Comparison of different Stackelberg solutions in a deterministic dynamic pollution control: the time inconsistency problem revisited

Thomas Vallée

Document de Travail

Working Paper

EA 4272

Comparison of different Stackelberg solutions in a deterministic dynamic pollution control: the time inconsistency problem revisited

Thomas Vallée*

2018/12

(*) LEMNA - Université de Nantes

Laboratoire d'Economie et de Management Nantes-Atlantique
Université de Nantes
Chemin de la Censive du Tertre – BP 52231
44322 Nantes cedex 3 – France
http://www.lemna.univ-nantes.fr/
Tél. +33 (0)2 40 14 17 17 – Fax +33 (0)2 40 14 17 49

UNIVERSITÉ DE NANTES

Comparison of different Stackelberg solutions in a deterministic dynamic pollution control: the time inconsistency problem revisited

Thomas Vallée
LEMNA, Université de Nantes
IAE de Nantes - Institut d'Économie et de Management
Chemin de la Censive du Tertre, BP 52231
44322 Nantes Cedex 3, France.
thomas.vallee@univ-nantes.fr

First version March 1998

Abstract

In this paper, we study the two well-known supposed consequences of time inconsistency: the suboptimality of the time consistent solution, and the assuming increase of the follower's cost. To achieve such a goal, we study different dynamic Stackelberg solutions within a pollution control problem framework. This study is made under the assumption of different information structures, mainly we assume open-loop, feedback and closed-loop structures of information. Some of the numerical results may appear counterintuitive. Hence, there may exist some situations where a time consistent solution is optimal in comparison of the time inconsistent one. Moreover, the perfect discretionary solution may be beneficial to everyone.

Keywords: Time inconsistency, Dynamic Stackelberg game, Pollution control
JEL Codes: C7

1 Introduction

When a firm pollutes while producing, it is known that this flow of pollution will negatively affect other economic agents. If the firm is not liable to directly compensate these agents for the nuisances it causes, the production and the associated pollution levels optimal for the firm will not be optimal for society as a whole. One of the main problems in environmental economics is to find ways for a regulator to force such a firm to make socially optimal decisions, for example, through a proper use of taxes.

The problem has been extensively treated for the static case (see for example [10]). However, regulatory taxes have both short and long term consequences on the social welfare and on the firm's behavior. Taking these properly into account makes an explicitly dynamic analysis imperative. As noted by Batabyal [5], among others, a natural way to conduct such an analysis is to model the interaction between the regulator and the firm as a dynamic Stackelberg game with the regulator as the leader.

Depending on the information structure many dynamic Stackelberg solutions do exist. In this paper, using a discrete time dynamic model of pollution control, we derive three of them, that is the open-loop, feedback and global (closed-loop) Stackelberg solutions and compare them. As simple as may seem the model, the derivation of the different dynamic Stackelberg solutions are not straightforward.

It is well-know, since the seminal works of Kydland and Prescott [8], and Barro and Gordon [2, 3], that open-loop Stackelberg solutions are time inconsistent. From this literature, two others conclusions have been generally admitted. First, the discretionary solution is worst for the follower than the open-loop one with commitment. Second, the time consistent solution is suboptimal. Using numerical simulations, we show that those two conclusions do not hold.

The plan of the paper is as follows. In the next section we define the pollution control model. Then in section 3, we derive the different dynamic Stackelberg solutions depending on the information's structure facing each player in the following order: first the open-loop one, second the feedback one, and third the global Stackelberg solution (that is a closed-loop structure of information). In section 4, using two numerical simulations, we compare these solutions. Finally we conclude.

2 The pollution control

2.1 The general model

We consider a discrete time version of the continuous time model of pollution suggested by Batabyal [5]. There are two players: the regulator (the leader, R) and a monopolist (the follower, F). The planning horizon is T periods, with $T \leq 20$. There is no discounting. The goal of the monopolist is to maximize its cumulated profits over the T periods with respect to its choice of output. In each period t, the monopolist's revenue is given by $P(q_t)q_t$, where q_t is its output in period t, and where $P(q_t)$ is the inverse demand curves it faces.

Following Batabyal [5], the monopolist is facing three kinds of costs associated with q_t. First, a production cost wq_t that is assumed to be proportional to the output. Second, the tax paid to the regulator $\tau_t q_t$. And third, a cost $c(x_t)q_t$ that depends on the current stock of pollution, x_t. This last cost reflects the fact that the production efficiency decreases as the environment becomes more polluted. It may be or not internalized by the firm.

The monopolist's optimization problem is thus given by

$$J^F = \sum_{t=1}^{T} P(q_t)q_t - wq_t - \tau_t q_t - c(x_t)q_t \to \{q_t\}_{t\in[1,T]} \max \tag{2.1.1}$$

We assume that $P^{'}(q_t) < 0$ and $P^{''}(q_t) \geq 0$, and that $c^{'}(x_t) > 0$, $c^{''}(x_t) < 0$ and $c(0) = 0$. Furthermore, we assume $w > 0$.

The regulator attempts to maximize, through its choice of tax rates, its cumulated payoff. Again, following Batabyal [5] this payoff depends on three components. First, a function $B(q_t)$ that represents a social benefit when tithe firm produces at the level q_t. Second, a function $D(x_t)$ which measures the damage from pollution. And finally the amount of money given by the tax $\tau_t q_t$. So, the cumulated regulator's payoff is

$$J^R = \sum_{t=1}^{T} B(q_t) + \tau_t q_t - D(x_t) \tag{2.1.2}$$

We assume that $B(.)$ and $D(.)$ are respectively at least C^2 and C^1 functions. Furthermore, [$B^{'}(q_t) > 0$, $B^{''}(q_T) < 0$, $D^{'}(x_t) > 0$ and $D^{''}(x_t) > 0$, that is the social costs of pollution are increasing in the pollution stock at an ever increasing rate. The strict concavity of $B(q_t) + \tau_t q_t$ is needed in order to insure the existence and uniqueness of a solution.

Finally, we suppose that x_t evolves according to

$$x_{t+1} = f(q_t, x_t) \tag{2.1.3}$$

with x_1 given, and where $f(q_t)$ is a differentiable function, with $f^{'}(q_t) > 0$ and $f^{''}(q_t) > 0$. We also have $f^{'}(x_t) > 0$ and $f^{''}(x_t) > 0$. Hence, the pollution stock in $t + 1$ is increasing in the pollution stock and in the firm's output in t.

For the purpose of the paper, we more specifically assume[1]:

$$P(q_t) \equiv a - bq_t, \tag{2.1.4}$$

$$c(x_t) \equiv \alpha x_t, \tag{2.1.5}$$

$$B(q_t) \equiv \gamma q_t - \frac{q_t^2}{2}, \tag{2.1.6}$$

$$D(x_t) \equiv \frac{\delta x_t^2}{2}, \tag{2.1.7}$$

$$x_{t+1} \equiv \beta q_t + \tilde{\beta} x_t. \tag{2.1.8}$$

where the coefficients a, b, α, γ, β and $\tilde{\beta}$ are supposed to be strictly positive and with $\beta < 1$ and $\tilde{\beta} < 1$. The functional forms, as well as the hypotheses made earlier on the different derivatives, are standard in economic theory and will not be further justified here. The assumption $\tilde{\beta} < 1$ captures the fact that there is a natural resorption of the current pollution stock, at the rate $(1 - \tilde{\beta})$.

We may now derive the different dynamic Stackelberg solutions.

3 The different solutions

We assume that there is no uncertainty and that the regulator knows perfectly the different parameters of the monopolist's profits, even his cost. Furthermore, the regulator, our leader, is strong enough to force the monopolist to take as given the level of taxation.

3.1 The open-loop Stackelberg solution

This solution was first introduced by Simaan and Cruz [12, 11] (for a more detail on it, see Başar and Olsder [1]). To achieve the solution, the following steps are required. First, to any fixed action of the leader τ_t, the reaction function of the follower is derived by maximizing the firm's payoff under the state constraint (2.1.3). Then, integrating this reaction function into the leader's payoff and minimizing again under the state constraint, gives the optimal action of the leader which induces an optimal action for the follower. As noticed by Simaan and Cruz [11], latter by Kydland [7] and popularized by Kydland and Prescott [8], this solution is time inconsistent.

Let the time interval be $[1, T]$. To any fixed τ_t, $t \in [1, T]$ the firm solves

$$\arg\max_{q_t \in \Re^*} \sum_{t=1}^{T} (a - bq_t)q_t - wq_t - \tau_t q_t - \alpha x_t q_t \tag{3.1.1}$$

subject to

$$x_{t+1} = \beta q_t + \tilde{\beta} x_t \tag{3.1.2}$$

Let define the firm's Hamiltonian by

$$H^F(q_t, x_t, p_{t+1}^F) \equiv J_t^F + p_{t+1}^F(\beta q_t + \tilde{\beta} x_t \tag{3.1.3}$$

By using the first order conditions required to maximize this Hamiltonian function, and after some algebras, we get

$$q = \frac{a - w - \tau_t - \alpha x_t + \beta p_{t+1}^F}{2b} \tag{3.1.4}$$

$$x_{t+1} = \frac{\beta(a - w - \tau_t - \alpha x_t + \beta p_{t+1}^F)}{2b} + \tilde{\beta} x_t \tag{3.1.5}$$

$$p_{t+1}^F = \frac{-\alpha(a - w - \tau_t - \alpha x_t)}{2b} + (\tilde{\beta} - \frac{\alpha\beta}{2b}) p_{t+1}^F \tag{3.1.6}$$

[1] Some others specifications are possible, see Batabyal [4, 5]

with initial and final condition $p^F_{T+1} = 0$ and x_1 given. The stock of pollution at the period $T+1$, x_{T+1} is free. One reason to let it free is that the regulator may not know what is or not an acceptable final level of pollution.

This above set of equations defines the reaction function of the monopolist (follower) to any announced tax path. Replacing (3.1.4) into J^L_t, and given (3.1.5) and (3.1.6), we may now solve the regulator's problem defined by the following Hamiltonian

$$\begin{aligned} H^R(\tau_t, p^L_{t+1}, p^F_{t+1}, x_t, \mu_t) \equiv & \frac{(\gamma+\tau_t)(a-w-\tau_t-\alpha x_t+\beta p^F_{t+1})}{2b} \\ & -\frac{1}{2}(\frac{a-w-\tau_t-\alpha x_t+\beta p^F_{t+1}}{2b})^2 - \frac{\delta x_t^2}{2} \\ & + p^R_{t+1}(\frac{\beta(a-w-\tau_t-\alpha x_t+\beta p^F_{t+1})}{2b} + \tilde{\beta} x_t) \\ & + \mu_t(\frac{-\alpha(a-w-\tau_t-\alpha x_t)}{2b} + (\tilde{\beta} - \frac{\alpha\beta}{2b})p^F_{t+1}) \end{aligned} \tag{3.1.7}$$

Then we know from Başar and Olsder [1] that the open-loop Stackelberg solution is given by the resolution of the following first-order conditions:

$$\begin{aligned} \frac{\partial H^R_t}{\partial \tau_t} = & \frac{-\gamma-\tau_t-\beta p^R_{t+1}+\alpha\mu_t}{2b} \\ & + \frac{(a-w-\tau_t-\alpha x_t+\beta p^F_{t+1})(1+2b)}{4b^2} = 0 \end{aligned} \tag{3.1.8}$$

$$x_{t+1} = \frac{\partial H^R_t}{\partial p^L_{t+1}} = \frac{\beta(a-w-\tau_t-\alpha x_t+\beta p^F_{t+1})}{2b} + \tilde{\beta} x_t \tag{3.1.9}$$

$$\begin{aligned} p^R_t = \frac{\partial H^R_t}{\partial x_t} = & \frac{p^R_{t+1}(2b\tilde{\beta}-\alpha\beta) - \alpha(\gamma+\tau_t) + \alpha^2\mu_t}{2b} \\ & + \frac{\alpha(a-w-\tau_t-\alpha x_t+\beta p^F_{t+1})}{4b^2} - \delta x_t, \end{aligned} \tag{3.1.10}$$

$$p^F_t = \frac{\partial H^R_t}{\partial \mu_t} = \frac{-\alpha(a-w-\tau_t-\alpha x_t)}{2b} + (\tilde{\beta} - \frac{\alpha\beta}{2b})p^F_{t+1} \tag{3.1.11}$$

$$\begin{aligned} \mu_{t+1} = \frac{\partial H^R_t}{\partial p^F_{t+1}} = & \frac{(\gamma+\tau_t)\beta + (2b\tilde{\beta}-\alpha\beta)\mu_t}{2b} \\ & + \frac{\beta(\beta p^R_{t+1} - 2b(a-w-\tau_t-\alpha x_t+\beta p^F_{t+1}))}{2b} \end{aligned} \tag{3.1.12}$$

$$\text{with } x_0 \text{ given, and } \mu_1 = 0 \tag{3.1.13}$$

The boundary condition $\mu_1 = 0$ is directly related to $p^F_{T+1} = 0$. Furthermore, we have $p^R_{T+1} = 0$. As known, the open-loop Stackelberg solution is time inconsistent, since a reoptimization latter in time, at period k for example, will give again to set $\mu_k = 0$ although initially calculated, at period 1, we have $\mu_k \neq 0$.

Anyway, these above necessary conditions, after some algebras and following Medanic [9] give us to solve an augmented discrete Hamiltonian matrix (i.e. with a tracking matrix) of the form:

$$\begin{bmatrix} \tilde{x}_{t+1} \\ \tilde{p}_t \end{bmatrix} = \begin{bmatrix} A & B \\ C & A \end{bmatrix} \begin{bmatrix} \tilde{x}_t \\ \tilde{p}_{t+1} \end{bmatrix} + \begin{bmatrix} D \\ E \end{bmatrix} \tag{3.1.14}$$

Where A, B, C are some 2×2 matrices, D is a 2×1 matrix and $\tilde{x}_t$ and $\tilde{p}_t$ are some 2×1 vectors defined by:

$$A \equiv \begin{bmatrix} \tilde{\beta} - \frac{\beta\alpha}{4b+1} & -\frac{\beta\alpha}{4b+1} \\ -\frac{\beta\alpha}{4b+1} & \tilde{\beta} - \frac{\beta\alpha}{4b+1} \end{bmatrix},$$

$$B \equiv \begin{bmatrix} \frac{\beta^2}{4b+1} & \frac{\beta^2}{4b+1} \\ \frac{\beta^2}{4b+1} & \frac{\beta^2}{4b+1} \end{bmatrix},$$

$$C \equiv \begin{bmatrix} \frac{\alpha^2}{4b+1} - \delta & \frac{\alpha^2}{4b+1} \\ \frac{\alpha^2}{4b+1} & \frac{\alpha^2}{4b+1} \end{bmatrix},$$

$$D \equiv \begin{bmatrix} \frac{\beta(a-w+\gamma)}{4b+1} \\ \frac{\beta(a-w+\gamma)}{4b+1} \end{bmatrix},$$

$$E \equiv \begin{bmatrix} \frac{-\alpha(a-w+\gamma)}{4b+1} \\ \frac{-\alpha(a-w+\gamma)}{4b+1} \end{bmatrix},$$

$$\tilde{x}_t \equiv \begin{bmatrix} x_t \\ \mu_t \end{bmatrix}, \text{ and } \tilde{p}_{t+1} \equiv \begin{bmatrix} p^R_{t+1} \\ p^F_{t+1} \end{bmatrix}.$$

3.1.1 Resolution

To solve this tracking problem defined above we use the sweep method (see Bryson and Ho [6]). That is, we assume a linear relation between the costate and the state vectors:

$$\tilde{p}_k = S_k \tilde{x}_k - g_k \tag{3.1.15}$$

Thus, using this into the augmented Hamiltonian matrix we first get an expression for x_{k+1}:

$$\tilde{x}_{k+1} = (I_{2\times 2} - BS_{k+1})^{-1}(A\tilde{x}_k - Bg_{k+1} + D) \tag{3.1.16}$$

Then using (3.1.16) and (3.1.15) into the definition of p_{k+1} as given by the augmented Hamiltonian matrix, and equating both sides we finally get the difference equations:

$$S_k = C + AS_{k+1}(I_{2\times 2} - BS_{k+1})^{-1}A, \tag{3.1.17}$$
$$g_k = AS_{k+1}(I_{2\times 2} - BS_{k+1})^{-1}(Bg_{k+1} - D) + Ag_{k+1} - E, \tag{3.1.18}$$

where the first equation is the so-called Riccati difference equation, and the second one defines a tracking difference equation.

The boundary conditions are:

$$\tilde{x}_1 = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}, \text{ and } \tilde{p}_{T+1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{3.1.19}$$

And then

$$S_{T+1} = 0_{2\times 2}, \text{ and } g_{T+1} = 0_{2\times 1}. \tag{3.1.20}$$

From the boundary conditions we get

$$S_T = C, \text{ and } , g_T = E. \tag{3.1.21}$$

and so on.

Once, the computation *off line*, backward in time, of the different values of S_k and g_k are made, the values of $\tilde{x}_t$ and $\tilde{p}_t$ follows. These values automatically give us the ones of x_t, μ_t, p^R_t and p^F_t, for all $t \in [1, T]$. The optimal open-loop Stackelberg actions are directly given after by (3.1.8) and (3.1.4).

3.1.2 The optimal discretionary open-loop Stackelberg solution

As we said, the open-loop Stackelberg solution is time inconsistent. That is, whatever is the optimal sequence of taxation $\{\tau^*\}_1^T$ calculated at time $t = 1$, it will not be optimal to pursue with it at time $t = 2$. Rather, the regulator is induced to solve the truncated problem starting at time $t = 2$ and ended at $t = T$ which give him a new optimal sequence of taxation $\{\tau^{**}\}_2^T$. Again at $t = 3$ this sequence will be suboptimal, and so on.

So, while $\{\tau^*\}_i^T$ is the optimal open-loop sequence of taxation for the problem starting at time $t = i$ and ended at time $t = T$, let $\{\tau_i^*\}_i^T$ be the first component of this sequence (and also unique one for the case where $i = T$). Then, the optimal discretionary sequence of taxation, realized *ex post*, is

$$\{\tau_t^{d*}\}_{t\in[1,T]} = (\{\tau_t^*\}_1^1, \{\tau_t^*\}_2^1, ..., \{\tau_t^*\}_{T-1}^1, \{\tau_t^*\}_T^1) \tag{3.1.22}$$

In the economic literature, such a discretionary policy is generally assumed to be worst for the follower regardless to the initial committed strategy that is $\{\tau^*\}_1^T$. But, as we will see, such a conclusion may not hold. In consequence, both players, monopolist and regulator, may gain by using such a discretionary policy. If so, the monopolist may rationally accept *ex post* any revisions of the initial strategy.

3.2 The feedback solution

To solve the game, under the feedback structure of information assumption, we use the dynamic programming method with appropriate value functions (see Başar and Olsder for more details [1]). Recall that this solution is time consistent by construction.

Let T be the last period of the problem. Since the level of pollution x_{T+1} is free, the reaction function of the monopolist at the last period is directly given by the resolution of

$$\arg\max_{q_T\in\Re} J_T^F \tag{3.2.1}$$

That is

$$q_T^* = \frac{a - w - \tau_T - \alpha x_T}{2b} \tag{3.2.2}$$

where x_T is a known fixed value. Then the problem facing the regulator is simply given by:

$$\arg\max_{\tau_T\in\Re} B(q_T^*) + \tau_T q_T^* - D(x_T) \tag{3.2.3}$$

where q_T^* is given by (3.2.2). The maximum is obtained when

$$\tau_T^* = \frac{(1 + 2b)(a - w - \alpha x_t) - 2b\gamma}{1 + 4b} \tag{3.2.4}$$

For the problem starting at the period $T - 1$, the value functions are defined by

$$V^F(T-1, T) = [\arg\max_{q_{T-1}} J_{T-1}^F] + J_T^{F*}, \tag{3.2.5}$$

$$V^R(T-1, T) = [\arg\max_{\tau_{T-1}} J_{T-1}^R] + J_T^{R*}, \tag{3.2.6}$$

where J_T^F and J_T^R are known and supposed to be defined in a linear quadratic form

$$J_T^{L*} = P_T x_T^2 + p_T x_T + n_T, \tag{3.2.7a}$$

$$J_T^{F*} = \tilde{P}_T x_T^2 + \tilde{p}_T x_T + n_T \tag{3.2.7b}$$

Using the definition $x_T = \beta q_{T-1} + \tilde{\beta} x_{T-1}$ into J_T^{F*} and maximize the value function for any fixed τ_{T-1} gives an optimal action for the monopolist for the period $T - 1$. Integrating this into the value function of the regulator, using again the state equation definition and maximizing over all possible τ_{T-1}, we can write the results in some general specific forms

$$\tau_t = K_t x_t + k_t, \tag{3.2.8}$$

$$q_t = \tilde{K}_t x_t + \tilde{k}_t, \tag{3.2.9}$$

$$x_{t+1} = \Omega_t x_t + \beta \tilde{k}_t, \tag{3.2.10}$$

$$J_t^R = P_t x_t^2 + p_t x_t + n_t, \tag{3.2.11}$$

$$J_t^F = \tilde{P}_t x_t^2 + \tilde{p}_t x_t + \tilde{n}_t, \tag{3.2.12}$$

for all $t \in [1, T]$ and with

$$K_t = \frac{-\alpha + 2\alpha(\beta^2(P_{t+1} + \tilde{P}_{t+1}) - b) + 2\beta\tilde{\beta}(\tilde{P}_{t+1} - 2b(P_{t+1} - \tilde{P}_{t+1}) - 2\beta^2\tilde{P}_{t+1}^2)}{1 + 4b - 2\beta^2(P_{t+1} + 2\tilde{P}_{t+1})}, \tag{3.2.13}$$

$$\begin{aligned} k_t = & \frac{-a - 2ab + 2b\gamma + 2b\beta p_{t+1} + 2a\beta^2 P_{t+1} - \beta\tilde{p}_{t+1}(1 + 2b - 2\beta^2 P_{t+1})}{-1 - 4b + 2\beta^2(P_{t+1} + 2\tilde{P}_{t+1})} \\ & + \frac{2\beta^2\tilde{P}_{t+1}(a - \gamma - w - \beta p_{t+1} + \beta\tilde{p}_{t+1}) + w - 2bw - 2\beta^2 w P_{t+1}}{-1 - 4b + 2\beta^2(P_{t+1} + 2\tilde{P}_{t+1}} \end{aligned} \tag{3.2.14}$$

$$\tilde{K}_t = \frac{\alpha + 2\beta\tilde{\beta}(P_{t+1} + \tilde{P}_{t+1})}{1 + 4b - 2\beta^2(P_{t+1} + 2\tilde{P}_{t+1})}, \tag{3.2.15}$$

$$\tilde{k}_t = \frac{a + \gamma - w + \beta(p_{t+1} - \tilde{p}_{t+1})}{1 + 4b - 2\beta^2(P_{t+1} + 2\tilde{P}_{t+1})}, \tag{3.2.16}$$

$$\Omega_t = \beta\tilde{K}_t + \tilde{\beta}, \tag{3.2.17}$$

$$P_t = \frac{-\tilde{K}_t^2}{2} + K_t\tilde{K}_t - \frac{\delta}{2}, \tag{3.2.18}$$

$$p_t = \gamma\tilde{K}_t - \tilde{K}_t\tilde{k}_t + K_t\tilde{k}_t + \tilde{K}_t k_t, \tag{3.2.19}$$

$$n_t = \frac{-\tilde{k}_t^2}{2} + \gamma\tilde{k}_t + k_t\tilde{k}_t, \tag{3.2.20}$$

$$\tilde{P}_t = -b\tilde{K}_t^2 - \alpha\tilde{K}_t - K_t\tilde{K}_t, \tag{3.2.21}$$

$$\tilde{p}_t = a\tilde{K}_t - 2b\tilde{K}_t\tilde{k}_t - w\tilde{K}_t - \alpha\tilde{k}_t - K_t\tilde{k}_t - \tilde{K}_t k_t, \tag{3.2.22}$$

$$\tilde{n}_t = a\tilde{k}_t - b\tilde{k}_t^2 - w\tilde{k}_t - k_t\tilde{k}_t. \tag{3.2.23}$$

where K_t and $\tilde{K}_T$ may be seen as some (1×1) matrices defined by the appropriate scalar Riccati difference equations (3.2.18) and (3.2.21). The terminal conditions are

$$K_T = \frac{-(1 + 2b)\alpha x_T}{1 + 4b}, \tag{3.2.24}$$

$$k_T = \frac{(1 + 2b)(a - w) - 2b\gamma}{1 + 4b}, \tag{3.2.25}$$

$$\tilde{K}_T = \frac{-\alpha}{1 + 4b}, \tag{3.2.26}$$

$$\tilde{k_T} = \frac{a - w + \gamma}{1 + 4b}. \tag{3.2.27}$$

To get the optimal feedback Stackelberg solutions, one must first solve *off-line* the set of equations (3.2.13)-(3.2.23) using the terminal conditions, and then compute *on line* the values of τ_t, q_t and x_t.

3.3 The global Stackelberg solution

Assume that the structure of the information facing each player is a closed-loop one. That is, the leader has a perfect knowledge of all the past and current values of the state and controls which lead him to also know the monopolist's actions. In such an information structure, the regulator may try to find an incentive strategy such that he can reach his global optimum (i.e. *optimum optimorum*).

This *optimum optimorum* is assumed to be unique. Let call (q_t^*, τ_t^*), $\forall t \in [1, T]$, the pair of actions that globally maximized J_t^R. Following Başar and Olsder [1], we know that this pair of actions is directly given by the resolution of

$$\max_{\{\tau\}_1^T, \{q\}_1^T} J^R(\{q\}_1^T, \{\tau\}_1^T) \tag{3.3.1}$$

A necessary condition to solve (3.3.1), by using the first order conditions, is the strict concavity of $J_t^R(q_t,\tau_t)$ in q_t and τ_t, $\forall t$. That is we need no existence of singularity[2]. Pity, $J_t^R(q_t,\tau_t)$ is singular in τ_t, $\forall t$. Hence, a direct optimization is not possible.

One way to avoid this problem, is to add a specific constraint on τ_t or q_t in order to reintroduce τ_t in the maximization's problem. Such a constraint may be the willingness that the monopolist's profits get close to a given level. For example, the regulator may want to reduce these profits to zero[3]. To require $J_t^F = 0$ involves that either

$$q_t = 0, \quad \text{or} \quad q_t = \frac{a - w - \alpha x_t - \tau_t}{b}. \tag{3.3.2}$$

Obviously, $q_t = 0$ must be excluded as a possible choice.

The regulator's problem may be defined by the Hamiltonian-Lagrangian function

$$L_t^R = J_t^R + p_{t+1}^R(\beta q_t + \tilde{\beta} x_t) + \lambda_t\left(\frac{a - w - \alpha x_t - \tau_t}{b} - q_t\right) \tag{3.3.3}$$

To maximize over τ_t and q_t this function, one must solve the following set of first order conditions

$$\frac{\partial L_t^R}{\partial \tau_t} = q_t - \frac{\lambda_t}{b} = 0, \tag{3.3.4}$$

$$\frac{\partial L_t^R}{\partial q_t} = \gamma - q_t + \tau_t + \beta p_{t+1}^R - \lambda_t = 0, \tag{3.3.5}$$

$$x_{t+1} = \frac{\partial L_t^R}{\partial p_{t+1}^R} = \beta q_t + \tilde{\beta} x_t, \tag{3.3.6}$$

$$p_t^R = \frac{\partial L_t^R}{\partial x_t} = -\delta x_t + \tilde{\beta} p_{t+1}^R - \alpha \lambda_t, \tag{3.3.7}$$

$$\frac{\partial L_t^R}{\partial \lambda_t} = \frac{a - w - \alpha x_t - \tau_t}{b} - q_t = 0. \tag{3.3.8}$$

After some algebras, we get

$$\lambda_t = \frac{b(a - w + \gamma - \alpha x_t + \beta p_{t+1}^R)}{2b+1}. \tag{3.3.9}$$

Using this, the following augmented Hamiltonian system has to be solved

$$\begin{bmatrix} x_{t+1} \\ p_t^R \end{bmatrix} = \begin{bmatrix} \tilde{\beta} - \frac{\alpha\beta}{2b+1} & \frac{\beta^2}{2b+1} \\ -\delta + \frac{\alpha^2 b}{2b+1} & \tilde{\beta} - \frac{\alpha b \beta}{2b+1} \end{bmatrix} \begin{bmatrix} x_t \\ p_{t+1}^R \end{bmatrix} + \begin{bmatrix} \frac{\beta(a-w+\gamma)}{2b+1} \\ \frac{-\alpha b(a-w+\gamma)}{2b+1} \end{bmatrix} \tag{3.3.10}$$

with the boundary conditions

$$p_{T+1}^R = 0, \quad \text{and} \quad x_1 \text{ given.} \tag{3.3.11}$$

By assuming a linear relationship between the co-state and the state, $p_t^R = K_t x_t - g_t$, the following scalar

[2] Namely the condition that $J_T^R(q_t,\tau_t)$ is at least a C^2 function in q_t and τ_t.

[3] We chose this zero profits' constraint for simplicity purpose. Hence, the constraint $J_t^R = 100$, $\forall t$, involves

$$q_t = \frac{a - \tau_t - w - \alpha x_t - \sqrt{-4\,b\,n + (-a + \tau_t + w + \alpha x_t)^2}}{2\,b}$$

$$\text{ou } q_t = \frac{a - \tau_t - w - \alpha x_t + \sqrt{-4\,b\,n + (-a + \tau_t + w + \alpha x_t)^2}}{2\,b}$$

which lead the regulator's problem to be a non linear optimal control one hard to solve analytically.

Riccati and tracking difference equations have to be solved *off line*, backard in time

$$K_t = -\delta + \frac{\alpha^2 b}{2b+1} + \frac{(\tilde{\beta} - \frac{\alpha b \beta}{2b+1})K_{t+1}(\tilde{\beta} - \frac{\alpha\beta}{2b+1})}{1 - \frac{\beta^2 K_{t+1}}{2b+1}}, \tag{3.3.12}$$

$$g_t = \frac{\alpha b(a-w+\gamma)}{2b+1} + \frac{(\tilde{\beta} - \frac{\alpha b\beta}{2b+1})K_{t+1}(\frac{\beta^2 g_{t+1} - \beta(a-w+\gamma)}{2b+1})}{1 - \frac{\beta^2 K_{t+1}}{2b+1}} + (\tilde{\beta} - \frac{\alpha b\beta}{2b+1})g_{t+1}. \tag{3.3.13}$$

with the terminal conditions

$$K_T = -\delta + \frac{\alpha^2 b}{2b+1}, \quad K_{T+1} = 0, \tag{3.3.14}$$

$$g_T = \frac{\alpha b(a-w+\gamma)}{2b+1}, \quad g_{T+1} = 0. \tag{3.3.15}$$

Once these *off line* values are found, the optimal sequences $\{x^*\}_1^T$, $\{p^{R*}\}_1^T$, $\{\lambda^*\}_1^T$, $\{\tau^*\}_1^T$ and $\{q^*\}_1^T$ can be calculated *on line*. Recall that $\{\tau^*\}_1^T$ and $\{q^*\}_1^T$ achieve the optimum optimorum of the regulator given the zero-profit constraint.

Next, the problem facing the regulator is to find and announce at the beginning of the game an optimal incentive strategy such that the monopolist implements the sequence $\{q^*\}_1^T$. Since the regulator know either directly q_t, $\forall t$, or at least may calculate it from its knowledge of x_t, and following Başar and Olsder [1], we know that a candidate incentive strategy, call it θ, is

$$\tau_t \equiv \theta_t(q_t) = \tau_t^* + k_t(q_t^* - q_t) \tag{3.3.16}$$

where τ_t^* and q_t^* are the desired actions from the viewpoint of the regulator, and are some known values. To proceed with θ, the regulator has to find the sequence $\{k\}_1^T$ such that the monopolist cannot do better than $\{q^*\}_1^T$ to which the regulator responds by $\{\tau^*\}_1^T$. If such a sequence of incentive strategies exists, the global Stackelberg solution is time consistent by hypothesis as it reaches the optimum optimorum of the regulator.

Since $\theta_t(q_t)$ is a known function, the problem facing the monopolist is a standard optimal control one. Furthermore, given that there is no uncertainty, the solution will be the same whatever the information structure, open-loop or feedback, faced by the monopolist is. For simplicity purpose, we derive the solution using the dynamic programming method.

So, let the incentive strategy for the last period be

$$\theta_T = \tau_T^* + k_T(q_T^* - q_T). \tag{3.3.17}$$

At this last period, the monopolist problem involves to solve

$$\arg\max_{q_T} J_T^F(q_T, \theta_T) \tag{3.3.18}$$

The first order condition is

$$q_T = \frac{a - \tau_T^* - k_T q_T^* - w - \alpha x_T}{2b - 2k_T}. \tag{3.3.19}$$

Recall that the equality $q_T = q_T^*$ is desired. Let k_T^* be an "incentive coefficient" such that this equality holds. Its value is given by

$$k_T^* = \frac{-(a - w - \alpha x_T - \tau_T^* - 2bq_T^*)}{q_T^*}. \tag{3.3.20}$$

We may easily guess the sign of k_T^*. It should be positive since the paire (τ_T^*, q_T^*) is calculated given a non-profit constraint which means that the monopolist, given τ_T^*, should not be able to produce more (i.e. $q_T \geq q_T^* \Rightarrow J_T^F(\tau_T^*, q_T) < 0$). Since the monopolist may only decide to produce less, a lower value of q_T should be associated to an increase of τ_T in order to induce the monopolist to choose q_T^*. In consequence, $k_T^* > 0$ is required.

Anyway, the last period payoff of the monopolist can be defined as follows

$$\begin{aligned} J_T^F &= (a - bq_T^*)q_T^* - wq_T^* - \alpha x_T q_T^* - \tau_T^* q_T^* \\ &= \tilde{P}_T x_T^2 + \tilde{p}_T x_T + \tilde{n}_T, \end{aligned} \tag{3.3.21}$$

where

$$\begin{aligned} \tilde{P}_T &= 0, \\ \tilde{p}_T &= -\alpha q_T^*, \\ \tilde{n}_T &= (a - bq_T^*)q_T^* - wq_T^* - \tau_T^* q_T^*. \end{aligned}$$

One may check that $\theta_T(k_T^*)$ also induces the regulator to implement τ_T^*. The regulator payoff can be rewritten as follows

$$J_T^R = P_T x_T^2 + p_T x_T + n_T \tag{3.3.22}$$

with

$$\begin{aligned} P_T &= \frac{-\delta}{2}, \\ p_T &= 0, \\ n_T &= \gamma q_T^* - \frac{q_T^{*2}}{2} + \tau_T^* q_T^*. \end{aligned}$$

Using a similar procedure we used to get the feedback Stackelberg solution, one can get the following general forms of the closed-loop Stackelberg solution

$$q_t = \frac{a - w - \alpha x_t + \beta \tilde{p}_{t+1} - \tau_t^* - k_t q_t^*}{2b - 2k_t}, \tag{3.3.23}$$

$$k_t^* = \frac{-(a - w - \alpha x_t + \beta \tilde{p}_{t+1} - \tau_t^* - 2bq_t^*)}{q_t^*}, \tag{3.3.24}$$

$$x_{t+1} = \beta q_t^* + \tilde{\beta} x_t, \tag{3.3.25}$$

$$J_t^R = P_t x_t^2 + p_t x_t + n_t, \tag{3.3.26}$$

$$J_t^F = \tilde{P}_t x_t^2 + \tilde{p}_t x_t + \tilde{n}_t. \tag{3.3.27}$$

where

$$\begin{aligned} P_t &= \frac{-\delta}{2}, \\ p_t &= 0, \\ n_t &= \gamma q_t^* - \frac{q_t^{*2}}{2} + \tau_t^* q_t^*. \\ \tilde{P}_t &= 0, \\ \tilde{p}_t &= -\alpha q_t^*, \\ \tilde{n}_t &= (a - bq_t^*)q_t^* - wq_t^* - \tau_t^* q_t^*. \end{aligned}$$

Remark: it is possible that for some values of the parameters, we have $k_t^* = b$ for some t. Then as easily seen from (3.3.19) or (3.3.23), the problem facing the monopolist becomes singular. In such a case, the optimal level of production may not be obtained by (3.3.19) or (3.3.23). In fact the optimal level of production is given by

$$q_t = \begin{cases} \dfrac{a - w - \alpha x_t + \beta \tilde{p}_{t+1} - \tau_t^* - k_t q_t^*}{2b - 2k_t} & \text{if } k_t^* \neq b, \\ q_t^* & \text{if } k_t^* = b. \end{cases} \tag{3.3.28}$$

4 Some numerical comparisons of the solutions

The results presented here were obtained for the following values of the parameters:

$$a = 150, b = 5, w = 2, \alpha = 2, \delta = 3, \text{ and } \gamma = 5.$$

The initial level of pollution is set to $x_1 = 1$. Two numerical simulations are run. In the first one, we set $\beta = 0.4$ and $\tilde{\beta} = 0.5$, and in tghe second one $\beta = 0.8$ and $\tilde{\beta} = 0.8$ are used.

4.1 First case: $\beta = 0.4$ and $\tilde{\beta} = 0.5$

Logically the best solution, from the regulator viewpoint, is the global one (table 1 and figure 1), and it is the worst for the monopolist since its profits reduce to zero (table1 and figure 2). This solution involves the higher levels of pollution[4], tax and production. Recall that his global Stackelberg solution is time consistent.

Quite surprising is that the time consistent feedback solution does also better than the open-loop one, with or without commitment (figure 1 and table 1). It is generally assumed that the problem of the time consistent solution is its suboptimality in respect of the discretionary one (cf. Kydland and Prescott [8], Barro and Gordon [2, 3]). What we learn from this simple model it's that there is no way it should be always the case when the follower has a real payoff function and not a very restrictive one[5].

Solutions	J_c^R	J_c^F
Open-loop (OL)	8.2256 10^3	3.7337 10^3
Optimal discretionary (OLd)	8.2363 10^3	4.1417 10^3
Feedback (Fd)	8.5344 10^3	3.9647 10^3
Closed-loop (CL)	1.5064 10^4	0

Table 1: Cumulated Payoffs

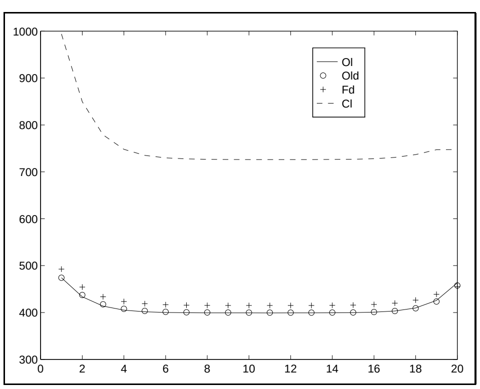

Figure 1: Evolution of J_t^R

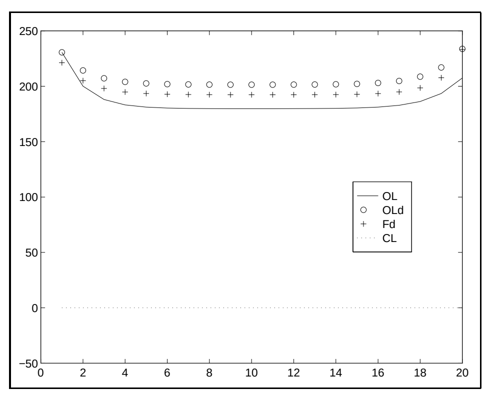

Figure 2: Evolution of J_t^F

The level of pollution is directly related to the regulator's welfare. And since all others variables are connected each others, we found the same order of the solutions in the figures. Hence, higher welfare will imply higher pollution, and so a higher price and production.

As the global solution involves zero-profits for the monopolist, one may wonder why the monopolist will still produce something ? Obviously, the regulator may accept some profit for the monopolist by allowing a little more pollution. That is our global solution is based on a non-profit constraint. All constraints that will involve

[4] The reader is implicitly refereed to the corresponding graphics that are shown in appendix.

[5] These literatures are based on some specific Stackelberg games where the follower has a kind of cheating aversion cost function.

a level of pollution between this one and the one obtained under the feedback solution will still allow the global Stackelberg solution to be the first one.

The two time consistent solutions mainly differ because of the level of taxation, this level is higher with the incentive solution (global Stackelberg) since the profit must be reduce to zero.

Another important conclusion is that, in an open-loop information structure, the discretionary solution is better for everyone than to stay committed to the initial announcement (figures 1 and 2 and table 1). In such a case, we don't see any reason why this discretionary solution should involve some loss of credibility, since the monopolist may be aware that to believe in a likely recalculated sequence of taxations will get him in a better position after. Then he may optimally believe an initial sequence of taxation knowing that the regulator will not continue with it latter.

4.2 Second case: $\beta = 0.8$ and $\tilde{\beta} = 0.8$

The simulation provides the same kinds of comments. That is, and the more important one, the monopolist will benefit from a not-committed regulator's policy to the open-loop initial solution (table 2 and figure 4).

For the regulator, the feedback time consistent solution is no more better than the optimal discretionary one (figure 3 and table 2). But these solutions are very closed. Finally, it seems that the gain from not staying committed to an initial open-loop solution (by using the optimal discretionary solution) is always quite small. So, the incentive to deviate is not very strong (tables 1 and 2).

Solutions	J_c^R	J_c^F
Open-loop (OL)	3.2652 10^3	1.0727 10^3
Optimal discretionary (OLd)	3.2725 10^3	1.2848 10^3
Feedback (Fd)	3.1988 10^3	1.3525 10^3
Closed-loop (CL)	4.9485 10^3	0

Table 2: Cumulated Payoffs

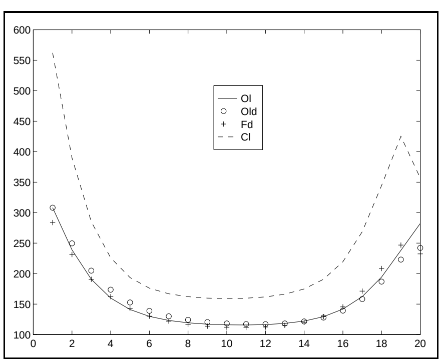

Figure 3: Evolution of J_t^R

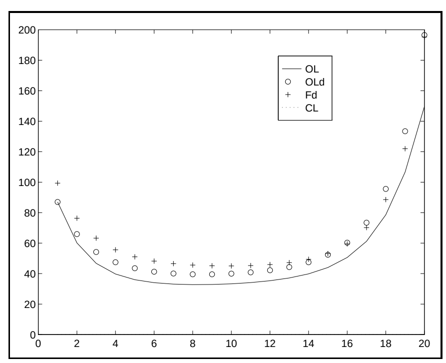

Figure 4: Evolution of J_t^F

5 Conclusion

In this paper, we derived the different possible Stackelberg solutions of a leader-follower pollution game. The different solutions are well-known, mainly because of the work of Başar and Olsder [1]. But despite this fact, some misunderstandings still exist concerning the comparison of these solutions. We underline the incorrectness of two of them: the suboptimality of the time consistent solution, and the assuming increased cost on the follower when the leader use a discretionary policy.

Hence, with one particular numerical simulation we presented, we found that the time consistent solution is the best one for the leader. Moreover, it is possible to find a simulation such as this conclusion also holds for the follower. The gain for both players of using optimal discretionary solution was underlined. This result is closely related to the fact that a cheating-by-second play strategy may also be a good strategy for both players (see Vallée, Deissenberg and Başar [13]). Finally, we concluded on the very small advantage of using such a solution.

Of course, those results were found with a very specific dynamic game model. Another one may give opposite results. Some more theoretical understandings of the different dynamic solutions are needed if we want, for example, to know exactly when and why a time consistent solution may be suboptimal or not. Such a project is a currently research.

References

[1] T. Başar and G.J. Olsder. *Dynamic Noncooperative Game Theory*. Academic Press, New York, 2nd edition, 1995.

[2] R. Barro and D. Gordon. A positive theory of monetary policy in a natural rate model. *Journal of Political Economy*, 91:589–610, 1983.

[3] R. Barro and D. Gordon. Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12:3–20, 1983.

[4] A. Batabyal. Consistency and optimality in a dynamic game of pollution control I: Competition. Economic Reserch Institue Study Paper, ERI 95-29, Utah State University, 1995.

[5] A. Batabyal. Consisteny and optimality in a dynamic game of pollution control II: Monopoly. *Environmental and Resource Economics*, 8:315–330, 1996.

[6] A.E. Bryson and Y.C. Ho. *Applied Optimal control*. John Wiley & Sons, New York, 1975.

[7] F. Kydland. Noncooperative and dominant player solutions in discrete dynamic games. *International economic review*, 16(2):301–335, 1975.

[8] F. Kydland and E. Prescott. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–492, 1977.

[9] J. Medanic and D. Radojevic. Multilevel Stackelberg strategies in linear-quadratic systems. *Journal of optimization theory and applications*, 24(3):485–497, march 1978.

[10] R. Pethig. *Conflicts and Cooperation in Managing Environmental Resources*. Springer-Verlag, 1992.

[11] M. Simaan and J.B. Cruz. Additional aspects of the Stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, 11:613–626, June 1973.

[12] M. Simaan and J.B. Cruz. On the Stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, 11:533–555, May 1973.

[13] T. Vallée, C. Deissenberg, and T. Başar. Optimal open loop cheating in dynamic reversed linear quadratic stackelberg games. *Annals of Operations Research*, 88:217–232, 1999.

A Graphics

We use the following abbreviations and notations for the graphics:

- ol (—) Open-loop solution,
- old (o) Open-loop discretionary solution,
- fd (+) Feedback solution,
- cl (- -) Closed-loop solution (myopic and nonmyopic cases),

A.1 First simulation: $\beta = 0.4, \tilde{\beta} = 0.5$

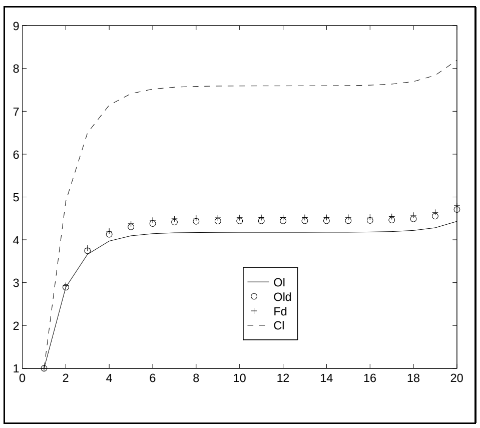

Figure 5: Pollution stock

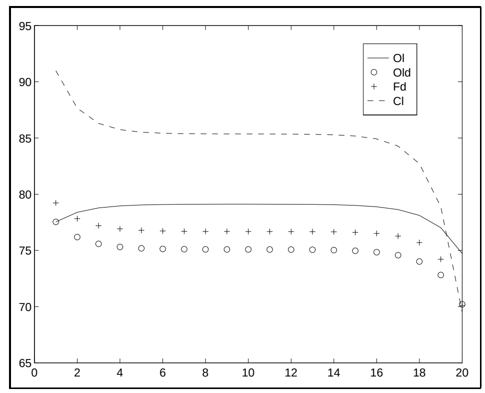

Figure 6: Taxation's level

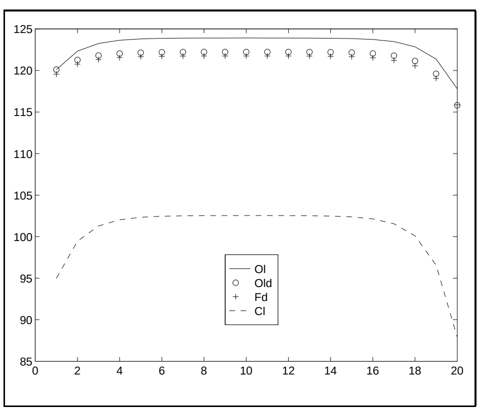

Figure 7: Price's level

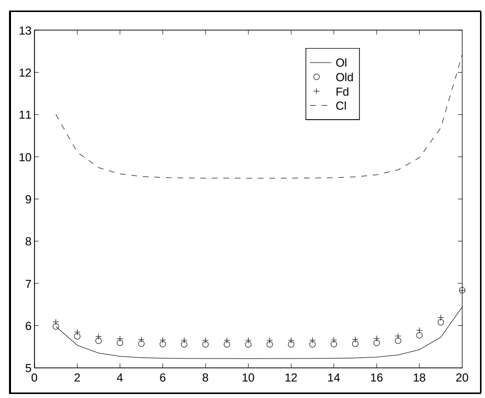

Figure 8: Production's level

A.2 Second simulation: $\beta = 0.8, \tilde{\beta} = 0.8$

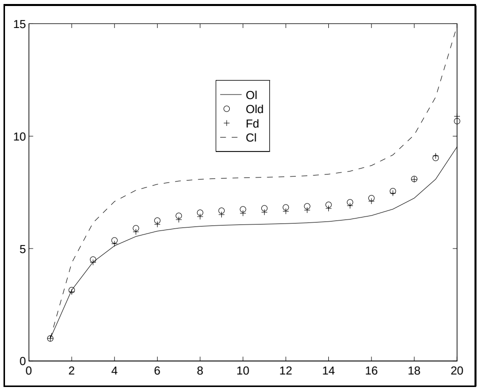

Figure 9: Pollution stock

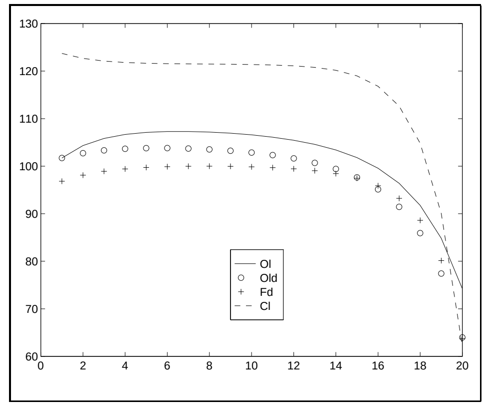

Figure 10: Tax's level

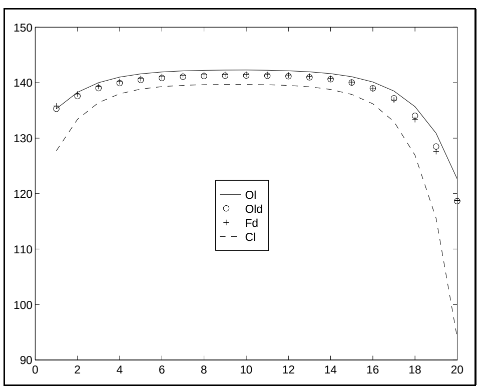

Figure 11: Price's level

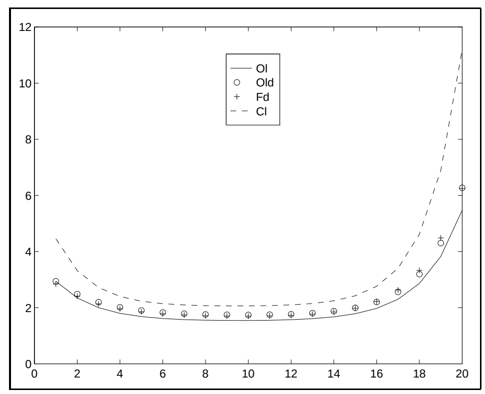

Figure 12: Production's level