objective function is

$$W(T,d) = \left(\frac{e^{-rT}}{r}\right)\left[v^L + \delta/r + (n-1)v^F(d)e^{-rd}\right] \tag{17}$$

which is maximized when $T = 0$ and $d = d^\circ$. ■

The intuition is straightforward: The firms can make a binding agreement in which they decide who is adopting first and the compensations the leader receives from the followers. A sectoral CDM project is an appropriate contractual framework to do so. In particular, the fact that the credits are jointly awarded makes utility transfers easy between firms[12].

Although such an agreement is beneficial to all parties, an exogenous rule requiring the participation of all adopters is necessary because full participation is not a Nash equilibrium: each follower has an incentive to deviate unilaterally as it would continue to benefit from the learning externality without the need to compensate the firm which takes the lead. However, the worst-case scenario−all firms engaged in free riding−is simply the scenario of the standard CDM. In other terms, the sectoral mechanism cannot do worse than the standard CDM

[12] Moreover a collective project also creates incentives to maximize the learning externalities by sharing information. This benefit is not modelled in our set up.

even in a setting allowing for non-cooperative behaviours.

The proposition also hinges on the assumption that there are no bargaining costs. In practice, sectoral projects may generate substantive transaction costs. But the standard CDM also entails fixed transaction costs, though, which sectoral solutions mitigate.

7 Conclusion

Project mechanisms such as the CDM are often depicted as powerful levers for the diffusion of environmental technologies in developing countries. In this paper, we explore this insight by developing a simple model that captures both the transfer of a technology into a developing country and its horizontal diffusion within the country.

As compared to other emissions trading schemes, the originality of the CDM lies in its additionality requirement, whereby credits are only granted to projects that would not be profitable otherwise. As a result, the CDM only yields a positive price signal to additional projects. By contrast, the price is uniform across all firms under other trading schemes (e.g., Cap and Trade, Baseline and Credit).

In order to investigate the role of additionality, we have compared a standard C&T system with the CDM. In the presence of learning spillovers we have shown that C&T schemes fail to implement the optimal diffusion path because the leading firm - which generates positive externalities - and the followers receive the same amount of credits. This leads to two inefficiencies: the standard underprovision problem and a coordination problem driven by the fact that adopting first is less profitable than following.

By design, the CDM either yields the same quantity of credits as C&T or, when the project is not additional, zero credits. Hence the CDM cannot reward the leader in order to internalize learning benefits as recommended in textbooks. But it can "punish" the followers. We show that this "punishment" may actually improve welfare. In fact, the CDM yields a higher welfare than C&T in the case where the leader receives credits and the followers do not. This does not solve the under-provision problem, but it does mitigate coordination costs.

This analysis suggests two improvements for CDM design: a relaxed additionality rule and collective CDM projects that gather all adopters. We show that removing the additionality requirement for the first adoption leads the CDM to outperform a C&T scheme for all parameters.

We also show that allowing the formation of collective CDM projects is an effective way to suppress strategic delays before the first adoption, thereby improving the overall efficiency of technology diffusion. This is a new argument in favour of collective (sectoral) CDM projects, whose potential is intensively debated in the policy arena. One of the key arguments is that they would reduce administrative burden. Our research shows that collective projects are not only more efficient socially for the implementation of new technologies and diffusion of those technologies, but such projects are also attractive for firms; if given the choice, firms would always opt for collective projects.

In post-Kyoto talks, discussion continues around the question of whether emitters located in such emerging economies as China, India, or Brazil should be covered by a Clean Development Mechanism-like scheme featuring additionality or by a Cap and Trade scheme. Our analysis stresses one advantage of the CDM over other emission trading schemes: The additionality requirement can be tailored to increase the speed of technology diffusion.

8 Appendix

8.1 Proof of Proposition 2

The firm i's expected payoff at any time t is given by (8). Using this equation we derive successively the conditions for the different equilibria to arise.

8.1.1 Case 1: No firm adopts ($x_i dt = 0$, $\forall i = 1, ..., n$)

If the other $(n-1)$ firms do not adopt, the expected payoff of firm i writes

$$V_i = \left(v^L + \delta/r\right) x_i dt + e^{-rdt} \prod_{k=1}^{n} \left(1 - x_k dt\right) V_i$$

$$V_i = \left(v^L + \delta/r\right) x_i dt + e^{-rdt} \left(1 - x_i dt\right) V_i$$

Since we consider infinitesimal values of dt, we can eliminate all terms in $(dt)^n$, $n > 1$. Noting moreover that $1 - e^{-rdt} \sim rdt$ and $e^{-rdt} \to 1$, the expression can write:

$$V_i = \frac{x_i \left(v^L + \delta/r\right)}{r + x_i} \tag{18}$$

This expression is decreasing in x_i if $v^L + \delta/r < 0$. Hence the equilibrium where no firm adopts exists when $v^L + \delta/r < 0$.

8.1.2 One firm j adopts immediately ($x_j dt = 1$).

In that case the expected payoff of the other firms $i \neq j$ write:

$$V_i = v^F(\hat{d}) + e^{-r\hat{d}}\delta/r + x_i dt \left[v^L + \delta/r - v^F(\hat{d}) + \left(1 - e^{-r\hat{d}}\right)\delta/r\right]$$

Recall that $v^L + \delta/r < v^F(\hat{d}) + e^{-r\hat{d}}\delta/r$ as $\hat{d} = d^* > 0$ by assumption. Hence the best reply for firm $i \neq j$ is clearly $x_i dt = 0$. Knowing this we have to check whether firm j will still play $x_j dt = 1$. From 18 we know that firm j's payoff is $V_j = x_j\left(v^L+\delta\right)/r/\left(r+x_j\right)$ and that firm j will play $x_j dt = 1$ only if $v^L > 0$. It follows that there are n equilibrium in which one firm adopts immediately ($x_j dt = 1$) while the others do no adopt ($x_i dt = 0$, $i \neq j$) if $v^F(\hat{d}) + e^{-r\hat{d}}\delta/r > v^L + \delta/r > 0$.

8.1.3 Case 3: all firms play mixed strategies

Consider again the expected payoff of firm i in (8). Since we consider infinitesimal values of dt, we can eliminate all terms in $(dt)^n$, $n > 1$. Noting moreover that $1 - e^{-rdt} \sim rdt$, the expression rewrites:

$$V_i = \frac{x_i\left(v^L+\delta/r\right) + \sum\limits_{k\neq i} x_k\left[v^F(\hat{d}) + e^{-r\hat{d}}\delta/r\right]}{r + \sum\limits_{k} x_k}$$

If $v^L + \delta/r \geq 0$, the expected profit V_i admits a maximum in x_i. The FOC of firm i's program rewrites into the following equation:

$$\sum_{k\neq i} x_k = \frac{rv^L+\delta}{v^F(\hat{d}) - v^L + \left(1-e^{-r\hat{d}}\right)\delta/r} \qquad (19)$$

It is clear from 19 that only one equilibrium is possible, where $\hat{x}_i = \hat{x}$ for all $i = 1, ..., n$. The equilibrium adoption strategy is then:

$$\hat{x} = \frac{rv^L+\delta}{[n-1]\left[v^F(d^*) - v^L + \left(1-e^{-rd^*}\right)\delta/r\right]} \qquad (20)$$

The strategy $\hat{x}$ followed by each firm defines a Poisson process of parameter $n\hat{x}$ for the first adoption. This allows us to calculate the expected delay until the first adoption:

$$E(T) = \int\limits_0^\infty tn\hat{x}e^{-n\hat{x}t}dt = \frac{n-1}{n}\frac{v^F(d^*) - v^L + \left(1-e^{-rd^*}\right)\delta/r}{rv^L+\delta} \qquad (21)$$

8.2 Proof of Proposition 5

We first compute the difference:

$$W'_{CDM} - W_{CDM} = \frac{n\,(\delta/r)\left[(n-1)^2\left(v^F\right)^2 + (\delta/r)\,e^{-rd\circ}(n-1)v^F\right] - v^L(v^L+\delta/r)}{\left[(n-1)v^F + v^L + \frac{\delta}{r}\right]\left[(n-1)v^F + v^L\right]}$$

where we omit d° for notational simplicity. The denominator is positive as we are in the case where v^L and v^F are positive. We can thus focus the analysis of the sign of the numerator

$$X \equiv \left[(n-1)^2\left(v^F\right)^2 + (\delta/r)\,e^{-rd\circ}(n-1)v^F\right] - v^L(v^L+\delta/r)$$

It is then obvious that $\partial X/\partial n = 2(n-1)\left(v^F\right)^2 + (\delta/r)\,e^{-rd\circ}v^F > 0$. Note also that $X > 0$ when n is sufficiently high. When $n = 2$, we have

$$X|_{n=2} \equiv \left(v^F\right)^2 - \left(v^L\right)^2 + (\delta/r)\left[e^{-rd^\circ}v^F - v^L\right]$$

of which sign is ambiguous. It is positive when δ is sufficiently low. When δ gets higher, what is key is the sign of $e^{-rd\circ}v^F - v^L$. In a supplementary material, we rely on simulations to show that X increases with λ and decreases with δ.

References

[1] Baron, R., Ellis, J. (2006) “Sectoral Crediting Mechanisms for GreenHouse Gas Mitigation: Institutional and Operational Issues” OECD/IEA Information Paper COM/ENV/EPOC/IEA/SLT(2006)4.

[2] Blackman A. (1999) "The Economics of Technology Diffusion: Implications for Climate Policy in Developing Countries", Discussion Paper 99-42, Resources For the Future: Washington DC.

[3] Dechezleprêtre A., M. Glachant, Y. Ménière (2008) "The Clean Development Mechanism and the international diffusion of technologies: An empirical study", Energy Policy, 36(4), pp 1273-1283.

[4] Fischer, C., Parry, I. and W. Pizer (2003) "Instrument choice for environmental protection when technological innovation is endogenous," *Journal of Environmental Economics and Management*, 45(3), pp 523-545.

[5] Fudenberg, D. and J. Tirole (1985) "Preemption and Rent Equalization in the Diffusion of New Technology" *Review of Economic Studies*, 52, 383-401.

[6] Fudenberg, D. and J. Tirole (1991) *Game Theory*, The MIT Press: Cambridge.

[7] Greiner, S. and A. Michaelowa (2003) "Defining Investment Additionality for CDM projects—practical approaches" *Energy Policy*, 31:10, pp 1007-1015.

[8] Hoppe (2002) "The Timing of New Technology Adoption: Theoretical Models and Empirical Evidence" *The Manchester School*, 70:1, pp 56-76.

[9] Kartha, S., M. Lazarus and M. LeFranc (2005) "Market penetration metrics: Tools for additionality assessment?", *Climate Policy*, 5:2.

[10] Jaffe A. and R. Stavins (1995) "Dynamic Incentives of Environmental Regulations: The Effects of Alternative Policy Instruments on Technology Diffusion," *Journal of Environmental Economics and Management*, 29(3), pp S43-S63.

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

[12] Mariotti, M. (1992) "Unused Innovations" *Economic Letters*, 38, 367-371.

[13] Michaelowa, A., Stronzik, M., Eckermann, F., and A. Hunt (2003) "Transaction costs of the Kyoto Mechanisms," *Climate Policy*, 3.

[14] Milliman, S. and R. Prince (1989) "Firm incentives to promote technological change in pollution control," *Journal of Environmental Economics and Management*, 17(3), pp 247-265.

[15] Millock, K. (2002) "Technology transfers in the Clean Development Mechanism: an incentive issue," *Environment and Development Economics*, 7.

[16] Montero, J.-P. (2002) "Permits, Standards, and Technology Innovation," *Journal of Environmental Economics and Management*, 44(1), pp 23-44.

[17] Reinganum, J. (1981) "Market Structure and the Diffusion of New Technology" *Bell Journal of Economics*, 12, 618-624.

[18] Requate, T. (1998) "Incentives to innovate under emission taxes and tradeable permits," *European Journal of Political Economy*, 14(1), pp 139-165.

[19] Rosendahl K.E. (2004) "Cost-effective environmental policy: implications of induced technological change, Journal of Environmental Economics and Management, 48, pp1089-1121.

[20] Schneider, M., Holzer, A. and V. Hoffmann (2008) "Understanding the CDM's contribution to technology transfer" *Energy Policy*, 36:8, pp 2930-2938.