CERNA WORKING PAPER SERIES

Project Mechanisms and Technology Diffusion in Climate Policy

Matthieu Glachant, Yann Ménière

Working Paper 2010-02

Cerna, Centre d'économie industrielle
MINES ParisTech
60, boulevard Saint Michel
75272 Paris Cedex 06 – France
Tél. : 33 (1) 40 51 90 00

March 2010

Kyoto Project Mechanisms and Technology Diffusion

Matthieu Glachant*, Yann Ménière†

June 1, 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 don't. This result lends support to the policy proposal of relaxing the CDM additionality constraint for projects which generate significant learning externalities.

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

JEL code: H87, O33, Q55, Q54

*Cerna, Ecole Nationale Supérieure des mines de Paris, France, glachant@ensmp.fr
†CORE, Université Catholique de Louvain, meniere@core.ucl.ac.be

1 Introduction

Due to economic growth, developing countries are expected to supersede industrialized countries as the leading source of Green House 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 Bush 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. The CDM is an arrangement under the Kyoto Protocol allowing industrialized countries with a greenhouse gas reduction commitment (so-called Annex 1 countries) or firms located in these countries to invest in emission reducing projects in countries that have not such commitments (or Annex 2 countries). These projects, usually carried out in developing countries, provide a cheaper alternative to costly emission reductions in industrialized countries. Besides, the CDM can contribute to technology transfer by financing projects using technologies not available in the host countries.[1] Such transfers have gradually gained in importance in policy debates, and are at the core of on-going 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. The focus is the CDM whose specificity lies in the additionality requirement: firms can implement a

[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 (in particular, in post-Kyoto talks).

CDM project only if it would not be profitable without credits. In order to investigate the impact of additionality, we compare the CDM with a traditional Cap and Trade program (C&T, hereafter) where any abatement - being privately profitable or not - makes emissions credits available.

The model describes n firms located in a host country which initially operate with an old technology. They can adopt a cleaner technology simultaneously or sequentially. The first adoption that occurs is the international technology transfer *per se* since the technology is not previously available in the country. The following adoptions correspond to the diffusion within the host country (sometimes referred to as horizontal diffusion).

Adoption entails a fixed cost. A key assumption of the model is that this cost endogenously decreases once the technology has been introduced in the host country. In reality, this can be so because observing the outcome of the first adoption may reduce the uncertainty on technology benefits for the following adopters. Or the first adopter accumulates learning-by-doing skills which diffuse to potential adopters through various channels (e.g., labor market).

These learning spillovers generate two types of inefficiencies. The first is the standard underprovision problem. The leader's propensity to adopt is too low as he does not take into account positive externalities, thereby hindering technology transfer. The second is a coordination problem which results from the dynamic character of the diffusion process. All firms would prefer to follow in order to enjoy a reduced adoption cost. But following requires that one firm takes the lead. This is a dynamic version of a "chicken game" where both firms derive a positive profit from adoption but have conflicting views on who should go first. As a result, one possible outcome is that the first adoption is delayed, albeit (privately and socially) profitable.

We show that a standard Cap and Trade scheme does not implement the first best diffusion path. One reason is that each adopter receives the same amount of credits whatever its adoption rank. Giving an additional premium to the leader is useful as it solves the externality problem and mitigates the coordination problem. However it only implements the first best optimum when the positive externality is low.

We turn next to the CDM. The CDM differs from C&T in that the credit price signal is not uniform across all firms-it is zero for non-additional projects. We show that the welfare impact of the additionality requirement is ambiguous as compared to C&T. Unsurprisingly, it is clearly detrimental when adoption by the leader is not additional, for technology diffusion is too slow in absence of credits. However, it unambiguously improves welfare when the leader receives credits while the others don't. It is so because it reduces the followers' advantage, thereby mitigating the coordination problem. This result is true whatever the parameters' value. Our analysis thus provides a strong case for granting CDM credits to non-additional projects that are expected to generate learning spillovers.

The economic literature on Kyoto project mechanims is extremely scarce. Michaelowa et al. (2003) evaluate the level of transaction costs that may impede the diffusion of project mechanisms. Millock (2002) studies the cost efficiency effect of technology transfers through bilateral CDM contracts when there is asymmetric information between the investor and the host party. In a recent paper, Dechezlepretre et al. (2008) develop an empirical study of a dataset describing about 600 CDM projects. They show that about 44% of the projects exhibit a transfer. This ratio increases with the project size, and varies across sectors.

Beside the specific literature on CDM, an important strand of theoretical literature has developed on environmental innovation and policy instruments[2].

[2]See for instance Jaffe & Stavins, 1995; Laffont & Tirole, 1996; Requate, 1998; Montero,

This literature has a much broader focus than our work: most papers compare different policy instruments. Moreover, excepting 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. Blackman (1999) surveys the general economic literature on technology diffusion in order to derive lessons for climate policy.

Our paper is also related to a strand of 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 characterize the socially optimal technology diffusion path. Section 3 characterizes diffusion patterns under a Cap and Trade scheme. In Section 4, we investigate the CDM and compare with Cap and Trade. Section 5 concludes.

2 Model and social optimum

In this section, we present a simple model in continuous time which describes the adoption of a GHG mitigation technology by n symmetric firms under emissions trading.

2002; Fischer, Parry, & Pizer, 2003.

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. The technology can increase the profit ($\pi > \pi^{\circ}$) or decrease it ($\pi \leq \pi^{\circ}$). For ease of presentation, we will maintain throughout that $\pi^{\circ} = 0$. Adoption also reduces GHG emissions. 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. Emissions generate a marginal damage δ. Under these assumptions, firm i 's profit function is:

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

where e_i is the firm's level of emissions[3].

Note that technology adoption by a given firm does not affect others' profit. This either means that firms operate in a perfectly competitive product market where a change in the production cost of one firm has a negligible impacts on other firms' level of output and profit. This assumption rules out strategic issues in the product market which are potentially associated with technology adoption. It greatly simplifies the analysis and allows to focus sharply on the issue of technology diffusion.[4]

Adopting the 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

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

[4] 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 CDM.

adoption. This differs from the assumptions made in most previous models of technology diffusion in which the adoption cost decreases with time for exogenous reasons (see for instance Reinganum,1981, and Fudenberg & Tirole,1985). Formally, c is the cost for the first adopter while a follower 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 $\lambda = 0$, there is no positive externality of adoption.

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

At $t = 0$, there exists a climate regime with an international Cap and Trade scheme to mitigate GHG emissions. The scheme is socially optimal so that the market price is equal to (constant) marginal damage δ. 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[5]. These credits can be sold at price δ on the market.[6]

- They can implement a CDM project where they get credits they can sold on the international carbon market, *but only if adoption is not profitable without credits*. In CDM terminology, the abatement project should be additional. Under this scenario, the amount of allowances is not adjusted as the additionality requirement preserves the environmental integrity of the international climate regime. We further assume that adoptions do not modify the price δ, meaning that adopters only represent a small subset of market participants.

The key difference between the two scenarios is that under C&T the n firms all face the same price signal δ whereas the CDM only yields a price signal to additional projects. 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

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

[6] 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 in the scheme.

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 i selects the adoption time $T+d_i$.

The fact that they act strategically influence dramatically the results. Importantly, this does not mean that the n firms operate in the same oligopolistic product market. In our game, the firms interact with other firms that could generate positive spillovers they would benefit from. To do so, firms must be similar from a technological point of view. But they are necessarily competitors. The fact that they operate on the same labor market is for instance 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 can now derive what should be the welfare-maximizing adoption path. Let W denote the social welfare function. It is the discounted flow of market profits plus environmental benefit:

$$\begin{aligned} W(T, d_2, .., d_{n-1}) &= \int_T^\infty \left(v^L + \delta\right) e^{-rt} dt + \sum_{i=2}^{n} \int_{T+d_i}^\infty \left(v^F(d_i) + \delta\right) e^{-rt} dt \\ &= \left(\frac{e^{-rT}}{r}\right) \left[v^L + \delta/r + \sum_{i=2}^{n} \left(v^F(d_i) + \delta/r\right) e^{-rd_i}\right] \end{aligned}$$

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

Note that this social welfare function is very 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.

It is then 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 $j \neq 1$. Hence we just need solve the following program:

$$\max_d v^F(d) + \delta e^{-rd}/r \tag{3}$$

Substituting (2) and solving for d yields:

$$d^* = \begin{cases} \frac{1}{\lambda} \ln \frac{c}{\pi+\delta} (r+\lambda) \text{ if } c > \frac{\pi+\delta}{r+\lambda} \\ 0, \text{ otherwise.} \end{cases} \tag{4}$$

Equation (4) essentially says that the higher the cost of adoption and the faster its decrease over time, the less likely should followers adopt immediately. In the case they should wait ($d^* > 0$), note that the socially optimal delay d^* decreases with the profit flow π.

Obviously, d^* characterizes the social optimal only if equilibrium welfare is positive. That is, if:

$$v^L + \delta/r + (n-1)\left[v^F(d^*) + \delta e^{-rd^*}/r\right] \geq 0 \tag{5}$$

which simplifies as follows

$$c \leq \beta \frac{\pi+\delta}{r} \quad with \quad \beta \equiv \frac{1+(n-1)\, e^{-rd^*}}{1+(n-1)\, e^{-(r+\lambda)d^*}} \geq 1. \tag{6}$$

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 simulatenously at* $T^* = 0$.

2. *If* $\frac{\pi+\delta}{r+\lambda} < c \leq \beta\frac{\pi+\delta}{r}$ *with* $\beta = \frac{1+(n-1)e^{-r\widehat{d}}}{1+(n-1)e^{-(r+\lambda)\widehat{d}}}$, *a first adoption should occur at* $T^* = 0$ *and the* $n-1$ *following adoptions at time* 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 paper that

Assumption: $c \leq \frac{\pi+\delta}{r+\lambda}$

This means that we forego the case where there is 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 study the scenario in which the firms are directly included in the C&T scheme. Recall that the C&T scheme yields a price δ, meaning that it is optimized only to account for the environmental externality. This suggests that the scheme will fail to implement the first best optimum given the existence of the learning externality. We will see that this positive externality induces two types of inefficiency in our dynamic setting: the traditional under provision problem and a coordination problem leading to socially-detrimental delays of adoption.

3.1 The second stage

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 as emissions fall to zero and the initial allocation of credits amounts to pre-adoption emissions. Hence, the welfare maximization program (3) and the followers' profit maximization program are exactly the same. This is not surprising: as followers' decisions entail zero externality, the decentralized outcome is socially optimal.

Denoting $\hat{d}(\delta)$ the equilibrium delay contingent on the credit price δ, we thus have $\hat{d}(\delta) = d^*$.

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 derive equilibria in mixed strategies. As we will see this setup allows 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 either consists in a probability $x_i dt = 1$ or $x_i dt = 0$. That is, firm i adopts (or not) in the short time interval $[t, t + dt]$. A mixed strategy is $0 < x_i dt < 1$.

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

$$V_i(t) = \left[v^L + \delta/r\right] x_i dt + (1 - x_i dt) \left[1 - \prod_{k \neq i} (1 - x_k dt)\right] \left[v^F(d^*) + \delta e^{-rd^*}/r\right] + \left[\prod_{k=1}^{n} (1 - x_k dt)\right] e^{-rdt} V_i(t + dt) \quad (7)$$

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 derivesV_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:

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* $\hat{T} = 0$ *and the others follow at* $\hat{T} + d^*$*.*

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

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

so that the expected delay until the first adoption is:

$$\hat{T} = \frac{n-1}{n} \frac{v^F(d^*) - v^L - (\delta/r)\left(1 - e^{-rd^*}\right)}{rv^L + \delta} \tag{8}$$

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 ($v^L + \delta/r < 0$). It is so either because the adoption cost c is high or because the growth 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 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 dynamic version of a "chicken game" where all firms are willing to adopt but have conflicting views on who should go first. As usual in chicken games, this gives birth to a coordination problem leading to multiple Nash equilibria.

The economic interpretation of the equilibria in pure strategies (2a) is problematic because all firms have an incentive to free ride on the first adoption, so that no firm wishes to adopt first at $T = 0$. In that case, we can reasonably expect strategic delays in the first adoption. This corresponds to the symmetric equilibrium in mixed strategies (3b) where the expected date of the first adoption$\hat{T}$ is strictly positive. In the rest of the paper, we keep focused on this equilibrium. Note that, given (8), 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.

3.3 Welfare properties

We are now able to investigate the welfare properties of the C&T regime. To begin with, recall that the followers' decision is optimal as it does not generate any externalities of adoption. Turning next to the leader, Proposition 1 tells us that the first adoption will take place iff

$$v^L + \delta/r \geq 0 \iff c \leq \frac{\pi + \delta}{r} \tag{9}$$

Unsurprisingly, the comparison of (9) with the optimality condition (6) 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.

Interestingly, there exists a second inefficiency: Proposition 2 predicts an equilibrium in mixed strategies which involves a delay in the first adoption while the optimal date is $T^* = 0$.

We summarize these findings in:

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 it never occurs.*

In summary, the social inefficiency exclusively concerns the leader: diffusion starts either too late or gets stuck. By contrast, the followers make efficient decisions. This suggests that offering a premium for the first adopter could solve the problem.[7] We consider now this policy solution.

[7] In a totally different setting, this idea has been explored by Rosendahl, 2004, in a case

3.4 Premium to the leader

We immediately rule out subsidies that would be based on adoption dates for realism: in practice, regulators cannot know when it is $T = 0$ - or any other date - as there is no clear beginning of the diffusion process.[8] We focus the analysis on subsidies based on the adoption rank which is more easily observable by the regulator.

Consider a scheme where all firms that adopt first (eventually simultaneously) enjoy the same premium. Let α denote this subsidy. We obtain a new participation constraint $v^L + \alpha \geq 0$ whereas the optimality condition writes $v^L + (n-1)\, v^F(d^*) \geq 0$. It is then immediate that a premium $\hat{\alpha}$ such that

$$\widehat{\alpha} \equiv (n-1)\, v^F(d^*) \tag{10}$$

solve the underprovision problem. This is the classical story where the leader which generates a positive externality should receive a subsidy internalizing the full social benefit of adoption by the $(n-1)$ followers.

Until now, we have ignored the coordination problem. What about the impact of $\widehat{\alpha}$ on the possible delay before the first adoption? From (8), we know that the longer the delay, the larger the gap between followers' and leader's payoff. Therefore, granting the premium $\widehat{\alpha}$ to the leader obviously mitigates the problem as it reduces the payoff difference. Whether it is sufficient to solve it completely depends on the subsidy level, and thus on the size of the externality. More precisely, two cases are possible:

1. If $v^L + \widehat{\alpha} \geq v^F(d^*)$ - which is equivalent to $v^L + (n-2)\, v^F(d^*) \geq 0$ -, the leader adopts at $T = 0$. However, all firms will do the same as waiting is

where the policy instrument is a pollution tax.

[8] Or at least, the beginning date is technology and sector-specific the date of adoption so that regulators cannot know it (or they might be eventually informed ex post which is useless for granting subsidies to leaders).

less profitable than adopting immediately. This is inefficient as followers should wait for a delay $d^* > 0$.

2. If $v^L + \widehat{\alpha} < v^F(d^*)$, there remains a gap between leader's and followers' payoffs so that $T > 0$. Then, followers make efficient decisions by adopting after a delay d^*.

In order to limit the distortions created by the premium in the first case, one can imagine to reward a single firm. Once this firm gets the premium, the other ones' best reply is to follow after an optimal delay d^*. However, being a follower will remain more attractive in the second case, meaning that the second inefficient diffusion path will not be eliminated in equilibrium. We state rigorously these results:

Proposition 3 *1) Granting a premium $\widehat{\alpha} = (n-1)\, v^F(d^*)$ to a firm that adopts first improves welfare as diffusion occurs iff this is socially optimal. Importantly, the premium should be granted to a unique firm in the case where several firms want to take the lead.*

2) This exclusive subsidy $\widehat{\alpha}$ implements the first best optimum under C&T except when $v^L + (n-2)\, v^F(d^*) < 0 \leq v^L + (n-1)\, v^F(d^*)$. In this case, the premium $\widehat{\alpha}$ induces a path where a first firm adopts with a strictly positive delay $(T > 0)$ whereas the optimal date is $T^* = 0$.

Proof. See the appendix. ■

The intuition is simple. When first adoption entails a loss ($v^L < 0$) and the social benefit of technology diffusion is so weak that it hinges on the last follower (it would be negative with one less follower), the premium is not large enough to compensate the opportunity cost of moving first. Hence, there remains a positive delay before the first adoption.

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. This means that a firm that adopts a new technology get credits − and thus enjoys a benefit δ per period after adoption − only if technology adoption is not profitable without. By definition, adoption by a leader is thus 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 $e^{\lambda d} > 1$, it is obvious that additionality subsists if followers adopt before a threshold delay $d^{\max}$ defined by $c \equiv \pi e^{\lambda d^{\max}}/r$.

To sum up,

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

It is then convenient to analyze separately diffusion when the first adoption is additional ($\pi/r < c$) and when it is not ($\pi/r \geq c$).

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

This case is extremely simple. When the initial adoption is not additional, the same is obviously true for subsequent ones. Hence nobody receives credits. 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 zero in Proposition 2 to derive the equilibrium diffusion paths. This leads to:

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

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

Once a firm has adopted, the others follow after the delay $\tilde{d} = \hat{d}(0) = \frac{1}{\lambda}\ln\frac{c}{\pi}(r+\lambda)$.

4.3 Additional adoptions ($\pi/r < c$)

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

In the first case, we know from (4) that the delay would be $d^* = \frac{1}{\lambda}\ln\left(\frac{c(r+\lambda)}{\pi+\delta}\right)$ under C&T. However, choosing the optimal delay d^* under the CDM implies losing additionality when $d^* > d^{\max}$. Hence, keeping additionality imposes a delay such that:

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

Calculations easily show that $d^* < d^{\max}$ is equivalent to

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

If this condition is met, the additionality constraint is not binding. The firms can thus select the optimal delay $\widehat{d}$ and get credits. Things are more ambiguous when the condition is not met (e.g., if $\frac{\pi}{r} > \frac{\delta}{\lambda}$). The followers may then decide either to choose a delay $d^{\max}$ in order to benefit from additionality, or not to implement a CDM project, and rather choose a delay $\widetilde{d} = d^*(0)$ as shown in Lemma 2. They need to compare $v_F(d^{\max}) + \frac{\delta}{r}e^{-rd^{\max}}$ with $v_F(d^*(0))$.

Appropriate substitutions yield:

$$\begin{array}{rcl} v_F\left(d^{\max}\right)+\dfrac{\delta}{r}e^{-rd^{\max}} & > & v_F\left(d^*\left(0\right)\right) \\ & \Leftrightarrow & \\ \dfrac{\pi}{r} & < & \dfrac{\delta}{\lambda}\left(\dfrac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} \end{array}$$

This is very intuitive: followers shorten their adoption delay to implement a CDM project when the credit price is high and/or post-adoption market profit is low.

We summarize the whole analysis in the following:

Lemma 4 *If adoption can be 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:*

$$\tilde{d} = \begin{cases} \widehat{d} \ \ \textit{if} \ \frac{\pi}{r} < \frac{\delta}{\lambda} \\ d^{\max} \ \ \textit{if} \ \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} > \frac{\pi}{r} \geq \frac{\delta}{\lambda} \\ d^*\left(0\right) \ \ \textit{if} \ \frac{\pi}{r} \geq \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} \end{cases}$$

Note that $d^{\max} < \hat{d} < d^*\left(0\right)$: as compared to the socially optimal delay $\hat{d}$, additionality can either induce too slow or too fast diffusion depending on the size of π/r. The welfare effect of additionality seems complex.

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 2 and the results of Lemma 3, we easily obtain:

Proposition 4 *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

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

where u^L *and* u^F *are leader's and followers' payoffs (including credits). They are given by:*

$$\begin{aligned} u^L &= v^L + \frac{\delta}{r} \\ u^F(\tilde{d}) &= v^F(\tilde{d}) + \frac{\delta}{r}e^{-r\tilde{d}} \end{aligned}$$

where

$$\tilde{d} = \begin{cases} \hat{d} \textit{ if } \frac{\pi}{r} < \frac{\delta}{\lambda} \textit{ (all followers get credits)} \\ d^{\max} \textit{ if } \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} > \frac{\pi}{r} \geq \frac{\delta}{\lambda} \textit{ (all followers get credits)} \\ \hat{d} \textit{ if } \frac{\pi}{r} \geq \frac{\delta}{\lambda}\left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} \textit{ (no follower gets credits)} \end{cases}$$

5 Welfare comparison

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

When $c < \frac{\pi+\delta}{r}$, diffusion occurs and social welfare consists of the private adoption benefits of the firms - v^L and $v^F(d)$ - plus the social benefit corresponding to avoided abatement costs by ther credit buyers - δ/r and $(\delta/r)\,e^{-rd}$ for the leader and the followers, respectively. Therefore, social welfare discounted at date $T = 0$ writes

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

which simplifies as follows

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

We can now use (12) to compute the equilibrium welfare in the different cases.

5.1 C&T

By substituting x^* and d^* given in Proposition 2 in (12) we obtain a very simple expression:

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

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

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

In fact, welfare under C&T is the same as if all firms adopt immediately and simultaneously. It means that 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 $v^L + \delta/r$ for the $n-1$ followers - is entirely dissipated by the delay before the first adoption. In other terms, the learning benefits and the coordination cost exactly cancel each other out. This is not that counter-intuitive: 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

Under CDM Lemma 2 and Proposition 5 give four possible diffusion paths which we consider in turn.

5.2.1 Case 1: $c \leq \pi/r$

In this case, adoptions are not additional and firms receive zero credits. Accordingly we substitute $\tilde{x}$ and $\tilde{d}$ from Lemma 1 in (12) leading to

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

Case 2: $\pi/r < c < (\pi + \delta)/r)$ **and** $\frac{\pi}{r} < \frac{\delta}{\lambda}$

In this case, all firms get credits and $\tilde{d} = \hat{d}$. Hence it is immediate that $W_{CDM} = W_{C\&T}$

Case 3: $\pi/r < c < (\pi + \delta)/r)$ **and** $\frac{\delta}{\lambda} \left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}} > \frac{\pi}{r} \geq \frac{\delta}{\lambda}$

In this case, all firms get credits as well but they adopt sooner at $\tilde{d} = d^{\max}$. Substituting $\tilde{x}$ and $d^{\max}$ in (12) yields $W_{CDM} = n(v^L + \delta/r) = W_{C\&T}$.

Case 4: $\pi/r < c < (\pi + \delta)/r)$ **and** $\frac{\pi}{r} \geq \frac{\delta}{\lambda} \left(\frac{r+\lambda}{r}\right)^{\frac{r+\lambda}{\lambda}}$

In this last case, credits are only granted to the leader so that $u^L = v^L + \delta/r$ and $u^F(\tilde{d}) = v^F(d^*(0))$. Substituting u^L and $u^F(\tilde{d})$ in x^* and then x^* and $d^*(0)$ in (12) leads to

$$W_{CDM} = nu^L \left[1 + \frac{(n-1)e^{-rd^*(0)}}{u^L + (n-1)u^F(d^*(0))} \frac{\delta}{r}\right] \tag{15}$$

Then, very simple calculations show that:

Proposition 5 *There are two cases where CDM & C&T are not welfare equivalent:*

1. *C&T dominates CDM when adoption by the leader is not additional under CDM so that no firms receive any credits (Case 1).*

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

Proof. By comparing (13), (14), and (15). ■

Let us comment on these results. To begin with, the fact that C&T outperforms CDM when all adoptions are not additional 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 ($d^* = \hat{d}$). 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.

The second result is very interesting. Like 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. Proposition 6 shows that the latter effect outweighs the former.

This 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. Instead, the CDM punishes the followers. Proposition 6 says that this solution partly mitigates the externality problem.

These results do not depend on the value of the parameters π, λ, c, or r. Proposition 6 thus allows to derive robust policy lessons: one should grant credits to non-additional projects when significant learning spillovers are expected.

Finally, the fact that CDM & C&T perform equally when everybody gets credits is not that intuitive as well. Recall that followers may adopt too early at $d^{\max} < \hat{d}$ to meet the additionality requirement under the CDM. The proposition says that this distortion is welfare neutral. One 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.

6 Conclusion

Kyoto mechanisms like the CDM are often depicted as a powerful lever for the diffusion of environmental technologies in developing countries. In this paper, we explore this insight by developing a simple model capturing 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 CDM originality is the additionality requirement, whereby credits are only granted to projects which 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 Cap and Trade system and the CDM. In the presence of learning spillovers we have shown that C&T fails to implement the optimal diffusion path for a classical reason: the leading firm - who generates positive externalities - and the followers receive the same amount of credits.

By design, the CDM either yields the same quantity of credits as C&T or zero credits when the project is not additional. Hence it 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 be useful. In fact, the CDM yields a higher welfare than C&T in the case where the leader receive credits whereas the followers don't. This does not solve the underprovision problem. But it mitigates coordination costs. The result is not ambiguous. It thus provides a strong case for relaxing the additionality requirement for non additional projects when significant learning spillovers are expected.

In post-Kyoto talks, whether emitters located in emerging economies like China, India or Brazil should be covered by a CDM-like mechanism featured by additionality or a Cap & Trade scheme is the subject of intense discussion. Our analysis stresses one advantage of the CDM: the additionality requirement can be tailored to speed up technology diffusion as compared to other emissions trading schemes.

Of course, technology policy solutions are also possible. In this regard, we have shown that combining CDM or C&T with adoption subsidies to leading firms is appealing. But the main focus of our analysis was to see whether trading mechanims *per se* could partly solve the positive externality problem.

7 Appendix

7.1 Proof of Proposition 2

The firm i's expected payoff at any time t is given by (7). Using this equation we derive successively the conditions for the different equilibria to arise.

7.1.1 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

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

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:

$$V_i = \frac{x_i v^L}{r + x_i} \tag{16}$$

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

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

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

$$V_i = v^F(d^*) + x_i dt \left[v^L - v^F(d^*)\right]$$

Recall that $v^L < v^F(d^*)$ as $d^* = \hat{d} > 0$ 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 16 we know that firm j's payoff is $V_j = x_j v^L / (r + x_j)$ 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 $v^F(d^*) > v^L > 0$.

7.1.3 Case 3: all firms play mixed strategies

Consider again the expected payoff of firm i in (7). 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:

$$V_i=\frac{x_i v^L+\sum\limits_{k\neq i} x_k v^F(d^*)}{r+\sum\limits_{k} x_k}$$

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

$$\sum_{k\neq i} x_k=\frac{rv^L}{v^F(d^*)-v^L} \tag{17}$$

It is clear from 17 that only one equilibrium is possible, where $x_i^*=x^*$ for all $i=1,...,n$. The equilibrium adoption strategy is then:

$$x^*=\frac{rv^L}{[n-1]\left[v^F(d^*)-v^L\right]} \tag{18}$$

The strategy x^* followed by each firm defines a Poisson process of parameter nx^* for the first adoption. This allows us to calculate the expected delay until the first adoption:

$$E(T)=\int_0^\infty tnx^*e^{-nx^*t}dt=\frac{n-1}{n}\frac{v^F(d^*)-v^L}{rv^L} \tag{19}$$

7.2 Proof of Proposition 4

We need to investigate precisely the impact of $\widehat{\alpha}$ when it is exclusively granted to a unique firm. This is not so straightforward as it is not possible to replace v^L by $v^L+\widehat{\alpha}$ in the Bellman equation (7) for all firms as just one obtains the premium. As a result, the first adoption game does not solve according to Proposition 1.

A firm's willingness to accept the premium depends on the difference between its payoff if it adopts at $T=0$ and its payoff if not. In turn the payoff of refusing the premium depends on whether another firm accepts it.

Assume that another firm would accept the premium and adopt at $T=0$. Then the best reply of the other firms is to wait a delay d^* before adopting in turn the technology, so that their payoff is $v^F\left(\widehat{d}\right)$. Knowing this, a firm will accept the premium if $v^L+\widehat{\alpha}\geq v^F\left(\widehat{d}\right)$. This condition thus implies that one firm will accept the premium and adopt at time $\widehat{T}=0$ while the other will follow after a delay $\widehat{d}$.

If, on the other hand, we have $v^L+\widehat{\alpha}<v^F\left(\widehat{d}\right)$, then being a follower (with $\widehat{d}>0$) is more profitable than accepting the premium. In this case the adoption game corresponds to the Bellman equation (**??**) in which v^L is replaced with $v^L+z\widehat{\alpha}$ where z denotes the firm i's probability to obtain the premium when it decides to adopt the technology. Ruling out pure strategies, the likelihood that two firms or more adopt simultaneously is a term in $(dt)^n<1$, with $n>1$. For small time increments, this term becomes negligible ($(dt)^n\sim 0, n>1$) such that $z\sim 1$. As a result, Proposition 2, point 2c, can apply.

References

[1] 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.

[2] 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.

[3] 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.

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

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

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

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

[8] 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.

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

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

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

[12] 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.

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

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

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

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

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