Environmental Policy and Uncertain Arrival of Future Abatement Technology: Extra Appendix

by Andreas von Döllen

Wingas Corporation, Friedrich-Ebert-Strasse 160, 34119 Kassel, Germany
Email: andreas-von.doellen@wingas.de

and

Till Requate (corresponding author)

Department of Economics, University of Kiel, Olshausenstrasse 40, 24118 Kiel, Germany
Email: requate@economics.uni-kiel.de

This appendix provides the proofs of the main results of the main paper (section A) and also includes some additional comparative statics results (section B). In particular in section B.1, we study how a change of a tax rate under long-run tax policy influences the firms' behavior, and in section B.2 we characterize the equilibria arising under different amounts of permits under long-term tax policies.

A Proofs

Proof of Proposition 3.1: Ad 1.-3.: Given n_a at stage 2 the social planner solves

$$\min_{\{n_b,n_0,e_0^2,e_a^2,e_b^2\}} \{\frac{1}{r}(n_0C_0(e_0^2) + n_aC_a(e_a^2) + n_bC_b(e_b^2) \tag{A.1}$$

$$+D(n_0e_0^2 + n_ae_a^2 + n_be_b^2)) + n_bF_b \tag{A.2}$$

subject to the constraints $n_0 \geq 0$, $n_b \geq 0$ and $n_0 + n_a + n_b = 1$ with corresponding Kuhn-Tucker multipliers μ_i for the non-negativity constraints $n_i \geq 0$ ($i = 0, b$) and Lagrange multiplier ν w.r.t. $n_0 = 1 - n_a - n_b$. The first-order conditions w.r.t. e_i^2, n_0 and n_b are given by

$$C_i'(e_i^2) + D'(E^2) = 0,\ i = 0, a, b \tag{A.3}$$

$$\frac{1}{r}(C_0(e_0^2) + e_0^2D'(E^2)) - \mu_0 - \nu = 0 \tag{A.4}$$

$$\frac{1}{r}(C_b(e_b^2) + e_b^2D'(E^2)) + F_b - \mu_b - \nu = 0 \tag{A.5}$$

Eliminating ν from equations (A.4) and (A.5) yields

$$\frac{1}{r}(C_b(e_b^2) - C_0(e_0^2) + (e_b^2 - e_0^2)D'(E^2)) + F_b - \mu_b + \mu_0 = 0 \tag{A.6}$$

Considering first the interior solutions (i.e. $\mu_0 = \mu_b = 0$), we differentiate the equation system (A.3) and (A.6) with respect to F_b . Employing the Envelope Theorem, we obtain

$$C_i''(e_i^2)\frac{\partial e_i^2}{\partial F_b} + D''(E)\frac{\partial E^2}{\partial F_b} = 0,\ i = 0, a, b \tag{A.7}$$

and

$$\frac{1}{r}D''(E)\frac{\partial E^2}{\partial F_b}(e_b^2 - e_0^2) + 1 = 0 \tag{A.8}$$

Solving for $\partial E^2/\partial F_b$ yields

$$\frac{\partial E^2}{\partial F_b} = \frac{r}{D''(E)(e_0^2 - e_b^2)} > 0$$

Substituting this equation into (A.7) yields $\partial e_i^2/\partial F_b < 0$ for each type of firm. We can write the total emissions in period 2 as

$$E^2(F_b) = (1 - n_a - n_b)e_0^2 + n_a e_a^2 + n_b e_b^2.$$

Differentiating this equation with respect to F_b and solving the result for $\partial n_b/\partial F_b$ yields

$$\frac{\partial n_b}{\partial F_b} = \frac{\frac{\partial E^2}{\partial F_b} - (1 - n_a - n_b)\frac{\partial e_0^2}{F_b} - n_a \frac{\partial e_a^2}{F_b} - n_b \frac{\partial e_b^2}{F_b}}{e_b^2 - e_0^2} < 0.$$

Note now that the expression $C_b(e_b^2) - C_0(e_0^2) + (e_b^2 - e_0^2)D'(E^2)$ decreases in E^2. Let us now use $\overline{E}^2$ and $\underline{E}^2$ to denote the maximum and the minimum feasible emission levels at the second stage occurring when no firm, or all remaining firms, respectively, adopt technology b. Furthermore we use $\underline{F}_b(n_a)$ and $\overline{F}_b(n_a)$ to denote the fixed investment costs for which complete or no adoption of technology b by the remaining $1 - n_a$ firms is optimal. Then for each $F \in (\underline{F}_b(n_a), \overline{F}_b(n_a)\,)$ we find an aggregate emission level E^2, satisfying

$$C_b(e_b^2) - C_0(e_0^2) + (e_b^2 - e_0^2)D'(E^2) + F = 0,$$

and a corresponding n_b ($0 < n_b < 1 - n_a$) such that

$$E^2 = n_a e_a^2 + n_b e_b^2 + (1 - n_a - n_b)e_0^2.$$

Conversely, every F_b for which an interior solution n_b of (A.1) exists must satisfy $F_b \in (\underline{F}_b(n_a), \overline{F}_b(n_a)\,)$ since for the corresponding E^2 we have $E^2 \in (\overline{E}^2, \underline{E}^2)$.

ad 4.: For the maximum feasible aggregate emission level $\overline{E}^2 = n_a e_a^2 + (1 - n_a)e_0^2$, where e_i^2 satisfies (A.3) we obtain

$$\frac{\partial \overline{E}^2}{\partial n_a} = (e_a^2 - e_0^2) \Big/ \left[1 + D'' \left(\frac{n_a}{\frac{\partial^2 C_a}{(\partial e_a^2)^2}} + \frac{1 - n_a}{\frac{\partial^2 C_0}{(\partial e_0^2)^2}} \right)\right] < 0.$$

Analogously, for the minimal feasible emission level we obtain

$$\frac{\partial \underline{E}^2}{\partial n_a} = (e_a^2 - e_b^2) \Big/ \left[1 + D'' \left(\frac{n_a}{\frac{\partial^2 C_a}{(\partial e_a^2)^2}} + \frac{1 - n_a}{\frac{\partial^2 C_b}{(\partial e_b^2)^2}} \right)\right] > 0.$$

Since $\overline{E}^2 = n_a e_a^2 + (1 - n_a)e_0^2$ and $\underline{E}^2 = n_a e_a^2 + (1 - n_a)e_b^2$, both boundary levels of E tend to the same aggregate emission level $E_M = e_a^*(1)$ as n_a tends to 1. Thus all firms use technology a. Now observe that $\underline{F}_b(n_a)$ and $\overline{F}_b(n_a)$ are given by the following two equations

$$\begin{aligned} -[C_b(e_b^2) - C_0(e_0^2) + (e_b^2 - e_0^2)D'(\overline{E}_2)]/r &= \overline{F}_b(n_a) \\ -[C_b(e_b^2) - C_0(e_0^2) + (e_b^2 - e_0^2)D'(\underline{E}^2)]/r &= \underline{F}_b(n_a) \end{aligned}$$

By differentiating these with respect to n_a we can easily show that $\underline{F}_b(n_a)$ increases and $\overline{F}_b(n_a)$ decreases in n_a. Since the LHSs of both equations obviously tend to the same value, as n_a tends to 1, both values $\underline{F}_b(n_a)$ and $\overline{F}_b(n_a)$ tend to the same level F_M.

ad 5.: Let $F_b \in (\underline{F}_b(n_a), \overline{F}_b(n_a))$ be fixed. Differentiating (A.3) and (A.6) with respect to n_a and applying the Envelope Theorem we obtain

$$C_i''(e_i)\frac{\partial e_i^2}{\partial n_a} + D''(E^2)\frac{\partial E^2}{\partial n_a} = 0,\ i = 0, a, b \tag{A.9}$$

and

$$\frac{1}{r}D''(E^2)\frac{\partial E^2}{\partial n_a}(e_b^2 - e_0^2) = 0 \tag{A.10}$$

Hence $\partial E^2/\partial n_a = 0$. Substituting this into (A.9) yields $\partial e_i^2/\partial n_a = 0$ for $i = 0, a, b$.

Proof of Proposition 3.2, parts 1-3: Note first that for given F_b four general possible scenarios exist. If $F_b \geq \overline{F}_b(0)$ for all n_a, no adoption of technology b will be socially optimal at the second stage since $\overline{F}_b$ decreases in n_a. If $F_b \leq \underline{F}_b(0)$ for all n_a, full adoption of technology b by the remaining firms will be socially optimal at the second stage. If $F_b \in [F_M, \overline{F}_b(0))$, by Proposition 3.1 we find a unique $\hat{n}_a \leq 1$ such that partial adoption of technology b is socially optimal for $n_a < \hat{n}_a$ and no adoption is optimal for $n_a \geq \hat{n}_a$. Conversely, if $F_b \in (\underline{F}_b(0), F_M]$, we know from Proposition 3.1 that a unique $\hat{n}_a \leq 1$ exists such that partial adoption of technology b is socially optimal for $n_a < \hat{n}_a$, and full adoption of technology b by the remaining firms will be optimal for $n_a \geq \hat{n}_a$. For the latter two cases we now solve the minimization problem under the restrictions $n_a \in (0, \hat{n}_a)$ and $n_a \in (\hat{n}_a, 1)$ and then add up the results.

Since obviously optimal total emissions $E^2(n_a)$ are continuous in n_a, we conclude that $n_a^*(F_a)$ is continuous. First we consider the case $F_b > \overline{F}_b(0)$. Hence for any n_a adoption of technology b will never be socially optimal. Therefore, in the case of $n_a = 0$, we can apply the proof of Proposition 3.1 i.) - iii.).

Next consider $F_b < \underline{F}_b(0)$. In this case, for any n_a it is socially optimal that the remaining firms will adopt technology b. Therefore, at the first stage, the social planner solves the problem

$$\begin{aligned}
&\min_{\{e_0^1, e_a^1, n_0, n_a\}} \frac{1}{\lambda + r} \left[n_0 C_0(e_0^1) + n_a C_a(e_a^1) + D(n_0 e_0^1 + n_a e_a^1) \right] + n_a F_a \\
&+ \frac{\lambda}{\lambda + r} \frac{1}{r} \left[n_0 C_b(e_b^2(n_a)) + n_a C_a(e_a^2(n_a)) + D(n_0 e_b^2(n_a) + n_a e_a^2(n_a)) \right] \\
&+ \frac{\lambda}{\lambda + r} n_0 F_b \qquad \text{(A.11)}
\end{aligned}$$

under the constraints $n_0 \geq 0$, $n_a \geq 0$, and $n_0 + n_a = 1$. Now we set $E^1 = n_0 e_0^1 + n_a e_a^1$ and $E^2 = n_0 e_b^2(n_a) + n_a e_a^2(n_a)$. Again μ_i, $i = 0, a$, is the Kuhn-Tucker multiplier for $n_i \geq 0$ and ν the Lagrange multiplier for $n_0 + n_a = 1$. Following the same procedure as in the proof of Proposition 3.1, i.e. differentiating equation (A.11) with respect to e_i^1 and n_i, $i = 0, a$, applying the Envelope Theorem, and eliminating ν we obtain:

$$\begin{aligned}
F_a =& \frac{1}{\lambda + r} (C_0(e_0^1) - C_a(e_a^1) + (e_0^1 - e_a^1) D'(E^1)) \\
&+ \frac{\lambda}{\lambda + r} \frac{1}{r} (C_b(e_b^2 - C_a(e_a^2)) + (e_b^2 - e_a^2) D'(E^2)) + \frac{\lambda}{\lambda + r} F_b + \mu_a - \mu_0 \qquad \text{(A.12)}
\end{aligned}$$

Considering first the interior solutions (i.e. $\mu_0 = \mu_a = 0$) and multiplying the last equation by $r(\lambda + r)$ we obtain:

$$\begin{aligned}
r(\lambda + r) F_a =& r(C_0(e_0^1 - C_a(e_a^1)) + (e_0^1 - e_a^1) D'(E^1)) \\
&+ \lambda(C_b(e_b^2 - C_a(e_a^2)) + (e_b^2 - e_a^2) D'(E^2)) + r\lambda F_b
\end{aligned} \qquad \text{(A.13)}$$

Furthermore, in the second period

$$C_i'(e_i^2) + D'(E^2) = 0, \ i = a, b \qquad \text{(A.14)}$$

still holds. Differentiating (A.14) with respect to n_a, we obtain

$$-C_i''(e_i^2) \frac{\partial e_i^2}{\partial n_a} = D''(E^2)[e_a^2 - e_b^2 + n_a \frac{\partial e_a^2}{\partial n_a} + (1 - n_a) \frac{\partial e_b^2}{\partial n_a}], \ i = a, b.$$

Thus $\partial e_a^2 / \partial n_a = [C_b''(e_b^2) / C_a''(e_a^2)] \cdot \partial e_b^2 / \partial n_a$. Substituting this expression into (A.13) we obtain

$$\frac{\partial e_b^2}{\partial n_a} = \frac{D''(E^2)(e_a^2 - e_b^2)}{-C_i''(e_i^2) - D''(E^2) n_a \frac{C_b''(e_b^2)}{C_a''(e_a^2)} - D''(E^2)(1 - n_a)} < 0$$

This implies $\partial e_a^2/\partial n_a < 0$. Since $D''(E^2) \cdot \partial E^2/\partial n_a = \partial^2 C/(\partial e)^2 \cdot \partial e_a^2/\partial n_a$ we obtain $\partial E^2/\partial n_a > 0$. Note that we can reinterpret the maximization problem above as a two stage procedure where we first determine e_a^1, e_0^1 for given n_a. These values will be delivered by (A.14) together with $E^1 = n_a e_a^1 + (1-n_a)e_0^1$. Now doing the analogous calculations as above, we derive $\partial e_i^1/\partial n_a > 0$, $i = 0, a$, and $\partial E^1/\partial n_a < 0$. With these results we are able to calculate $\partial E^1/\partial F_a$ and $\partial n_a/\partial F_a$. For that purpose we differentiate (A.13) and (A.14) w.r.t. F_a yielding

$$C_i''(e_i^1)\frac{\partial e_i^1}{\partial F_a} + D''(E^1)\frac{\partial E^1}{\partial F_a} = 0,\ i = 0, a,$$

and

$$r(\lambda + r) = r(e_0 - e_a)D''(E^1)\frac{\partial E^1}{\partial F_a} + \lambda(e_b^2 - e_a^2)D''(E^2)\frac{\partial E^2}{\partial F_a}.$$

Using $\partial E^1/\partial F_a = \partial E^1/\partial n_a \cdot \partial n_a/\partial F_a$ and solving for $\partial n_a/\partial F_a$ yields

$$\frac{\partial n_a}{\partial F_a} = \frac{r(\lambda + r)}{r(e_0^1 - e_a^1)D''(E^1)\frac{\partial E^1}{\partial n_a} + \lambda(e_b^2 - e_a^2)D''(E^2)\frac{\partial E^2}{\partial n_a}} < 0$$

From this we can derive $\partial e_i^1/\partial F_a < 0$, $\partial e_i^2/\partial F_a > 0$, $\partial E^1/\partial F_a > 0$, and $\partial E^2/\partial F_a < 0$. Furthermore by differentiating the RHS of equation (A.12) with respect to n_a, still assuming an interior solution (i.e. $\mu_0 = \mu_a = 0$), we verify that the RHS of (A.12) decreases in n_a. Thus, analogously to the proof of Proposition 3.1, we can show that an interval $[\underline{F}_a, \overline{F}_a]$ exists such that $\mu_0 = \mu_a = 0$ for $F_a \in [\underline{F}_a, \overline{F}_a]$, $\mu_0 > 0$ for $F_a > \overline{F}_a$, and $\mu_a > 0$ for $F_a < \underline{F}_a$.

Finally let $F_b \in [\underline{F}_b(0), \overline{F}_b(0)]$. Depending on whether $F_b > F_M$ or $F_b < F_M$, an $\overline{n}_a < 1$ exists such that $F_b = \underline{F}_b(\overline{n}_a)$ or $F_b = \overline{F}_b(\overline{n}_a)$. The social planner's problem can then be written as

$$\begin{aligned}
&\min_{\{e_0,e_a,n_0,n_a\}} \{\frac{1}{\lambda + r}(n_0 C_0(e_0^1) + n_a C_a(e_a^1) + D(n_0 e_0^1 + n_a e_a^1)) + n_a F_a \\
&+ \frac{\lambda}{\lambda + r}\frac{1}{r}((1 - n_a - n_b(n_a))C_0(e_0^2(n_a)) + n_b(n_a)C_b(e_b^2(n_a)) + n_a C_a(e_a^2(n_a)) \\
&+ D((1 - n_a - n_b(n_a))e_0^2(n_a) + n_b(n_a)e_b^2(n_a) + n_a e_a^2(n_a)) + r n_b(n_a)F_b)\}
\end{aligned}$$

subject to the constraints $0 \le n_a \le \overline{n}_a$ and $n_0 + n_a = 1$. Set $E^1 := n_0 e_0^1 + n_a e_a^1$ and $E^2 := (1 - n_a - n_b(n_a))e_0^2(n_a) + n_b(n_a)e_b^2(n_a) + n_a e_a^2(n_a)$. Analogously, as in the case above, we can derive from the first-order conditions

$$\begin{aligned}
0 \;=\; &\frac{1}{\lambda + r}(C_a(e_a^1) - C_0(e_0^1) + (e_a^1 - e_0^1)D'(E^1)) \qquad \text{(A.15)} \\
&+ \frac{\lambda}{\lambda + r}\frac{1}{r}(C_a(e_a^2) - C_0(e_0^2) + (e_a^2 - e_0^2)D'(E^2)) + F_a
\end{aligned}$$

Now we can apply Proposition 3.1 iii.) to the case of partially adopting technology b. Hence both e_i^2, $i = 0, a, b$, and E^2 only depend on F_a but neither on n_a nor on F_a. Thus, by differentiating (A.14) and (A.15) w.r.t. F_a, we obtain

$$\frac{\partial^2 C_i}{(\partial e_i^2)^2}\frac{\partial e_i^2}{\partial F_a} + D''(E^1)\frac{\partial E^1}{\partial F_a} = 0, \; i = 0, a,$$

and

$$\frac{1}{\lambda + r}\left((e_a^1 - e_0^1)D''(E^1)\frac{\partial E^1}{\partial F_a}\right) + 1 = 0$$

Therefore

$$\frac{\partial E^1}{\partial F_a} = \frac{\lambda + r}{(e_0^1 - e_a^1)D''(E^1)} > 0$$

Following the proof of Proposition 3.1 we derive $\partial e_i^1/\partial F_a < 0$, for $i = 0, a$, and $\partial n_a/\partial F_a < 0$. Thus we find a minimum value F_{min} and a maximum value F^{max} such that for each $F_a \in (F_{min}, F^{max})$ we will have an interior solution for n_a. Now, if $F_b = F_M$, then obviously $\underline{F}_a = F_{min}$ and $\overline{F}_a = F^{max}$. For the case that no or full adoption of technology b is optimal, the result follows immediately. Otherwise, as derived above, two intervals $[\underline{F}, \overline{F}]$ and $[\underline{F}', \overline{F}']$ exist such that for all $F_a \in (\underline{F}, \overline{F})$ we have $0 < n_a < \hat{n}_a$, and for $F_a \in (\underline{F}', \overline{F}')$ we have $\hat{n}_a < n_a < 1$, and either no adoption (in case of $F_b > F_M$) or full adoption of technology b (in case of $F_b < F_M$) follows. Obviously $\overline{F} = \underline{F}'$. Thus $[\underline{F}, \overline{F}']$ is the interval we are looking for.

Part 4: Let us first consider the case where partial adoption is socially optimal at the second stage. Thus (A.15) is relevant. If we differentiate (A.15) with respect to λ, solve for $\partial n_a/\partial\lambda$, and apply that E^2 is independent of n_a, we obtain

$$\begin{aligned}\frac{\partial n_a}{\partial \lambda} &= (C_a(e_a^1) - C_0(e_0^1) + (e_a^1 - e_0^1)D'(E^1) - (C_a(e_a^2) - C_0(e_0^2) \\ &\quad +(e_a^2 - e_0^2)D'(E^2)))/(\lambda + r)(e_a^1 - e_0^1)D''(E^1)\frac{\partial E^1}{\partial n_a}\end{aligned}$$

The second part of the numerator, representing the second stage, is smaller than the first part in absolute values. Thus the numerator is negative. The denominator is clearly positive.

If we differentiate (A.15) with respect to F_b and solve for $\partial n_a/\partial F_b$ we obtain:

$$\frac{\partial n_a}{\partial F_b} = \frac{\frac{\lambda}{r}(e_a^2 - e_0^2)D''(E^2)\frac{\partial E^2}{\partial F_b}}{(e_0^1 - e_a^1)D''(E^1)\frac{\partial E^1}{\partial n_a}}$$

Both the numerator and the denominator are negative, since $\partial E^1/\partial n_a < 0$ and $\partial E^2/\partial F_b > 0$.

In case of full adoption at the second stage, we differentiate (A.12) with respect to λ and again solve for $\partial n_a/\partial\lambda$ to obtain:

$$\begin{aligned}\frac{\partial n_a}{\partial\lambda} &= \frac{C_a(e_a^1)-C_0(e_0^1)+(e_a^1-e_0^1)D'(E^1)}{(\lambda+r)[(e_a^1-e_0^1)D''(E^1)\frac{\partial E^1}{\partial n_a}+\frac{\lambda}{r}(e_a^2-e_b^2)D''(E^2)\frac{\partial E^2}{\partial n_a}]}\\ &\quad-\frac{C_a(e_a^2)-C_0(e_0^2)+(e_a^2-e_0^2)D'(E^2)}{(\lambda+r)[(e_a^1-e_0^1)D''(E^1)\frac{\partial E^1}{\partial n_a}+\frac{\lambda}{r}(e_a^2-e_b^2)D''(E^2)\frac{\partial E^2}{\partial n_a}]}\\ &\quad+\frac{(C_b(e_b^2)-C_0(e_0^2)+(e_b^2-e_0^2)D'(E^2)+rF_b}{(\lambda+r)[(e_a^1-e_0^1)D''(E^1)\frac{\partial E^1}{\partial n_a}+\frac{\lambda}{r}(e_a^2-e_b^2)D''(E^2)\frac{\partial E^2}{\partial n_a}]}\end{aligned}$$

The denominator is clearly positive. The difference of the first two parts of the numerator is negative for the same reasons as in case of partial adoption at the second stage. The third part basically represents the first-order condition with respect to n_b. Since, however, the optimum is adopted at the upper boundary for n_b (complete adoption), this term must be negative. Analogously we obtain:

$$\frac{\partial n_a}{\partial F_b} = \frac{\lambda}{(e_a^1-e_0^1)D''(E^1)\frac{\partial E^1}{\partial n_a}+\lambda\frac{1}{r}(e_a^2-e_b^2)D''(E^2)\frac{\partial E^2}{\partial n_a}} < 0$$

Part 5: It is obvious that neither $\underline{F}_a$ nor $\overline{F}_a$ depends on F_b if no adoption of technology b is optimal at the second stage, which for $\underline{F}_a$ is the case if and only if $F_b > F_M$, while for $\overline{F}_a$ this is the case if and only if $F_b > \overline{F}_b(0)$. In case of partial adoption at the second stage, we only have to consider $\overline{F}_a$. Partial adoption will be optimal if and only if $\underline{F}_b(0) < F_b < \overline{F}_b(0)$. In that case $\overline{F}_a(F_b)$ is given by equation (A.15) where $E^1 = e_0^1$ and $E^2 = n_b e_b^2 + (1-n_b)e_0^2$ and n_b is the optimal share of firms using technology b. From these equations we obtain

$$= \frac{\lambda}{(\lambda+r)r}(e_a^2-e_0^2)D''(E^2)\frac{\partial E^2}{\partial F_b} + \frac{\partial\overline{F}_a}{\partial F_b}.$$

Since $\partial E^2/\partial F_b = r/(D''(E^2)(e_b^2-e_0^2))$ we get $\partial\overline{F}_a/\partial F_b = \lambda/(\lambda+r)\cdot[(e_a^2-e_0^2)/(e_b^2-e_0^2)$. For $\overline{F}_a(F_b)$ full adoption of technology b can only occur if and only if $F_b < F_M$. For $\overline{F}_a(F_b)$ full adoption of technology b is the relevant scenario if and only if $F_b < \underline{F}_b(0)$. Thus to derive the effect on $\overline{F}_a(F_b)$ we have to differentiate (A.12) considering $E^1 = e_0^1$ and $E^2 = e_b^2$. Thus we obtain:

$$0 = \frac{\partial\overline{F}_a}{\partial F_b} - \frac{\lambda}{(\lambda+r)}.$$

Analogously we can compute that $\frac{\partial \underline{F}_a}{\partial F_b} = \frac{\lambda}{(\lambda+r)}$. Finally note that on $[\underline{F}_b(0), F_M]$ we have $\partial \underline{F}_a/\partial F_b > \partial \overline{F}_a/\partial F_b$ since $(e_a^2 - e_0^2)/(e_b^2 - e_0^2) < 1$.

Proof of Proposition 4.1: The tax case: First we will analyze the second stage where n_a is given. The regulator will set $\tau^2 = D'(E^{2^*}(n_b))$ where $E^{2^*}(n_b)$ is the socially optimal aggregate emission level corresponding to n_b. To find the equilibrium value for n_b, we have to compare τ to the threshold level $\overline{\tau}_b$ satisfying $\Delta_{0b}(\overline{\tau}_b) = 0$. Looking at the proof of Proposition 3.1, $\Delta_{0b}(\overline{\tau}_b) = 0$ corresponds to the first-order condition for the socially optimal n_b if we replace $\overline{\tau}_b$ by $D'(E^2)$. Therefore $\overline{\tau}_b = D'(E^{2^*}(n_b^*))$ holds if the optimal share n_b^* satisfies $0 < n_b^* < 1 - n_a$. In this case the resulting share of firms adopting technology b must be socially optimal. Otherwise $\tau > D'(E^{2^*}(n_b^*))$ or $\tau < D'(E^{2^*}(n_b^*))$ must hold by the *ex post* regulation rule and due to the fact that $E^{2^*}(n_b)$ decreases in n_b by Proposition 3.1. If $n_b^* = 0$, then we conclude from the proof of Proposition 3.1 that $D'(E^{2^*}(n_b^*)) \leq \overline{\tau}_b$ holds. Since E^{2^*} decreases in n_b, it must be the case that $D'(E^{2^*}(n_b)) < \overline{\tau}_b$ holds for any share $n_b > 0$. Thus partial or full adoption cannot occur in equilibrium. An analogous argument applies to the case $n_b^* = 1 - n_a$.

Now we analyze the **first stage**. First, for each n_a we define the threshold tax rate $\overline{\tau}(n_a)$. If no adoption of technology b happens at the second stage under *ex post* regulation, $\overline{\tau}(n_a)$ is given by

$$\Delta C_{0a}(\overline{\tau}(n_a), D'(E^{2^*}(0))) = 0.$$

If partial adoption of technology b happens, $\overline{\tau}(n_a)$ is given by

$$\Delta C_{ab}(\overline{\tau}(n_a), D'(E^{2^*}(n_b^*(n_a)))) = 0,$$

while if full adoption of technology b follows, the threshold tax rate is given by

$$\Delta C_{ab}(\overline{\tau}(n_a), D'(E^{2^*}(1 - n_a))) = 0.$$

Relating these equations to those in the proof of Proposition 3.2 we see that these correspond to the first-order conditions of the socially optimal share at the first stage. If in particular partial adoption of both technologies is optimal, then $\Delta C_{ab}(D'(E^{1^*}), D'(E^{2^*})) = 0$ will hold. Thus, if for given F_a and F_b we have $0 < n_a^* < 1$, then $\overline{\tau}(n_a^*) = D'(E^{1^*}(n_a^*))$ will be the case. Furthermore $n_a^* = 1$ is equivalent to $\overline{\tau}(1) \leq D'(E^{1^*}(1))$, and $\overline{\tau}(0) \geq D'(E^{1^*}(0))$ is equivalent to $n_a^* = 0$. Thus the socially optimal outcome is also an equilibrium with the optimal tax rate. We still need to show that it is unique. Thus we assume first that $\overline{\tau}(n_a)$ is non-decreasing in n_a: First consider the case $0 < n_a^* < 1$. Since E^{1^*} decreases in n_a for all $n_a < n_a^*$, we must have $\overline{\tau}(n_a) < D'(E^{1^*}(n_a))$, and for all $n_a > n_a^*$ we must have $\overline{\tau}(n_a) > D'(E^{1^*}(n_a))$. Therefore, in the first case, the remaining firms have a further incentive to adopt technology a, while in the second case, too many firms would have adopted technology a. In case of $n_a^* = 0$, for each share $n_a > 0$ we would have $\overline{\tau}(n_a) > D'(E^{1^*}(n_a))$. Thus partial or full adoption of technology a cannot be a market equilibrium. An analogous argument applies to the case of $n_a^* = 1$. Thus it only remains to show that indeed $\overline{\tau}(n_a)$ is non-decreasing. For this purpose we distinguish three cases:

Case 1: No adoption will be expected in equilibrium at the second stage: Then $\overline{\tau}(n_a)$ is given by $\Delta C_{0a}(\overline{\tau}(n_a), D'(E^{2^*}(0))) = 0$. Obviously $\partial E^{2^*}/\partial n_a < 0$. Thus

$$\frac{\partial \bar{\tau}}{\partial n_a} = \frac{(e_a^2 - e_0^2) D''(E^{2^*}) \frac{\partial E^{2^*}}{n_a}}{e_0^1 - e_a^1} > 0.$$

Case 2: Partial adoption at the second stage: By Proposition 3.1 the socially optimal emission level at the second stage is independent of n_a. Since the socially optimal emission level is also the market outcome at the second stage, $\Delta C_{ab}(\overline{\tau}(n_a), D'(E^{2^*}(n_b^*(n_a)))) = 0$ holds. Differentiating this equation w.r.t. n_a we obtain $\partial \overline{\tau}/\partial n_a = 0$.

Case 3: Complete adoption will be expected in equilibrium at the second stage: In this case $\overline{\tau}(n_a)$ is given by $\Delta C_{ab}(\overline{\tau}(n_a), D'(E^{2^*}(1 - n_a))) = 0$. But it immediately follows that $\partial E^{2^*}(1 - n_a)/\partial n_a > 0$ and thus

$$\frac{\partial \bar{\tau}}{\partial n_a} = \frac{(e_a^2 - e_b^2) D''(E^{2^*}) \frac{\partial E^{2^*}(1-n_a)}{n_a}}{e_0^1 - e_a^1} > 0.$$

The case of tradable permits works analogously if we substitute the tax rate by the permit price.

B Comparative statics results for long-term policy commitment

In this subsection we study how a change in the level of the tax or permit policy, respectively, affects the firms' decisions to adopt the available technology immediately or to postpone investment. For this purpose we assume that the regulator moves first and makes a long-term commitment to a particular level of his tax rate τ, or to the quantity of permits respectively.

B.1 Long-term tax Commitment

We start with the tax case. Thus, according to section 4.1 in the main paper, we have the special case where $\sigma^1 = \sigma^2 = \tau$. Thus we can rewrite (8), (6), and (7), respectively, as

$$\Delta C_{0a}(\tau) \equiv \frac{1}{r}[C_a(e_a^1) - C_0(e_0^1) + \tau(e_a^1 - e_0^1)] + F_a,$$

$$\Delta C_{0b}(\tau) \equiv \frac{1}{r}[C_b(e_b^1) - C_0(e_0^1) + \tau(e_b^1 - e_0^1)] + F_b,$$

and

$$\Delta C_{ab}(\tau) \equiv \frac{1}{r+\lambda}[C_a(e_a^1) - C_0(e_0^1) + \tau(e_a^1 - e_0^1)]$$
$$+ \frac{\lambda}{r+\lambda}\frac{1}{r}[C_a(e_a^2) - C_b(e_b^2) + \tau(e_a^2 - e_b^2)] + F_a - \frac{\lambda}{r+\lambda}F_b.$$

When these expressions are zero, the firm is indifferent between the two technologies 0 and a, 0 and b, or a and b, respectively. The following lemma characterizes the roots of these expressions:

Lemma B.1.

Lemma B.2. *1. For $i = a, b$, the function ΔC_{0i} is decreasing in τ. If F_i is not too large, a unique solution $\bar{\tau}_{0i}$ of $\Delta C_{0i}(\tau) = 0$ exists such that firms are indifferent between technologies 0 and a. If the solution exists, staying with the conventional technology 0 is more profitable for $\tau < \bar{\tau}_{0i}$, while for $\tau > \bar{\tau}_{0i}$ the firm is better off adopting technology i. Furthermore, $\frac{\partial \bar{\tau}_{0i}}{\partial F_i} > 0$.*

2. The slope of ΔC_{ab} is ambiguous. If ΔC_{ab} is monotonic in τ and if a solution $\bar{\tau}_{ab}$ for $\Delta C_{ab}(\tau) = 0$ exists, then it will be unique.[5] If ΔC_{ab} is decreasing (increasing), the firm will prefer to invest in technology a if $\tau < \bar{\tau}_{ab}$ (if $\tau > \bar{\tau}_{ab}$) and to postpone investment and adopt technology b as soon as the latter is available if $\tau > \bar{\tau}_{ab}$ (if $\tau < \bar{\tau}_{ab}$).

For the proof see section B.3. The intuition is as follows. The greater τ is, the higher the cost difference will be and the more it matters what technology a firm uses. Thus, if the investment costs F_i are not too large, we will always find a tax rate sufficiently high for both technologies a and b to become more attractive compared to the conventional technology 0. As can be seen from

[5] Whether a solution exists depends on the cost difference $F_a - \frac{\lambda}{r+\lambda}F_b$.

the ambiguity of the slope of ΔC_{ab}, the impact of τ on the decision between technology a and b is less clear-cut. Investing in the first technology may lead to a decrease of the abatement costs as long as technology b is not available. As soon as technology b *is* available, however, the firm faces opportunity costs caused by using a less efficient technology. It is not clear which of the two technologies leads to the larger cut in the present value of total future costs. The following example shows that this ambiguity can arise even for a simple quadratic cost function:

Example B.3. Let $C(\theta, e) := (e^{-\alpha\theta} + A - \beta e)^2/2\beta$. One can easily verify that given a tax rate τ, the cost minimizing emission level for a firm with technology θ is given by $e(\theta, \tau) = (A + e^{-\alpha\theta} - \tau)/\beta$. This yields

$$\frac{\partial \Delta C_{ab}}{\partial \tau} = \frac{1}{\lambda + r}\frac{e^{-\alpha\theta_a} - e^{-\alpha\theta_0}}{\beta} + \frac{\lambda}{r(\lambda + r)}\frac{e^{-\alpha\theta_a} - e^{-\alpha\theta_b}}{\beta}$$

Therefore we get $\Delta C_{ab} \geq (\leq) 0$ if and only if $r(e^{-\theta_0} - e^{-\theta_a}) \leq (\geq) \lambda(e^{-\theta_a} - e^{-\theta_b})$. We see that the slope of ΔC_{ab} basically depends on both the efficiency parameter θ_i and the parameters λ and r.

Lemma B.4. *Assume that $\Delta C_{ab}(\tau)$ is monotonic in τ and that a (unique) tax rate $\overline{\tau}_{ab}$ exists, leaving the firm indifferent between technology a and b. Then the following holds: If $\Delta C_{ab}(\tau)$ is decreasing (increasing), then $\partial\overline{\tau}_{ab}/\partial F_a < (>) 0$ and $\partial\overline{\tau}_{ab}/\partial F_b > (<) 0$.*

Intuition suggests that a firm's decision will depend on the order of these three threshold tax rates. However, not every combination of these three values is feasible. In particular we will prove the following result:

Proposition B.5. If F_a and F_b are such that all three threshold taxes exist, then:

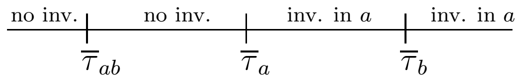
1. For every $\overline{\tau}$ satisfying $\Delta C_{ab}(\overline{\tau}) = 0$ we get $\overline{\tau} \notin [\bar{\tau}_i, \bar{\tau}_j]$ for $i, j = a, b$ and $i \neq j$. Thus if $\bar{\tau}_a < \bar{\tau}_b$, then $\Delta C_{ab} < 0$ for all $\tau \in [\bar{\tau}_a, \bar{\tau}_b]$. If $\bar{\tau}_b < \bar{\tau}_a$, then $\Delta C_{ab} > 0$ for all $\tau \in [\bar{\tau}_b, \bar{\tau}_a]$.

2. If ΔC_{ab} is monotonically decreasing, then $\overline{\tau}_{ab} < \bar{\tau}_a < \bar{\tau}_b$ or $\bar{\tau}_b < \bar{\tau}_a < \overline{\tau}_{ab}$.

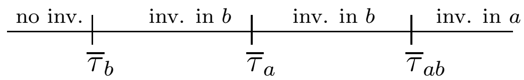
3. If ΔC_{ab} is monotonically increasing, then $\overline{\tau}_{ab} < \bar{\tau}_b < \bar{\tau}_a$ or $\bar{\tau}_a < \bar{\tau}_b < \overline{\tau}_{ab}$.

The last two results can be illustrated as follows:

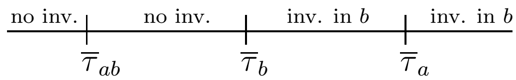
1. If ΔC_{ab} is decreasing, we obtain



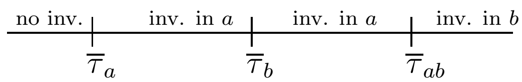
or



2. If ΔC_{ab} is increasing, we obtain

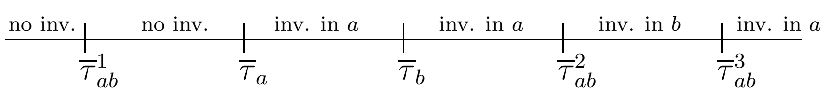


or



The intuition is straightforward. If ΔC_{ab} decreases and $\overline{\tau}_b < \tau$, a firm's optimal decision at the second stage would be to adopt technology b. Therefore, at the first stage, the relevant decision has to be made between adopting technology a, on the one hand, and waiting for technology b, on the other. Thus the relevant threshold tax rate is $\overline{\tau}_{ab}$, so e.g. $\overline{\tau}_{ab} < \tau$ would induce investment in technology a.

If we do not assume ΔC_{ab} to be monotonic, then multiple switches may occur between intervals where all firms adopt technology a and intervals where all firms adopt technology b for $\tau > \max\{\overline{\tau}_a, \overline{\tau}_b\}$. Note, however, that no switches will occur between $\overline{\tau}_a$ and $\overline{\tau}_b$, so we only need to know the sign of ΔC_{ab} either at $\overline{\tau}_a$ or at $\overline{\tau}_b$. The next figure illustrates what the investment pattern may look like in that case:



B.2 Long-Term Commitment to a Quantity of Permits

In this section we analyze how the firms make their investment decisions when they are regulated by tradable permits and the regulator has made a long-term commitment to a constant quantity of permits for issue. We also study what amounts of permit supply create what kinds of equilibria, which is an interesting question in itself.

As in the tax case, we can define price thresholds for permits where a firm is indifferent between any two strategies. According to the different strategies, these threshold prices can be derived from (6), (7), or (8) as the prices where these terms are equal to zero. Thus for any $\sigma^2 > 0$ we define $\overline{\sigma}_{0a}(\sigma^2)$ by

$$\Delta C_{0a}(\overline{\sigma}_{0a}(\sigma^2), \sigma^2) = 0,$$

$\overline{\sigma}_{ab}(\sigma^2)$ by

$$\Delta C_{ab}(\overline{\sigma}_{ab}(\sigma^2), \sigma^2) = 0,$$

and $\overline{\sigma}_{0b}$ by

$$\Delta C_{0b}(\overline{\sigma}_{0b}) = 0.$$

Accordingly, the threshold prices at the first stage depend on the expected permit price at the second stage. By backward induction we analyze the second stage first, where the share of firms n_a adopting the technology a is given. Once we have done that, we can write the threshold prices at the first stage as a function of n_a since we know the permit price at the second stage corresponding to n_a.

After these preparations we are now ready to characterize the permit price in the second stage depending on the supply of permits.

Proposition B.6. [Equilibrium at stage 2] Let L be the amount of permits issued by the regulator. Let $0 \leq n_a < 1$ be given, and let σ^2 denote the market price for permits at the second stage. Then the following holds:

1. If $L > \overline{E}^2(n_a)$, then $\sigma^2 < \overline{\sigma}_{0b}$ and none of the remaining firms invest.

2. If $L < \underline{E}^2(n_a)$, then $\sigma^2 > \overline{\sigma}_{0b}$ and all of the remaining firms invest.

3. If $L \in [\underline{E}^2(n_a), \overline{E}^2(n_a)]$, then $\sigma^2 = \overline{\sigma}_{0b}$ and a share of firms

$$n_b = \frac{e_0(\overline{\sigma}_{0b}) + n_a \cdot [e_a(\overline{\sigma}_{0b}) - e_0(\overline{\sigma}_{0b})] - L}{e_0(\overline{\sigma}_{0b}) - e_b(\overline{\sigma}_{0b})}$$

   will adopt the second technology.

4. If n_a increases, $\underline{E}^2(n_a)$ will increase and $\overline{E}^2(n_a)$ will decrease. Furthermore, both $\underline{E}^2$ and $\overline{E}^2$ converge to $E_M = e_a(\overline{\sigma}_{0b})$ as n_a goes to 1.

For the proof see section B.3. Note that for $n_a = 0$ the claims 1.) - 3.) correspond to Lemma 2 in Requate and Unold [2003]. The reason for 4.) is that a greater number n_a means that a greater abatement level has already been achieved.

We can now determine both the firms' decision and the pattern of investment at the first stage. For this purpose it is useful to first investigate the relationship between the number of firms n_a investing in technology a and the first-stage equilibrium price for permits. Since for given L and n_a the equilibrium permit price $\sigma^2(n_a)$ at the second stage is uniquely determined, we write $\overline{\sigma}^1(n_a)$ as the relevant threshold price at the first stage, where the remaining $1-n_a$ firms are indifferent between adopting technology a, on the one hand, and not adopting it and behaving optimally at the second stage, on the other. Thus if no firm wants to adopt the second technology at the second stage, $\overline{\sigma}^1(n_a)$ is given by $\overline{\sigma}_{0a}(\sigma^2(n_a))$. Otherwise $\overline{\sigma}^1(n_a)$ is given by $\overline{\sigma}_{ab}(\sigma^2(n_a))$. If partial adoption results at the second stage, we obtain $\overline{\sigma}_{ab}(\sigma^2(n_a)) = \overline{\sigma}_{0a}(\sigma^2(n_a))$.

The next result establishes how both threshold price $\overline{\sigma}^1$ and equilibrium price σ^1 depend on n_a.

Proposition B.7. The threshold price $\overline{\sigma}^1(n_a)$ is continuous and non-decreasing in n_a, while the market price σ^1 is decreasing and continuous in n_a.

To determine the pattern of investment given the supply of permits L, we proceed as in the second stage. We define $\overline{E}^1 := e_0^1(\overline{\sigma}^1(0))$ as the aggregate emission level that induces no firm to adopt technology a but where each firm is indifferent between adopting and not adopting technology a, and thus leading to a permit price equal to $\overline{\sigma}^1(0)$. Analogously, we define $\underline{E}^1 := e_a^1(\overline{\sigma}^1(1))$ as the aggregate emission level inducing an outcome where all firms adopt technology a but each firm is indifferent between adopting and not adopting technology a, thus leading to a permit price equal to $\overline{\sigma}^1(1)$. As we will see in the next result, $(\underline{E}^1, \overline{E}^1)$ is the interval of permits where partial adoption of technology a occurs.

Proposition B.8. [Equilibrium at stage 1]

1. For $L > \overline{E}^1$ none of the firms will adopt technology a and $\sigma^1 < \overline{\sigma}^1(0)$.
2. If $L < \underline{E}^1$ all firms will adopt technology a and $\sigma^1 > \overline{\sigma}^1(1)$.

3. If $L \in [\underline{E}^1, \overline{E}^1]$ then a share $0 < n_a < 1$ of firms will adopt technology a and the permit price is equal to $\overline{\sigma}^1(n_a)$. Moreover, n_a is given by the solution of the equation $L = (1 - n_a)e_0^1(\overline{\sigma}^1(n_a)) + n_a e_a^1(\overline{\sigma}^1(n_a))$.

The intuition for the result is as follows. At the threshold price $\overline{\sigma}^1(0)$ firms are indifferent between adopting and not adopting technology a, but no firm in fact adopts that technology, so total emissions amount to $\overline{E}^1$. If the number of permits L exceeds $\overline{E}^1$, the price for permits must therefore be lower than $\overline{\sigma}^1(0)$, and all firms will strictly prefer not to adopt technology a. At the threshold price $\overline{\sigma}^1(1)$ firms are again indifferent between adopting and not adopting technology a, but all firms in fact adopt technology a, so total emissions amount to $\underline{E}^1$. Therefore, if the number of permits L is smaller than $\underline{E}^1$, the price for permits must exceed $\overline{\sigma}^1(1)$ and all firms strictly prefer to adopt technology a. If L is between $\underline{E}^1$ and $\overline{E}^1$, a share $0 < n_a < 1$ adopts the new technology, and the permit price is $\sigma^1(n_a) = \overline{\sigma}^1(n_a)$ with $\overline{\sigma}^1(0) < \overline{\sigma}^1(n_a) < \overline{\sigma}^1(1)$. Moreover, $\overline{\sigma}^1(n_a)$ is increasing in n_a.

Figure 4 illustrates the relationship between the threshold price and n_a for the cases where either full or no adoption of technology b occurs at the second stage. The curves AA, BB, and CC each depict the equilibrium price for permits if we vary n_a but treat it as an exogenous variable. Along each curve L is fixed. If we increase L, we move from AA to CC. Along AA, the quantity of permits is larger than $\overline{E}^1$, and the equilibrium prices for each fixed n_a are always larger than the threshold price $\overline{\sigma}^1(n_a)$. Hence full adoption will occur. Along CC, the quantity of permits is smaller than $\underline{E}^1$, and the equilibrium prices for each fixed n_a are always smaller than the threshold price $\overline{\sigma}^1(n_a)$. Hence no adoption will occur. Along BB, where $\underline{E}^1 < L < \overline{E}^1$, the equilibrium price curve $\sigma^1(n_a)$ for any exogenously given n_a, intersects the threshold price curve $\overline{\sigma}^1(n_a)$. At this point, the equilibrium price for fixed n_a is exactly equal to the threshold price where firms are indifferent between adopting and not adopting technology a. This intersection determines the equilibrium number of adopting firms.

Why is the curve $\overline{\sigma}^1(n_a)$ increasing? Recall that Figure 4 covers the case where either full or no adoption of technology b occurs at the second stage. Consider first the case where none of the firms adopt technology b. Here a higher n_a induces less demand for permits and hence a lower price at stage 2, making technology a less attractive at stage 1. Therefore the threshold price $\overline{\sigma}^1(n_a)$ at which it begins to be attractive to invest in technology a must be increasing in n_a.

Consider now the case where all the remaining firms $1-n_a$ adopt technology b at stage 2. Here, a higher n_a triggers more demand for permits and thus

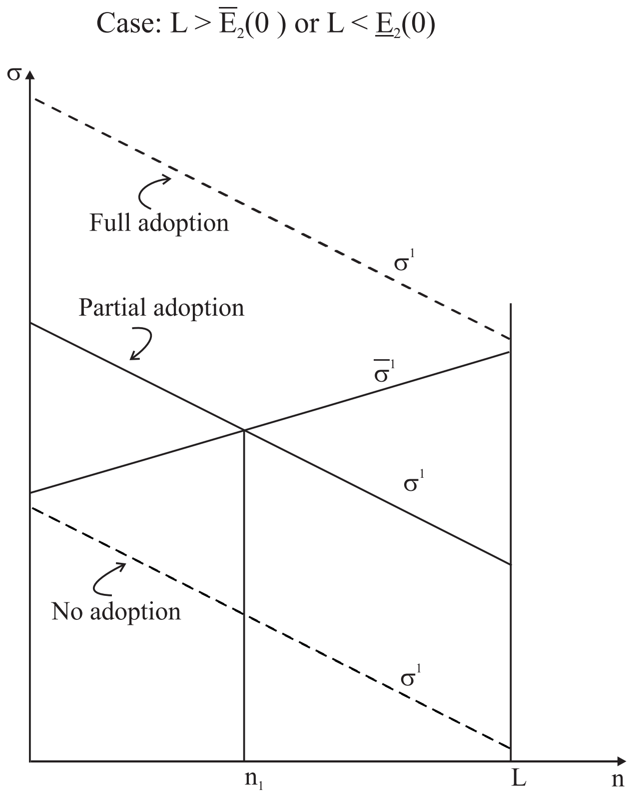


Figure 4: *Permit market (first stage) in case of no or full adoption at the second stage. The case:* $L > \overline{L}_2(0)$ *or* $L < \underline{E}_2(0)$.

induces the permit price to rise at stage 2, which makes it more attractive to wait for technology b instead of adopting technology a at stage 1. Put differently, higher n_a raises the option value of postponing the investment decision. Therefore again, threshold price $\overline{\sigma}^1(n_a)$ must increase with n_a to compensate for that effect. Note the fundamental difference over and against the case where no second technology will be come up in the future (as in the Requate-Unold [2003] model) and where the threshold price is always constant in n_a.

Figure 5 captures the case where partial adoption may occur at the second stage. Again, the curves AA through DD depict the equilibrium prices if we vary n_a but treat it as an exogenous variable. Along each curve L is constant, and if we increase L, we move from AA to DD. If n_a is sufficiently large, say greater than some $\hat{n}_a$ as depicted in the figure, we know from Proposition 5.5 (part 4) that either no or full adoption will occur at stage 2. Therefore to the right of $\hat{n}_a$ we obtain the same picture as in Figure 4. To the left of $\hat{n}_a$ partial

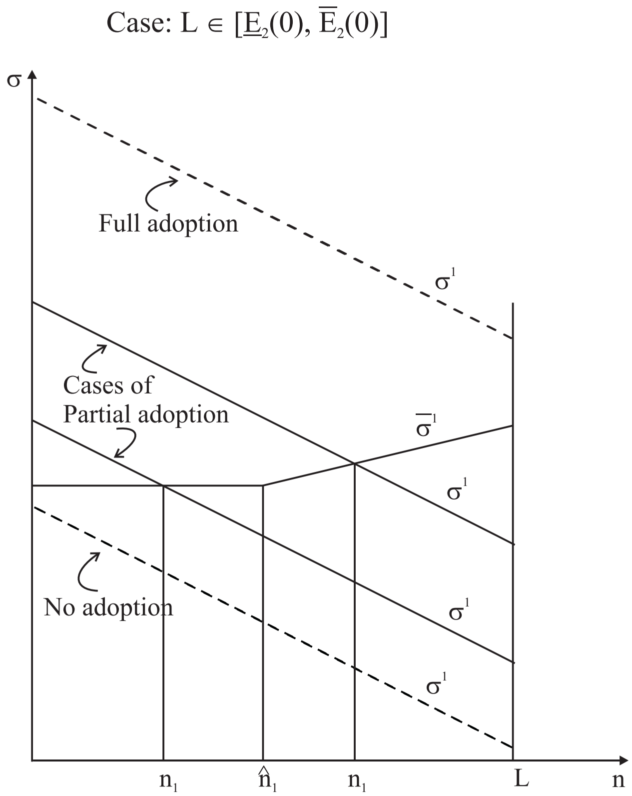


Figure 5: *Permit market (first stage) in case of no or partial adoption at the second stage. Case:* $L \in [\underline{E}_2(0), \overline{E}_2(0)]$.

adoption of technology b will occur at stage 2. In this case, higher n_a will be compensated for by lower n_b at the second stage so that the threshold price stays constant.

Long-term commitment to a constant tax or permit policy is, of course, not optimal in general. In an extended version of this paper, von Döllen & Requate [2007] show that myopic long-term commitment to a tax policy, though optimal with respect to conventional technology 0, will typically induce over-investment, while myopic long-term commitment to a particular quantity of permits (socially optimal with respect to technology 0) typically triggers under-investment.[6]

[6]Note that in this model the converse can also be true for both types of regulation, i.e. for special cases, under-investment (over-investment) may occur under taxes (permits).

B.3 Further proofs of additional results

Proof of Lemma B.1: *ad 1.:* Differentiating ΔC_{0i} with respect to τ and using the Envelope Theorem yields

$$\frac{\partial \Delta C_{0i}}{\partial \tau} = \frac{1}{r}(e_i - e_0) < 0.$$

For $\tau = 0$ we have $\Delta C_{0i} = F_i$. If the term $C_i(e_i) - C_0(e_0)$ is bounded (which in particular is the case if C is bounded), a sufficient condition for a solution of $\Delta C_{0i}(\tau) = 0$ to exist is that F_i is smaller than $\frac{1}{r}[C_i(0) - C_0(0)]$. If $C_i(e_i) - C_0(e_0)$ goes to $-\infty$ as e_i and e_0 tend to 0, a root for $\Delta C_{0i}(\tau) = 0$ exists for each F_i. Differentiating $\Delta C_{0i}(\tau) = 0$ implicitly with respect to F_i, using the Envelope Theorem, and solving for $\partial\bar{\tau}_i/\partial F_i$ yields

$$\frac{\partial \bar{\tau}_i}{\partial F_i} = \frac{r}{(e_0 - e_i)} > 0.$$

ad 2.: Differentiating $\Delta C_{ab}(\tau) = 0$ with respect to τ and using the Envelope Theorem yields

$$\frac{\partial \Delta C_{ab}}{\partial \tau} = \frac{1}{\lambda + r}(e_a^1(\tau) - e_0^1(\tau)) + \frac{\lambda}{\lambda + r}\frac{1}{r}(e_a^2(\tau) - e_b^2(\tau)).$$

Obviously, the sign and its development is ambiguous without further restrictions. Let us therefore assume that the sign is unique and let us consider the case where ΔC_{ab} increases (decreases). Now observe that $\Delta C_{ab}(0) = F_a - \frac{\lambda}{r+\lambda}F_b$. Thus if F_a and F_b are such that $\Delta C_{ab}(0) < 0$ (> 0) holds and such that $\Delta C_{ab}(\tau)$ becomes positive (negative) for τ sufficiently large, by reasons of continuity a tax rate $\bar{\tau}$ must exist such that $\Delta C_{ab}(\bar{\tau}) = 0$. Since Δ_{ab} is monotonic, this value is unique.

Proof of Lemma B.4: Differentiating $\Delta C_{ab}(\tau) = 0$ with respect to F_a and using the Envelope Theorem we obtain:

$$\frac{\partial \overline{\tau}_{ab}}{\partial F_a}\left(\frac{1}{\lambda + r}(e_a^1 - e_0^1) + \frac{\lambda}{\lambda + r}\frac{1}{r}(e_a^2 - e_b^2)\right) + 1 = 0.$$

This yields

$$\frac{\partial \overline{\tau}_{ab}}{\partial F_a} = -\frac{r(\lambda + r)}{r(e_a^1 - e_0^1) + \lambda(e_a^2 - e_b^2)} = \frac{1}{\frac{\partial \Delta C_{ab}}{\partial \tau}}.$$

Analogous we can derive

$$\frac{\partial \overline{\tau}_{ab}}{\partial F_b} = \frac{r\lambda}{r(e_a^1 - e_0^1) + \lambda(e_a^2 - e_b^2)} = \frac{\lambda}{\lambda + r}\frac{1}{\frac{\partial \Delta C_{ab}}{\partial \tau}}.$$

Proof of Proposition B.5 *ad 1.:* Note first that a firm's emission level depends on the size of the tax rate only but not on the other firms' technology choices. We can therefore omit superscripts referring to stages, i.e. we write e_0, e_a,and e_b instead of e_0^i, e_a^i,and e_b^i.

First we consider the case $\overline{\tau}_a < \overline{\tau}_b$. Assume first $\tau \in (\overline{\tau}_a, \overline{\tau}_b)$. Because ΔC_i $(i = a, b)$ decreases, we obtain

$$\Delta C_{0a}(\tau) = \frac{1}{r}[C_a(e_a) - C_0(e_0) + \tau(e_a - e_0)] + F_a < 0$$

and

$$\Delta C_{0b}(\tau) = \frac{1}{r}[C_b(e_b) - C_0(e_0) + \tau(e_b - e_0)] + F_b > 0.$$

This gives us the following inequality chain:

$$\begin{aligned}
0 &> \frac{1}{r}[C_a(e_a) - C_0(e_0) + \tau(e_a - e_0)] + F_a \\
&= \frac{1}{r+\lambda}[C_a(e_a) - C_0(e_0) + \tau(e_a - e_0)] \\
&\quad + \frac{\lambda}{r+\lambda}\frac{1}{r}[C_a(e_a) - C_0(e_0) + \tau(e_a - e_0)] + F_a \\
&> \frac{1}{r+\lambda}[C_a(e_a) - C_0(e_0) + \tau(e_a - e_0)] \\
&\quad + \frac{\lambda}{r+\lambda}\frac{1}{r}[C_a(e_a) - C_b(e_b) + \tau(e_a - e_b)] + F_a \\
&\quad - [\frac{\lambda}{r+\lambda}\frac{1}{r}[C_b(e_b) - C_0(e_0) + \tau(e_b - e_0)] + \frac{\lambda}{r+\lambda}F_b] \\
&= \Delta C_{ab}(\tau).
\end{aligned}$$

Next we consider the case $\overline{\tau}_b < \overline{\tau}_a$. Assume that $\tau \in (\overline{\tau}_b, \overline{\tau}_a)$. Following the same argument as above we obtain

$$\Delta C_{0a}(\tau) = \frac{1}{r}[C_a(e_a) - C_0(e_0) + \tau(e_a - e_0)] + F_a > 0$$

and

$$\Delta C_{0b}(\tau) = \frac{1}{r}[C_b(e_b) - C_0(e_0) + \tau(e_b - e_0)] + F_b < 0.$$

A similar inequality chain yields $0 < \Delta C_{ab}(\tau)$.

ad 2.: First assume $\overline{\tau}_a < \overline{\tau}_b$. Since ΔC_{ab} decreases, $\Delta C_{ab}(\tau) < 0$ for each $\tau \in (\overline{\tau}_a, \overline{\tau}_b)$ and since $\bar{\tau}_{ab}$ is given by $\Delta C_{ab}(\bar{\tau}_{ab}) = 0$, it must be the case that

$\bar{\tau}_{ab} < \overline{\tau}_a$. In a similar way, for $\overline{\tau}_a > \overline{\tau}_b$ it follows that $\bar{\tau}_{ab} > \overline{\tau}_a$.

ad 3.: Analogously.

Proof of Proposition B.6 We will prove the first three claims via contradiction:

ad 1.: Let $L > \overline{E}^2(n_a)$ and suppose $\sigma > \overline{\sigma}^2$. Hence all $(1 - n_a)$ firms using the old technology will adopt technology b, and aggregate emissions will fall below $\underline{E}^2(n_a)$. But then the permit market would not clear because $E(\sigma) < L$, a contradiction. Now suppose $\sigma = \overline{\sigma}^2$. In this case the emission level can be no larger than $\overline{E}^2(n_a)$. But $\overline{E}^2(n_a) < L$, again a contradiction. Analogously we can prove part 2. and 3. If $L \in [\underline{E}^2(n_a), \overline{E}^2(n_a)]$, the share of firms investing can easily be derived from the market clearing condition $n_a e_a + n_b e_b + (1 - n_a - n_b)e_0 = L$.

ad 4.: It is obvious that $\overline{E}^2(n_a)$ is strictly decreasing in n_a and tends to $e_a(\overline{\sigma}^2)$. Similarly $\underline{E}_2(n_a)$ is strictly increasing in n_a and also tends to $e_a(\overline{\sigma}^2)$.

To prove Proposition B.8 we use the following Lemma which characterize the market price at the first stage:

Lemma B.9. *The permit price at the first stage $\sigma^1(n_a)$ is strictly decreasing in n_a.*

Afterwards we will prove that the threshold price $\overline{\sigma}^1(n_a)$ is non-decrea-sing and continuous in n_a. For this purpose we will prove the following Lemma:

Lemma B.10. *The threshold price $\overline{\sigma}^1(n_a)$ is characterized as follows:*

1. If $L > \overline{E}^2(0)$, then $\overline{\sigma}^1(n_a)$ will be equal to $\overline{\sigma}_{0a}(\sigma^2(n_a))$ for all n_a. Moreover $\overline{\sigma}^1(n_a)$ increases in n_a.

2. If $L < \underline{E}^2(0)$, then $\overline{\sigma}^1(n_a)$ will be equal to $\overline{\sigma}_{ab}(\sigma^2(n_a))$ for all n_a. Moreover $\overline{\sigma}^1(n_a)$ increases in n_a.

3. For $L \in [\underline{E}^2(0), \overline{E}^2(0)]$ we obtain the following three subcases:

   (a) For $L = E_M$, $\overline{\sigma}^1(n_a)$ equals $\overline{\sigma}_{ab}(\overline{\sigma}^2)$ for all n_a and is therefore constant.

   (b) For $L < E_M$ let $\overline{n}_a$ be the share of firms where $L = \underline{E}^2(\overline{n}_a)$. Then:

      i. For $n_a \leq \overline{n}_a$ the threshold price $\overline{\sigma}^1(n_a)$ equals $\overline{\sigma}_{ab}(\overline{\sigma}^2)$ and is therefore constant in n_a.

      ii. For $n_a > \overline{n}_a$ the price $\overline{\sigma}^1(n_a)$ equals $\overline{\sigma}_{ab}(\sigma^2(n_a))$ and is increasing.

iii. For $n_a = \overline{n}_a$ we have $\overline{\sigma}^1(n_a) = \overline{\sigma}_{ab}(\overline{\sigma}^2) = \overline{\sigma}_{ab}(\sigma^2(\overline{n}_a))$.

(c) If $L > E_M$, let $\overline{n}_a$ be the share of firms where $L = \overline{E}^2(\overline{n}_a)$. Then:

i. For $n_a \leq \overline{n}_a$ the threshold price $\overline{\sigma}^1(n_a)$ equals $\overline{\sigma}_{ab}(\overline{\sigma}^2)$ and is therefore constant in n_a.

ii. For $n_a > \overline{n}_a$ the price $\overline{\sigma}^1(n_a)$ equals $\overline{\sigma}^0 < (\sigma^2(n_a))$ and is increasing.

iii. For $n_a = \overline{n}_a$ we have $\overline{\sigma}^1(n_a) = \overline{\sigma}_{ab}(\overline{\sigma}^2) = \overline{\sigma}_{0a}(\sigma^2(\overline{n}_a))$.

Proof of Lemma B.9: We differentiate the two market equilibrium equations, $-\frac{\partial C_0}{\partial e} = \sigma^1$, and $L = n_a e_a^1 + (1-n_a)e_0^1$ implicitly with respect to n_a. This yields $-\partial^2 C_i/(\partial e)^2 \cdot \partial e_i^1/n_a = \partial\sigma^1/\partial n_a$, $i = 0, a$, and $0 = e_a^1 - e_0^1 + n_a \partial e_a^1 \partial n_a + (1-n_a)\partial e_0^1/\partial n_a$. Thus we can derive

$$\frac{\frac{\partial^2 C_a}{(\partial e)^2}}{\frac{\partial^2 C_0}{(\partial e)^2}} \frac{\partial e_a}{\partial n_a} = \frac{\partial e_0}{\partial n_a},$$

and therefore

$$\frac{\partial e_a}{\partial n_a} = \frac{e_0^1 - e_a^1}{n_a + (1-n_a)\frac{\frac{\partial^2 C_a}{(\partial e)^2}}{\frac{\partial^2 C_0}{(\partial e)^2}}} > 0.$$

Obviously this implies $\partial\sigma^1 \partial n_a < 0$.

Proof of Lemma B.10: *ad 1.:* For any n_a it is optimal for all remaining firms not to adopt technology b at the second stage. Moreover, the market price at the second stage equals the one at the first stage since none of the remaining firm adopts technology b. Differentiating $\Delta C_{0a}(\overline{\sigma}^1(n_a), \sigma^1) = 0$ with respect to n_a (note that $\overline{\sigma}^1(n_a) = \overline{\sigma}^{0a}(\sigma^1)$) and employing both the Envelope Theorem and Lemma B.9, we obtain:

$$\frac{\partial\overline{\sigma}^1}{\partial n_a} = \frac{\lambda \frac{\partial \sigma^1}{\partial n_a}(e_a^1(\sigma^1) - e_0^1(\sigma^1))}{r(e_0^1(\overline{\sigma}^1) - e_a^1(\overline{\sigma}^1))} > 0.$$

ad 2.: In that case, for any n_a it is optimal for the remaining firms to adopt technology b at the second stage. Thus at the first stage each firm decides between adopting technology a immediately and adopting technology b at the second stage. Thus $\overline{\sigma}^1(n_a) = \overline{\sigma}^{ab}(\sigma^2)$. Differentiating $\Delta C_{ab}(\overline{\sigma}^1(n_a), \sigma^2) = 0$ with respect to n_a and using the Envelope Theorem, we obtain:

$$\frac{\partial\overline{\sigma}^1}{\partial n_a} = \frac{\lambda \frac{\partial \sigma^2}{\partial n_a}(e_a^2(\sigma^2) - e_b(\sigma^2))}{r(e_0(\overline{\sigma}^1) - e_a(\overline{\sigma}^1))}.$$

Thus to prove $\partial\overline{\sigma}^1/\partial n_a > 0$ it suffices to show that $\partial\sigma^2/\partial n_a > 0$. Differentiating the market equilibrium equations $-\partial C_i/\partial e_i = \sigma^2$, $i = a, b$, and $L = n_a e_a^2 + (1-n_a)e_b^2$ with respect to n_a we obtain $-\partial^2 C_i/(\partial e)^2 \cdot \partial e_i^2/\partial n_a = \partial\sigma^2/\partial n_a$, $i = a, b$, and $0 = e_a^2 - e_b^2 + n_a \partial e_a^2/\partial n_a + (1-n_a)\partial e_b^2/\partial n_a$, yielding $[\frac{\partial^2 C_a}{(\partial e)^2}/\frac{\partial^2 C_i}{(\partial e)^2}] \cdot [\partial e_a^2/\partial n_a] = \partial e_b^2/\partial n_a$ and therefore

$$\frac{\partial e_a^2}{\partial n_a} = \frac{e_b^2 - e_a^2}{n_a + (1-n_a)\frac{\frac{\partial^2 C_a}{(\partial e)^2}}{\frac{\partial^2 C_b}{(\partial e)^2}}} < 0.$$

Obviously this implies $\partial\sigma^2/\partial n_a > 0$.

ad 3.: (a) By Proposition B.6 we know that, if $L = E_M$, then for all $n_a < 1$ partial adoption of technology b will occur at the second stage. This implies that in equilibrium each firm that has not adopted technology a at the first stage is indifferent between whether or not to adopt technology b at the second. Therefore $\sigma^2 = \overline{\sigma}^2$, and at the first stage $\overline{\sigma}^1 = \overline{\sigma}_{ab}(\overline{\sigma}^2) = \overline{\sigma}_{0a}(\overline{\sigma}^2)$ must hold in equilibrium.

(b) Since $L < E_M$, we know by Proposition B.6 that we can find an $\overline{n}_a$ satisfying $L = \underline{E}(\overline{n}_a)$. Furthermore, for all $n_a < \overline{n}_a$ partial adoption will occur while for all $n_a \geq \overline{n}_a$ full adoption of technology b will happen. Thus in each case we can apply the arguments from above and therefore we only need to show that $\overline{\sigma}^1(n_a)$ is continuous at $\overline{n}_a$. By Proposition B.6 we know, however, that all remaining $(1-\overline{n}_a)$ firms will invest in technology b, and σ^2 must be equal to $\overline{\sigma}^2$. By definition of $\overline{n}_a$, at the second stage all $(1-\overline{n}_a)$ firms are indifferent between investing and not investing in technology b. It therefore directly follows from $\Delta C_{0b}(\overline{\sigma}^2) = 0$ that for $\overline{\sigma}^1$ and $\overline{\overline{\sigma}}^1$ satisfying $\Delta C_{0b}(\overline{\sigma}^1, \overline{\sigma}^2) = 0$ and $\Delta C_{ab}(\overline{\overline{\sigma}}^1, \overline{\sigma}^2) = 0$, respectively, the equality $\overline{\sigma}^1 = \overline{\overline{\sigma}}^1$ must hold. Hence the threshold price $\overline{\sigma}^1$ is continuous in $\overline{n}_a$.

(c) Analogously to (b).

Proof of Proposition B.8: First note that in case of an equilibrium with partial adoption the threshold price $\overline{\sigma}$ must equal the market price.

ad 1.: If $L > \overline{E}^1$ the market price is lower than the lowest possible threshold price. Thus no firm will adopt technology a.

ad 2.: If $L < \underline{E}^1$ the market price is higher than the highest possible threshold price. Therefore all firms will adopt technology a.

ad 3.: In this case we find n_a satisfying $E(\sigma^1(n_a)) = L = E(\overline{\sigma}(n_a))$. For this n_a the permit market clears and no fimr has an incentive to deviate from its decision.