Optimal Pollution Trading without Pollution Reductions: A Note

Jorge H. García
Matthew T. Heberling
Hale W. Thurston

Abstract

Many kinds of water pollution occur in pulses, e.g., agricultural and urban runoff. Ecosystems, such as wetlands, can serve to regulate these pulses and smooth pollution distributions over time. This smoothing reduces total environmental damages when the “instantaneous” damage function is convex. This paper introduces a water quality trading model between a farm (a pulse-pollution source) and a firm (a more steady pollution source) where the object of exchange is the “temporary” retention of runoff as opposed to total runoff reductions. The optimal trading ratio requires firm emissions to be offset by more than a proportional retention of the initial agricultural runoff pulse. The reason is twofold: a) emissions are steady over time and -in this sense- have relatively larger environmental impact, and b) certain kinds of runoff management cause otherwise inexistent delayed environmental damages.

Key terms: water quality trading; flow pollution; wetlands; trading ratio

JEL: D21; Q53; Q58

---

1 Respectively, Assistant Professor (García), Departamento de Economía, Pontificia Universidad Javeriana, Edificio Gabriel Giraldo SJ, Calle 40 No 6-23 P7, Bogota, Colombia (E-Mail: jgarcia-l@javeriana.edu.co); Economist (Heberling, Thurston), USEPA National Risk Management Research Laboratory (NRMRL), Cincinnati, Ohio.

We are grateful to Carlos Chavez for insightful comments on an earlier version. We would also like to thank Jake Beaulieu and Charles Lane for useful discussion on wetlands. García would like to acknowledge that part of this research was conducted under a postdoctoral research associateship provided by the National Academy of Sciences (NAS), National Research Council (NRC). The views expressed herein are strictly the opinion of the authors and in no manner represent current or planned policy by U.S. Environmental Protection Agency.
Intercambio de contaminación óptimo sin reducciones de contaminación: Una nota

Jorge H. García
Matthew T. Heberling
Hale W. Thurston

Resumen

Varios tipos de contaminación ocurren en pulsos, e.g., la contaminación hídrica proveniente del sector agrícola. Algunos ecosistemas como los humedales pueden actuar como reguladores de estos pulsos y suavizar las distribuciones de contaminación en el tiempo. Esta redistribución de emisiones reduce los daños ambientales totales cuando la función instantánea de daños es convexa. Este artículo introduce un modelo de permisos de emisiones negociables entre una finca agrícola (una fuente contaminación en pulsos) y una industria (una fuente de contaminación continua) donde el bien objeto de intercambiado es la “retención temporal” de contaminantes –a diferencia de su reducción total-. La tasa óptima de intercambio demanda que las emisiones de la industria sean compensadas más que proporcionalmente por reducciones en el pulso inicial de contaminación agrícola dado que: a) las emisiones industriales son continuas en el tiempo, y en este sentido, tienen un impacto ambiental relativamente mayor y b) la implementación de humedales por parte de la finca agrícola produce daños ambientales rezagados inicialmente inexistentes.

Palabras Clave: sistemas de permisos de emisiones negociables; contaminantes flujo; humedales; tasa optima de intercambio

JEL: D21; Q53; Q58
1. Introduction

Some ecosystems constitute cost-effective pollution sinks and play an important role in environmental policy such as cap-and-trade programs. Reforestation, for example, generates tradable carbon credits under the Kyoto Protocol and will most likely be included in future climate agreements. Similarly, wetland creation and restoration is currently being considered as a possible source of nitrogen and phosphorous credits in local water quality trading programs or nutrient markets in the United States (USEPA 2007). These watershed level programs have usually taken the following form: a source of nutrient loading, typically an industrial or a sewage treatment plant, facing a Total Maximum Daily Load (TMDL) restriction on nutrients, has the option of building more treatment capacity or purchasing nutrient reductions from the agricultural sector (Morgan and Wolverton 2005; Pharino 2007; Woodward et al. 2002). Farmers have a variety of options to mitigate nutrient runoff such as reducing fertilizer use on their crops, implement no-till techniques, or constructing and restoring wetlands to act as runoff sinks on their property. For the agents participating in these markets, natural ecosystems that act as final or long-term repositories of pollution are, in terms of their functions, no different from other abatement activities or technologies. The general framework used in the economics literature to analyze emissions” markets (e.g., Baumol and Oates 1988) thus also extends to ecosystems as abatement technologies.

A significant fact that has been given little attention in the trading literature is the idea that ecosystems often act as “regulators” of environmental variables. For instance, a basic

---

2 Water quality trading is to a large extent still in experimental stage. Only 19 out of a total of 47 existing initiatives as of 2007 presented trading activities. The existence of high transaction costs (Nguyen and Shortle 2008; Stavins 1995) and the impossibility to transfer liability along pollution allowances in existing programs (USEPA 2003; Shortle 2009) are some of the main reasons preventing trading in some programs.
function of wetland ecosystems is the regulation of water regimes. Wetlands are often described as sponges that absorb water during wet periods and then release it in dry periods (e.g., Mitsch and Gosselink 2007). Thus, they provide a natural source of flood, agricultural, and urban runoff control (e.g. Person et al. 1999). Runoff exports of phosphorous and nitrogen into waterways typically occur in pulses following rain and storm events (e.g., Tucker and Brass 2000).

Consider an agricultural runoff pulse that enters a wetland. It is absorbed and dispersed slowly following different paths towards one or several outlets into a stream or a river. The residence time of agricultural runoff in the wetland is given by the time difference between its entrance and the time where is completely released into the river. While the runoff pulse may occur over a few hours, its release can take as long as a few days (e.g., Kadlec and Wallace 2009). The retention time varies depending on a number of factors such as the size and shape of the wetland, its vegetation, soil type, etc. (e.g., Persson et al. 1999). Since a typical precipitation distribution comprises a series of spikes that are preceded and followed by relatively dry periods, a direct consequence of wetland implementation is the reduction of runoff load discharge “per unit of time.” This has important consequences for water quality and human health. Pollution pulses are associated with high concentrations and loads in short periods of time. It is a well accepted fact that humans, animals and plants are particularly sensitive to bursts, rather than gradual, intakes of foreign and poisonous substances. In practice, the smoothing properties of wetlands are useful in meeting TMDLs.

---

3 While this is a general attribute of wetlands, the particulars of this and other environmental services may vary greatly across different regions (Mitsch and Gosselink 2007). Apart from providing water quality enhancement and flood control services, wetlands are also hosts of biodiversity/wildlife and support recreation activities (Mitsch et al. 2009). It has been estimated that about half of the United States’ original wetlands had been converted to other uses by 1970 (Dahl, 2000). There exists a local, and worldwide, growing interest in preserving, creating and restoring wetlands ecosystems (Knight, 1997).
This paper introduces a model of water quality trading between a farm, some times referred here as a pulse-pollution source, and a firm, a more steady or constant source of pollution. The farm may supply pollution credits to the firm via wetland implementation. The object of exchange in this context is the temporary retention of pollution as opposed to traditional pollution reductions. As mentioned earlier, wetlands also act as permanent repositories of pollution. Total agricultural runoff abatement measured from a baseline, which depends on wetlands and a number of other management practices, has been studied by other authors (e.g., Shortle and Horan 2008). This paper focuses on the welfare gains derived from the regulation of pollution flows and the implications for environmental policy.

We propose a trading ratio to be used in the trading program that specifically acknowledges the pulse nature of the abatement offered by wetlands. A trading ratio is the rate at which firms’ emissions ought to be offset by runoff reductions from farming so that the optimal level of environmental damage is maintained. Factors already considered in the literature that affect the optimal trading ratio include uncertainty associated with runoff and enforcement costs (e.g., Malik et al. 1993). It has also been noted that local environmental agencies do not have control over the number of permits in the market and call for the use of second-best trading ratios (Horan and Shortle 2005). A well-established result in the water quality trading literature states that the optimal trading ratio is higher than 1 for perfectly mixed pollution (Shortle 1987; Malik et al. 1993). Runoff is stochastic due to unpredictable weather and other factors. When uncertainty is costly to society (i.e., the damage function is convex) a unit of firm emissions can only be offset by more than a unit of (expected) runoff reduction. Our results on the other hand, indicate that

---

4 Due to the diffuse nature of runoff, farming is usually described as a nonpoint source of pollution, while municipal sewage treatment plants and industrial plants are defined as point sources of pollution. We on the other hand are interested on pollution flows over time and differentiate sources as pulse and steady. While swage treatment plants operate under relatively steady flows, some industrial plant may be thought as following pulse regimes. These, however, are typically lengthier and more frequent than rainfall events and runoff pulses.
when eschewing focus on the effects of stochastic events on runoff, a trading ratio greater than one may also be optimal but for different reasons. As suggested earlier, the trading ratio reflects the relative environmental impact of emissions and runoff. In our case, management of the initial runoff pulse via wetland implementation is not a perfect substitute for an increase of a (constant) emission flow overtime. A unit of firm emissions should be offset by more than a proportional reduction of the initial runoff pulse since, unlike runoff pulses, emissions generate a steady stream of environmental damages. Also wetland implementation only holds the runoff pulse temporarily and facilitates delayed environmental damages.

2. Model

Consider a watershed with two polluting agents, a farm and a firm, and two distinct periods, each one possibly comprising different time lengths. A storm event occurs in period 1 while period 2 is relatively dry. Farming, coupled with the storm of period 1 generates runoff $r$ in that period. The farmer can, however, temporarily mitigate runoff exported into the river through the implementation of wetlands. We initially assume that no runoff is permanently retained in the wetland. That is, there is no significant infiltration. The consequences of including this aspect in the model are discussed later in the paper. Let $\alpha(w)$ be the proportion of runoff detained in wetland area $w$ in period 1 with $\alpha(0) = 0$, $\alpha' > 0$ and $\alpha'' > 0$. While farming activities may continue in period 2, weather conditions are not conducive for new production of runoff in that
period. Total exports of runoff into the waterway with wetlands are \( r_1 = (1 - \alpha(w)) \times r \) and \( r_2 = \alpha(w) \times r \) for periods 1 and 2, respectively.\(^5\)

The firm produces emissions \( e \) in each period and derives benefits, \( B(e) \), with \( B' > 0 \) \( B'' < 0 \).

Total pollution exports into the waterway are \( 2 \times e + r \), although they may not be equally distributed over the two periods. Let \( p_w \) denote the cost of implementing wetlands per unit of area. With instantaneous convex damage function \( D \) (\( D' > 0 \) and \( D'' > 0 \)), social welfare may be expressed as \( W(e, w) = 2 \times B(e) - p_w w - D(e, (1 - \alpha(w)) \times r) - D(e, \alpha(w) \times r) \).

The presence of marginally increasing damages implies that small reallocations of runoff from the period with higher levels of nutrients to the period with lower levels of nutrients always generates a reduction in total damages. In this narrow sense, smoother pollution profiles are always preferred by society. Maximizing total social welfare, we obtain the following conditions:

\[
2 \times B'(\hat{e}) = D_e(\hat{e}, (1 - \alpha(\hat{w})) \times r) + D_e(\hat{e}, \alpha(\hat{w}) \times r) \tag{1}
\]

\[
\frac{p_w}{\alpha'(\hat{w}) \times r} = D_{r1}(\hat{e}, (1 - \alpha(\hat{w})) \times r) - D_{r2}(\hat{e}, \alpha(\hat{w}) \times r) \tag{2}
\]

where \( \hat{\alpha} \) and \( \hat{w} \) denote the optimal choices of emissions and wetland area. \( D_e, D_{r1} \) and \( D_{r2} \), with \( D_{r1} = D_{r2} \), represent instantaneous marginal damages of pollution and runoff. Condition 1 equates marginal benefits and damages of emissions. Condition 2, on the other hand, weights benefits and damages of emitting an extra unit of runoff higher in period 1 than in period 2. The

\(^5\) Assuming discrete periods suffices the purpose of this paper. The residence time of \( r \) in the wetland can typically be characterized by a distribution function, \( r(t) \) (e.g., Kadlec and Wallace 2009). A mass balance with no runoff loss may be expressed as \( \int r(t) dt = r \). At any given time, \( t_x \), between the precipitation event and the time when \( r \) has been released, \( r(t_x) < r \). In our model this is equivalent to \( r_1, r_2 < r \).
negative sign in condition 2 reflects the fact that released runoff in period 1 has negative effects on water quality in that period and positive effects in period 2. Figure 1 illustrates social choice for the special case of perfectly mixed pollutant, i.e., \(D(e,r)=D(e+r)\). The x-axis represents firm emissions \(e\) in period 1 and the y-axis represents farm runoff in period 1, \(r_1\). Farm runoff in period 2, \(r_2\), is given by the difference between \(r\) and \(r_1\). As suggested earlier, minimum damage occurs at \((0, 1/2r)\) for mixed pollution, and iso-damage curves are concentric to this point.

![Graph](image)

**Figure 1:** Emissions \((e)\) vs Runoff \((r_1)\) Exports for Perfectly Mixed Pollutants and \(p_w>0\)

Note that for a given level of emissions there are typically two levels of runoff in period 1 for which the same environmental damage is generated. This contrasts with earlier treatments of pollution where the iso-damage functions are concentric to the origin (e.g., Malik *et al.* 1993; Horan and Shortle 2005). Higher levels of emissions in both periods and runoff in period 1 provide larger private benefits and this is the driving force northeastwards in Figure 1. Since
wetland implementation is costly, larger marginal damages in period 1 than in period 2 are optimal. As mentioned in page 3, wetlands ecosystems often constitute rich wildlife habitats, and subsidies are often discussed on this basis (e.g., Heberling et al. 2009). If a large enough wetland subsidy existed in the economy, it would be the case that \( p_w < 0 \). In this case, the iso-benefit curves would have positive slopes (as opposed to negative slopes as in Figure 1) and \( \hat{r}_1 < r/2 \).

For a given level of environmental damage, optimality is achieved when the iso-benefit curve is tangent to the iso-damage curve (i.e., where conditions 1 and 2 are met). The optimal marginal rate of substitution between profiles of emissions \((e, e)\) and runoff \((r_1, r - r_1)\), is given by

\[
\frac{2 \times B'(\hat{\varepsilon})}{p_w / \alpha'(\hat{\omega}) \times r} = \frac{D_e(\hat{\varepsilon}, (1 - \alpha(\hat{\omega})) \times r) + D_e(\hat{\varepsilon}, \alpha(\hat{\omega}) \times r)}{D_e(\hat{\varepsilon}, (1 - \alpha(\hat{\omega})) \times r) - D_{e2}(\hat{\varepsilon}, \alpha(\hat{\omega}) \times r)}
\]

(3)

Although perfectly mixed pollutants are often associated with perfect substitution, this is not the case when the flow-regulating properties of wetland ecosystems are taken into account. Under perfectly mixed pollution \( D_e = D_{e1} = D_{e2} \), and the right hand side of condition (3) can be expressed as \( (D_e(\hat{\varepsilon} + \hat{r}_1) + D_e(\hat{\varepsilon} + \hat{r}_2))/(D_e(\hat{\varepsilon} + \hat{r}_1) - D_e(\hat{\varepsilon} + \hat{r}_2)) \). This is clearly greater than 1.

Thus, in order to maintain the socially optimal level of environmental damage, a unitary increase in emissions \( e \) may be offset by more than a unitary decrease of runoff in period 1. The reason is twofold: a) unlike runoff, emissions entail a continuous, and therefore relatively larger, flow of environmental damages and b) runoff management causes delayed environmental damages. Note that if no emissions existed in period 2 or runoff was permanently retained, the respective optimal trading ratios would be \( D_e(\hat{\varepsilon} + \hat{r}_1)/(D_e(\hat{\varepsilon} + \hat{r}_1) - D_{e2}(\hat{r}_2)) \) and \( (D_e(\hat{\varepsilon} + \hat{r}_1) + D_e(\hat{\varepsilon})) / D_e(\hat{\varepsilon} + \hat{r}_1) \) which are both greater than one.
**Water quality trading**

We now turn to a situation wherein a regulatory agency seeks to implement an emissions trading market between the farm and the firm. The agency controls two variables: the total number of permits and the rate at which these permits may be exchanged, that is the trading ratio. As explained earlier, the marginal damages from the two sources of pollution differ, and substitution in the market may not occur on an one-to-one basis (e.g., Tietenberg 2006). Two types of permits could be issued by the regulator: “traditional” permits that allow the firm to emit \( e \) units of pollution in each period, and permits that allow the farm produce \( r \) units of runoff in period 1 (and \( r \) in period 2). Following existing water quality trading markets, we assume that the farm receives all pollution allowances so that \( \tilde{e} = 0 \). As suggested in the introduction, history shows the agricultural sector in the United States has owned pollution rights (Horan and Shortle 2005). The object of exchange in this setting becomes the regulation of runoff flow as opposed to traditional runoff reductions. Let \( p \) be the equilibrium price in the emissions trading market.

With \( \tilde{t} \) denoting the trading ratio set by the environmental agency, the firm’s and the farm’s profit functions are 

\[
2 \times B(e) - p \times e \times \tilde{t} \quad \text{and} \quad - p_w \times w + p \times (\tilde{r} - (1 - \alpha(w)) \times r),
\]

respectively. The associated first order conditions are:

\[
2 \times B'(e) = p \times \tilde{t} \tag{4}
\]

\[
\frac{p_w}{\alpha'(w) \times r} = p \tag{5}
\]

with the market clearing conditions given by:

\[
\tilde{r} = (1 - \alpha(w)) \times r + e \times \tilde{t} \tag{6}
\]
Equations (4), (5) and (6) define a decentralized solution for emissions and wetland levels given parameters \( \tilde{r}_i \) and \( \tilde{t} \). Now assume that the regulator sets initial permits according to:

\[
\tilde{r}_i = (1 - \alpha(\hat{w})) \times r_i + \hat{e} \times \hat{t}
\]

That is, the total number of permits, in the denomination of runoff reductions in period 1 as opposed to period 2, equals the social optimal choice of pollution. Dividing equation (4) by equation (5) and replacing \( e \) from conditions (6) and (7), we obtain:

\[
2 \times B'((1 - \alpha(\hat{w})) \times r \times (1/\tilde{t}) + \hat{e} - (1 - \alpha(w) \times r \times (1/\tilde{t}))) = \tilde{t}
\]

This can be interpreted as an implicit function of the market solution for \( w \) in terms of \( \tilde{r} \). When the trading ratio is set according to the social marginal rate of substitution (right hand side of equation (3)), it directly follows that the induced choice of wetlands by the farm in the market coincides with the social optimum, \( \hat{w} \). Consequently, the market level of firm emissions is given by \( \hat{e} \). Figure 1 helps illustrate how the water quality market leads to this result. The market clearing condition (6) can be expressed as a linear function of \( e \) in terms of \( r_i \). When \( \tilde{t} \) and \( \tilde{r}_i \) are chosen optimally, this corresponds to the tangent to the iso-damage function evaluated at the social optimum. The emergent market price equals marginal private benefits of emissions and runoff in period 1 equals the corresponding optimal marginal damages.

The objective of our analysis was to study the role of wetlands as flow-regulators in water quality policy. We, therefore, isolated the retention effect and assumed that no runoff was permanently infiltrated into the wetland. This feature can be incorporated in our model by setting period 2 runoff as \( r_2 = \beta \times \alpha(w) \times r \), with \( \beta \in [0,1] \). The marginal damage of delayed runoff is given by \( \beta \times D_{r2}(e, \beta \times \alpha(w) \times r) \). Note that this is 0 for \( \beta = 0 \). As expected, including the water treatment
properties of wetlands decreases the optimal trading ratio (equation 3) with a subsequent increase in the firm’s demand for credits (equation 4).

3. Conclusion

Water quality impairment and wetland loss represent two of the most pressing environmental problems affecting watersheds in the United States. Since the implementation of water quality trading has been rather limited in practice, we need to determine how new policy can account for all relevant socio-economic, ecological and institutional factors. Establishing benchmark scenarios, like this paper and earlier work, plays an important role in informing policy. This research looks at the role of wetland creation and restoration in the context of water quality trading. Wetlands are generally considered effective agricultural and urban runoff sinks (e.g., Hey et al. 2005; Crumpton et al. 2008). However, their flow-regulating properties are often neglected in water quality policy. It was illustrated here, that even in the (hypothetical) case where no runoff is permanently retained in wetlands, these ecosystems can generate pollution credits. Runoff typically occurs in pulses that follow rainfall and storm events. By reallocating runoff over time and smoothing overall pollution profiles, wetlands reduce total environmental damages. Our model consists of a farm, referred as a pulse-pollution source, and a firm or non-pulse pollution source. The farm may supply pollution credits to the firm via wetland implementation. Our results show that the optimal trading ratio is greater than one for perfectly mixed pollutants. In particular, a unit of firm emissions (which are steady) should be offset by more than a proportional reduction of the initial runoff pulse. The former induces a stream of environmental damages whereas the latter reduces damages only during periods where rainfall or storm event occurs. Furthermore, management of the runoff pulse via wetland implementation
generates delayed environmental damages. Including the water treatment properties of wetlands reduces differences in environmental damages across sources and, naturally, decreases the optimal trading ratio.
References


