2001

Three essays on environmental incentives: dynamics, asymmetric information, and dual policy goals

Hongli Feng
Iowa State University

Three essays on environmental incentives: dynamics, asymmetric information, and dual policy goals

by

Hongli Feng

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Catherine L. Kling, Co-major Professor
Jinhua Zhao, Co-major Professor
John C. Beghin
Brent M. Hueth
Jean-Didier Opsomer

Iowa State University

Ames, Iowa

2001

UMI Number: 3034180

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Hongli Feng

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Major Program

To my grandma.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to express my deep appreciation to Dr. Catherine Kling and Dr. Jinhua Zhao for their guidance of my research efforts. Their knowledge of economic theory and advising skill have been particularly helpful. Discussions with them have always been inspiring, enlightening, and enjoyable.

I would like to express my most sincere gratitude to Dr. Brent Hueth for serving on my committee and being so generous with his time and ideas. His expertise on contract theory and his unwavering quest for quality research have greatly improved the fourth chapter of my dissertation.

I also wish to thank Dr. John Beghin and Dr. Jean-Didier Opsomer for serving on my dissertation committee.

Finally, I would like to thank my parents and my brother for their love and support. In particularly, I am thankful for their understanding, sacrifice, and encouragement throughout my years of formal education far away from home.

CHAPTER 1.

GENERAL INTRODUCTION

Improved scientific understanding of the natural world has led scientists to better comprehend the dynamic complexity and interrelationship of many environmental problems. A clear example of this evolution of understanding is the case of global warming and greenhouse gas emissions, where the best scientific consensus of the dynamic nature of the global warming problem and its ramifications is summarized in the third report of the International Panel on Climate Change in 2000.

As the scientific understanding evolves, so too does environmental policy. Specifically, environmental economists and policy analysts continue to design and assess efficient mechanisms that both address the inherent risks of environmental concerns, such as climate change, and are politically feasible. To accomplish these two objectives, economists need to understand and incorporate the complexities of the underlying science and policy concerns into the design and implementation of economic incentives.

This dissertation is comprised of three essays that address fundamental features of policy-relevant incentive design for complex environmental problems. In the first essay, I study the efficient design of policies for storing carbon (an important greenhouse gas) in agricultural soils. In the second essay, I consider the optimal intertemporal trading of emissions permits in the important case of asymmetric information and in the third, I focus on multiple policy goals, using the case of green payments as a case study.

Although in each essay/chapter, a separate problem is addressed, the unifying theme of this thesis is the importance of understanding and incorporating both the relevant scientific information about the fundamental aspects of an environmental problem and the goals of policy in designing incentives

In the second chapter of this dissertation, I investigate the design of efficient incentives for farmers to sequester carbon, that is, take carbon out of the atmosphere and store it in the soil. Carbon sequestration reduces the carbon stock in the atmosphere and thus may ameliorate global warming. The essay begins with an analysis of the dynamic path of carbon in the atmosphere and in the soil. The fact that carbon taken out of the atmosphere can later be released into the atmosphere almost costlessly plays a critical role in the design of mechanisms. In particular, in the absence of appropriate incentives farmers may release the carbon sequestered in their land, mechanisms must be carefully designed if farmers are to store carbon for the optimal length of time.

In the third chapter, I consider the design of optimal intertemporal permit trading systems under asymmetric information. Regardless of whether the Kyoto Protocol or any other international treaty is ratified, it is likely that a market-based trading approach will be a part of any U.S. policy considered for global warming. In addition, emissions permit trading has been a part of policies addressing other environmental problems such as acid rain. The importance and potential for intertemporal trading has not been well addressed in the environmental economics literature.

When there is asymmetric information, regulators have at least two choices to improve social welfare (relative to using standards or taxes): one is to give agents more flexibility, the other is to design mechanisms that properly separate agents with heterogeneous benefits or costs related to the externalities. When agents have more flexibility, social welfare might be improved since agents know more than the regulators and thus are capable of making better decisions. A well-known application of this concept is to allow firms to trade emission permits. Typically, regulators do not know firms' abatement costs and so it is difficult, if not impossible, for the regulators to set the right standard for each firm. Permit trading enables firms to equate the marginal costs of abatement among themselves. Otherwise, firms with low abatement costs may emit more than those with high abatement costs, which is not efficient.

However, there is a cost of allowing agents more flexibility if they disregard the externalities of their decisions. In the case of permit trading, if pollutants from each firm are uniformly mixed such that firms' emissions are perfect substitutes for each other, then trading among firms does not affect total damages of an externality. However, if one unit of emission from firm A produces more or less damage than that from firm B and firms trade permits at a one-to-one ratio, then total damages from emissions will differ from a no-trading situation, even though total emissions remain the same. In such a case, the flexibility of trading has two effects: it saves abatement costs for firms but it also may generate additional social damages. The efficient degree of flexibility in this case will depend on the trade-off. In the third chapter, I study this trade-off in the context of intertemporal permit trading.

The second approach to dealing with asymmetric information draws from the large literature on designing mechanisms to separate agents. The basic idea is to design incentive compatible mechanisms under which agents of a certain type find it optimal to behave as their own type. For example, different farmers may have different costs of providing environmental services such as adopting conservation tillage and reducing fertilizer usage. If the government is to purchase these services at at least cost, then farmers with low costs will be paid less than those with high costs. However, this is often infeasible either because the costs of providing such services are private information or because it is not "fair" to pay farmers differently for the same amount of services. Thus, proper mechanisms have to be used to induce farmers to reveal their true type by self-selection. The design of such mechanisms becomes more complicated if the government also wants to make positive net transfers (over and above farmers' conservation costs) to a certain group of farmers.

In the fourth chapter, this issue of asymmetric information and dual policy goals in the context of green payments, i.e, paying farmers for the provision of environmental goods is studied. In discussions of the next farm bill, interest in green payments has been high for two reasons. First, in recent years, the public has become increasingly aware of agriculture's multiple outputs. Agriculture not only produces food and fiber, it affects water quality, provides wildlife habitat, and may contribute to solving global warming. Paying farmers to provide these environmental goods can make economic sense. Second, green payments may be a vehicle for transferring income to the farm sector and thus replace traditional income support programs that distort production. The fourth chapter considers the efficient design of a green payments program in the presence of each of these complications.

The last chapter provides general conclusions and discusses possible future areas of research.

CHAPTER 2.

THE TIME PATH AND IMPLEMENTATION OF CARBON SEQUESTRATION

A paper accepted by American Journal of Agricultural Economics

Hongli Feng, Jinhua Zhao, Catherine Kling

Abstract

We develop a dynamic model to investigate the optimal time paths of carbon emissions, sequestration and the carbon stock. We show that carbon sinks should be utilized as early as possible, and carbon flow into sinks should last until the atmospheric carbon concentration is stabilized. We rule out any cyclical patterns of carbon sequestration and release. We propose and assess three mechanisms to efficiently introduce sequestration into a carbon permit trading market: a pay-as-you-go system, a variable-length-contract system and a carbon annuity account system. Although the three mechanisms may not be equally feasible to implement, they are all efficient.

Introduction

Under the Kyoto Protocol, industrialized countries have pledged to reduce their carbon emissions to below their 1990 emission levels over the period 2008-2012. To fulfil their commitment, some countries, including the U.S., have proposed the inclusion of three broad land management activities pursuant to Article 3.4 of the Protocol, including forest, cropland and grazing land management.[1] These activities can reduce atmospheric carbon stock by sequestering, or removing, carbon from the atmosphere and storing it in soil or biomass. For land rich countries like the U.S., Canada and Russia, carbon sequestration by these activities could potentially account for significant their emission reductions. For example, estimates indicate that the total carbon sequestration potential of U.S. cropland through improved management is 75-208 MMTC/year (Lal *et al*). Soil sinks, combined with forest sinks, could potentially be used by the U.S. to meet half of its emission reduction commitment (USDOS). However, skepticism remains among environmental groups who argue that "While preventing the emission of carbon dioxide is permanent, sequestering carbon pollution is a cheap, short-term fix that fails to address a long-term problem" (WWF).

The concerns raised by environmentalists and others relate specifically to the fact that sinks may be short run in nature and consequently, do not provide the same benefits as permanent emission reductions. This non-permanence issue is one of the focal points in post Kyoto negotiations on carbon sinks (IPCC, 2000a),[2] and disagreement over sinks was a major impediment to progress at the Sixth Conference of the Parties to the Framework Convention on Climate Change in Hague in November 2000 (IISD).[3]

At the heart of the debate lie two inter-related difficulties of carbon sinks due to the non-permanence feature. The first difficulty has to do with accounting and implementation. If sequestered carbon can be easily released, governments must find ways of properly accounting for the "net value" of possibly *temporary* storage, and design mechanisms to implement carbon sinks that correctly reflect this value. For example, if a permit trading system is devised for carbon abatement, which permanently reduces carbon released into the atmosphere, the system cannot be directly applied to carbon sequestration if the sequestered carbon is only temporarily kept out of the atmosphere. The second issue concerns when carbon sinks should be utilized. While the U.S., Canada and some other countries, arguing for cost effectiveness, prefer earlier inclusion, the EU has argued for later usage, stressing the key role of improving energy efficiency and shifting toward renewable energy resources. A related but deeper question is the optimal time path of carbon sequestration. Given that stored carbon can be easily released, thereby providing opportunities for future sequestration, the optimal time path may possibly have a cyclical pattern: sequestration and release, followed by sequestration and release, and so on.

In this paper, we develop a stylized model of carbon emissions (or abatement) and carbon sequestration to investigate the optimal time patterns of sequestration, emissions and carbon stock, and to propose three mechanisms that can efficiently implement carbon sinks in a permit trading system based on emissions and abatement. We show that a cyclical pattern is not optimal for soil sinks. In particular, both carbon emissions and stock are monotone in time: depending on the starting carbon stock level, they either increase or decrease through time monotonically. There are two possibilities for sequestration, depending again on the starting point. In one scenario, carbon is sequestered first and partially released later. In the more realistic scenario, carbon is continually being sequestered, although eventually, approaching the steady state, the scale of additional sequestration goes to zero. In both cases, we find that, if sinks are to be used at all, we should start to use them now.

We then propose three systems to implement the optimal sequestration levels: a pay-as-you-go (PAYG) system, a variable-length-contract (VLC) system, and a carbon annuity account (CAA) system. Each could be used in conjunction with a well functioning (emission reduction based) carbon permit trading system to efficiently include the sequestration of carbon. We show that each system can be efficient, but requires different conditions to be so. Further, the systems are likely to differ in the transaction costs associated with their implementation. Consequently, one or more may be desirable in practice and under different circumstances. These systems also indicate the proper way of accounting for the value of (possibly temporary) sequestration.

There are two studies in the literature on the optimal time patterns of carbon sequestration and emission. Van Kooten *et al.* investigate optimal carbon sequestration for an exogenously given time path of emissions. Richards studies the optimal emission levels and carbon stock without explicitly introducing sequestration as one of the control variables. Our paper extends this literature by modeling emissions and sequestration simultaneously, and studies an optimal control problem of two state and two control variables. These complications are important since abatement and sequestration are two ways of reducing the carbon concentration in the atmosphere, and their optimal time paths are bound to be inter-dependent.

A few studies have discussed various aspects of the implementation of carbon sinks. Recognizing the difference between abatement and sequestration, and between sequestration projects, Fearnside and Chomitz advocate a "ton-years" accounting method, which distinguishes between, say, one ton of carbon sequestered for one year and the same amount sequestered for five years. McCarl and Schneider, and Marland, McCarl and

Schneider, discussing soil carbon sinks, suggest that incentive programs for sequestration have to address the issue of "preservation of gains over time" or "longevity of agricultural carbon." Marland, McCarl, and Schneider. also argued that if sequestered carbon becomes a commodity, then credits could be issued for carbon sequestered but there must be subsequent debits if the carbon is later released. To our knowledge, our paper presents the first systematic study of the efficient implementation of carbon sinks that formally accounts for the non-permanence of sinks.

Throughout the paper, we use the term "abatement" to refer to reductions in carbon loadings and "sequestration" to mean the storage of carbon in soils or terrestrial biosphere in general. Thus, abatement by its nature is permanent. If a ton of carbon is not produced and emitted into the atmosphere today, it will not be present in the atmosphere at a later date. In contrast, a ton of carbon stored in a sink today may be only temporarily out of the atmosphere as it might be released in a future period.

An important issue that we do not address, but nevertheless provides the justification for this study, is the cost effectiveness of carbon sequestration compared with abatement. Stavins, reviewing a large body of the existing studies and providing his own analysis, argues that growing trees to sequester carbon has lower marginal costs than emission abatement for a considerable range of stored carbon. Antle *et al.*, Pautsch *et al.*, and Mitchell *et al.* assess the cost and potential of carbon sequestration by changing management practices within the agricultural sector. Although their results vary for different practices, all show there is economic potential for soil carbon sequestration.

Model Setup

Consider the social planner's problem of maximizing the benefits of carbon emissions minus the cost of sequestration and the damage caused by global warming. Let $e(t)$ be the society's emission rate at time t and $B(e(t))$ the benefits of emissions, with $B(0) = 0$, $B'(\cdot) > 0$, $B''(\cdot) < 0$, and $\lim_{e \to 0} B'(e) = \infty$. Higher emissions represent lower levels of abatement in the economy, and the benefits are equivalent to the saved abatement costs. The monotonicity and concavity of $B(\cdot)$ then is a result of the monotonicity and convexity of the abatement cost function (in the level of abatement)[4]. The marginal abatement cost approaches infinity if all of the society's emissions are to be abated, leading to the last condition on $B(\cdot)$.[5]

The emitted carbon accumulates in the atmosphere causing global warming effects. Let $C(t)$ be the total carbon stock in the atmosphere. The pollution damage of the carbon stock is $D(C(t))$ with $D'(\cdot) > 0$ and $D''(\cdot) > 0$.

We assume that the carbon stock $C(t)$ decays at an exponential rate $\delta \geq 0$. By decay we mean the process by which atmospheric carbon is "sunk" into the ocean. There is a constant process of carbon flow between the atmosphere and the ocean, the direction and speed of which depends on the temperature and carbon concentration in both media. Typically carbon flow is not exponential. Our assumption simplifies the model and captures the notion that carbon flow into the ocean increases as the carbon stock rises in the atmosphere. The assumption also indicates that carbon is not a pure stock pollutant, and theoretically, its concentration level *can* go down, and thus can be stabilized (through sequestration and reduced emissions).

Let $A(t)$ be the total units of land that are enrolled in carbon sequestration programs at time t. To simplify notation, we define one unit of land as the acreage that is needed to sequester one ton of carbon. For example, if one acre of land can sequester α tons of carbon, one unit of land is equal to $1/\alpha$ acres. Let $Q(A(t))$ be the cost of enrolling $A(t)$ with $Q'(\cdot) > 0$, $Q(0) = 0$ and $Q''(\cdot) > 0$. The cost of carbon sequestration can be interpreted in two ways. If sequestration requires changing agricultural production practices, the cost may be the agricultural profit foregone for doing so. For example, switching from conventional to conservation tillage may reduce a farmer's profit (Pautsch et al., Antle and Mooney), and some amount of profit may also be lost if cropland is converted to forestland (Plantinga, Mauldin, and Miller). In the case of improved management of an existing forest stand, the cost of carbon sequestration is the expenditure incurred to enhance management, e.g., fertilization (Hoen and Solberg, and Boscolo Buongiorno, and Panayotou).

The cost function $Q(\cdot)$ can be convex for a variety of reasons. Different land may incur different sequestration costs: some highly productive land is best kept in conventional tillage and some land can be converted to forest without much economic loss. Typically, land with low sequestration cost is converted first. As the land area A increases, the cost $Q(A)$ will increase at a faster rate when land of higher sequestration cost is converted.[6] Let $\bar{A}$ be the total land units. We assume that $\lim_{A \to \bar{A}} Q(A) = \infty$, implying that all land will never be converted.

Let $a(t)$ be the units of land newly enrolled ($a(t) > 0$) or withdrawn ($a(t) < 0$) in period t. For simplicity, we assume that when land is newly enrolled, carbon is *immediately* removed from the atmosphere, up to its full capacity (of one ton per unit). Likewise, all of the stored carbon is completely and *immediately* released if the land is converted back to its original use. In truth, soil carbon sequestration is a gradual process, and it may take up to fifty years for certain soil to reach its full sequestration capacity. Our assumption simplifies the model, and incorporates a key feature of sequestration: a piece of land can only hold a certain amount of carbon, all of which could be released back to the atmosphere during a very short period. To capture in a simple way the fact that there are costs (or physical limits) of converting land, we place bounds on the amount of land that can be converted each period: $\underline{a} \leq a(t) \leq \bar{a}$, with $\underline{a} < 0$ and $\bar{a} > 0$.

The equations of motion for $C(t)$ and $A(t)$ are

$$\dot{C}(t) = e(t) - a(t) - \delta C(t), \quad C(0) = C_0 > 0, \tag{1}$$

$$\dot{A}(t) = a(t), \quad A(0) = A_0 \geq 0, \quad 0 \leq A(t) \leq \bar{A}, \quad \underline{a} \leq a(t) \leq \bar{a}. \tag{2}$$

Equation (1) indicates that the change in the stock of carbon each period equals new emissions less the amount sequestered and the amount of natural decay. Let r be the social discount rate. Then the social planner's net payoff function is

$$V^0(A, C, e, a) = \int_0^\infty e^{-rt}[B(e(t)) - D(C(t)) - Q(A(t))]dt. \tag{3}$$

Maximizing (3) subject to (1) and (2) yields the optimal carbon sequestration and emission levels over time.

Optimal Paths of Sequestration and Emissions

Since $\lim\limits_{A \to \bar{A}} Q(A) = \infty$, the constraint $A(t) \leq \bar{A}$ is never binding along the optimal path.

So is the constraint $A(t) \geq 0$, as we will show later (Remark 3). This observation is intuitive: since the marginal cost of sequestration $Q'(A)$ is low when A is close to zero, and is typically lower than the marginal cost of emission reduction $B'(e)$, it makes economic sense to use some sinks to store a positive amount of carbon.

The current value Hamiltonian for the social planner's problem is

$$\begin{aligned} H(C, A, e, a, \lambda, \mu) &= B(e(t)) - D(C(t)) - Q(A(t)) \\ &\quad -\lambda(t)\,[e(t) - a(t) - \delta C(t)] - \mu(t)a(t), \end{aligned} \tag{4}$$

where $\lambda(t)$ and $\mu(t)$ are the negative of the costate variables which are continuously differentiable, and are assumed to be twice continuously differentiable almost everywhere. The necessary conditions are

$$\frac{\partial H}{\partial e} = -\lambda(t) + B'(e(t)) = 0, \quad \text{or} \quad \lambda(t) = B'(e(t)), \tag{5}$$

$$\max_a H \quad \text{or} \quad a(t) \begin{cases} = \bar{a} \\ = \underline{a} \\ \in [\underline{a}, \bar{a}] \end{cases} \quad \text{if} \quad \lambda(t) - \mu(t) \begin{cases} > 0, \\ < 0, \\ = 0, \end{cases} \tag{6}$$

$$\dot{\lambda}(t) = r\lambda(t) + \frac{\partial H}{\partial C} = (r + \delta)\lambda(t) - D'(C(t)), \tag{7}$$

$$\dot{\mu}(t) = r\mu(t) + \frac{\partial H}{\partial A} = r\mu(t) - Q'(A(t)), \tag{8}$$

$$\lim_{t \to \infty} e^{-rt}\lambda(t) = 0, \qquad \lim_{t \to \infty} e^{-rt}\mu(t) = 0. \tag{9}$$

From (7) and (8), we know

$$\lambda(t) = \int_t^{\infty} e^{-(r+\delta)(s-t)} D'(C(s))ds, \tag{10a}$$

$$\mu(t) = \int_t^{\infty} e^{-r(s-t)} Q'(A(s))ds. \tag{10b}$$

Thus $\lambda(t)$ measures the total discounted (to period t) future damages caused by one more unit of atmospheric carbon in period t. Notice the symmetric role played by the natural decay rate, δ, and the social discount rate, r. Equation (5) simply says that the marginal benefit of emitting one unit of carbon must equal its marginal cost. Similarly, $\mu(t)$ measures the discounted costs of maintaining one unit of land in sequestration, and (6) indicates that land should be converted at its maximum speed whenever the benefit of land conversion $\lambda(t)$ is different from the cost $\mu(t)$. This feature of land conversion is due to the linearity of the Hamiltonian in $a(t)$.

The Steady State

We assume that a steady state exists (thus the transversality conditions in (9) are naturally satisfied). Later we will show that the steady state is a saddle point. Setting $\dot{C} = 0$, $\dot{A} = 0$, $\dot{\lambda} = 0$ and $\dot{\mu} = 0$, from the conditions in (7) - (8) and the two state equations (1) - (2), we obtain the following description of the steady state:

$$\begin{aligned}(11)\qquad &(i)\quad e^* = \delta C^*, \quad (ii)\quad B'(e^*) = \frac{D'(C^*)}{r+\delta}, \quad (iii)\quad \lambda^* = \frac{D'(C^*)}{r+\delta}, \\ &(iv)\quad \mu^* = \lambda^* = \frac{Q'(A^*)}{r}, \quad \text{and} \quad (v)\quad a^* = 0.\end{aligned}$$

The steady state levels of emission and stock e^* and C^* are uniquely determined by (11-i) and (11-ii), and are *independent* of the sequestration activities. In particular, they are independent of the cost of sequestration $Q(\cdot)$. That is, once the steady state emission and stock levels are attained, there is no role for additional carbon sequestration activities. In the very long run, emissions have to be kept at such a level that it is just offset by the reduction in carbon stock due to the natural decay. Thus, in setting the *targets* for the long-run control of global warming, the government only needs to consider the costs and benefits of emission abatement and atmospheric carbon concentration. The option of carbon sequestration should not matter.

However, from (11-iii) - (11-iv), we know $\lambda^* > 0$ and $A^* > 0$. Thus, a certain amount of carbon in fact is sequestered in the biomass in the steady state. This amount is higher if the marginal cost of sequestration $Q'(\cdot)$ is lower, or as the sequestration becomes more effective. Then the positive stock A^* must be the result of using sequestration during the transition path toward the steady state. That is, sequestration does affect the *process* of reaching the long-run targets. We will investigate this process in the rest of this section. In summary,

Remark 1 *The long-run targets of controlling global warming are independent of the sequestration possibilities. That is, carbon sequestration cannot efficiently provide a long-run solution to global warming on its own. However, it may be efficiently employed "in the process," resulting in a permanent level of sequestered carbon.*

The Transition Paths

Analyzing the transition paths of the system and the stability of the steady state is complicated because there are two state variables and the problem is singular in one of the control variables. In principle we have to study a system of four differential equations, and the Eigenvalues are difficult to characterize without assuming special functional forms for the cost and benefit functions. We develop an alternative method to analyze the system, and partly rely on the (quasi) phase diagrams in the space of $e(t)$

and $C(t)$, shown in Figure 1. Setting $\dot{C} = 0$ in (1), we get

$$e(t) = a(t) + \delta C(t). \tag{12}$$

Thus the $\dot{C} = 0$ locus is linear and upward sloping, and its location depends on the value of $a(t)$. This locus is shown for three constant levels of $a(t)$, $\bar{a}$, 0, and $\underline{a}$, in Figure 1.

To derive the equation of motion for $e(t)$, we differentiate both sides of (5), and get

$$\dot{\lambda}(t) = B''(e(t))\dot{e}(t). \tag{13}$$

Plugging (5) and (13) into (7) and rearranging, we know

$$\dot{e}(t) = \frac{r+\delta}{-B''(e(t))}\left[\frac{D'(C(t))}{r+\delta} - B'(e(t))\right]. \tag{14}$$

Setting $\dot{e}(t) = 0$ leads to

$$\frac{D'(C(t))}{r+\delta} = B'(e(t)). \tag{15}$$

Finally, totally differentiating equation (15), we obtain

$$\frac{de}{dC} = \frac{D''(C(t))}{(r+\delta)B''(e(t))} < 0. \tag{16}$$

Thus the $\dot{e} = 0$ locus is downward sloping and is independent of the level of $a(t)$. In Figure 1, from (13), we know $\dot{e}(t) < 0$ when $\{C(t), e(t)\}$ is to the left of the $\dot{e} = 0$ locus, and $\dot{e}(t) > 0$ when $\{C(t), e(t)\}$ is to the right, independent of the value of $a(t)$.

Figure 1 Here

The dotted lines with arrows in Figure 1 represent the stable and unstable branches of the system for *fixed* levels of a. It is obvious that when there is no or constant rate of sequestration, the system of $e(t)$ and $C(t)$ should approach the steady state on a saddle path.

The derivation of the optimal paths (characterized in Propositions 1-5) when $a(t)$ is endogenous is given in . Here we only present the results and discuss the intuition. The rate of sequestration $a(t)$ is determined by (6). The system is thus on a singular path when $\lambda(t) = \mu(t)$. This path is important because it contains the steady state, which is associated with $\underline{a} < a(t) = 0 < \bar{a}$.

Intuitively, once the system $\{C(t), A(t), e(t), a(t)\}$ reaches the singular path, it should stay on the path until the steady state is reached in the limit. The reason is that before the singular path is reached, the marginal benefit and cost of sequestration $\lambda(t)$ and $\mu(t)$ are not equal. Thus land should be converted (either into or out of the sequestration programs) at the maximum rate in order to mitigate the inequality. Once they are equal, the social planner has no incentive to break the equality because any future inequality can be avoided by adjusting the conversion rate $a(t)$ now while on the singular path to improve welfare.

This intuition suggests that the transition path resembles a most rapid approach path, except that the system "quickly" approaches the singular path, rather than the steady state. Before the singular path is reached, $a(t)$ equals either $\bar{a}$ or $\underline{a}$. The values of $e(t)$ and $C(t)$ are then determined jointly by (1) and (14) with appropriate boundary conditions (discussed later). The following Proposition shows that our intuition is indeed correct.

Proposition 1 *Given the starting point* $\{C_0, A_0\}$, *the optimal path will move to the singular path as soon as possible, by setting* $a(t)$ *to be either* $\bar{a}$ *or* $\underline{a}$ *and choosing* $e(t)$ *accordingly. The system will then stay on the singular path forever, approaching the steady state.*

The proposition describes what happens before the singular path is reached, we next characterize the features of the singular path.

Proposition 2 *The land conversion rate on the singular path is given by*

$$a(t) = \frac{-\delta(r+\delta)B'(e(t)) + \delta D'(C(t)) + D''(C(t))(e(t) - \delta C(t))}{Q''(A(t)) + D''(C(t))}. \tag{17}$$

As emissions, carbon stock and land stock $A(t)$ change overtime, the value of $a(t)$ is likely to change as well. In fact, it is generic that $a(t)$ is not a constant on any time interval on the singular path. This observation highlights the limitations of studying carbon emissions assuming fixed levels of sequestration.

Next we show that the emission and carbon levels are monotonic.

Proposition 3 *Along the singular path, both* $e(t)$ *and* $C(t)$ *are monotone. That is, except at the steady state, neither* $\dot{e}(t)$ *nor* $\dot{C}(t)$ *can be zero.*

The Proposition says that, if the steady state carbon stock level is lower (higher) than the starting level, then the carbon stock will increase (decrease) steadily through time. Since the damage function $D(\cdot)$ of the stock is convex, the planner has an incentive to "smooth out" $C(t)$ overtime and avoid cyclical variations such as raising the stock for a while only to abate it later. Similarly, the benefit function of $B(e)$ is concave, and emission smoothing implies monotone $e(t)$. The Proposition implies that depending on the starting point, the transition path can only stay on one side of the isocline $\dot{e} = 0$ in Figure 1 (recall that $\dot{e}(t)$ is a function of $e(t)$ and $C(t)$ only). The next Proposition shows that, under a certain condition, $A(t)$ is monotone as well, i.e., $a(t)$ does not change signs on the singular path.

Proposition 4 *$A(t)$ is monotone along the singular path if and only if*

$$\ddot{\lambda}(t) / \dot{\lambda}(t) < r. \tag{18}$$

When (18) is satisfied, if the singular path approaches the steady state from the left of the isocline $\dot{e} = 0$ in Figure 1, $a(t) > 0$ and $A(t)$ monotonically rises. If the path approaches the steady state from the right, $a(t) < 0$ and $A(t)$ monotonically decreases.

Condition (18) requires that the marginal emission damage $\lambda(t)$ is not "too convex" in time. For example, suppose the system approaches the steady state from the left so that $\dot{\lambda}(t) > 0$. Then condition (18) requires that either $\dot{\lambda}(t)$ is decreasing over time, i.e., $\lambda(t)$ is increasing but concave in t, or is not increasing at a rate higher than the discount rate r. In other words, eventually the rate of increase of $\lambda(t)$ cannot be too high.

From (5), we know $r\dot{\lambda} - \ddot{\lambda} = r\dot{e}B'' - \ddot{e}B'' - (\dot{e})^2 B'''$. Again, consider the system to the left of the steady state. Then sufficient conditions for (18) are $\ddot{e} \geq 0$ and $B''' \leq 0$, or the absolute values of these two variables are low regardless of their signs. The following Proposition presents two additional sufficient conditions.

Proposition 5 *(i) The condition (18) is satisfied when the system is sufficiently close* *to the steady state. (ii) It is also satisfied if the rate of increase of the marginal damage of atmospheric carbon,* $dD'(C(t))/dt = D''(C)\dot{C}$, *is constant or decreasing overtime.*

Sudden and drastic reductions in the emission level may occur at time zero as the system moves to the optimal trajectory. Afterwards, emissions tend to stabilize toward the steady state level. That is, the absolute value of $\ddot{e}$ tends to be small. Further, as the system moves toward the steady state, carbon stock will slowly approach its steady state level, or $\dot{C}$, and $D''\dot{C}$, will be decreasing. Thus, in general, we expect that (18) is satisfied. We assume this is the case in the paper.[7] Propositions 3 and 4 rule out any cyclical patterns in the transition path and any spiral (the steady state being a spiral point) or orbital stability (such as limit cycles). In fact, since there is a unique singular path passing through the steady state, the steady state must be a saddle point.

Remark 2 *The optimal emission level and atmospheric carbon concentration are not cyclical: they should monotonically increase or decrease overtime. Further, under a rather general condition, the optimal path does not involve any cyclical patterns of carbon sequestration, or repeated sequestration and release activities.*

Proposition 4 and Remark 2 imply that the constraint $A(t) \geq 0$ is never binding. Since A_0 is low, $a(t) = \bar{a}$ before the singular path is reached. Thus, if the system approaches the steady state from the left of $\dot{e} = 0$, $a(t) > 0$ for all t, and thus $A(t) > 0$ for all t. If the system starts from the right of $\dot{e} = 0$, carbon is sequestered first, and since $A^* > 0$, only part of it is released later. Thus $A(t) > 0$ for all t. Therefore,

Remark 3 *On the optimal trajectory,* $A(t) > 0$ *for all* $t \geq 0$.

Effects of Carbon Sequestration

To completely characterize the paths of carbon emissions and sequestration, and to evaluate the effects of the availability of sinks on the optimal emission and stock, we need to specify the starting point, or the levels of $\{C_0, A_0, e(0), a(0)\}$, in particular their relative positions to the steady state. It is safe to assume that $A_0 < A^*$: since no mechanism exists to encourage carbon sequestration yet, the current use of carbon sinks is likely below the socially optimal long run level. In addition, as discussed in the introduction, the marginal cost of carbon sequestration is low (close to zero) if only a small amount of carbon is sequestered. We assume a low current rate of land conversion $a(0)$. This rate may even be negative given widespread deforestation in many parts of the world.

We refer to the recent IPCC reports to specify C_0 and $e(0)$.[8] IPCC (2000b, 2001) projects the atmospheric CO_2 concentrations by year 2100 to be about $540-970$ ppm for a wide range of emission scenarios. In contrast, the current CO_2 concentration is about 360 ppm. In these scenarios, IPCC lists 450, 650 and 1000 ppm as possible alternative targets of CO_2 concentration levels in the long run. These numbers seem to indicate that $C_0 < C^*$.[9] The IPCC (2001) further noted that to reach the three targets, CO_2 emissions have to "drop below the 1990 levels within a few decades, about a century, or about two centuries, respectively, and continue to decrease steadily thereafter. Eventually, CO_2 emissions have to decline to a very small fraction of current emissions." We therefore assume that $e_0 > e^*$, and that e_0 is above the optimal emission level given C_0 and A_0.

With our specification of the starting condition, the optimal transition path is represented by the heavy solid line in Figure 1, with the arrows indicating the direction of movement. Before the singular path is reached at time t_1, $a(t) = \bar{a}$, and the motion of the system is dictated by the locus of $\dot{e} = 0$ and $\dot{C} = 0$ for $a = \bar{a}$. To guarantee that the path reaches the singular path, $e(0^+)$ must fall below the stationary arm of the steady state associated with $a = \bar{a}$. Along the entire transition path, $e(t)$ decreases and $C(t)$ increases.

Land is converted at its maximum rate $\bar{a}$ before t_1, and is converted at a lower, but positive rate afterwards. Sequestered carbon is never released. Sequestration will be utilized as early and as extensively as possible. The intuition for this is as follows. At the beginning, the marginal cost of sequestration is low, lower than the marginal damage of an additional unit of carbon in the atmosphere, thus it makes economic sense to reduce (or eliminate) the difference between the marginal cost and marginal benefit of sequestration. The sooner this is done the better. However, the amount of carbon that can be sequestered at any point of time is constrained. So the best we can do is to sequester the maximum amount of carbon that can be sequestered to bridge the difference between the marginal cost and benefit of sequestration.[10] Of course, early use of sinks should also be accompanied by (possibly drastic) emissions reduction. After the system reaches the singular path, $a(t)$ should be set so as to maintain the equality of marginal cost and benefit of sequestration, as the carbon stock approaches its steady state level.

To further study the effects of carbon sinks on the optimal emission levels, consider the optimal emission trajectory when sinks are not available, or when $a(t) = 0 \, \forall t$, denoted by $\tilde{e}(t)$. At any stock level C, the fact that $a(t) > 0$ when sinks are available indicates that $\lambda(C) < \tilde{\lambda}(C)$, where $\tilde{\lambda}(C)$ is the marginal damage of emissions without the sinks.[11]

The reason is that sequestration offers an additional way of reducing the carbon stock, thereby reducing the marginal damage of emissions. Since $B(\cdot)$ is concave, we know from (5) that $e(C) > \tilde{e}(C)$ for all $C < C^*$. Figure 1 shows the relative positions of the two paths: the optimal trajectory with sinks lies strictly above that without sinks, before reaching the steady state. In summary,

Remark 4 *(i) Sequestration should be utilized as early as possible, accompanied by a reduction of the emissions. (ii) The availability of carbon sinks raises the optimal emissions, or decreases the degree of emission reduction that is needed to reach the steady state level of carbon stocks.*

The Remark further shows the role of carbon sequestration: sinks only affect the *processes*, but not the steady state levels, of carbon emission and stock. This result, of course, is consistent with the steady state analysis in Remark 1.

Implementation Mechanisms of Carbon Sinks

We have shown above that sequestration can be used to reduce the pressure on emission abatement. In this section, we propose and assess three distinct trading mechanisms, each of which can implement the socially optimal level of carbon sequestration. We refer to the three mechanisms as Pay-As-You-Go (PAYG), Variable Length Contract (VLC), and Carbon Annuity Account (CAA). All three mechanisms are designed to be implemented within a well functioning permit market for carbon emission reductions. Thus, we assume there is a carbon permit trading system, and that the permit price in the system is efficient: $P(t) = B'(e(t)) = \lambda(t)$. We analyze how trade between sources and sinks can take place efficiently, yielding the optimal amount of sequestration. We also discuss some of the potential advantages and drawbacks of the three mechanisms in terms of ease of implementation. Throughout this discussion, one "carbon credit" means a unit of carbon that is permanently removed from the atmosphere and the carbon price is the payment for one full carbon credit.

The Pay-As-You-Go (PAYG) System

In a PAYG system, owners of sinks sell (and repurchase) emission credits based simply on the permanent reduction of carbon. For example, in the first year, a farmer who adopts conservation tillage practices on 100 acres may earn 200 permanent carbon reduction credits which he can then sell at the going rate. If, in the fifth year, the farmer plows the field and releases all of his stored carbon, he would be required to purchase carbon credits from the market at the going price to cover his emissions.

In a world of certainty, the price trajectory $P(t)$ is known. Suppose there is perfect competition in the sink credit market. Then the competitive solution is equivalent to the problem of maximizing the present discounted revenue from carbon sequestration, $a(t)P(t)$, minus the sequestration cost $Q(A(t))$. Mathematically, the problem can be written as,

$$\text{(19)} \qquad \underset{a(t)}{Max} \int_0^\infty [P(t)a(t) - Q(A(t))]e^{-rt}dt$$

$$\text{s.t.} \quad \dot{A}(t) = a(t), \quad 0 \leq A(t) \leq \bar{A}, \quad \underline{a} \leq a(t) \leq \bar{a}.$$

As in the last section, we first ignore the constraint $A(t) \geq 0$, and derive the optimality conditions. We show that these conditions replicate the social planner's problem. Then by Remark 3, we know the constraint is not binding. Thus in the balance of this section, we will ignore the constraint $A(t) \geq 0$.

The Hamiltonian is $H^1 = P(t)a(t) - Q(A(t)) - \mu(t)a(t)$, and the first order necessary conditions are

$$\text{(20)} \qquad \max_a H^1 \quad \text{or} \quad a(t) \begin{cases} = \bar{a} \\ = \underline{a} \\ \in [\underline{a}, \bar{a}] \end{cases} \quad \text{if} \quad P(t) - \mu(t) \quad \begin{cases} > 0, \\ < 0, \\ = 0, \end{cases}$$

$$\dot{\mu}(t) = r\mu(t) - \frac{\partial H^1}{\partial A} = r\mu(t) - Q'(A(t)),$$

$$\lim_{t \to \infty} e^{-rt}\mu(t) = 0.$$

The first order conditions are the same as (6), (8), and the transversality condition for μ in (9). Together with the efficient permit price $P(t) = \lambda(t)$, these conditions exactly replicate the social planner's choice of sequestration together with abatement. Therefore, given that the permit price equals the present discounted value of marginal damage, the PAYG is efficient. Given the obvious practical difficulties (i.e., the obligation of purchasing credits upon release, intentional or accidental) of implementing and enforcing such a system, we present the efficiency results in large part as a basis of comparison for the following two systems.

The Variable Length Contract (VLC) System

The VLC system might evolve through independent broker arrangements. If a broker wishes to buy permits from sink sources and sell them to emitters, the broker must contract with sink sources to achieve a permanent reduction in carbon. This could be accomplished by making a contract with one farmer to adopt conservation tillage for, say 3 years before plowing the field, contracting with a second farmer to plant trees beginning in year 4 for a certain number of years and so on. In each period, the broker might offer farmers a menu of prices associated with different contract lengths. In this system, private brokers provide the service of generating "permanent" carbon reductions from a series of separate temporary reductions.

Formally, suppose that a broker offers farmers a menu of prices for different contract lengths in each period. Let $q(t,\tau)$ be the price offered at time t for a contract with length τ. Then given this price menu, a farmer's decision is to maximize the net gain from carbon sequestration by choosing units of land for contracts of different lengths. Let $a(t,\tau)$ be the units of land enrolled at time t for a contract length of τ periods. The farmer's problem is

$$\text{(21)} \qquad \underset{a(t,\tau)}{Max} \int_0^\infty e^{-rt} \left[\int_0^\infty q(t,\tau)a(t,\tau)d\tau - Q(A(t)) \right] dt$$

$$\text{s.t.} \quad \dot{A}(t) = \int_0^\infty a(t,\tau)d\tau - \int_0^t a(t-\tau,\tau)d\tau, \quad 0 \le A(t) \le \bar{A},$$

$$\underline{a} \le \dot{A}(t) \le \bar{a},$$

where $\int_0^\infty [q(t,\tau)a(t,\tau)]\, d\tau$ is the sum of total revenue at time t from contracts of all lengths; $\int_0^\infty a(t,\tau)d\tau$ is the total units at time t of newly enrolled land under contracts of all lengths, and, $\int_0^t a(t-\tau,\tau)d\tau$ is the total unit of contracts expiring at time t.

Proposition 6 *The VLC system is efficient if*

$$\text{(22)} \qquad q(t,\tau) = P(t) - e^{-r\tau}P(t+\tau).$$

The proof is given in . The condition in (22) is intuitive: for the VLC to be efficient, the price paid to the temporary storage should equal the difference between the damage that is reduced when carbon "flows into" the sinks and the added (discounted) damage when carbon is released into the atmosphere.

The condition in (22) will always be satisfied if there is no arbitrage in the trading of VLCs and emission permits. To see this, suppose a certain contract, say $\tilde{q}(t,\tau)$, is offered that is different from (22), and without loss of generality, suppose $\tilde{q}(t,\tau) > q(t,\tau)$. Then a broker can earn strictly positive profits by buying at time t an emission permit at $P(t)$, selling at t a VLC for the length of τ at $\tilde{q}(t,\tau)$, and selling at $t+\tau$ the emission permit at $P(t+\tau)$. The strategy clearly covers the broker's position: at each moment, the broker's balance of net emission is zero. However, the broker's loss in buying and selling the emission permit, $-P(t) + e^{-r\tau}P(t+\tau) = -q(t,\tau)$, is more than covered by the gain in selling the VLC, $\tilde{q}(t,\tau)$, leading to the arbitrage opportunity.

Arbitrage opportunities are not likely to arise if the emission permit and VLC trading markets are perfectly competitive. For a global pollutant like carbon with countless emission sources, the emission permit market is likely to be competitive. The nature of the VLC market will depend on the geographical distribution of the sinks and the brokers. It can be competitive if multiple brokers operate in each geographical area of carbon sinks. Since the owners of the sinks (i.e. farmers) do not have to directly "pay out" when carbon is released, the VLC approach is likely to be more feasible to implement compared with the PAYG system.

The "ton-years" accounting method mentioned in the introduction section can be made equivalent to the VLC if the correct discount factor is used. According to the "ton-years" accounting method, the amount of carbon sequestered is directly discounted, while in the VLC system, the price of sequestration is discounted. In both methods, the "correct "discount factor (either for quantity or price), depends on the duration of sequestration, the discount rate for future damage and the natural decay rate of carbon.

The Carbon Annuity Account (CAA) System

Finally, a CAA system may be the most straightforward to implement of all three systems. Similar to the PAYG system, in a CAA system, the generator of a sink is paid the full value of the permanent reduction in the GHG's stored in the sink. However, CAA is also different from PAYG in that the payment, in stead of being paid to a farmer (or whoever sequesters carbon), is put directly into an annuity account. The payment deposited in the annuity account works as a "bond"—with the money in the account, the farmer is discouraged to release her stored carbon, and if she releases it, it is guaranteed that she will be able to pay at least partly for the released carbon. As long as the sink remains in place, the owner can access the earnings of the annuity account, but not the principal. The principal is reduced at the on-going permit price when and if the sink is removed (e.g. the soil is tilled or other change is made to release the stored carbon). If the sink remains permanently, the sink owner eventually earns all of the interest payments, the discounted present value of which equals the principal itself - the permanent permit price. We now show that a CAA system is efficient.

Let $M(t)$ be the balance in the CAA account. Then in each period, $M(t)r$ will be the farmer's revenue, and $Q(A(t))$ will be her cost. The farmer's objective is to maximize the present discounted value of net revenue.

(23)
$$\underset{a(t)}{Max} \int_0^\infty [M(t)r - Q(A(t))]e^{-rt}dt$$

$$\text{s.t.} \quad \dot{A}(t) = a(t), \quad 0 \leq A(t) \leq \bar{A}, \quad \underline{a} \leq a(t) \leq \bar{a},$$

$$\dot{M}(t) = a(t)P(t).$$

Let $\theta(t)$ be the costate variable for $M(t)$. Again, we first ignore the constraint $A(t) \geq 0$. Then the current value Hamiltonian is $H^2 = M(t)r - Q(A(t)) + \theta(t)a(t)P(t) - \mu(t)a(t)$, and the necessary conditions are,

(24)
$$\max_a H^2 \quad \text{or} \quad a(t) \begin{cases} = \bar{a} \\ = \underline{a} \\ \in [\underline{a}, \bar{a}] \end{cases} \quad \text{if} \quad \theta(t)P(t) - \mu(t) \quad \begin{cases} > 0, \\ < 0, \\ = 0, \end{cases}$$

(25)
$$\dot{\theta}(t) = r\theta(t) - \frac{\partial H^2}{\partial M} = r\theta(t) - r = r(\theta(t) - 1),$$

(26)
$$\dot{\mu}(t) = r\mu(t) - \frac{\partial H^2}{\partial A} = r\mu(t) - Q'(A(t)).$$

Rearranging (25), we know $\frac{d}{dt}[\theta(t) - 1] = r[\theta(t) - 1]$, which implies that $\theta(t) - 1 = e^{rt}[\theta(0) - 1]$. But since $\theta(0) = 1$, that is, the marginal value of money in period zero is equal to one, we know $\theta(t) = 1$ for all t. Then the necessary conditions are the same as those in the PAYG system. Thus the CAA system is efficient.

Discussion and Final Remarks

Resolving the permanence issue will be key to introducing carbon sequestration into the Kyoto Protocol or any other international agreement concerned with global warming. In this paper, we have addressed this issue directly with a model of carbon emissions and sequestration dynamics. Several valuable policy insights come directly from the framework. First, the view that carbon sequestration should not be used to address global warming is not warranted from a theoretical perspective. Ultimately, as long as there is less carbon in the air, it does not matter whether the reduction is done by sequestration or emission abatement. We showed that carbon sequestration should be used as early as possible (as long as it is ever efficient to use it) to reduce the pressure on emission abatement, and the carbon flow into sinks lasts until the atmospheric carbon concentration is stabilized. We also ruled out any cyclical patterns of carbon sequestration and release in the utilization of sinks.

The insights concerning the efficient and early use of sequestration shown in this paper are particularly interesting in light of the current policy forum about global warming. Some businesses and even some nations, including the U.S., are very reluctant to take actions to reduce carbon emissions. Sequestration can reduce the pressure on emission abatement in current periods, providing time to develop political support for and the technological capability to reduce carbon emissions.

However, despite the clear theoretical role for carbon sequestration, it is equally clear that it should not be treated the same as carbon emission reductions. Sequestration, by its nature, always has the potential to be temporary; consequently, it cannot be attributed the same value that emission reductions have if an efficient solution is to be obtained. The correct view is that sequestration has value, but the value is different from (and less than) the value of direct emission reduction. Therefore, special mechanisms should be used to address the difference. We define three such systems and demonstrate the efficiency properties of each of them.

To properly implement any of the three systems, we will need accurate approaches to measure the amount of carbon stored in sinks. Likewise, for carbon trading to occur between sinks and emission sources, all three systems need price information from outside the agricultural and forest sector. PAYG and CAA both require the current permit prices and VLC requires prices of temporary carbon storage for all lengths of duration. Note that there is nothing preventing the simultaneous use of all systems.

Given that all three systems can be demonstrated to yield the theoretically efficient solution, the choice between which, if any, of these systems to actually implement may largely depend on the costs involved of implementation as well as the general acceptability of the approach to all involved. On this score, we suspect that the repayment obligations inherent in the PAYG system will render it politically infeasible. The CAA system might be more appealing because it partially solves the repayment problem. The comparison of the systems also depends on how their efficiency will be altered when the carbon prices are not efficient. This issue is an interesting topic for future research.

Appendix A Derivation of the Optimal Transition Paths

We prove the propositions used in deriving the optimal transition paths.

Proof of proposition 1. We follow a similar approach used by Tsur and Zemel, who extend the Spence and Starrett methodology to more than one state variables.

Suppose the time path of the costate variable $\lambda(t)$ is given. Thus from the necessary conditions (5) and (7), we know $e(t)$ and $C(t)$ are completely determined by $\lambda(t)$. In particular, we write them as $e(\lambda(t))$ and $C(\lambda(t), \dot{\lambda}(t))$.

Then, by (3) and (4), we can rewrite the optimal control problem for $a(t)$ as

(27)

$$\max_{a} \int_{0}^{\infty} e^{-rt} \left[B(e(\lambda)) - D\left(C(\lambda(t), \dot{\lambda}(t))\right) - \lambda\left[e(\lambda) - a - \delta C(\lambda(t), \dot{\lambda}(t))\right] - Q(A)\right] dt$$

$$\text{s.t.} \qquad \dot{A} = a, \quad \underline{a} \leq a \leq \bar{a}.$$

We can check that the necessary conditions in (27) replicate the original necessary conditions for a in (6) and (8).

Replacing a by $\dot{A}$ in the objective function, and integrating $\int_0^\infty e^{-rt}\lambda\dot{A}dt$ by parts, we know (27) can be rewritten as

$$\max_{A} \int_{0}^{\infty} e^{-rt} \left[B(e(\lambda)) - \lambda e(\lambda) - D\left(C(\lambda(t), \dot{\lambda}(t))\right) + \lambda\delta C(\lambda(t), \dot{\lambda}(t))\right] dt.$$

$$+\int_{0}^{\infty} \left[-A(\delta\lambda - D'(C(\lambda, \dot{\lambda}))) - Q(A)\right] dt - \lambda(0)A_0 \tag{28}$$

Thus given $\lambda(t)$, the objective function depends only on A. We only need to choose $A(t)$ to maximize the second integrand in (28) for each time t, and the first order condition is

$$-\delta\lambda + D'(C) = Q'(A). \tag{29}$$

With $\lambda(t) = \mu(t)$ and $\dot{\lambda} = \dot{\mu}$ on the singular path, we know from (7) and (8) that condition (29) is satisfied on the singular path. Thus, the objective function in (28) is maximized on the singular path.

Since (28) depends only on A, we would want to move to the optimal path of A, or the singular path, as soon as possible. Thus the optimal solution involves choosing $a(t)$ to be either $\underline{a}$ or $\bar{a}$ until the singular path is reached. Once the arc is reached, the system stays on it forever. ■

Proof of proposition 2. From (7) and (8), we know $\ddot{\lambda} = (r+\delta)\dot{\lambda} - D''\dot{C}$ and $\ddot{\mu} = r\dot{\mu} - Q''a$. However, $\ddot{\lambda} = \ddot{\mu}$ and $\dot{\lambda} = \dot{\mu}$ on the singular path. Thus we know $Q''a = -\delta\left[(r+\delta)\lambda - D'\right] + D''(e - a - \delta C)$, which implies (17). ■

Proof of proposition 3. Part (i). We first prove that $e(t)$ is monotonic. Suppose this is not true. In particular suppose there exists a time $t < \infty$ such that $\dot{e}(t) = 0$. In the phase diagram, if the path crosses $\dot{e} = 0$ from the left, $e(t)$ must be convex in time: it first decrease and then increases. Similarly, if the path crosses $\dot{e} = 0$ from the right, $e(t)$ must first increase and then decrease. The two possibilities are described in Figure A1 by the four short arrowed curves.

Figure A1. Here

Therefore, if a path has ever crossed the $\dot{e} = 0$ curve, there are only two scenarios in which the path will approach the steady state. In the first case, the *last* crossing of $\dot{e} = 0$ is from the right, and thus the system approaches the steady state from the left. Since $\dot{e}(t) < 0$ before reaching the steady state, the last crossing must have occurred above the steady state along the $\dot{e} = 0$ line. This path is depicted as Path One in Figure A1. In the second scenario, the last crossing is from the left, and the path finally approaches the steady state from the right. Thus $e(t)$ is increasing after the last crossing, and the crossing point is below the steady state on the $\dot{e} = 0$ isocline.

Consider Path One. Since the $\dot{e} = 0$ isocline is downward slopping, the carbon stock $C(t_1)$ at the crossing time t_1 is lower than the steady state level C^*. Thus after t_1, $C(t)$ will eventually increase. However, immediately after t_1, $C(t)$ is decreasing since the path crossed the $\dot{e} = 0$ isocline from the right. Therefore, there exists a time $t_2 > t_1$ at which $\dot{C}(t_2) = 0$ Further, at this time, $\{C(t_2), e(t_2)\}$ is to the northwest of the steady state. This observation implies that $a(t_2) > 0$, since the $\dot{C} = 0$ isoclines are higher as a is higher.

Taking a time derivative of (7) and (8), and using $\lambda(t) = \mu(t)$, $\dot{\lambda} = \dot{\mu}$, and $\ddot{\lambda} = \ddot{\mu}$ on the singular path, we know $\delta\dot{\lambda}(t_2) = -Q''(A(t_1))a(t_2)$. However, since $\dot{e}(t_2) < 0$, we know $\dot{\lambda}(t_2) > 0$, violating the fact that $a(t_2) > 0$ (since $Q'' > 0$). Thus Path One never arises.

Similarly, we can show that Path Two should not arise either. Establishing the monotonicity of $e(t)$ on the singular path.

If there are infinite number of crossings so that there is not a *last* crossing, a point like $\{C(t_2), e(t_2)\}$ (or its counterpart to the right of $\dot{e} = 0$) still exists when the system is close to the steady state. Thus, our proof still carries through.

Part (ii): We next prove that $C(t)$ is monotone. Without loss of generality, consider the singular path that is to the left of the $\dot{e} = 0$ isocline so that $\dot{e}(t) < 0$. (we have just shown that the path should never cross the isocline). Suppose there exists a time $\infty > t_3 > 0$ such that $\dot{C}(t_3) = 0$. Differentiating (7) and (8) and adjusting, using $\lambda(t) = \mu(t)$, we know $\delta\dot{\lambda}(t_3) = -Q''a(t_3)$. But since $\dot{\lambda} = B''\dot{e} > 0$, we know $a(t_3) < 0$.

Figure A2. Here.

Therefore, in Figure 1, the system at t_3 must be at a point on the $\dot{C} = 0$ isocline for a negative a level. Since this isocline is upward slopping, we know this point must be below the steady state. The point is represented by x in Figure A2. Thus, to reach the steady state, $e(t)$ must eventually increase. But this condition contradicts the fact that we are to the left of the $\dot{e} = 0$ isocline, or $e(t)$ cannot increase. ■

Proof of proposition 4. From (8), we know $Q''a = r\dot{\mu} - \ddot{\mu}$. Substituting λ for μ along the singular path, we know $a = (r\dot{\lambda} - \ddot{\lambda})/Q''$. To the left of the steady state, $\dot{\lambda} > 0$. Thus $a > 0$ if and only if $r\dot{\lambda} > \ddot{\lambda}$, or $\ddot{\lambda}/\dot{\lambda} < r$. To the right of the steady state, $\dot{\lambda} < 0$. Then $a < 0$ if and only if $r\dot{\lambda} < \ddot{\lambda}$, or $\ddot{\lambda}/\dot{\lambda} < r$. ■

Proof of proposition 5. Without loss of generality, we consider the case when the optimal path is to the left of the $\dot{e} = 0$ isocline, or when $\dot{\lambda} > 0$.

Differentiating (7) with respect to time, we know $\ddot{\lambda}/\dot{\lambda} = r + \delta - D''(C)\dot{C}/\dot{\lambda}$. Thus (18) is true if and only if

(30) $$\delta\dot{\lambda} < dD'(C(t))/dt.$$

At the steady state, $(r + \delta)\lambda^* = D'(C^*)$ (cf. (11)). Before the steady state, $\dot{\lambda} > 0$, or $(r + \delta)\lambda(t) > D'(C(t))$. Therefore, when the system is sufficiently close to the steady state, the left hand side, $(r + \delta)\lambda(t)$, must be increasing at a smaller rate than the right hand side $D'(C(t))$. That is, $(r+\delta)\dot{\lambda}(t) < dD'(C(t))/dt$, which implies (30). This proves the first part of the Proposition.

Differentiating (10a) with respect to t, and integrating the resulting right hand side by parts, we can show that

$$\dot{\lambda} = \int_t^{\infty} e^{-(r+\delta)(s-t)} \frac{d[D'(C(s))]}{ds} dt. \tag{31}$$

If $\frac{d[D'(C(s))]}{ds} = D''(C)\dot{C}(s)$ is constant, then the right hand side equals $\frac{d[D'(C(t))]}{dt}/(r+\delta)$, which together with (31) implies (18). If $\frac{d[D'(C(s))]}{ds}$ is decreasing, then (31) implies $\dot{\lambda} < \frac{d[D'(C(t))]}{dt}/(r+\delta)$, and (18) follows from (30). ■

Appendix B Proof of Proposition 6

Proof. Define $\tilde{a}(t)$ as follows,

$$\tilde{a}(t) \equiv \dot{A}(t) = \int_0^{\infty} a(t,\tau)d\tau - \int_0^{t} a(t-\tau,\tau)d\tau. \tag{B1}$$

We show below Problem (21) is just the same as Problem (19) with $\tilde{a}(t)$ in place of $a(t)$. We have shown the solution to Problem (19) is efficient. If both problems are the same, then the solution to Problem (21) also has to be efficient. Plug (22) into the objective function of problem (21), we get

$\int_0^{\infty} e^{-rt} \left[\int_0^{\infty} q(t,\tau)a(t,\tau)d\tau - Q(A(t))\right] dt$

$= \int_0^{\infty} e^{-rt} \left[\int_0^{\infty} [P(t) - e^{-r\tau}P(t+\tau)]\, a(t,\tau)d\tau - Q(A(t))\right] dt$

$= -\int_0^{\infty} e^{-rt} Q(A(t))]dt + + \int_0^{\infty} e^{-rt} \left[\int_0^{\infty} (P(t)a(t,\tau)d\tau\right] dt$

$- \int_0^{\infty} e^{-rt} \left[\int_0^{\infty} e^{-r\tau}P(t+\tau))a(t,\tau)d\tau\right] dt$

$= -\int_0^{\infty} e^{-rt} Q(A(t))]dt$

$$+\int_0^\infty e^{-rt}\left[\int_0^\infty (P(t)a(t,\tau)d\tau\right] dt - \int_0^\infty e^{-rt}\left[\int_0^t P(t)a(t-\tau,\tau)d\tau\right] dt$$

$$= -\int_0^\infty e^{-rt}Q(A(t))]dt + \int_0^\infty e^{-rt}\left[\int_0^\infty (P(t)a(t,\tau)d\tau - \int_0^t P(t)a(t-\tau,\tau)d\tau\right] dt$$

$$= -\int_0^\infty e^{-rt}Q(A(t))]dt + \int_0^\infty e^{-rt}P(t)\tilde{a}(t)dt$$

It is easy to see that the above expression is just the same as that in Problem (19) with $a(t)$ replaced by $\tilde{a}(t)$.

The third line follows because of the following,

$$\int_0^\infty e^{-rt}\left[\int_0^\infty e^{-r\tau}P(t+\tau))a(t,\tau)d\tau\right] dt$$

$$= \int_0^\infty \int_0^\infty e^{-r(t+\tau)}P(t+\tau))a(t,\tau)dtd\tau \text{(By change of integration order)}$$

$$= \int_0^\infty \int_\tau^\infty (e^{-rx}P(x))a(x-\tau,\tau)dxd\tau \text{(By change of variable, } x = t+\tau, y=\tau)$$

$$= \int_0^\infty \int_0^x e^{-rx}P(x))a(x-\tau,\tau)d\tau dx \text{(By change of integration order)}$$

$$= \int_0^\infty e^{-rt}\left[\int_0^t P(t)a(t-\tau,\tau)d\tau\right] dt. \blacksquare$$

References

[1] Antle, J.M., and S. Mooney. "Economics and Policy Design for Carbon Sequestration in Agriculture", Research Discussion Paper No. 36, October 1999.

[2] Antle, J.M., S.M., Capalbo, S. Mooney, E. Elliot and K. Paustian. "Economics of Agricultural Soil Carbon Sequestration in the Northern Plains", Research Discussion Paper No. 38, October 1999, Revised March 2000.

[3] Boscolo, M., J. Buongiorno, and T. Panayotou. "Simulation options for carbon sequestration through improved management of a lowland tropical rainforest", *Environ. and Develop. Econ.* 2(1997): 241-263.

[4] Chomitz, K.M. "Baselines for Greenhouse Gas Reductions: Problems, Precedents, Solutions". Prepared for the Carbon Offsets Unit, World Bank, 16 July 1998.

[5] Fearnside, Philip M. "Monitoring Needs to Transform Amazonian Forest Maintenance into a Global Warming-Mitigation Option", *Mitigation and Adaptation Strategies for Global Change* 2 (1997): 285-302.

[6] Hoen, H.F., and B. Solberg "Potential and Economic Efficiency of Carbon Sequestration in Forest Biomass Through Silvicultural Management", *Forest Science,* vol. 40. no. 3 (August 1994): 429-451.

[7] International Institute for Sustainable Development (IISD). "Summary of the Sixth Conference of the Parties to the Framework Convention on Climate Change", *Earth Negotiations Bulletin,* Vol. 12, No. 163, Nov, 2000.

[8] Intergovernmental Panel on Climate Change (IPCC). "Summary for Policymakers—Land Use, Land-Use Change, and Forestry", May, 2000a.

[9] Intergovernmental Panel on Climate Change (IPCC). "Summary for Policymakers—Special Report on Emissions Scenarios", A special report of IPCC Working Group III, 2000b.

[10] Intergovernmental Panel on Climate Change (IPCC). "Summary for Policymakers—Climate Change 2001: The Scientific Basis", the Third Assessment Report, January, 2001.

[11] Lal, R., J.M. Kimble, R.F. Follett, and C.V. Cole. "The Potential of U.S. Cropland to Sequester Carbon and Mitigate the Greenhouse Effect", Ann Arbor, Michigan: Sleeping Bear Press, 1998.

[12] Marland, G., B.A. McCarl and U. Schneider. "Soil Carbon: Policy and Economics", *Climatic Change*, in press.

[13] McCarl, B.A. and U.A. Schneider. "U.S. Agriculture's Role in a Greenhouse Gas Emission Mitigation World: an Economic Perspective", *Rev. Agr. Econ.* 22 (spring/summer, 2000): 134-59.

[14] Mitchell, P.D., P.G. Lakshminarayan, T. Otake, and B.A. Babcock. "The Impact of Soil Conservation Policies on Carbon Sequestration in Agricultural Soils of the Central United States", in *Management of Carbon Sequestration in Soil.* R. Lal, J.M. Kimble, R.F. Follett, and B.A. Stewart. Boca Raton, FL: CRC Press, 1997, 125-42.

[15] Montgomery, W.D. "Markets in Licenses and Efficient Pollution Control Program," *J. Econ. Theory* 5(1972), 395-418 .

[16] Pautsch, G.R., L.A. Kurkalova, B. Babcock and C.L. Kling. "The Efficiency of Sequestering Carbon in Agricultural Soils", *Contemporay Policy Issues*, vol. 19, No. 2, April 2001, 123-134.

[17] Plantinga, A.J., T. Mauldin, and D.J. Miller. "An Econometric Analysis of the Costs of Sequestering Carbon in Forests", *Amer. J. Agr. Econ. 81 (November 1999): 812-824.*

[18] Richards, K.R. "The Time Value of Carbon in Bottom-up Studies", *Critical Reviews in Environmental Science and Technology,* 27(1997, Special): S279-S292.

[19] Spence, M. and D. Starrett. "Most Rapid Approach Paths in Accumulation Problems," *Int. Econ. Rev.* 16, no. 2, (June 1975), 388-403.

[20] Stavins, R. N. "The Costs of Carbon Sequestration: A Revealed-Preference Approach" *Amer. Econ. Rev.* 89(4): 994-1009 (September 1999).

[21] Tsur, Y. and A. Zemel. "Optimal Transition to Backstop Substitutes for Nonrenewable Resources" Working Paper, 2000, Department of Agricultural Economics and Management, The Hebrew University of Jerusalem.

[22] van Kooten, G.C., A. Grainger, E. Ley, G. Marland, and B. Solberg. "Conceptual Issues Related to Carbon Sequestration: Uncertainty and Time," *Critical Reviews in Environmental Science and Technology,* 27 (1997, Special): S67-S80.

[23] United States, Department of States (USDOS). "United States Submission on Land-Use, Land-Use Change, and Forestry", August 1, 2000.

[24] World Wildlife Fund (WWF). "Critical Decisions Could Let Nations Keep Polluting and Still Reach Kyoto Targets", 4 May, 2000, WWF Press Release.

Notes

[1]The Kyoto Protocol currently allows only forests as carbon sinks, but left the door open for soil carbon sequestration through changes in land management practices.

[2]Other issues related to sinks include measurement, accounting rules, verification procedures and leakage.

[3]The Umbrella Group, a loose alliance of Annex I Parties that includes the US, Canada, Australia, Japan, Norway, the Russian Federation, Ukraine, and New Zealand, urged the development of simple procedures that facilitate the widespread use of mechanisms across a broad range of practices (including sequestration), while the European Union insisted on imposing limitations on the use of sinks, including the exclusion of "additional activities" in the first commitment period and quantitative limit for the use of sinks in Clean Development Mechanism projects.

[4]Montgomery formally establishes the monotonicity and convexity of the abatement cost function.

[5]We can relax the monotonicity assumption by allowing $B'(e)$ to be negative. Then in our paper, the relevant domain of $B(\cdot)$ is $[0,\ \bar{e}]$, where $\bar{e}$ is the optimal emission level in the absence of any regulation, i.e., $B'(\bar{e}) = 0$. We ignore this domain restriction because it is never binding.

[6]If a substantial amount of land is diverted from agricultural production, agricultural output prices may increase and the profit reduction $Q(A)$ would be even greater. Then $Q(A)$ is likely to be convex even with homogeneous land.

[7]In the rare cases, where (18) is not satisfied, $A(t)$ may not be monotone and there may be cyclical patterns of carbon sequestration and release. Such patterns need further study.

[8]IPCC (Intergovernmental Panel on Climate Change) was established in 1988 by the World Meteorological Organization (WMO) and the United Nations' Environmental Program (UNEP). It organizes scientists from all over the world to conduct rigorous surveys of the latest technical and scientific literature on climate change. The IPCC's assessment reports are widely recognized as the most credible sources of information on climate change.

[9]On the other hand, IPCC (2001) cites an increasing body of observations supporting the notion that global warming is already happening (and that most of the warming observed over the last 50 years is attributable to human activities). Depending on the damage of the warming (which may takes some years to realize) and the costs of reducing the current emissions (which we do not consider in this paper), it is also possible that the steady state C^* should be lower than the current C_0. Our analysis can be easily extended to analyze this situation.

[10]If the system starts from the right of $\dot{e} = 0$, $a(t) = \bar{a}$ before the singular path is reached, after which $a(t) < 0$. However, since $A^* > 0$, only part of sequestered carbon is released. In this situation, sinks are utilized early and to a great extent, so much so that part of the stored carbon has to be released.

[11]To simplify notation, we write λ as a function of C , instead of t, because the systems with and without carbon sequestration will arrive at C at different times. Strictly speaking, we should write $\lambda(t_1) < \tilde{\lambda}(t_2)$, where t_1, t_2 are such that $C(t_1) = C$ and $\tilde{C}(t_2) = C$. In this paragraph, we will use similar notations for $e(\cdot)$ and $\tilde{e}(\cdot)$.

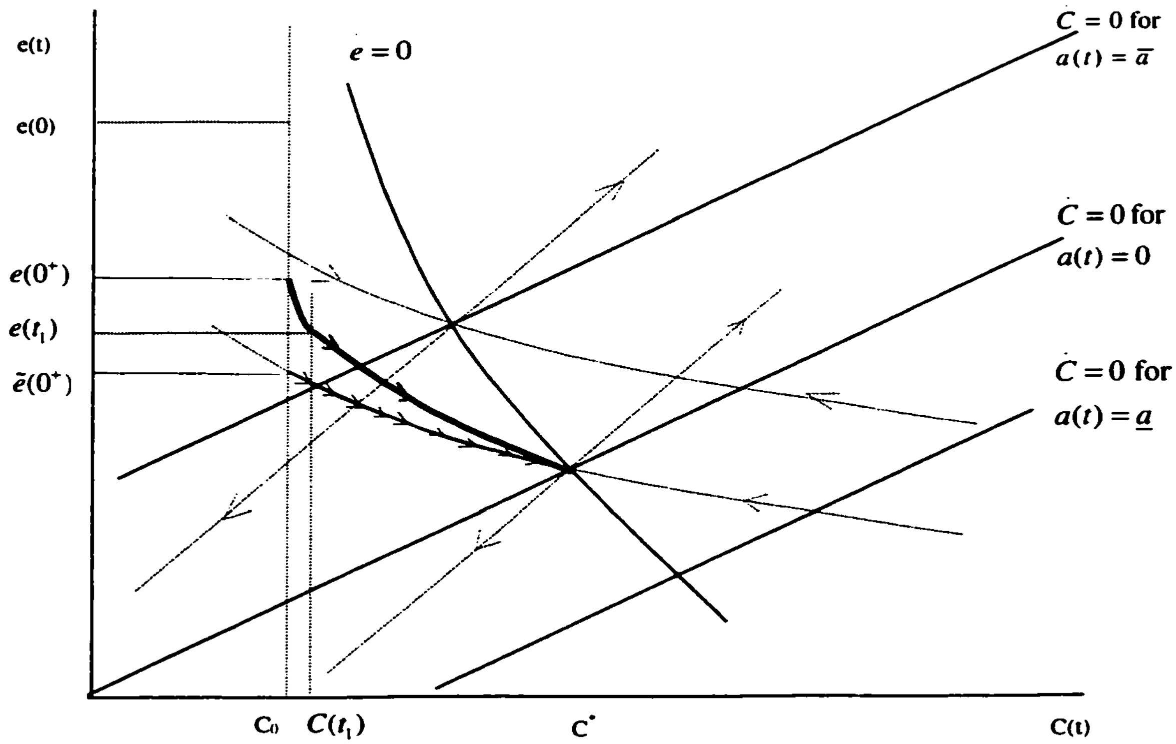

Figure 1: Phase Diagram for Carbon Emission and Stock

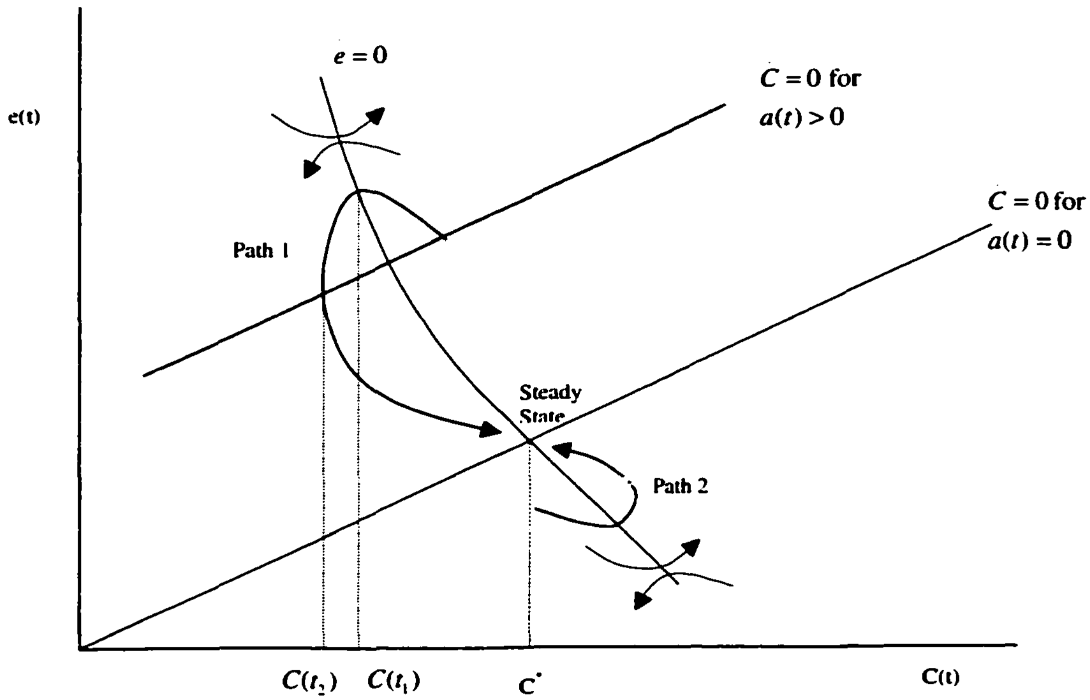

Figure A1: Possible Singular Paths

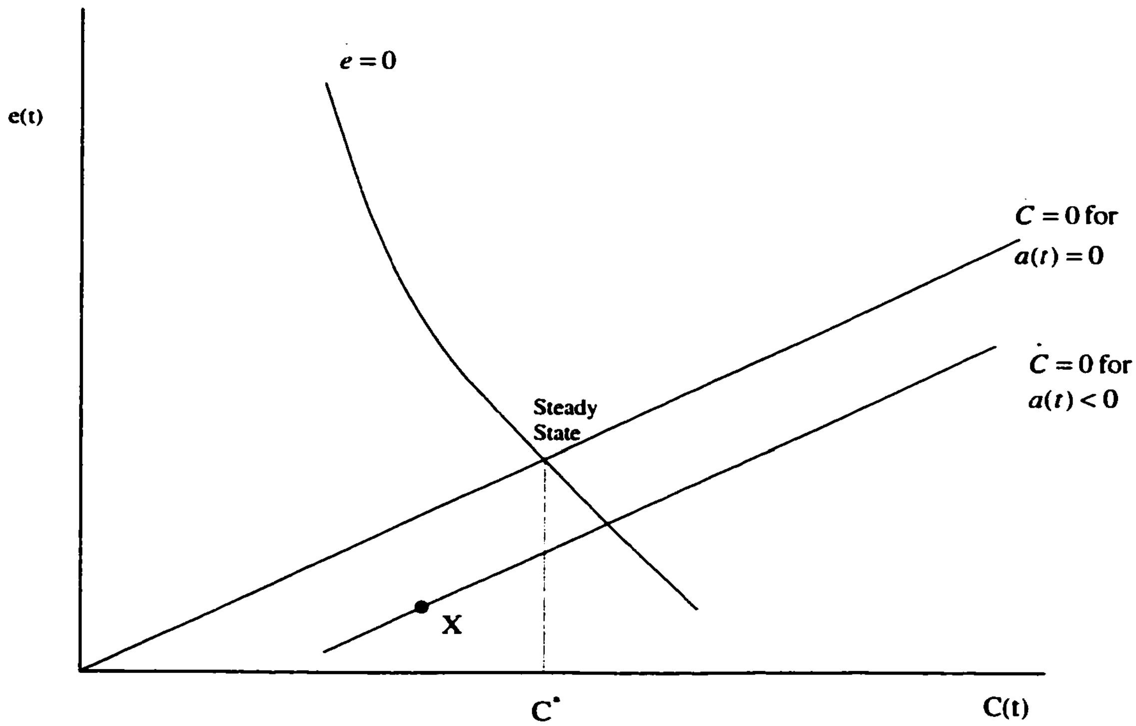

Figure A2: The Position of the System at t_2

CHAPTER 3.

ALTERNATIVE INTERTEMPORAL PERMIT TRADING REGIMES WITH STOCHASTIC ABATEMENT COSTS

A paper to be submitted to

Journal of Environmental Economics and Management

Abstract

When there are large number of firms, permit trading within one period tends to absorb firm-specific shocks in that period. In the presence of industry-wide shocks, however, allowing trade across time can attain a higher welfare level than a no-banking system. Bankable permit regimes with a 1-to-1 or non-unitary intertemporal trading ratios (ITRs) are examined. When banking is welfare improving, the optimal ITR is always less than 1+r, the ITR for monetary values. The more industry-wide shocks vary, and/or the more they are negatively correlated across time, the more efficient a bankable permit regime. Bankable permits with ITR=1 or ITR=1+r can both do better than a no banking regime. However, which one is better depends on the covariance structure of the shocks and the benefit and damage functions.

Introduction

Recent years have witnessed increasing interest in the use of tradable permit systems, which have been adopted for pollution control both in the U.S. and by an increasing number of nations (Stavins, 2000). While most permit systems focus on the flexibility provided by trading among or within emitting sources, the flexibility provided by trading across time has also been considered. Temporal permit trading may include banking and borrowing. Banking occurs when permits authorized for the current period are saved for use in some subsequent period. Borrowing occurs when permits authorized for some future period are instead used now. Temporal trading can lower compliance costs by allowing hedging against risks in emissions patterns and smoothing out fluctuations in abatement costs over time. Stavins (2000), Kolstad and Toman (2000) and Tietenberg (2000) all recognized that the temporal dimension can be a key component of a permit trading system.

In fact, banking has played an important role in some pollution control programs. For example, banking has likely enhanced the performance of the SO_2 allowance trading program (Ellerman *et al.* 1997), the U.S. lead rights trading program a decade earlier (Kerr and Maré, 1997) and the control of automobile hydrocarbon emissions in California (Rubin and Kling, 1993). Other examples that have made use of bankable permits include the Corporate Average Fuel Economy standards for automobiles and light trucks, which allowed banking and, in some cases, borrowing (Farrell *et al.* 1999); the Ozone Transport Region NO_x and VOCs emission trading program, which allowed banking; and, as an example of state-level programs, the Delaware NO_x and VOCs emission trading program, which also allowed banking.[1]

In spite of the potential for application of bankable permits, and the extensive studies on permit trading (see Tietenberg 1985, and Cropper and Oates 1992 for a review), there

[1]For a comprehensive description of permit trading programs, see Stavins (2000).

is limited research on the efficiency of bankable permits. Much of the literature on tradable permit systems has focused on the cost-effectiveness of these pollution control mechanisms. Most economists now agree that permit trading, including bankable permit programs can be cost-effective (Tietenberg 1985, Cronshaw and Kruse 1996, and Rubin 1996).

While separating means (cost-effective instruments) from ends (efficiency) highlights a strength of permit trading systems, there are limitations to this wisdom. As Stavins (1998) notes, "one risks designing a fast train to the wrong station". Kling and Rubin (1997) demonstrate the risks from focusing on cost-effectiveness by showing that, in a bankable permit system, firms will suboptimally choose excessive emissions in early periods and correspondingly too few in later periods. Leiby and Rubin (2000) extend their study to stock pollutants. Neither of these two models considers the consequences of incomplete information.

However, with complete information, there is no real advantage to permit trading, either across time or across firms, since the regulator can set the optimal number of permits for each firm in each period. Thus, it is important to analyze bankable permits in a framework with incomplete information. Yates and Cronshaw (2001) provides a careful analysis of bankable permits when polluting firms have better information about their abatement costs than a regulator. They investigate what is the optimal intertemporal trading ratio (ITR) and whether allowing bankable permits is welfare improving given that the bankable permit system is optimally designed.

Our work builds upon the Yates and Cronshaw study in several important aspects.

First, we recognize that what really drives intertemporal trading is industry wide shocks, not firm specific shocks. When there are a large number of firms, firm specific shocks in any period tend to be absorbed by permit trading among firms in that period. However, since every firm is affected by the same industry wide shocks, these shocks cannot be absorbed by static permit trading. This is where allowing banking or borrowing may be welfare improving.

Second, instead of focusing on the optimal ITR, we examine two special ITRs that are likely to be considered by policymakers due to their simplicity: the unitary ITR under which permits in every period are treated the same, and a non-unitary ITR where the interest rate on banked (or borrowed) permits are the same as the monetary interest rate. Using a two-period model, we find that a unitary bankable permit regime can dominate a no-banking regime if the marginal benefit curve is steeper than the marginal damage curve *and* uncertainties in the two periods are adequately negatively correlated. We also find that allowing intertemporal trading with an $ITR = 1+r$ is always welfare improving as long as the marginal benefit curve is steeper than the marginal damage curve.

Third, in Yates and Cronshaw (2001), firms are assumed to know the shocks in both periods, while in our model, firms know the shocks in the first period but not the second. This is an important distinction as we show that as firms' information advantage decreases, the gain from allowing intertemporal trading becomes smaller. To the extent that firms in general do not actually have complete information about second period shocks, this result is important to consider whether banking should be a part of permit trading system.

The rest of this paper is organized as follows. We lay out the basic elements of the model in the next section. In the third section, we examine firms' behaviors in alternative bankable permit regimes. Section 4 examines the unitary and non-unitary bankable permit regimes. Section 5 concludes.

Model Setup

Assume a two period situation, $t = 1, 2$, with multiple firms, $i = 1, 2, ...n$. A pollutant is emitted in both periods. Let e_t^i be firm i's emissions and $e_t = \sum_i e_t^i$ be the total emissions of all the firms in period t. Since higher emissions represent lower levels of abatement, firms benefit from emissions and the benefits are equivalent to the saved abatement costs. In each period, random shocks occur to firms' benefit functions. This could be technological progress or changes in the market environment.[2] Let $\mu_t^i \backsim i.i.d.(0, \sigma^2)$ be the firm specific shock for firm i in period t, and $\mu_t^0 \backsim (0, \sigma_t^2)$ be the industry wide shock. Firm-specific shocks may arise from a firm's internal production or management process, while industry-wide shocks could be due to fluctuations in the demand for the industry's products or the supply of its inputs. Shocks may be persistent, i.e., μ_t^i and μ_t^0 could be correlated across time.

[2]For example, the current marginal abatement costs for SO_2 are much lower than were estimated ten years ago. Over the decade preceding 1995, a typical unit's marginal abatement cost function was lowered by almost \$50 dollars per ton of SO_2 by technical improvements including advances in the ability to burn low-sulfur coal at existing generators, as well as improvements in overall generating efficiency. Moreover, the decline in fuel costs lowered the marginal abatement costs by about \$200 per ton (Carlson et al. 2000).

Since we wish to study the consequences of stochastic influence on trading independent of heterogeneity, we assume firms have the same benefit function, $B(e_t^i, \mu_t^i, \mu_t^0)$, with $B_e(\cdot) > 0$, $B_{ee}(\cdot) < 0$. This simplifies the modeling allowing us to focus on the stochastic nature of the problem. The monotonicity and concavity of $B(\cdot)$ is a result of the monotonicity and convexity of the abatement cost function (in the level of abatement)[3]. Since every firm has the same industry wide shock μ_t^0, we know if $B_e(e_t^i, \mu_t^i, \mu_t^0) = B_e(e_t^j, \mu_t^j, \mu_t^0)$, then $B_{e\mu_t^0}(e_t^i, \mu_t^i, \mu_t^0) = B_{e\mu_t^0}(e_t^j, \mu_t^j, \mu_t^0)$ for any μ_t^0. That is, μ_t^0 affects the marginal benefit equally across firms. The pollutant generates social damages, which for simplicity are assumed to depend on the total pollution by all firms. Let $D(e_t)$ be the damage function[4] with $D'(\cdot) > 0$, $D''(\cdot) > 0$.

Before the beginning of the first period, the regulator, without knowing the shocks in each period, has to determine the total number of permits to be issued for both periods and the intertemporal trading rules. We assume the regulator is able to commit. That is, after the number of permits and the trading ratio are set, they are then written into law and the regulator cannot change them.

At the beginning of period 1, the shocks, μ_1^i and μ_1^0 occur and firms observe their realizations. They then distribute their permits between the two periods. When firms make their decisions in period 1, they do not know what the shocks will be in the second

[3] Montgomery (1972) formally establishes the monotonicity and convexity of the abatement cost function.

[4] Uncertainty in the damage function will affect our results only if it is correlated with uncertainty in the benefit function. Stavins (1996) provided an analysis of policy instrument choice when uncertainties are correlated.

period. However, they could update their information about the distributions of μ_2^i, μ_2^0 based on what they have observed. When shocks are correlated across time, firms, after knowing μ_1^i, μ_1^0, may have better information about μ_2^i, μ_2^0 than the regulator had when the parameters of the market were set. If the shocks are perfectly correlated, then firms will know exactly what μ_2^i, μ_2^0 will be one period before their actual realizations.

Firms' Problem

In designing an optimal bankable permit regime, the regulator needs to take into account how firms behave in such a regime. So, we first derive firms' optimal decisions for a given regime, where the number of permits issued for each period is $\bar{e}_1$ and $\bar{e}_2$, and the trading ratio for banked permits is set at $1+\theta$. For every permit banked (borrowed), $1+\theta$ will be available for later use (be repaid), i.e., the interest rate on permits is θ. We will discuss two regimes. In the first, $\theta = 0$, i.e., a unitary intertemporal trading ratio is used. This case is of interest because it is simple to implement and is being used in actual bankable permits programs. In the second, θ is chosen optimally. This case has been discussed in previous studies, as noted in the introduction.

Montgomery (1972) first demonstrates that the decentralized behavior of firms in a permit trading system leads to the solution attainable under joint-benefit maximization. Yates and Cronshaw (2001) extend the result to an environment with intertemporal trading. Thus, to derive $e_t^i(\bar{e}_1, \bar{e}_2, \theta)$, we do not have to study the decentralized permit trading market; we can simply maximize firms' joint-benefits from emissions subject to the total permit cap:

$$\max_{e_t^i} \quad \sum B(e_1^i, \mu_1^i, \mu_1^0) + \frac{1}{1+r} \sum E\left[B(e_2^i, \mu_2^i, \mu_2^0) \mid \mu_1^i, \mu_1^0\right] \tag{1}$$

$$\text{such that} \qquad e_1 + e_2 \leq \bar{e}_1 + \bar{e}_2 + \theta(\bar{e}_1 - e_1), \quad e_t^i \geq 0. \tag{2}$$

where r is the financial interest rate and $\theta(\bar{e}_1 - e_1)$ is the interest for banked (or borrowed, if $\bar{e}_1 - e_1 < 0$) permits. The optimal conditions are, (second order conditions are satisfied by the curvature of the functions),

$$(1+r)\frac{\partial B(e_1^i, \mu_1^i, \mu_1^0)}{\partial e_1^i} = (1+\theta)\frac{\partial E\left[B(e_2^i, \mu_2^i, \mu_2^0) \mid \mu_1^i, \mu_1^0\right]}{\partial e_2^i}. \tag{3a}$$

$$\frac{\partial B(e_1^i, \mu_1^i, \mu_1^0)}{\partial e_1^i} = \frac{\partial B(e_1^j, \mu_1^j, \mu_1^0)}{\partial e_1^j}, \tag{3b}$$

$$\frac{\partial E\left[B(e_2^i, \mu_2^i, \mu_2^0) \mid \mu_1^i, \mu_1^0\right]}{\partial e_2^i} = \frac{\partial E\left[B(e_2^j, \mu_2^j, \mu_2^0) \mid \mu_1^j, \mu_1^0\right]}{\partial e_2^j}; \tag{3c}$$

Where $i, j = 1, 2, ...n$. Equation (3a) requires that the (adjusted) marginal benefit in the two periods be equal for all firms. The marginal benefit in the first period is multiplied by $(1 + r)$ because benefits in the first period are worth $(1 + r)$ more than those in the second period. The reason that the expected marginal benefits in the second period is multiplied by $(1 + \theta)$ is that one unit of permit saved in the first period will be worth $(1 + \theta)$ units of permit in the second period. Equations (3b)-(3c) are the usual marginal conditions for cost-effectiveness in a static permit market: the (expected) marginal benefits should be equal across firms at any point of time. Equation (3a) determines the distribution of permits between periods for any firm, while (3b)-(3c) determines how permits will be distributed across firms at any point of time.

Remark 1 *Industry wide shocks do not generate within period permit trading. That is,* *given an initial equilibrium, an industry wide shock will not change (3b) and (3c).*

The reason follows directly from the definition of industry wide shocks. Starting from an equilibrium, if (3b) and (3c) remain the same for the initial equilibrium upon any industry wide shock, then firms will have no reason to trade permits among themselves because their marginal benefit is the same. More specifically, since industry wide shocks affect every firm, if one firm has excessive demand for permits because it turns out to be expensive to abate emissions, then all other firms will also have excessive demand. When firms are identical and receive the same amount of initial permits, in the absence of firm-specific shocks, all firms will have the same amount of excessive demand (supply), thus, no trading takes place.

While industry wide shocks maintain the homogeneity of firms, firm specific shocks tend to break this homogeneity and thus make trading profitable. With firm specific shocks, one firm may find it costly to abate emissions and have excessive demand for permits, while another finds it cheap to do so and has excessive supply of permits. Thus, it is profitable for the first firm to buy some permits from the second, and both firms share the gains from intra-period trading.

Remark 2 *When there are a large number of firms, firm specific shocks tend to cancel each other. Formally, let* $\tilde{e}_1^i(\mu_1^1, \mu_1^2, ...\mu_1^n; \mu_1^0)$ *and* $\tilde{e}_2^i(\mu_1^1, \mu_1^2, ...\mu_1^n; \mu_1^0)$ *be the solution to (3a)-(3c), then, for any given industry wide shock* (μ_1^0), *we have*

$\lim_{n \to \infty} \Pr \left[\left| \frac{\sum_i \tilde{e}_t^i(\cdot)}{n} - \tilde{c}_t(\mu_1^0) \right| < \varepsilon \right] = 1$, *where* $\tilde{c}_t(\mu_1^0)$ *is independent of firm specific shocks.*

Since our focus is on intertemporal trading, the explanation for the remark is given in the appendix. Given the above remark and our focus on permit trading across time, we work now with a single firm, representing all regulated firms. Then, the firms' problem and the corresponding optimal conditions can be rewritten as follows,

$$\max_{e_1,e_2} \; B\left(e_1,\mu_1^0\right) + \frac{1}{1+r} E\left[B(e_2,\mu_2^0) \mid \mu_1^0\right] \tag{4a}$$

$$\text{Such that} \qquad e_1 + e_2 \leq \bar{e}_1 + \bar{e}_2 + \theta(\bar{e}_1 - e_1), \quad e_1, e_2 \geq 0; \tag{4b}$$

$$\frac{\partial B_1\left(e_1,\mu_1^0\right)/\partial e_1}{\partial E\left[B_2(e_2,\mu_2^0) \mid \mu_1^0\right]/\partial e_2} = \frac{1+\theta}{1+r}. \tag{4c}$$

Yates and Cronshaw (2001) have shown the parameters that matter are θ and $\bar{e}$, which is the present discounted value of $\bar{e}_1$ and $\bar{e}_2$, i.e., $\bar{e} = \bar{e}_1 + \frac{1}{1+\theta}\bar{e}_2$. Given that firms are allowed to move emissions freely between periods, the number of permits issued for any period has little real meaning. From (4b)-(4c), we know, in a bankable permit regime, firms' emissions in both periods are functions of $\bar{e}$ and θ, and in general will be different from $\bar{e}_1$ and $\bar{e}_2$. In the following analysis, we will examine the efficiency of alternative bankable permit regimes.

Optimal Bankable Permit Regimes

To design an optimal bankable permit regime, the regulator chooses θ and $\bar{e}$ to maximize the expected benefits minus costs, knowing that firms choose their emission levels

$e_1(\bar{e},\theta)$ and $e_2(\bar{e},\theta)$ according to (4b)-(4c). Thus, the regulator's problem is,

$$\max_{\bar{e},\theta} \quad E\left[W(e_1(\bar{e},\theta), e_2(\bar{e},\theta))\right]$$

$$\equiv E\left[B(e_1(\bar{e},\theta),\mu_1^0) + \tfrac{1}{1+r}B(e_2(\bar{e},\theta),\mu_2^0) - D(e_1(\bar{e},\theta)) - \tfrac{1}{1+r}D(e_2(\bar{e},\theta))\right]$$

$$\text{such that} \quad \bar{e} \geq 0,\ \theta \geq 0. \tag{5}$$

The firms' emissions for each period are fixed when banking is not allowed. In this case, the condition for social welfare maximizing is,

$$\frac{\partial E\left[B(e_t^{nb},\mu_t^0)\right]}{\partial e_t} = \frac{\partial D(e_t^{nb})}{\partial e_t} \quad \forall\ t = 1,2;\ i = 1,2,...,n, \tag{6}$$

where the superscript, "nb", stands for the ex ante no-banking social optimal level. Equation (6) says that emission standards should be set such that the expected marginal benefits equal the marginal damages in each period.

We next solve e_1 and e_2 when they are a function of $(\bar{e},\theta)$. To facilitate the comparison between no banking and banking, we use second-order Taylor expansions of the benefit and damage functions around the no banking social optimal[5], e_t^{nb}, i.e.,

$$B(e_t,\mu_t^0) \doteq B_0 + \left(B_1 + \mu_t^0\right)(e_t - e_t^{nb}) - \tfrac{1}{2}B_{11}(e_t - e_t^{nb})^2, \tag{7a}$$

$$D(e_t) \doteq D_0 + D_1(e_t - e_t^{nb}) + \tfrac{1}{2}D_{11}(e_t - e_t^{nb})^2, \tag{7b}$$

where B_0, B_1, B_{11} and D_0, D_1, D_{11} are fixed coefficients with $B_1 > 0$, $B_{11} \geq 0$, $D_1 > 0$, $D_{11} \geq 0$. Uncertainty is assumed to affect the marginal benefit function by shifting it up or down, while keeping its slope unchanged. By (6), we know $B_1 = D_1$.

[5] The "accurate local approximation" is very similar to what was originally used in Weitzman (1974) and since then has been widely used in other studies; for example, Kolstad (1987), Hoel and Karp (1999 and 2000), Newell and Pizer (1998).

To simplify notation, we define the following.

Definition 1 *Firms' demand for banking*[6] $(\Delta e_t(\bar{e}, \theta))$ *is the difference between the demand for permits in a banking regime and the no-banking optimum in period* t *given* $\bar{e}$ *and* $(1+\theta)$, *that is* $\Delta \bar{e}_t \equiv e_t(\bar{e}, \theta) - e_t^{nb}$.

The difference between $\bar{e}$ *and the no-banking optimal total permits is given by* $\Delta \bar{e} \equiv \bar{e} - \bar{e}^{nb}$, *where* $\bar{e} = \bar{e}_1 + \frac{1}{1+\theta}\bar{e}_2$, *and* $\bar{e}^{nb} = e_1^{nb} + \frac{1}{1+\theta}e_2^{nb}$.

Solving (4) with (7a), we get,

$$\Delta e_1(\bar{e}, \theta) = \frac{(1+r)\left(B_1 + \mu_1^0\right) - (1+\theta)\left(B_1 + E\left[\mu_2^0 | \mu_1^0\right]\right)}{B_{11}\left[(1+r) + (1+\theta)^2\right]} + \frac{(1+\theta)^2 \Delta \bar{e}}{(1+r) + (1+\theta)^2}, \tag{8a}$$

$$\frac{\Delta e_2(\bar{e}, \theta)}{(1+\theta)} = \frac{-(1+r)\left(B_1 + \mu_1^0\right) + (1+\theta)\left(B_1 + E\left[\mu_2^0 | \mu_1^0\right]\right)}{B_{11}\left[(1+r) + (1+\theta)^2\right]} + \frac{(1+r)\Delta \bar{e}}{(1+r) + (1+\theta)^2}, \tag{8b}$$

where $\frac{1}{(1+\theta)}$ on the left of the second equation is used to convert the second period emissions into its present discounted value, making it comparable to $\Delta e_1(\bar{e}, \theta)$.

Given $1+\theta$ and $\Delta \bar{e}$, firms' demand for banking in each period depends on a four factors. The first is the slope of the marginal benefit curve (B_{11}). The flatter the marginal benefit curve is, the more emissions in a bankable permit regime deviate from the no banking optimal emissions. The second is the difference between total permits in a bankable permit regime and the no banking optimal total emissions$(\Delta \bar{e})$. Whenever $\Delta \bar{e} > 0$ (or < 0), firms will split the difference between the two periods to equalize

[6] Firms' demand for banking in any period has a one-to-one relationship with the number of permits actually issued for this period. If the number of permits increases by one then the demand for banking decreases by one. Thus, to focus on other factors affecting banking, we study demand for banking when the number of permits issued is equal to the no banking optimum.

marginal benefits across time. The third is the relative magnitudes of (expected) marginal benefits evaluated at the no banking optimum, i.e., $B_1+\mu_1^0$, and $B_1+E\left[\mu_2^0|\mu_1^0\right]$. If marginal benefits in the first period adjusted by $(1+r)$ are higher than the expected marginal benefits in the second period adjusted by $(1+\theta)$, then first period emissions tend to be higher. Lastly, if μ_1^0 and μ_2^0 are negatively correlated, i.e., they tend to have opposite signs, their effects tend to enhance each other. Otherwise, their effects tend to cancel each other.

Optimal Total Permits

Substituting (8a)-(8b) into (5), we can derive the optimal $\bar{e}$ (or $\Delta\bar{e}$) and θ. We approach the problem in two steps. We first solve the optimal $\Delta\bar{e}$ for a given θ, and then discuss the optimal θ.

Solving (5) for a any given θ, we have the following,

Proposition 1 *The optimal total permits in a bankable permit regime with any intertemporal trading ratio equals the total permits in a no-banking regime. That is,*

$$\Delta\bar{e} = 0, \forall\theta. \tag{9}$$

The result is due to the way firms distribute the difference $\Delta\bar{e}$ between the two periods. Firms distribute the extra permits across the two periods such that marginal benefits in the two periods remain equal. From (8), we know for any $\Delta\bar{e}$, no matter what the realizations of the shocks are, firms allocate $\frac{(1+\theta)^2}{(1+r)+(1+\theta)^2}\Delta\bar{e}$ to the first period and $\frac{(1+r)}{(1+r)+(1+\theta)^2}\Delta\bar{e}$ to the second period. This implies that, setting $\Delta\bar{e} \neq 0$ does not help adjust firms' distribution of emissions toward social optimal. So there is no reason to set $\Delta\bar{e}$ different from zero.

Proposition 1 has one interesting policy implication: at least in terms of the total number of permits, the design of a bankable permit regime is no more complicated than a no banking regime. We next discuss the other parameter of a bankable permit regime, the intertemporal trading ratio, and compare the welfare of different permit trading regimes.

Unitary Bankable Permit Regime For a unitary bankable permit regime, θ is set at zero implying that firms can bank or borrow emissions across time periods at a one for one rate. As to the number of total permits, from Proposition 1, we know it is equal to the total permits in the optimal no banking regime. Substituting (9) back into the demand functions with $\theta = 0$, we get firms' demand for banking in a unitary bankable permit regime,

$$\Delta e_1^{ub}(\bar{e}, 0) = \frac{rB_1 + (1+r)\mu_1^0 - E\left[\mu_2^0 \mid \mu_1^0\right]}{(2+r)B_{11}}, \tag{10a}$$

$$\Delta e_2^{ub}(\bar{e}, 0) = \frac{-rB_1 - (1+r)\mu_1^0 + E\left[\mu_2^0 \mid \mu_1^0\right]}{(2+r)B_{11}}. \tag{10b}$$

The explanation for (10) are the same as in (8) except that here the weight on the marginal benefit in the second period is 1 instead of $1+\theta$. When there is no uncertainty, the demand for banking is,

$$\Delta e_1^{ub}(\bar{e}^{ub}, 0) = \frac{rB_1}{(2+r)B_{11}}, \quad \Delta e_2^{ub}(\bar{e}^{ub}, 0) = \frac{-rB_1}{(2+r)B_{11}}, \text{ for } \mu_1^0 = \mu_2^0 = 0. \tag{11}$$

Thus in the absence of uncertainty, if permits for each period are set at the no-banking optimum, firms will desire borrowing. This point is made by Kling and Rubin (1997), who find that in many cases firms would suboptimally choose excessive emission levels in early periods and correspondingly too few in later periods given the opportunity to freely move emissions between time periods. This is because firms discount future benefit streams and disregard the social damages of their emissions. Equation (11) quantifies this effect. For comparison with later analysis, we illustrate the case in Figure 1.

In Figure 1, MB and MD are the marginal benefit and damage functions. When there is no uncertainty, social optimality, which is also the no banking optimum, requires that $MB(e_1^{nb})=MD(e_1^{nb})$ and $MB(e_2^{nb})=MD(e_2^{nb})$, implying $e_1^{nb} = e_2^{nb} \left(= e^{nb}\right)$. However, if the regulator is going to issue permits equal to the no-banking optimal levels and then let firms trade with ITR=1, then the emission levels are e_1^{ub} and e_2^{ub}, with $e_1^{ub} > e_1^{nb}$, $e_2^{ub} < e_2^{nb}$ and $e_1^{ub} - e^{nb} = e^{nb} - e_2^{ub}$. When firms have the freedom to adjust their emissions through time, they will find e_1^{nb} and e_2^{nb} suboptimal, because by moving some emissions from period 2 to period 1 the additional benefits in period 1 will outweigh the reduced benefits in period 2.

When there is uncertainty, from $\Delta e_1^{ub}(\cdot)$ and $\Delta e_2^{ub}(\cdot)$ in (10), we know the higher the (conditional) expectation of second period marginal benefits for firms, the lower the first period emissions will be. However, it is not clear how firms' demand for permits will differ from the no-banking optimum because the relative magnitude of the uncertain terms is not known ex ante.

Definition 2 *The relative efficiency of a bankable permit regime ($\Delta E\left[W(\bar{e},\theta)\right]$) is the welfare difference between a bankable permit regime and the no-banking regime, i.e.,* $\Delta E\left[W(\bar{e},\theta)\right] \equiv E\left[W(e_1(\bar{e},\theta), e_2(\bar{e},\theta)\right] - E\left[W(e_1^{nb}, e_2^{nb})\right]$.

Substituting (10) back into the welfare function (5) and then comparing it with the no banking welfare level, we get,

$$\Delta E\left[W^{ub}(\bar{e}^{ub},0)\right]=\frac{-(B_{11}+D_{11})(rB_1)^2}{2(1+r)(2+r)B_{11}^2}+\frac{(B_{11}-D_{11})E\left[(1+r)\mu_1^0-E\left(\mu_2^0|\mu_1^0\right)\right]^2}{2(1+r)(2+r)B_{11}^2}. \tag{12}$$

The first term is negative, which represents the welfare loss (compared to a no banking regime). The sum of the two shaded areas on Figure 1 represents this loss. The slope of marginal benefit has two effects. On the one hand, as we discussed before, a flatter marginal benefit function means that a bigger emissions adjustment is needed to equalize marginal benefits across time. Since permits issued for each period are optimal in the absence of uncertainty, we want the adjustment of emissions as small as possible. In this sense, a steeper marginal benefit function will result in less loss, which is captured by B_{11}^2 in the denominator. On the other hand, for any given deviation of emissions from the optimal, we want both the marginal benefit and damage functions to be flatter, which is indicated by the presence of $(B_{11}+D_{11})$ in the numerator. From (11), we also know the higher the marginal benefits are, the more emissions in the two periods will differ, which in turn means more loss will occur. This explains the term $(rB_1)^2$ in the numerator.

The second term captures the welfare effects of uncertainty, which cannot be unambiguously signed (and is not represented in Figure 1). It will be positive if marginal benefit curve is steeper than marginal damage curve. Taking expectations, we have $E\left[(1+r)\mu_1^0-E\left(\mu_2^0|\mu_1^0\right)\right]^2=(1+r)^2\sigma_1^2-2(1+r)\sigma_{12}+\sigma_{2|1}^2$, where σ_1^2, σ_{12}, and $\sigma_{2|1}^2$, are the variance of μ_1^0, the covariance of μ_1^0 and μ_2^0, and the variance of the conditional expectation of μ_2^0, respectively. That is, $\sigma_1^2=E\left[\mu_1^0\right]^2$, $\sigma_{12}=E\left[\mu_1^0\mu_2^0\right]$, and $\sigma_{2|1}^2=E\left[E\left(\mu_2^0|\mu_1^0\right)\right]^2$. Thus, the absolute value of the second term will be larger if μ_1^0 and μ_2^0 have larger variances and/or are negatively correlated.

The factor $B_{11} - D_{11}$ in the second term is very similar to a term in Weitzman (1974). The no banking regime is a quantity tool since it fixes emissions in each period, just as a standard fixes the emissions for each firm. A bankable permit regime is akin to a price system in that emission levels in each period can deviate from the "standard" for the period[7]. In this sense, a bankable permit regime resembles a price tool. Weitzman (1974) showed that whether a price tool dominates a quantity tool depends on the slopes of the marginal benefit and damage functions. Similarly, whether a bankable permit regime dominates a no banking regime depends on the slopes of marginal benefit and damage functions. In both cases, how much one is preferred to the other depends on the covariance structure of the shocks.

Thus, a unitary bankable permit regime could be welfare enhancing because it gives firms who have better information about the random shocks the flexibility of adjusting to shocks. However, since firms ignore the social damages, their adjustment may be suboptimal. Therefore, how a unitary permit regime performs relative to a no-banking regime depends on how much there is to gain from flexibility (the second term) relative to how severely firms' redistribution of permits differs from the social optimal (the first term). In other words, directly from (12), we have

Proposition 2 *A unitary bankable regime dominates a no-banking regime if*

[7]Another way to think about this is that given total permits available, once firms know the shock in the first period and make their best guess about the shock in the second period, the permit prices in the two periods are given. Thus, when firms make their emission decisions, it is as if they face a fixed price.

$$(B_{11} - D_{11}) \, E\left[(1+r)\mu_1^0 - E\left(\mu_2^0 | \mu_1^0\right)\right]^2 > (D_{11} + B_{11}) \, (rB_1)^2 .$$

For a unitary bankable regime to dominate a no-banking regime, the condition that the benefit function is steeper than the damage function is necessary but not sufficient. In particular, the covariance structure of the shocks are important. The more shocks are negatively correlated, the more likely that a unitary bankable regime will dominate a no-banking regime.

Figure 2 illustrates a case where a unitary bankable regime dominates a no-banking regime. In the figure, the no-banking optimum requires that $e_1^{nb} = e_2^{nb} = e^{nb}$. In the first period, firms observe the realization of $\mu_1^0 > 0$, and expect the shock in the second period to be $E\left[\mu_2^0 | \mu_1^0\right] < 0$. Then they distribute their emissions in the two periods as e_1^{ub} and e_2^{ub}, such that expected marginal benefits are equalized across the two periods, i.e., $E\left[MB\right] + \mu_1^0 = \frac{E[MB] + E\left[\mu_2^0 | \mu_1^0\right]}{1+r}$. Since it turns out that abatement costs in the first period are very high and relatively low in the second, firms emit more in the first and correspondingly fewer in the second period. In the first period, to achieve the expected social optimal, which requires $E\left[MB\right] + \mu_1^0 = MD$ and $\frac{E[MB] + E\left[\mu_2^0 | \mu_1^0\right]}{1+r} = \frac{MD}{1+r}$, the emissions would have been e_1^* and e_2^*. Without intertemporal trading, the social loss would be the areas of the two dotted triangles due to too few emissions in the first period and too many in the second. With intertemporal trading, the social loss would be the areas of the two bold-line bordered triangles because too many emissions in the first and too few in the second period. Neither regime attains e_1^* and e_2^*. Which of the two is better depends on the slopes of the marginal benefit and damage functions and the covariance structure of the shocks. When the condition in proposition 2 are satisfied, the loss of a no-banking regime is greater than a bankable permit system, as is shown in the figure.

There is also an alternative interpretation for Figure 2. Compared to the no-banking regime, firms gain from additional emissions in the first period and lose due to reduced emissions in the second. The opposite is true in the damage side, there are more (fewer) damages in the first (second). In the figure, the benefit gain outweighs the increase in damages.

Non-unitary Bankable Permit Regimes

In this case, the regulator has two instruments to maximize social welfare. Instead of fixing θ at zero, the regulator can choose both $\bar{e}$ and $1+\theta$, given $\Delta e_t(\bar{e},\theta)$, to maximize social welfare. From Proposition 1, we know $\bar{e} = e^{nb}$, or $\Delta\bar{e} = 0$.

Substituting (9) into (8), we get firms' demand for banking in a non-unitary bankable permit regime,

$$\Delta e_1^{gb}(\bar{e},\theta) = \frac{(r-\theta)\,B_1 + (1+r)\,\mu_1^0 - (1+\theta)E\left[\mu_2^0 \mid \mu_1^0\right]}{\left[(1+r)+(1+\theta)^2\right]B_{11}}, \tag{13a}$$

$$\Delta e_2^{gb}(\bar{e},\theta) = \frac{(1+\theta)\left[-(r-\theta)\,B_1 - (1+r)\,\mu_1^0 + (1+\theta)E\left[\mu_2^0 \mid \mu_1^0\right]\right]}{\left[(1+r)+(1+\theta)^2\right]B_{11}}. \tag{13b}$$

where the superscript "gb" means optimal solutions for a general bankable permit regime with optimally set ITR. Substituting these demand functions back into the welfare function, we get,

$$\Delta E\left[W^{gb}(\bar{e}^{nb},\theta)\right] = \Delta W_c^{gb}(\bar{e}^{nb},\theta) + \Delta W_{uc}^{gb}(\bar{e}^{nb},\theta), \text{ with} \tag{14}$$

$$\Delta W_c^{gb}(\bar{e}^{nb},\theta) = \frac{-\left(B_{11}+D_{11}\right)(\theta-r)^2 B_1^2}{2(1+r)\left[(1+r)+(1+\theta)^2\right]B_{11}^2},$$

$$\Delta W_{uc}^{gb}(\bar{e}^{nb},\theta) = \frac{\left(B_{11}-D_{11}\right)E\left[(1+r)\mu_1^0 - (1+\theta)E\left(\mu_2^0|\mu_1^0\right)\right]^2}{2(1+r)\left[(1+r)+(1+\theta)^2\right]B_{11}^2},$$

where subscripts "c" and "uc" stand for certain and uncertain, respectively. By maximizing $\Delta E\left[W^{gb}(\bar{e}^{nb},\theta)\right]$ with respect to θ, we can get the optimal θ, θ^{gb}. The following proposition described the magnitude of the optimal θ relative to r,

Proposition 3 *If the optimal θ^{gb} exists, then,*

$\theta^{gb} < r$ *if* $(B_{11} - D_{11}) > 0$,

$\theta^{gb} > r$ *if* $(B_{11} - D_{11}) < 0$,

$\theta^{gb} = r$ *if* $(B_{11} - D_{11}) = 0$.

The intuition behind the proposition is as follows. There is a trade-off in choosing the optimal θ. On the one hand, the value of θ is set to correct firms' discounting behavior, i.e., θ is set to obtain a higher value of $\Delta W_c^{gb}(\bar{e}^{nb},\theta)$. This purpose can be best served by setting $\theta = r$. On the other hand, θ is set to best utilize information on uncertainty, i.e., to maximize $\Delta W_{uc}^{gb}(\bar{e}^{nb},\theta)$. Setting $\theta = r$ in general does not maximize $\Delta W_{uc}^{gb}(\bar{e}^{nb},\theta)$. If $(B_{11} - D_{11}) > 0$, we know from $\Delta W_{uc}^{gb}(\bar{e}^{nb},\theta)$ the benefit gain from adjusting to uncertainty dominates the increase in damage. Since benefits in the first period are more valuable than those in the second, it makes sense to have a small θ to encourage first period emissions. A trade-off between maximizing the first and the second terms would require $\theta^{gb} \leq r$. If $(B_{11} - D_{11}) < 0$, then the damage loss from adjusting to uncertainty dominates the benefit gain. Since second period damage is less valued than first period, the optimal policy should encourage permit saving and discourage borrowing. Thus a trade-off would require $\theta^{gb} \geq r$. If $(B_{11} - D_{11}) = 0$, then emissions in neither period are more beneficial than the other, so θ is set to maximize just $\Delta W_c^{gb}(\bar{e}^{nb},\theta)$ and $\theta^{gb} = r$.

We will next discuss the welfare levels of non-unitary bankable regime relative to unitary bankable regime and no-banking regime, given some value of θ. First, we have the following,

Proposition 4 *A bankable regime with* $\theta = r$, $\bar{e} = \bar{e}^{nb}$ *always weakly dominates a no-banking regime given that* $(B_{11} - D_{11}) > 0$.

The proof is trivial. Substituting $\theta = r$ into $\Delta E\left[W^{gb}(\bar{e}^{nb}, \theta)\right]$, we get $\Delta E\left[W^{gb}(\bar{e}^{nb}, \theta)\right]$ $= \Delta W^{gb}_{uc}(\bar{e}^{nb}, \theta)$. If $(B_{11} - D_{11}) > 0$, then $\Delta E\left[W^{gb}(\bar{e}^{nb}, \theta)\right] \geq 0$.

Since it may be hard to find the optimal θ in reality, this finding may have very important practical implications. If we are not sure what the optimal θ is, or just for simplicity, the regulator may set the interest rate on banked permits equal to the interest rate on monetary values. By doing so, she can be sure a bankable permit regime with $ITR = 1+r$ still performs better than a no-banking regime. The intuition underlying this is that, by setting $\theta = r$, the regulator offsets firms' tendency of suboptimally distributing permits and provides firms the flexibility of adjusting emissions. Given that marginal benefit function is steeper, the benefit gain from adjusting to uncertainty outweighs the damage loss. Thus, setting $\theta = r$ is welfare improving.

Remark 3 *An optimal bankable permit regime always weakly dominates a no banking regime if* $(B_{11} - D_{11}) > 0$.

The remark is a direct result of Proposition 4 given that a bankable permit regime with $\theta = r$, $\bar{e} = \bar{e}^{nb}$ is not necessarily the optimal choice. This remark reconfirms one of the findings of Yates and Cronshaw (2001).

Remark 4 *Given $(B_{11} - D_{11}) > 0$, a non-unitary bankable regime with $\theta = r$ does not necessarily dominate a unitary bankable regime.*

If $(B_{11} - D_{11}) > 0$, the regime with $ITR = 1 + r$ is always better than a no-banking regime while a unitary regime can be better than a no-banking regime. However, it is not easy to see which is better, since both of them are suboptimal settings. The answer depends on the structure of the benefit and damage functions and the covariance structure of the shocks. From (12) and (14), we expect the regime with $ITR = 1 + r$ to be better if the optimal θ is close to r, and the loss from not fully correcting firms' discounting behavior is big.

From (12) and (14), we know,

Remark 5 *Ceteris paribus, the more shocks are negatively correlated, and/or the more they vary, the higher (lower) the welfare of a bankable permit regime if $(B_{11} - D_{11}) > 0$ ($(B_{11} - D_{11}) < 0$). However, the covariance structure of the shocks does not impact whether a bankable regime dominates a no-banking regime.*

It is easy to see that the value of $[(1+r)^2\sigma_1^2 - 2(1+r)(1+\theta)\sigma_{12} + (1+\theta)^2\sigma_{2|1}^2]$ is higher, if μ_1^0 and μ_2^0 are negatively correlated. Intuitively, the flexibility provided by intertemporal trading has higher value if it turns out the abatement costs (potential benefits from permits for firms) are high in one period and low in another period. Otherwise, there is not much gain from emission smoothing for firms. When $(B_{11} - D_{11}) > 0$, the gain for firms from banking outweighs the social loss from the damages of emissions. However, if $(B_{11} - D_{11}) < 0$, then the opposite is true. In this case, the more firms trade, the more the damage from emissions, which outweighs firms' gain from trading. Thus, the more firms trade, the more the social loss.

We have analyzed the case where the regulator does not know shocks in both periods when designing a permit regime while firms know the first period shock when making emissions decisions. We may consider the situation without asymmetric information, i.e., firms and the regulator have the same information, including: both sides know the shocks in each period, know the first period shock but not the second, or know the shocks in neither period. For the situation where firms have better information than the regulator, in addition to the case we have considered in previous sections, we may also consider the case where firms know the shocks of both periods while the regulator knows neither of them. For bankable permits in these various cases, we have the following remark,

Remark 6 *With or without uncertainty, if the regulator and firms have the same information about the shocks in the two periods, then the optimal θ equals r and there is no welfare gain from banking. If firms have better information than the regulator, then whether there is a gain from banking depends on the slopes of the marginal benefit and damage functions, and the magnitude of gain or loss is increasing with the degree of asymmetric information.*

For a detailed explanation see the appendix. The motivation for the flexibility of trading permits across time is to let firms adjust to situations which they know better than the regulator. If firms have no better information than the regulator when they are making their emissions decisions, giving flexibility to them will not result higher welfare and may well result in lower welfare since firms disregard the externality caused by their emissions.

Conclusions

This paper studies the design and efficiency of alternative bankable permit regimes. Similar to Yates and Cronshaw, we find that a necessary condition for banking to dominate no banking is that the marginal benefit curve is steeper than the marginal damage curve. When this necessary condition is satisfied, whether banking dominates no banking depends on the ITR. Two of our findings are of particular interest.

One is related to bankable permit regimes with ITR=1+r. When the slope condition is satisfied, a bankable permit regime with total permits equal to the no banking optimum and ITR=1+r always dominates no banking. We consider this result interesting because it demonstrates that the design of a welfare improving bankable permit regime is no more complicated than the no banking regime.

The other is related to bankable permit regimes with ITR=1. Previous studies (e.g. Rubin and Kling, 1997) have shown that such regimes are suboptimal. We find that this does not necessarily mean that a unitary bankable permit regime should never be used. As long as the flexibility of banking provides more benefit gains than damage losses, unitary banking dominates no banking. This case is more likely to be true when firms have big information advantage over the regulator about their benefits and benefits tend to be negatively correlated across time. We consider this result interesting because it says something of a permit regime which is often actually used.

Appendix

Detailed explanation for Remark 2 The remark can be better explained by studying the decentralized permit trading market. For any firm i, the benefit maximization problem is

$$\max_{e_t^i} \; B(e_1^i, \mu_1^i, \mu_1^0) + \frac{1}{1+r} E\left[B(e_2^i, \mu_2^i, \mu_2^0) \mid \mu_1^i, \mu_1^0\right] - p_1 x_1^i - \frac{1}{1+r} p_2 x_2^i$$

$$\text{Such that} \qquad e_1^i + \frac{1}{1+\theta} e_2^i \leq \bar{e}_1^i + \frac{1}{1+\theta}\bar{e}_2^i + x_1^i + \frac{1}{1+\theta} x_2^i, \quad e_t^i \geq 0,$$

where p_t is the permit price in period t, and x_t^i is the amount of permits firm i buys from the market in period t. From the above problem, we can derive this firm's demand for permits in each period as a function of permit prices, i.e., $e_1^i(p_1, \mu_1^i, \mu_1^0)$ and $e_2^i(p_2, \mu_2^i, \mu_2^0)$ with $p_2 = \frac{1+r}{1+\theta} p_1$. Setting total demand for permits equal to the supply (the total amount of permits issued), we may obtain the equilibrium prices. Since μ_t^i, μ_t^0 are random variables, firms' demand for permits is also random. Also, the demand by firm i is independent of the demand by firm j, because μ_t^i are independent across firms for any given t. If $e_t^i(\cdot)$ has finite variance, then by the law of large numbers,

$\lim_{n\to\infty} \Pr\left[\left|\frac{\sum_i e_t^i(p_1, \mu_1^i, \mu_1^0)}{n} - c_t(p_1, \mu_1^0)\right| < \varepsilon\right] = 1$, where $c_t(p_1, \mu_1^0)$ is independent of firm specific shocks.

Thus, the industry's demand for permits will be the same no matter what firm-specific shocks are in any given period. Given that the supply of permits is fixed, the industry's equilibrium emissions in each period will also be the same with probability one.

Proof for Proposition 3 Differentiating the first term of $\Delta E\left[W^{gb}(\bar{e}^{nb}, \theta)\right]$ with respect to θ, we know $\frac{d\left(\Delta W_c^{gb}(\bar{e}^{nb}, \theta)\right)}{d\theta} = 0$ for $\theta = r$, and $\frac{d\left(\Delta W_c^{gb}(\bar{e}^{nb}, \theta)\right)}{d\theta} > 0 \; (< 0)$ for $\theta < r$

$(> r)$.

Proof. We next determine the sign of the derivative of the second term for $\theta \leq r$. Expanding the expectation in the second term, we get

$\Delta W_{uc}^{gb}(\bar{e}^{nb}, \theta) = \frac{(B_{11}-D_{11})}{2(1+r)B_{11}^2} \frac{[(1+r)^2 var(\mu_1^0) - 2(1+r)(1+\theta)cov(\mu_1^0, \mu_2^0) + (1+\theta)^2 var(E(\mu_2^0|\mu_1^0))]}{[(1+r)+(1+\theta)^2]}$. Differentiating with respect to θ, we get

$$\begin{aligned}
& \frac{d\left(\Delta W_{uc}^{gb}(\bar{e}^{nb}, \theta)\right)}{d\theta} \\
= \; & A * \left[(1+r)var\left(\mu_1^0\right) - var\left(E\left(\mu_2^0|\mu_1^0\right)\right) - \left(\tfrac{2\theta+\theta^2-r}{1+\theta}\right) cov\left(\mu_1^0, \mu_2^0\right)\right], \\
= \; & A * \left[E\left[var\left(\mu_2^0|\mu_1^0\right)\right] + \left(\tfrac{(1+\theta)r - \left(2\theta+\theta^2-r\right)\rho}{1+\theta}\right) var\left(\mu_1^0\right)\right], \qquad (15)
\end{aligned}$$

where $A \equiv \frac{(B_{11}-D_{11})}{2(1+r)B_{11}^2} \frac{-2(1+\theta)(1+r)}{[(1+r)+(1+\theta)^2]^2}$ and ρ is the correlation coefficient of μ_1^0 and μ_2^0. In the second step, several relations are used, $var\left(\mu_1^0\right) = E\left[var\left(\mu_2^0|\mu_1^0\right)\right] + var\left(E\left(\mu_2^0|\mu_1^0\right)\right)$ and $cov\left(\mu_1^0, \mu_2^0\right) = \rho var\left(\mu_1^0\right)$, $var\left(\mu_1^0\right) = var\left(\mu_2^0\right)$. If $\theta \leq r$, then $\left(\frac{(1+\theta)r - \left(2\theta+\theta^2-r\right)\rho}{1+\theta}\right) > 0$. Thus, for $\theta \in [0, \ r]$, sign $\left(\frac{d\left(\Delta W_{uc}^{gb}(\bar{e}^{nb}, \theta)\right)}{d\theta}\right) = \text{sign}\ (A) = -\ \text{sign}\ (B_{11} - D_{11})$.

Consider the derivatives of the two terms together. If $(B_{11} - D_{11}) = 0$, then $\Delta E\left[W^{gb}(\bar{e}^{nb}, \theta)\right] = \Delta W_c^{gb}(\bar{e}^{nb}, \theta)$, thus the optimal θ, $\theta^{gb} = r$. If $(B_{11} - D_{11}) > 0$, then $\frac{d\left(\Delta W_{uc}^{gb}(\bar{e}^{nb}, \theta)\right)}{d\theta} < 0$, for $\theta \in [0, \ r]$. Given that $\theta = r$ is the global maximum for $\Delta W_c^{gb}(\bar{e}^{nb}, \theta)$, we have $\theta^{gb} < r$, if θ^{gb} exists. On the contrary, If $(B_{11} - D_{11}) < 0$, then $\frac{d\left(\Delta W_{uc}^{gb}(\bar{e}^{nb}, \theta)\right)}{d\theta} > 0$ for $\theta \in [0, \ r]$, and so $\theta^{gb} > r$,if θ^{gb} exists. ■

Detailed explanation for Remark 6 *(i) When the regulator and firms have the* *same information* By the definitions of Δe_t, and $\Delta E[W]$, we know

$$\begin{aligned}\Delta\left[W^{gb}(\bar{e}^{nb},\theta)\right] &= (B_1+\mu_1^0)\Delta e_1^{gb} - \frac{1}{2}B_{11}\left(\Delta e_1^{gb}\right)^2 - D_1\Delta e_1 - \frac{1}{2}D_{11}\left(\Delta e_1^{gb}\right)^2 + \\ &\quad \tfrac{1}{1+r}\left[(B_2+\mu_2^0)\Delta e_2^{gb} - \tfrac{1}{2}B_{11}\left(\Delta e_2^{gb}\right)^2 - D_1\Delta e_2^{gb} - \tfrac{1}{2}D_{11}\left(\Delta e_2^{gb}\right)^2\right], \\ &= \frac{-1}{2(1+r)}(B_{11}+D_{11})\left[(1+r)+(1+\theta)^2\right]\left(\Delta e_1^{gb}\right)^2.\end{aligned}$$

The second step requires some explanation. We first show that the terms that are linear in Δe_t^{gb} will vanish when the regulator and firms have the same information. If μ_t^0 is known to both parties, then $B_1+\mu_1^0 = D_1$ and $(B_2+\mu_2^0) = D_1$. These two equations are just the first order conditions for the optimal no-banking permit policies. When there is uncertainty, proper expectations have to be used in place of μ_t^0. If only μ_1^0 is known to both parties, then $E(\mu_2^0|\mu_1^0)$ replaces μ_2^0. We will have $B_1+\mu_1^0 = D_1$ and $(B_2+E(\mu_2^0|\mu_1^0)) = D_1$. If both μ_1^0 and μ_2^0 are unknown to both parties, then μ_t^0 will be replaced by its expectation (zero) and we have $B_1 = D_1$ and $B_2 = D_1$. Rearranging the square terms using the relation $\Delta e_2^{gb}(\bar{e},\theta) = (1+\theta)\Delta e_1^{gb}(\bar{e},\theta)$, we will get the second line.

From (13a), we know $\left(\Delta e_1^{gb}\right)^2$ can be minimized if $\theta = r$, since $B_1+\mu_1^0 = B_2 + E(\mu_2^0|\mu_1^0)$. Given than $\Delta\left[W^{gb}(\bar{e}^{nb},\theta)\right] \le 0$, setting $\theta = r$ maximizes it.

(ii) When firms have better information Most of the analysis in previous sections deals with the case where the regulator does not know the shocks in both periods while firms know the shocks in the first but not that of the second. So here we will only analyze the case where the regulator still does not know the shock in either period while firms know the shocks in both periods. In this case, we say the degree of asymmetric information increases because firms' information has improved while the regulator's stays

the same. When firms know μ_2^0, the conditional expectation term in (8), $E\left[\mu_2^0|\mu_1^0\right]$ will be replaced by μ_2^0. And $E[(1+r)\mu_1^0 - (1+\theta)E\left(\mu_2^0|\mu_1^0\right)]^2$, in $\Delta E\left[W^{gb}(\bar{e}^{nb},\theta)\right]$ will be replaced by $E[(1+r)\mu_1^0 - (1+\theta)\mu_2^0]^2$. Since $E[(1+r)\mu_1^0 - (1+\theta)E\left(\mu_2^0|\mu_1^0\right)]^2 - E[(1+r)\mu_1^0 - (1+\theta)\mu_2^0]^2 = \sigma_{2|1}^2 - \sigma_2^2$ and $\sigma_{2|1}^2 - \sigma_2^2 \leq 0$, we know $E[(1+r)\mu_1^0 - (1+\theta)E\left(\mu_2^0|\mu_1^0\right)]^2 < E[(1+r)\mu_1^0 - (1+\theta)\mu_2^0]^2$. The second half of the remark follows directly from this.

References:

Carlson, Curtis, Dallas Burtraw, Manreen Cropper and Karen Palmer, "Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade?", *Journal of Political Economy*, 2000, vol. 108, no. 6.

Cronshaw, Mark B., and Jamie Brown Kruse, "Regulated Firms in Pollution Permit Markets With Banking", *Journal of Regulatory Economics*, 9, 179-189, (1996).

Ellerman, D., R. Schmalensee, P. Joskow, J. Montero, and E. Bailey, *Markets for Clean Air: the U.S. Acid Rain Program*, Cambridge University Press, (2000).

Farrell, Alex, Robert Carter and Roger Raufer, "The NO_x Budget: Market-Based Control of Tropospheric Ozone in the Northeastern United States", *Resource and Energy Economics*, vol 21 (2), pp. 103-124, May (1999).

Hoel, Michael and Larry Karp, "Taxes Versus Quotas for a Stock Pollutant", University of California, Berkeley, Working Paper, (1999).

Hoel, Michael and Larry Karp, "Taxes Versus Quotas for a Stock Pollutant with Multiple Uncertainty", *Journal of Public Economics*, Forthcoming, (2001).

Kerr, S. and D. Maré, "Efficient Regulation Through Tradeable Permit Markets: The United States Lead Phasedown", WP 96-06, Department of Agricultural and Resource Economics, University of Maryland (1997).

Kling, Catherine and Rubin, Jonathan D., 1997, "Bankable Permits for the Control of Environmental Pollution", *Journal of Public Economics*, 64, 101-115, (1997).

Kolstad, Charles D., "Uniformity versus Differentiation in Regulating Externalities", *Journal of Environmental Economics and Management, (14), 386-399, (1987).*

Kolstad, Charles D., and Michael Toman, "The Economics of Climate Policy", Discussion Paper 00-40, Resource For the Future, (2000).

Leiby, Paul and Jonathan Rubin, "Bankable Permits for the Control of stock and Flow Pollutants: Optional Intertemporal Greenhouse Gas Emission Trading," *Environmental and Resource Economics*, 19(3): 229-256, July 2001.

Montgomery, W. David, "Markets in Licenses and Efficient Pollution Control Programs", *Journal of Economic Theory*, 5, 395-418, (1972).

Newell, Richard G., and William A. Pizer, "Regulating Stock Externalities Under Uncertainty", Discussion Paper 99-10, Resources For the Future, (1999).

Rubin, Jonathan D., "A Model of Intertemporal Emission Trading, Banking, and Borrowing", *Journal of Environmental Economics and Management,* 31, 269-286, (1996).

Rubin, Jonathan D., and Catherine Kling, "An Emission Saved is an Emission Earned: an Empirical Study of Emission Banking for Light-Duty Vehicle Manufacturers", *Journal of Environmental Economics and Management*, 25, 257-274 (1993).

Stavins, N. Robert, "Correlated Uncertainty and Policy Instrument Choice", *Journal of Environmental Economics and Management, 30*, 218-232, (1996).

Stavins, N. Robert, "What Can We Learn From the Grand Policy Experiment? Lessons from SO_2 Allowance Trading", *Journal of Economic Perspectives, Vol. 12, No. 3, 69-88*, Summer (1998).

Stavins, N. Robert, "Experience with Market-Based Environmental Policy Instruments", Prepared as Chapter 21 of *The Handbook of Environmental Economics*, edited by Karl-Göran Mäler and Jeffrey Vincent; General Editors: Kenneth Arrow and Michael

Intriligator, Amsterdam: North-Holland/Elsevier Science, (2000).

Tietenberg, Tom, "Editor's Introduction" in "The Evolution of Emission Trading: Theoretical Foundations and Design Considerations", (2000).

Weitzman, M. L., "Prices versus Quantities", *Review of Economic Studies,* 41, 477-491 (1974).

Yates, Andrew J. and Mark B. Cronshaw. "Pollution Permit Markets with Intertemporal Trading and Asymmetric Information", *Journal of Environmental Economics and Management, 42,* 104-118, (2001).

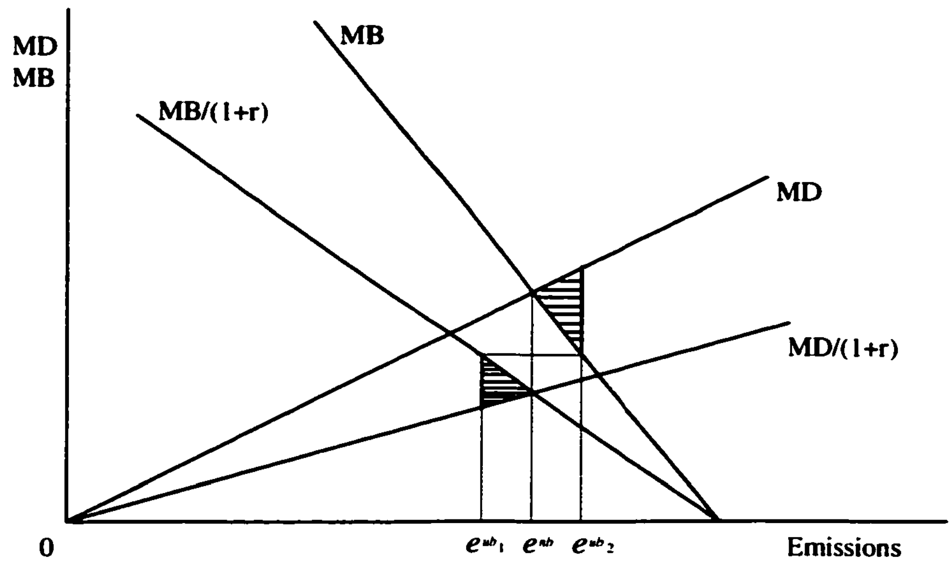

**Figure 1: Emissions and Welfare—
Banking with ITR=1 and No Uncertainty**

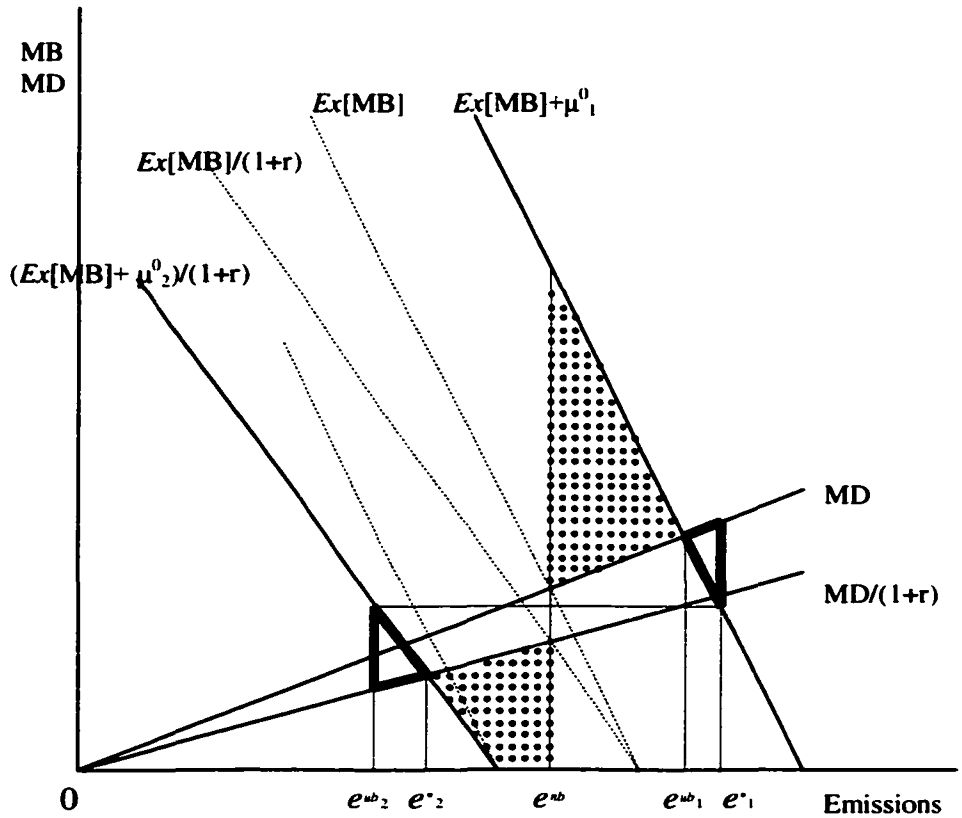

Figure 2: Emissions and Welfare—
Banking with ITR=1 and Uncertainty

CHAPTER 4.

GREEN PAYMENTS AND DUAL POLICY GOALS

A paper to be submitted to

American Journal of Agricultural Economics

Abstract

This paper analyzes the efficient design of green payments. Green payments may generate environmental benefits and support the income of small farmers. We find that if the government intends to achieve both of these two goals, then the decoupling of green payments and farm size is not optimal when information is limited. Moreover, the effectiveness of green payments critically depends on the correlation between conservation efficiency and farm size.

Introduction

Green payments, which are also called conservation payments or green support programs, are payments government provides for farmers for voluntarily maintaining or adopting conservation practices that enhance the environment, natural resources and wildlife habitat. As we move closer to the 2002 bill, which will replace the 1996 FAIR Act, we have seen increased interest in green payment programs. In May 2001, Senator Tom Harkin introduced in the Senate the Conservation Security Act, under which farmers would receive payments for providing environmental services. A similar Act was also introduced in the house at the same time. There are two basic reasons for the strong interest in green payments.

First, it provides a foundation for farm support by society at large. Since Fiscal Year 1998, $70 billion has flowed to agriculture either as "emergency aid" or aid automatically triggered by low prices (Babcock 2001). If agriculture is to continue to receive the direct payments it has been receiving in recent years, it is likely to need to provide more substantial justification for them (Babcock 2000, Claassen et al 2001). Paying farmers to maintain beneficial impacts of agriculture (such as rural landscape amenities, wildlife habitat) or to mitigate adverse environmental impacts (such as soil erosion, runoff form nutrient and pesticide) can be an appealing rationale for farm support (Potter 1998).

Second, green payments can address agri-environmental problems that have not been adequately addressed. The Conservation Reserve Program and the Wetland Reserve Program provide environmental services by taking land out of production but do not have direct effects on land in production. Cost-share or incentive payment programs, such as Environmental Quality Incentive Program and the Wildlife Habitat Incentives Program, pay farmers for environment-friendly farming practices. However, when the cost share is less than 100 percent, farmers have no incentive to participate unless the targeted practices also provide private benefits. Further, Cox (2001) and Veneman (2001) provide evidence that there is demand for additional environmental management on land in production. Green payments are well positioned to meet such demands.

Lynch and Smith (1994) argue "what distinguishes green support programs from most traditional agricultural conservation and environmental programs is that they would explicitly support participants' farm incomes at the same time they purchase environmental benefits (page ii)". If green payments are to provide income support, they continue to argue green payment levels must exceed the private cost of taking environmental protection action (page iii). Horan, Shortle and Abler (1999) also suggest that green payments can provide a degree of income support if they more than cover the costs of providing the required environmental services.

Despite the strong interest shown for green payments, they are often just discussed but seldom analyzed as policy instruments (Heimlich and Claassen, 1998). While Wu and Babcock (1995, 1996) provide excellent analyses of the optimal design of green payments programs, income support is not considered as a goal of such programs. Claassen et al. (2001) provide some empirical evidence on how green payments may meet the dual goals of environmental protection and income support. They show that targeting payments to support the incomes of any specific group of farmers is unlikely to solve any given agri-environmental problem, and that targeting any specific agri-environmental problem may exclude many producers that qualify for income support.

In this paper we analyze the efficient design of green payments programs explicitly considering dual policy goals. Formally, we represent the two policy goals as two components in the policymaker's objective function. We begin by noting in a setting where the government is unrestricted in the set of variables that are contractible, efficient program design requires targeting policies to each type of farm and green payments can achieve its dual goals.

However, in practice it may be difficult to observe the true cost of engaging in conservation activities for a given farmer. Likewise, it may be politically infeasible to contract directly on farm size. These two potential noncontractibilities affect the feasible design of a green payments program. To model these noncontractibilities as constraints, we extend our analysis by adopting a mechanism design framework (e.g., Bourgeon and Chambers, 2000) where farm operations are characterized by two parameters: one representing conservation efficiency, and the other farm size.

We first examine the case when only farm size is contractible; this allows us to analyze welfare losses from the government's inability to contract directly on farm size. We find that for any farm size, relatively more efficient farmers obtain whatever income support the least conservation efficient farmers obtain. This is a direct result of incomplete information: farmers can always present themselves as any conservation type that brings them the highest payment. Thus, if large farmers also tend to be conservation efficient, providing them income support is unavoidable. Moreover, since the policymaker knows that on average farmers get positive rent and every farmer gets at least the income support of the least efficient farmer, net payments for the least (most) efficient farmers will be lower (higher) than in the complete information case. The policymaker adjusts payments to equalize the marginal cost of income support with its expected marginal benefit.

We next examine the case where both farm size and conservation efficiency are noncontractible but where the policymaker still wants to use a single policy instrument. Since the policymaker can no longer target a specific farm size, the policymaker has to rely on the correlation of conservation efficiency and farm sizes to make green payments more effective as an income support tool. Specifically, if small farmers also tend to be conservation efficient, it is better to pay conservation efficient farmers more than their conservation costs, since these extra payments will also act as income support for them. On the contrary, if small farmers tend to be conservation inefficient, then conservation inefficient farmers should be paid relatively more.

The rest of this paper is organized as follows. We lay out the basic elements of the model in the next section. In the third section, we examine green payments with complete information. Section 4 analyzes the situation when green payments can be targeted at small or large farmers, but conservation efficiency is private information. In the fifth section, we study the optimal design of green payments when both farm size and conservation efficiency are not contractible. Section 6 concludes this paper.

Model Setup

Farmers produce two types of goods: a market good q and an environmental good e. The market good generates revenue pq, where p is the market price of output. There is no market for the environmental good. It is costly to provide a positive level of e, so e equals zero in the absence of any external conservation incentives.

Farmers are characterized by two variables: farm size ϕ and conservation efficiency θ. For simplicity, we assume ϕ and θ have two levels: $\phi \in \Phi \equiv \{\phi_L, \phi_S\}$ and $\theta \in \Theta \equiv \{\theta_h, \theta_l\}$. We denote their joint and marginal distributions as P_{ij}, P_i, and P_j, respectively, where $\phi_i \times \theta_j \in \Phi \times \Theta$. The two variables may be correlated. For example, positive correlation may occur if large farmers may be able to adopt conservation practices more efficiently because they have efficient management. Negative correlation may also occur if small farmers can achieve large environmental benefits at relatively low cost because their land is very environmentally sensitive.

Denote the cost function of providing e and q as $c(q, e; \phi, \theta)$, then a higher θ is associated with lower total and marginal costs for a given level of e and q, i.e., $c_\theta(\cdot) < 0$, $c_{e\theta}(\cdot) < 0$. Following Bourgeon and Chambers (2000), the difference between smaller and larger farmers is assumed to be purely technical with larger farmers having everywhere lower marginal costs of production than smaller farmers, or formally, $c_\phi(\cdot) < 0$, $c_{q\phi}(\cdot) < 0$.

Let $\pi(e; \phi, \theta) = \max_q \{pq - c(q, e; \phi, \theta)\}$ be a farmer's profit given that she provides e. When no conservation services are provided, farmers with the same ϕ are assumed to have the same cost, then we have $\pi(0; \phi, \theta_l) = \pi(0; \phi, \theta_h)$. That is, when $e = 0$, small (or large) farmers have the same profit regardless of their conservation efficiency. Moreover, $\pi(0; \phi_S, \theta) < \pi(0; \phi_L, \theta)$. For income support to be relevant, we assume there is a target income $\bar{\pi}$, deemed desirable for a farmer by the policymaker, with $\pi(0; \phi_S, \theta) < \bar{\pi}$, and $\pi(0; \phi_H, \theta) \geq \bar{\pi}$, where $\theta \in \Theta$. Thus, small farmers do not achieve the target income even with zero environmental services, while large farmers do.

The policymaker intends to make transfers to farmers as incentives for the provision of environmental benefits and as a way of supporting farmers whose status quo income is below $\bar{\pi}$. We refer to such transfers as green payments and denote them as $t(\phi_i, \theta_j)$. The "non-market" benefits the policymaker derives from supporting small farmers' income are represented as $\tilde{w}(y)$, with

$$\tilde{w}(y) = \begin{cases} w(y), & \text{if } y \leq \bar{\pi}, \\ 0, & \text{if } y > \bar{\pi}, \end{cases} \tag{1}$$

where y is a farmer's income which is the sum of profits and green payments, i.e., $y \equiv \pi(e;\phi,\theta) + t(\phi_i,\theta_j)$; w is concave with $w_y \geq 0$, $w_{yy} \leq 0$ and $w_y(\bar{y}) = 0$. Thus, the policymaker only derives utility from supporting small farmers. The closer a farmer's income is to the parity income, the less benefit the policymaker derives from supporting them. Once a farmer's income exceeds the parity income, the policymaker will not want to support her any more.

Social surplus from farming consists of two parts: farmers' profit $\pi(\cdot)$ and environmental improvement. The social benefit of environmental improvement is denoted as $v(e)$ with $v' \geq 0$, $v'' \leq 0$. In a green payments program, the policymaker's problem is to choose levels for the environmental benefit and green payments to maximize the sum of social surplus and the "non-market" benefits from income support, minus the social cost of funding green payments. Funds for green payments are usually financed with some sort of distortionary tax. We denote the unit deadweight loss from such distortion as $\lambda > 0$, that is, for each dollar given to farmers, the social cost is λ dollars due to efficiency loss. An alternative explanation for λ is that it is the multiplier of the policymaker's budget constraint.

We will model green payments program as a truthful direct revelation mechanism $[e(\phi_i,\theta_j),\ t(\phi_i,\theta_j)]$. With such mechanism, the government offers farmers a menu of conservation levels and green payments and farmers can pick any one choice from the menu. This way of modelling enables us to focus on the main subject of this paper which is how green payments achieve the dual goals. In the next three sections, we will see that the policymaker's optimal decisions depend not only on farm sizes and conservation efficiency but also on how much information she has and can base policies on. Before we proceed, we list below the definitions of a few terms to avoid confusion,

Definition 1 *In a green payments program, the income $y(\phi, \theta)$ for a farmer of type (ϕ, θ) is the sum of her profits $\pi(e; \phi, \theta)$ and green payments. The net payment $\tau(\phi, \theta)$, is the transfer that is over and above her profit loss due to the provision of environmental services e, i.e.,*

$$y(\phi, \theta) = \pi(e; \phi, \theta) + t(\phi_i, \theta_j), \tag{2}$$

$$\begin{aligned} \tau(\phi, \theta) &= t(\phi, \theta) - [\pi(0; \phi, \theta) - \pi(e, \phi, \theta)] \\ &= y(\phi, \theta) - \pi(0; \phi, \theta). \end{aligned} \tag{3}$$

Green Payments With Complete Information

To design a green payments program with complete information, the policymaker can specify levels of e and t for each farmer to maximize net social surplus and the non-market benefits from supporting small farmers, i.e.,

$$\max_{e,t} \quad \sum_i \sum_j \left[v(e(\phi_i, \theta_j)) + \pi(e(\phi_i, \theta_j) + \tilde{w}\left(y(\phi_i, \theta_j)\right) - \lambda t(\phi_i, \theta_j)\right] P_{ij}, \tag{4a}$$

$$s.t. \quad y(\phi_i, \theta_j) \geq \pi(0; \phi_i, \theta_j), \quad \forall \;\; \phi_i \times \theta_j \in \Phi \times \Theta. \tag{4b}$$

The constraint requires voluntary participation since it is in general politically infeasible to require farmers to provide environmental services without compensation. By (1), for large farmers, $\tilde{w}\,(y) = 0$ and for small farmers, $\tilde{w}\,(y)$ is a concave function, $w(y)$. The optimal $e(\cdot)$, and $t(\cdot)$ are characterized below,

$$v_e(\hat{e}(\phi_i,\theta_j)) = -(1+\lambda)\pi_e\left(\hat{e}(\phi_i,\theta_j);\phi_i,\theta_j\right), \tag{5a}$$

$$\hat{t}(\phi_L,\theta_j) = \pi(0;\phi_L,\theta_j) - \pi_e\left(\hat{e}(\phi_L,\theta_j);\phi_L,\theta_j\right), \tag{5b}$$

$$\lambda = w_t\left[\hat{t}(\phi_S,\theta_j) + \pi(\hat{e}(\phi_S,\theta_j);\phi_S,\theta_j)\right], \tag{5c}$$

where "^" indicates the optimal green payments program with complete information. The marginal condition (5a) requires that marginal benefit equals marginal cost for conservation, which is the lost profit for providing conservation services. The optimal conditions for green payments to small and large farmers are different. In particular, large farmers' green payments just cover their conservation costs while small farmers' green payments are such that the marginal benefit from income support equals the marginal cost of transfer. If $\lambda < w_t(\pi(0;\phi_S,\theta_j))$, which we assume to be the case, then we will have $t(\phi_S,\theta_j) > \pi(0;\phi_S,\theta_j) - \pi(\hat{e}(\phi_S,\theta_j);\phi_S,\theta_j)$, that is, it is desirable for the policymaker to boost small farmers' income.

Remark 1 *With complete information, we have a separation of income support and conservation efficiency: a farmer's conservation level only depends on her conservation efficiency and her net payment only depends on her need for income support.*

Since the policymaker does not care about income support for large farmers, their conservation payments just cover their cost. However, for small farmers, their average net payments are positive because the policymaker is concerned with boosting their income.

Remark 2 *With complete information, green payments can achieve its dual policy goals efficiently by*

(1) specifying e for each farmer such that the marginal benefit from conservation equals its marginal cost;

(2) compensating farmers' conservation services not only according to their conservation efficiency but also according to farm size.

This result provides the base case for comparing the more realistic situation when one or both of the parameters are not contractible.

Green Payments With ϕ Contractible But θ Not

We turn now to the case where the policymaker may not always know farmers conservation efficiency. For example, how the adoption of conservation tillage affects a farmer's profit depends on many factors: natural resource endowments of the field, weather conditions, the farmer's years of experience and days of off-farm work, and what kind of equipment the farmer already has, etc. Although past studies (Chambers 1992, Bourgeon and Chambers 2000, and Hueth 2000) suggest that the policymaker may not be able to explicitly discriminate between small and large farmers, we study the case with ϕ contractible for two reasons. First, farm size, which can be measured by production acres, animal numbers, etc., is potentially targetable. Second, if targeting farm size is not feasible, it is interesting to understand the efficiency loss due to the policymaker's inability to target farm size.

If farm size is contractible, the policymaker can design policies for small and large farmers separately. Thus, we first discuss optimal green payments for large farmers and then discuss those for small farmers.

For Large Farmers

Since the policymaker is not constrained to support this group of farmers' income, her problem is to choose a menu of conservation levels and the corresponding green payments to maximize net social surplus,

$$\max_{e,t} \sum_j \left[v(e(\phi_L,\theta_j)) + \pi(e(\phi_L,\theta_j) - \lambda t((\phi_L,\theta_j)\right] P_{Lj} \tag{6a}$$

$$s.t. \;\; \pi\left(e(\phi_L,\theta_j),\phi_L,\theta_j\right) + t(\phi_L,\theta_j) \geq \pi(0;\phi_L,\theta_j), \tag{6b}$$

$$\pi\left(e(\phi_L,\theta_j),\phi_L,\theta_j\right) + t(\phi_L,\theta_j) \geq \pi\left(e(\phi_L,\theta_k),\phi_L,\theta_j\right) + t(\phi_L,\theta_k), \tag{6c}$$

where θ_j, $\theta_k \in \Theta$. The first set of constraints are the individual rationality constraints (IRs) for type (ϕ_i,θ_j), which we denote as (IR_{ij}). The IRs ensure that farmers' participation in the program is voluntary. By the revelation principle, an optimal mechanism requires truth-telling: for farmers to tell the truth, doing so has to maximize their income. This is specified by (6c), the incentive compatibility constraints (ICs) for type (ϕ_i,θ_j), which we denote as (IC_{ij}).

We can rewrite the problem as a two-stage problem,

$$\max_{e} \sum_j \left[v(e(\phi_L,\theta_j)) + \pi(e(\phi_L,\theta_j)\right] P_{Lj} + z_L^*\left(e(\phi_L,\theta_l),e(\phi_L,\theta_h)\right), \tag{7}$$

$$\text{where } z_L^*(\cdot) = \max_t \left\{ z_L(\cdot) \equiv -\lambda \textstyle\sum_j t((\phi_L,\theta_j)P_{Lj} : \text{ (6b) and (6c)} \right\}.$$

In the first stage, the optimal green payments are obtained for given conservation services and in the second stage the optimal conservation services are determined. Since this is a standard one-dimensional adverse selection model, we will not discuss it in detail. However, for comparison with other cases, we present a graphical solution to the first stage optimization.

In Figure 1, we depict all of the relevant constraints. Since (IR_{Lh}) will be satisfied if (IR_{Ll}) and (IC_{Lh}) are, only (IR_{Ll}) is drawn. The green payments that satisfy (IC_{Lh}) are along the line IC_{Lh} and the area above it, and those satisfying (IC_{Ll}) are along the line IC_{Ll} and the area below it. So, the green payments that satisfy both IRs and ICs are the shaded area. Graphically $z_L(\cdot)$ is a negatively-sloped line. Thus, the lowest point in the shaded area in Figure 1 is the unique point that satisfies both constraints and maximizes $z_L(\cdot)$. In the second stage, the optimal conservation levels are solved by substituting the optimal green payments into (7) and then maximizing with respect to $e(\phi_L, \theta_h)$ and $e(\phi_L, \theta_l)$.

Proposition 1 *For any given* $\{e(\phi_L, \theta_l), e(\phi_L, \theta_h)\}$ *, the solution to the first stage of (7) can be characterized as,*

$$t^*(\phi_L, \theta_l) = \pi(0; \phi_L, \theta_l) - \pi\left(e(\phi_L, \theta_l); \phi_L, \theta_l\right), \tag{8a}$$

$$t^*(\phi_L, \theta_h) = \pi(0; \phi_L, \theta_h) - \pi\left(e(\phi_L, \theta_h); \phi_L, \theta_h\right) + I(e(\phi_L, \theta_l)), \tag{8b}$$

The solution to the second stage of (7) can be characterized as,

$$v_e(e^*(\phi_L, \theta_h)) = -(1+\lambda)\pi_e\left(e^*(\phi_L, \theta_h); \phi_L, \theta_h\right), \tag{9}$$

$$v_e(e^*(\phi_L, \theta_l)) = -(1+\lambda)\pi_e\left(e^*(\phi_L, \theta_l); \phi_L, \theta_l\right) + \lambda I_e(e^*(\phi_L, \theta_l))\frac{P_{Lh}}{P_{Ll}}, \tag{10}$$

$$\textit{where,} \quad I(e) \equiv \pi\left(e; \phi, \theta_h\right) - \pi\left(e; \phi, \theta_l\right).$$

Here "*" indicates the optimal green payments program when farm size is targetable. From (8a) and (8b), we can obtain the net payments to each conservation type,

$$\tau^*(\phi_L, \theta_l) = 0, \qquad \tau^*(\phi_L, \theta_h) = I(e^*(\phi_L, \theta_l)).$$

We see that there is now no longer a separation of net payments and conservation efficiency: among large farmers, the net payments to conservation inefficient farmers is zero, while the net payments to conservation efficient farmers is the information rent they earn.

The results on conservation levels are standard: there is no distortion for the conservation efficient type, but the conservation level for the conservation inefficient type has to be adjusted down according to the information rent $I(\cdot)$. The probability ratio, $\frac{P_{Lh}}{P_{Ll}}$, is the hazard rate. The larger $\frac{P_{Lh}}{P_{Ll}}$ is, the larger the effect of information rent will be for any given level of e. In other words, when there are relatively more conservation efficient large farmers, the conservation service by conservation inefficient large farmers will be more expensive because more farmers can receive information rent.

For Small Farmers This case is different from the previous case because here the policymaker is concerned with boosting farmers' income. Otherwise the two problems are the same,

$$\max_{e,t} \sum_j \left[v(e(\phi_S,\theta_j)) + \pi(\phi_S,\theta_j) + w(y(\phi_S,\theta_j)) - \lambda t(\phi_S,\theta_j)\right] P_{Sj} \tag{11a}$$

$$s.t.\ \ \pi\left(e(\phi_S,\theta_j),\phi_S,\theta_j\right) + t(\phi_S,\theta_j) \geq \pi(0;\phi_S,\theta_j), \tag{11b}$$

$$\pi\left(e(\phi_S,\theta_j),\phi_S,\theta_j\right) + t(\phi_S,\theta_j) \geq \pi\left(e(\phi_S,\theta_k),\phi_S,\theta_j\right) + t(\phi_S,\theta_k), \tag{11c}$$

where θ_j, $\theta_k \in \Theta$. The policymaker may not want to push the left hand side of (11b) to the lower bound due to its presence in the objective function. The incentive constraints are essentially the same as the case for large farmers.

We again rewrite the problem as a two-stage optimization problem:

$$\max_{e} \sum_j \left[v(e(\phi_S,\theta_j)) + \pi(e(\phi_S,\theta_j)\right] P_{Sj} + z_S^*\left(e(\phi_S,\theta_l), e(\phi_S,\theta_h)\right), \tag{12}$$

$$\text{where } z_S^*(\cdot) = \max_t \left\{z_S(\cdot) : \text{ (11b) and (11c)}\right\},$$

$$\text{and } z_S(\cdot) = \sum_j \left[w(y(\phi_S,\theta_j)) - \lambda t(\phi_S,\theta_j)\right] P_{Sj}.$$

This problem is the same as (7) except that here the "non-market" benefits from income support are added to the objective function of the first stage. The solution to the first stage optimization is illustrated in Figure 2 and more details are provided in the appendix. The IR and IC constraints are the same as those in Figure 1. However, while in Figure 1 the isoquant for $z_L(\cdot)$ is a straight line, here the isoquant for $z_S(\cdot)$ is a concave curve due to the curvature of $w(\cdot)$. The green payments that would maximize $z_S(\cdot)$ without constraints (11b)-(11c) are denoted as $t^0 \left(= \{t^0(\phi_S,\theta_l), t^0(\phi_S,\theta_h)\}\right)$. If t^0 is below the feasible set[1], i.e., the shaded area, then the optimal green payments $\{t^*(\phi_S,\theta_l), t^*(\phi_S,\theta_h)\}$ must satisfy $t^*(\phi_S,\theta_h) > t^0(\phi_S,\theta_h)$. By the definition of $z_S(\cdot)$, we also know that $t^*(\phi_S,\theta_l) \leq t^0(\phi_S,\theta_l)$. Otherwise, we can increase the value of $z_S(\cdot)$ by reducing $t(\phi_S,\theta_l)$ and keeping $t(\phi_S,\theta_h)$ unchanged. Thus, the optimal green payments lie to the northwest of t^0. As we move further away from t^0 in the northwest direction, the

[1]There are three possible locations for t^0, below, in, or above the feasible set. We consider the first case because it is the most likely case. Conservation inefficient farmers have high cost of conservation, so relatively more green payments have to be provided for them to provide conservation services. Thus, when the optimal green payments (t^0) are designed without regard to incentive compatibility constraints, conservation efficient farmers would have incentive to misrepresent themselves, i.e., t^0 lies below the feasible set. The other two cases could be similarly analyzed.

value of $z_S(\cdot)$ decreases. The green payments that maximize $z_S(\cdot)$ are the tangent point between $z_S(\cdot)$ and the feasible set. Solving for the green payments at the tangent point and then using them to solve the second stage of (12), we get the following proposition,

Proposition 2 *For any given $\{e(\phi_S, \theta_l), e(\phi_S, \theta_h)\}$, the solution to the first-stage optimization of (12) satisfies*

$$\lambda = w_t(y^*(\phi_S, \theta_h))\frac{P_{Sh}}{P_S} + w_t(y^*(\phi_S, \theta_l))\frac{P_{Sl}}{P_S}, \tag{13a}$$

$$t^*(\phi_S, \theta_h) = \pi(0; \phi_S, \theta_h) - \pi\left(e(\phi_S, \theta_h); \phi_S, \theta_h\right) + I(e(\phi_S, \theta_l)) + \tau^*(\phi_S, \theta_l), \tag{13b}$$

$$t^*(\phi_S, \theta_l) = \pi(0; \phi_S, \theta_l) - \pi\left(e(\phi_S, \theta_h); \phi_S, \theta_l\right) + \tau^*(\phi_S, \theta_l). \tag{13c}$$

The solution to the second-stage optimization of (12) satisfies

$$v_e(e^*(\phi_S, \theta_h)) = -(1+\lambda)\pi_e\left(e^*(\phi_S, \theta_h); \phi_S, \theta_h\right), \tag{14}$$

$$v_e(e^*(\phi_S, \theta_l)) = -(1+\lambda)\pi_e\left(e^*(\phi_S, \theta_l); \phi_S, \theta_l\right) + \left[\lambda - w_t(y^*(\phi_S, \theta_h))\right] I_e(e^*(\phi_S, \theta_l))\frac{P_{Sh}}{P_{Sl}}. \tag{15}$$

From (13b)-(13c) and the individual rationality constraints, we have,

$$\tau^*(\phi_S, \theta_l) \geq 0, \qquad \tau^*(\phi_S, \theta_h) = I(e^*(\phi_S, \theta_l)) + \tau^*(\phi_S, \theta_l).$$

Remark 3 *Conservation efficient small farmers' net payment is the sum of information rent and income support for conservation inefficient small farmers. Just because a farmer gets information rent does not mean her income support should be less than those who do not get any information rent.*

This is a direct result of asymmetric information: by definition, information rent is due to private information which the policymaker cannot make use of. The income support levels $y^*(\phi_S, \theta_h)$ and $y^*(\phi_S, \theta_l)$ are chosen such that marginal cost of income support equals its expected marginal benefit. From (5c) and (13a) we know, for $\theta_j \in \Theta$,

$$w_t(\hat{y}(\phi_S, \theta_j)) = \lambda = w_t(y^*(\phi_S, \theta_h))\frac{P_{Sh}}{P_S} + w_t(y^*(\phi_S, \theta_l))\frac{P_{Sl}}{P_S}. \tag{16}$$

Since $y^*(\phi_S, \theta_h) \geq y^*(\phi_S, \theta_l)$, we have the following remark,

Remark 4 *When conservation cost is not known and income support can be targeted at small farmers, the income of conservation efficient small farmers is higher and the income of conservation inefficient small farmers is lower than the complete information case, i.e.,* $y^*(\phi_S, \theta_h) \geq \hat{y}(\phi_S, \theta_h)$, *and* $y^*(\phi_S, \theta_l) \leq \hat{y}(\phi_S, \theta_l)$, *or in terms of net payments,* $\tau^*(\phi_S, \theta_h) \geq \hat{\tau}(\phi_S, \theta_h)$, *and* $\tau^*(\phi_S, \theta_l) \leq \hat{\tau}(\phi_S, \theta_l)$.

Intuitively, the policymaker knows that conservation efficient small farmers receive information rent in addition to income support which is the net payment for conservation inefficient small farmers. Thus, if she makes marginal benefit from supporting conservation inefficient small farmers equal to marginal cost, λ, then marginal benefit from supporting conservation efficient small farmers will be lower than marginal cost, which is not optimal. So she will reduce conservation inefficient small farmers' income support to equalize the expected marginal benefit and marginal cost of income support.

Remarks 3 and 4 imply one important policy consequence of green payments when farm size is contractible but conservation cost is not: conservation efficient small farmers will receive excessive (comparing to first best) net payments, while conservation inefficient small farmers will receive inadequate net payments.

The condition for the conservation level of the efficient type, (14), is standard. Equation (15) differs from (10) by an additional term, $-w_t(y^*(\phi_S,\theta_h))I_e(\cdot)\frac{p_{Sh}}{p_{Sl}}$, which is due to the effect of information rent on income support. Now information rent has two impacts. On the one hand, it increases the cost of conservation by conservation inefficient small farmers, i.e., $\lambda I_e(\cdot)\frac{p_{Sh}}{p_{Sl}}$, which is the effect in standard adverse selection models and which we refer to as cost effect; on the other hand, it boosts conservation efficient small farmers' income, i.e., $w_t(y^*(\phi_S,\theta_h))I_e(\cdot)\frac{p_{Sh}}{p_{Sl}}$, which we refer to as income effect. Cost effect is not desirable for the policymaker while income effect is. How asymmetry of information will affect conservation inefficient small farmers' conservation level depends on the magnitudes of income and cost effects. If $[\lambda - w_t(y^*(\phi_S,\theta_h))]\, I_e(\cdot)\frac{p_{Sh}}{p_{Sl}} > 0$, then cost effect dominates and $e^*(\phi_S,\theta_l)$ will be less than the first best level. If the opposite is true, then income effect dominates, and $e^*(\phi_S,\theta_l)$ will be higher than the first best. Since $y^*(\phi_S,\theta_h) \geq y^*(\phi_S,\theta_l)$, equation (13a) implies $\lambda - w_t(y^*(\phi_S,\theta_h)) > 0$. Also, from the definition of $I(\cdot)$, we know $I_e(e) = \pi_e(e;\phi,\theta_h) - \pi_e(e;\phi,\theta_l) > 0$. Thus, $[\lambda - w_t(y^*(\phi_S,\theta_h))]\, I_e(\cdot)\frac{p_{Sh}}{p_{Sl}} > 0$, i.e., cost effect always dominates. In summary,

Remark 5 *When the policymaker derives utility from supporting small farmers' income, information rent has two effects: cost effect,* $\lambda I_e(e^*(\phi_S,\theta_l))\frac{p_{Sh}}{p_{Sl}}$, *and income effect,* $w_t(y^*(\theta_l,\pi_l))I_e(e^*(\phi_S,\theta_l))\frac{p_{Sh}}{p_{Sl}}$. *In the case where income is contractible but conservation cost is not, cost effect always dominates income effect, and so we always have* $e^*(\phi_S,\theta_l) < \hat{e}(\phi_S,\theta_l)$.

The intuition that cost effect always dominates is as follows. When income support can be targeted at small farmers, conservation efficient small farmers receive excessive income support, and the marginal benefit from supporting them is less than the marginal cost (λ). Thus $\lambda - w_t(y^*(\phi_S, \theta_h)) > 0$, that is an extra unit of conservation by conservation inefficient small farmers incurs more information cost than its contribution to income support.

Green Payments With Both ϕ and θ Uncontractible

When policies discriminating between small and large farmers are not politically feasible, the policymaker can no longer use policies that directly target farm sizes. Instead, she will have to explicitly treat all farmers the same. That is, farmers who provide the same level of conservation services will have to be paid the same regardless of their sizes. However, the policymaker may still design mechanisms that make farmers reveal their conservation efficiency truthfully. Thus, her problem is still to maximize social surplus plus the non-market benefits from income support subject to the individual rationality and incentive compatibility constraints, i.e.,

$$\max_{e,t} \sum_j \left[v(e(\theta_j)) + \tilde{\pi}\left(e(\theta_j); \theta_j\right) + w\left(y(\phi_S, \theta_j)\right) \frac{P_{Sj}}{P_j} - \lambda t(\theta_j) \right] P_j \tag{17a}$$

$$s.t. \qquad \pi\left(e(\theta_j); \phi_i, \theta_j\right) + t(\theta_j) \geq \pi(0; \phi_i, \theta_j), \tag{17b}$$

$$\pi\left(e(\theta_j); \phi_i, \theta_j\right) + t(\theta_j) \geq \pi\left(e(\theta_k); \phi_i, \theta_j\right) + t(\theta_k), \tag{17c}$$

where $\phi_i \in \Phi$, θ_j, $\theta_k \in \Theta$, and $\tilde{\pi}\left(e(\theta_j); \theta_j\right) = \sum_i \frac{P_{ij}}{P_j} \pi\left(e(\theta_j); \phi_i, \theta_j\right)$, which is the average income of farmers with conservation efficiency θ_j.

Although (17a)-(17c) look very similar to (11a)-(11c), they have a few important differences. The parameter ϕ does not appear in $e(\cdot)$ and $t(\cdot)$, because it is no longer targetable. For the same reason, the probability term P_{Sj} outside of the square brackets in

(11a) is replaced by P_j in (17a). The ratio $\frac{P_{Sj}}{P_j}$ in the square brackets gives us the proportion of small farmers among those with conservation efficiency θ_j. Then, $w\left(y(\phi_S, \theta_j)\right) \frac{P_{Sj}}{P_j}$ in the objective function is the expected benefits from income support for conservation type θ_j.

Before we characterize the optimal green payments program when farm size is not targetable and conservation efficiency is not known, we introduce one assumption to simplify analysis.

Assumption 1: $c_{e\phi}\left(e, q; \phi, \theta\right) = 0$.

The assumption assumes that marginal cost of conservation is independent of farm size, that is, conservation efficiency θ solely determines the marginal cost of conservation. The assumption is reasonable for the following reasons. First, we can always define θ in such a way that it ranks farmers' relative conservation efficiency regardless of the value of ϕ. Second, the notion that large farmers may be more efficient or otherwise is captured by the correlation between ϕ and θ. Third, when the assumption does not hold, the analysis and results will be similar, although a different set of constraints may become binding.

There are two participation and incentive constraints for each $\phi \in \Phi$. However, with assumption 1, the constraints for small and large farmers with the same conservation efficiency either both hold or both do not hold. Formally, since $\pi_{\phi e} = -c_{\phi e} = 0$, we know $[\pi\left(e(\theta_i), \phi_L, \theta\right) - \pi\left(e(\theta_j), \phi_L, \theta\right)] - [\pi\left(e(\theta_i), \phi_S, \theta\right) - \pi\left(e(\theta_j), \phi_S, \theta\right)] = \int_{\phi_S}^{\phi_L} \pi_\phi\left(e(\theta_i), \phi, \theta\right) - \pi_\phi\left(e(\theta_j), \phi, \theta\right) d\theta = 0$. Thus, for ϕ_i, $\phi_k \in \Phi$, and $\theta \in \Theta$, if ($\text{IR}_{i\theta}$) holds, then ($\text{IR}_{k\theta}$) must also hold. Likewise, if ($\text{IC}_{i\theta}$) holds, then ($\text{IC}_{k\theta}$) must also hold. Therefore, in the following, we will just use θ, implying it could either be $(\phi_L,\ \theta)$ or $(\phi_S,\ \theta)$.

The two-stage version of (17) is as follows,

$$\max_{e} \sum_j \left[v(e(\theta_j)) + \bar{\pi}\left(e(\theta_j), \theta_j\right)\right] P_j + z^{**}\left(e(\theta_l), e(\theta_h)\right) \tag{18}$$

$$\text{where } z^{**}(\cdot) = \max_{t} \left\{z(\cdot) : \text{ (17b) and (17c)}\right\},$$

$$\text{and } z(\cdot) = \sum_j \left[w\left(y(\phi_S, \theta_j)\right) \frac{P_{Sj}}{P_j} - \lambda t(\theta_j)\right].$$

Because of the similarity between (12) and (18), we can use the same procedures in the previous section. The graphics for the first stage optimization of the two problems look almost the same except the differences in the labels. Thus we will state the solution to (18) without further discussion on their derivations.

Proposition 3 *The solutions to the first stage of (18) can be characterized as follows,*

$$\lambda = P_S \left[w_t(y^{**}(\phi_S, \theta_h)) \frac{P_{Sh}}{P_S} + w_t(y^{**}(\phi_S, \theta_l)) \frac{P_{Sl}}{P_S}\right], \tag{19a}$$

$$t^{**}(\theta_h) = \pi(0; \phi, \theta_h) - \pi\left(e^{**}(\theta_h); \phi, \theta_h\right) + I(e^{**}(\theta_l)) + \tau^{**}(\theta_l), \tag{19b}$$

$$t^{**}(\theta_l) = \pi(0; \phi, \theta_l) - \pi\left(e^{**}(\theta_l); \phi, \theta_l\right) + \tau^{**}(\theta_l). \tag{19c}$$

The solution to the second stage of (18) satisfies

$$v_e(e^{**}(\theta_h)) = -(1+\lambda)\pi_e\left(e^{**}(\theta_h); \phi, \theta_h\right), \tag{20}$$

$$v_e(e^{**}(\theta_l)) = -(1+\lambda)\pi_e\left(e^{**}(\theta_l); \phi, \theta_l\right) + \left[\lambda - \tfrac{P_{Sh}}{P_h} w_t(y^{**}(\phi_S, \theta_h))\right] I_e(e^{**}(\theta_l)) \tfrac{P_h}{P_l}. \tag{21}$$

Here "**" indicates the optimal green payments program when both ϕ and θ are not targetable. Equation (19a) requires the marginal cost of income support equals its expected marginal benefit. Equations (13a) and (19a) are the same except the additional term P_S in (19a), which is the proportion of small farmers. The additional term is due to the policymaker's inability to target income level. The term in the brackets is the expectation of marginal benefits from supporting small farmers' income. Multiplying it by P_S gives us the expectation of marginal benefit from income support when both ϕ and θ are not contractible.

From (13a) and (19a), we have

$$\begin{aligned}
& w_t(y^{**}(\phi_S, \theta_h))\frac{P_{Sh}}{P_S} + w_t(y^{**}(\phi_S, \theta_l))\frac{P_{Sl}}{P_S} \\
= \; & \frac{\lambda}{P_S} \geq \lambda \\
= \; & w_t(y^{*}(\phi_S, \theta_h))\frac{P_{Sh}}{P_S} + w_t(y^{*}(\phi_S, \theta_l))\frac{P_{Sl}}{P_S}.
\end{aligned}$$

The first and the last expressions are the averages of marginal benefit from different levels of income support. Define the average income support as the weighted average of income support for the two conservation types. Then, by the concavity of $w(\cdot)$, we have the following remark,

Remark 6 *For the case where conservation efficiency is not contractible, when $w_t(\cdot)$ is linear (i.e., $w(\cdot)$ is quadratic), average income support is higher when farm size is contractible than when it is not.*

The intuition is that, when farm size is not contractible, income support becomes more expensive because part of it will be "wasted" on large farmers whose income support does not generate benefit. As a result, the marginal benefit of income support is lower relative to the case where green payments can target farm size. Thus, average income support is decreased to equalize the marginal benefit and marginal cost of transfer.

From equations (19b) and (19c), we can derive the net payments to each conservation type of farmers,

$$\tau^{**}(\phi,\theta_l) \geq 0, \qquad \tau^{**}(\phi,\theta_h) = I(e^{**}(\theta_l)) + \tau^{**}(e^{**}(\theta_l),\theta_l).$$

Thus, in a dual goal green payments program, farmers who have high conservation efficiency benefit the most, while farmers who have low conservation efficiency benefit the least.

Remark 7 *In the case where neither farm size nor conservation efficiency is contractible, a green payments program is most effective as a tool of income support when small farmers also tend to be conservation efficient.*

As to the optimal conservation level, (20) says that there is no distortion for efficient farmers' conservation level. Equation (21) differs from (15) by two terms, both of which are due to the policymaker's inability to target farm size. The first is the additional term $\frac{P_{Sh}}{P_h}$ in the square brackets, and the second is the replacement of $\frac{P_{Sh}}{P_{Sl}}$ by $\frac{P_h}{P_l}$. When farm size is not contractible, the policymaker can only take expectations of marginal benefit of income support for any conservation type. The term $\frac{P_{Sh}}{P_h}$ is the probability that a farmer is small given that she is conservation efficient. Multiplying marginal benefit by $\frac{P_{Sh}}{P_h}$ gives us the expected marginal benefit from conservation efficient farmers. Also, the policymaker cannot use the relative proportion of a conservation type within either small or large farmers, and so $\frac{P_{Sh}}{P_{Sl}}$ is replaced by $\frac{P_h}{P_l}$.

From (19a), we have $w_t(y^{**}(\phi_S,\theta_h)) \leq \frac{\lambda}{P_S} \leq w_t(y^{**}(\phi_S,\theta_l))$, so $\lambda \geq \frac{P_{Sh}}{P_h} w_t(y^{**}(\phi_S,\theta_h))$ if $\frac{P_{Sh}/P_h}{P_S} \leq 1$. If $\frac{P_{Sh}/P_h}{P_S} > 1$, then there is an interval between and $\lambda\frac{P_h}{P_{Sh}}$ and $\lambda\frac{1}{P_S}$. If

$w_t(y^{**}(\phi_S, \theta_h))$falls into this interval, then $\lambda < \frac{P_{Sh}}{P_h} w_t(y^{**}(\phi_S, \theta_h))$.

Remark 8 *If* $\frac{P_{Sh}/P_h}{P_S} \leq 1$, *then cost effect dominates income effect. If* $\frac{P_{Sh}/P_h}{P_S} > 1$, *and* $w_t(y^{**}(\phi_S, \theta_h)) \in \left[\lambda \frac{P_h}{P_{Sh}},\ \lambda \frac{1}{P_S}\right]$, *then income effect dominates.*

The term, $\frac{P_{Sh}/P_h}{P_S}$, is the ratio of proportion of conservation efficient small farmers to the proportion of all small farmer. The above remark basically says when small farmers also tend to be conservation efficient, income effect is more likely to dominate; otherwise, cost effect is more likely to dominate. So, the joint distribution of ϕ and θ, especially, the correlation between these two parameters is very important in determining the conservation level for conservation inefficient farmers, information rent received by conservation efficient farmers, and the total income for each group. In particularly, we have the following remark,

Remark 9 *For given* P_h, P_l, P_S, *and* P_L, *i.e., for given marginal distributions of* ϕ *and* θ,

(i) $e^{**}(\theta_l)$ *increases with* P_{Sh};

(ii) Information rent increases with P_{Sh};

(iii) $y^{**}(\phi_S, \theta_l)$ *decreases with* P_{Sh}.

For a proof, see appendix. From (21), ceteris paribus, we know the higher $\frac{P_{Sh}}{P_h}$ is, the more likely income effect will dominate, and the higher $e^{**}(\theta_l)$ will be. In other words, when conservation efficient farmers also tend to be small farmers, they will provide more conservation service. This is because, the information rent received by them acts as income support, which may offset part or all of its cost effect. Remark (ii) directly comes from (i). As $e^{**}(\theta_l)$ increases, conservation efficient farmers will have a greater incentive to pretend as inefficient farmers. Greater information rent is provided to counteract this incentive. As conservation efficient farmers receive more information rent, they drag down the expected marginal benefit of income support and make income support more expensive. As a result, less payments go to conservation inefficient small farmers.

Conclusions

This paper provides answers to the question how green payments programs may meet the dual goal of environmental protection and income support. Given that the government provides payments to farmers for conservation services, it makes sense to pay farmers more than just their conservation costs so as to boost their income. Thus, green payments appear to be able to achieve the dual goals efficiently. This is the case when there is complete information.

When there is asymmetric information, things are different. First of all, if the government does not know farmers' conservation costs, which is generally true, it cannot decide how much green payments should be. In such situations, farmers who can do conservation at a lower cost will benefit more from such a program. If large farmers who are not supposed to receive income transfer are also more conservation efficient, then it is unavoidable that they will benefit more from a green payments program.

Second, if means test cannot be used, every farmer's green payments will have to exceed her conservation cost. Inevitably, farmers who are not intended to benefit from green payments, namely large farmers, will derive benefit from them. Since the budget for green payments is limited, the income support for small farmers will be affected.

Thus, with limited information or targeting tools, how green payments can achieve the dual goals hinges on the overlapping of small farmers and conservation efficient farmers. Green payments are likely to be more efficient when small farmers also tend to be more conservation efficient.

Appendix

Details on Figure 2 and the solution to first-stage optimization of (12)

Proof. By the definition of $z_S(\cdot)$, we know the green payments that maximize $z_S(\cdot)$ without any constraints satisfy

$$w_t(\pi(e(\phi_S,\theta_l);\phi_S,\theta_l)+t^0(\phi_S,\theta_l))=\lambda=w_t(\pi(e(\phi_S,\theta_h);\phi_S,\theta_h)+t^0(\phi_S,\theta_h)),$$

that is, the marginal benefit of income support equals the marginal cost for either conservation type. As we discussed in the text, if t^0 is below the feasible set, then the optimal green payments in the feasible set must lie to the northwest of t^0, that is among those payments satisfying

$$t(\phi_S,\theta_h)>t^0(\phi_S,\theta_h) \quad \text{and} \qquad t(\phi_S,\theta_l)\leq t^0(\phi_S,\theta_l). \tag{22}$$

We can totally differentiate $z_S(\cdot)=\sum_j\left[w(y(\phi_S,\theta_j))-\lambda t(\phi_S,\theta_j)\right]P_{Sj}$, for green payments satisfying (22), we get

$$\begin{aligned}\frac{dt(\phi_S,\theta_h)}{dt(\phi_S,\theta_l)} &= \frac{P_{Sl}}{P_{Sh}}\frac{-[w_t(y(\phi_S,\theta_l))-\lambda]}{[w_t(y(\phi_S,\theta_h))-\lambda]}>0,\\ \frac{d^2t(\phi_S,\theta_h)}{dt^2(\phi_S,\theta_l)} &= \frac{P_{Sl}}{P_{Sh}}\frac{w_{tt}(y(\phi_S,\theta_h))[w_t(y(\phi_S,\theta_l))-\lambda]-w_{tt}(y(\phi_S,\theta_l))[w_t(y(\phi_S,\theta_h))-\lambda]}{[w_t(y(\phi_S,\theta_h))-\lambda]^2}<0\end{aligned}$$

The above derivatives indicate that the isoquant of $z_S(\cdot)$ is concave in the green payments space. Also, as we move toward the northwest, i.e., as $t(\phi_S,\theta_h)$ increases or $t(\phi_S,\theta_l)$ decreases, or both, the value of $z_S(\cdot)$ decreases. This is because for green payments satisfying (22), $\frac{dz_S(\cdot)}{dt(\phi_S,\theta_h)}=[w_t(y(\phi_S,\theta_h))-\lambda]\,P_{Sh}<0$ and $\frac{dz_S(\cdot)}{dt(\phi_S,\theta_l)}=[w_t(y(\phi_S,\theta_l))-\lambda]\,P_{Sh}>0$, by the concavity of $w(\cdot)$. Thus the optimal green payments must be the closest to

t^0. The tangent point in Figure 2 is the closest to t^0 in the feasible set and so it is the optimal point. ■

Proof for Remark 9

Proof. (i) Since $e^{**}(\theta_h)$ is independent of $e^{**}(\theta_l)$ and $t^{**}(\theta_l)$. We may split the problem into two maximization problems, one for $e^{**}(\theta_h)$, and the other for $e^{**}(\theta_l)$ and $t^{**}(\theta_l)$, ($t^{**}(\theta_h)$ is substituted by a function of $t^{**}(\theta_l)$). The maximization problem with respect to $e^{**}(\theta_l)$ and $t^{**}(\theta_l)$ is,

$$\max_{e,t} \tilde{U}\left(t(\theta_l), e(\theta_l); P_{Sh}\right) \equiv \left[v(e(\theta_l)) + \pi\left(e(\theta_l); \phi_S, \theta_h\right) - \lambda t(\theta_l)\right] g_l$$

$$+ \left[-\lambda\left(t(\theta_l) + \pi\left(e(\theta_l); \phi_S, \theta_h\right)\right)\right] g_h$$

$$+ w\left(t(\theta_l) + \pi\left(e(\theta_l); \phi_S, \theta_l\right)\right) P_{Sl} + w\left(t(\theta_l) + \pi\left(e(\theta_l); \phi_S, \theta_h\right)\right) P_{Sh}.$$

By comparative statics, we have,

$$\frac{de(\theta_l)}{dP_{Sh}} = \frac{\tilde{U}_{et}\tilde{U}_{Pt} - \tilde{U}_{tt}\tilde{U}_{eP}}{|H|} \tag{23}$$

where $\tilde{U}_{et} \equiv \frac{\partial \tilde{U}(t(\theta_l), e(\theta_l))}{\partial e(\theta_l)\partial t(\theta_l)}$, $\tilde{U}_{Pt} \equiv \frac{\partial \tilde{U}(t(\theta_l), e(\theta_l))}{\partial P_{Sh}\partial t(\theta_l)}$, $\tilde{U}_{tt} \equiv \frac{\partial \tilde{U}(t(\theta_l), e(\theta_l))}{\partial t(\theta_l)\partial t(\theta_l)}$, $\tilde{U}_{eP} \equiv \frac{\partial \tilde{U}(t(\theta_l), e(\theta_l))}{\partial e(\theta_l)\partial P_{Sh}}$, and $|H| = \begin{vmatrix} \tilde{U}_{ee} & \tilde{U}_{et} \\ \tilde{U}_{et} & \tilde{U}_{tt} \end{vmatrix}$. From second order conditions, we know $|H| > 0$. Substituting the derivatives into (23), and then simplifying, we obtain,

$$\frac{de(\theta_l)}{dP_{Sh}} = \pi_e\left(e^{**}(\theta_l); \phi, \theta_l\right) - \pi_e\left(e^{**}(\theta_l); \phi, \theta_h\right) \times \left[w'_{Sh}w''_{Sl}P_{Sh} + w'_{Sl}w''_{Sh}P_{Sl}\right] > 0$$

where $w'_{Sh} = w_t(y^{**}(\phi_S, \theta_h))$, $w''_{Sh} = w_{tt}(y^{**}(\phi_S, \theta_h))$, $w'_{Sl} = w_t(y^{**}(\phi_S, \theta_l))$, and $w''_{Sl} = w_{tt}(y^{**}(\phi_S, \theta_l))$.

(ii) The definition of information rent is, $I(e; \phi) \equiv \pi\left(e; \phi, \theta_h\right) - \pi\left(e; \phi, \theta_l\right)$. Then, $\frac{dI(e;\phi)}{de} = \pi_e\left(e; \phi, \theta_l\right) - \pi_e\left(e; \phi, \theta_h\right) > 0$, because $\pi_{e\theta} = -c_{e\theta} > 0$.

(iii) Suppose, as P_{Sh} increases, $y^{**}(\phi_S, \theta_l)$ increases or does not change. When P_{Sh} increases, P_{Sl} decreases by the same degree since the marginal distributions remain the same, in particular, P_S does not change. Thus, given that $y^{**}(\phi_S, \theta_l)$ increases or does not change, we need $w'(y^{**}(\phi_S, \theta_h))$ to increase for (19a) to hold. However, when $y^{**}(\phi_S, \theta_l)$ increases or does not change, $y^{**}(\phi_S, \theta_h) = y^{**}(\phi_S, \theta_l) + I(e; \phi)$ increases because of (ii), which in turn implies $w'(y^{**}(\phi_S, \theta_h))$ decreases. Thus, $y^{**}(\phi_S, \theta_l)$ must decrease. ∎

References

[1] Babcock, Bruce A. "Positioning U.S. Agriculture for the Next Farm Bill", Presented at the American Farm Bureau Federation's convention and annual meeting, January 8, 2001.

[2] Babcock, Bruce A. "A Farm Policy Objective: the Search Goes On". Iowa Ag Review, Fall 2000, Vol. 6 No.4, pp. 1-3.

[3] Bourgeon, Jean-Mark and Robert G. Chambers. "Stop-and-Go Agricultural Policies". *Amer. J. Agr. Econ.* 82 (February 2000): 1-13.

[4] Chambers, Robert. "On the Design of Agricultural Policy Mechanisms". *Amer. J. Agr. Econ.* 74 (August 1992): 646-53.

[5] Claassen, R., L. Hansen, M. Peters, V. Breneman, M.Weinberg, A. Cattaneo, P. Feather, D. Gadsby, D. Hellerstein, J. Hopkins, P. Johnston, M. Morehart, and M. Smith. "Agri-Environmental Policy at the Crossroads: Guideposts on a Changing Landscape," ERS/USDA, Agricultural Economic Report No. 794, January 2001.

[6] Cox, Craig A., "New Authorities and Additional Funding Needed to Ensure Improved Stewardship of Agricultural Resources", Presented at the conference, *Fixing the Farm Bill*, March 27, 2001, The National Press Club Washington, D.C.

[7] Heimlich, Ralph, E. and Roger Claassen. "Agricultural Conservation Policy at a Crossroads", *Agricultural and Resource Economics Review*, v27, n1 (April 1998): 95-107.

[8] Horan, Richard D., James S. Shortle, and David G. Abler. "Green Payments for Nonpoint Pollution Control", *Amer. J. Agr. Econ.* 81 (Number 5, 1999): 1210-121.

[9] Hueth, Brent. "The Goals of U.S. Agricultural Policy: A Mechanism Design Approach". *Amer. J. Agr. Econ.* 82 (February 2000): 14-24.

[10] Lynch, Sarah and Katherine R. Smith. "Lean, Mean and Green... Designing Farm Support Programs in a New Era", Policy Studies Program Report No. 3, December 1994, Henry A. Wallace Institute For Alternative Agriculture.

[11] Potter, Clive. *Against the Grain: Agri-environmental Reform in the United States and the European Union*", New York: CAB International, 1998.

[12] Veneman, Ann M. Remarks at the Sparks Companies 9th Annual Food & Agriculture Policy Conference, April 17, 2001 Washington, D.C., Release No 0065.01. http://www.usda.gov/news/releases/2001/04/0065.htm, accessed on May 22, 2001.

[13] Wu, Junjie and Bruce A. Babcock. "Optimal Design of a Voluntary Green Payment Program under Asymmetric Information", *Journal of Agricultural and Resource Economics* 20(2): 316-327.

[14] Wu, Junjie and Bruce A. Babcock. "Contract Design for the Purchase of Environmental Goods from Agriculture", *Amer. J. Agr. Econ.* 78 (November, 1996): 935-945.

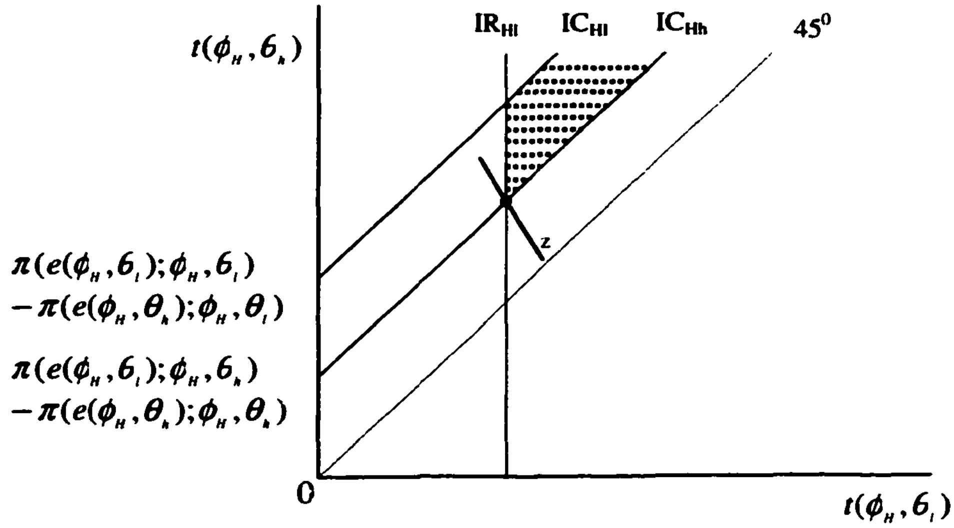

Figure 1: Transfer optimization for given $e(\phi_H, \theta_l)$ and $e(\phi_H, \theta_l)$ ------large farmers

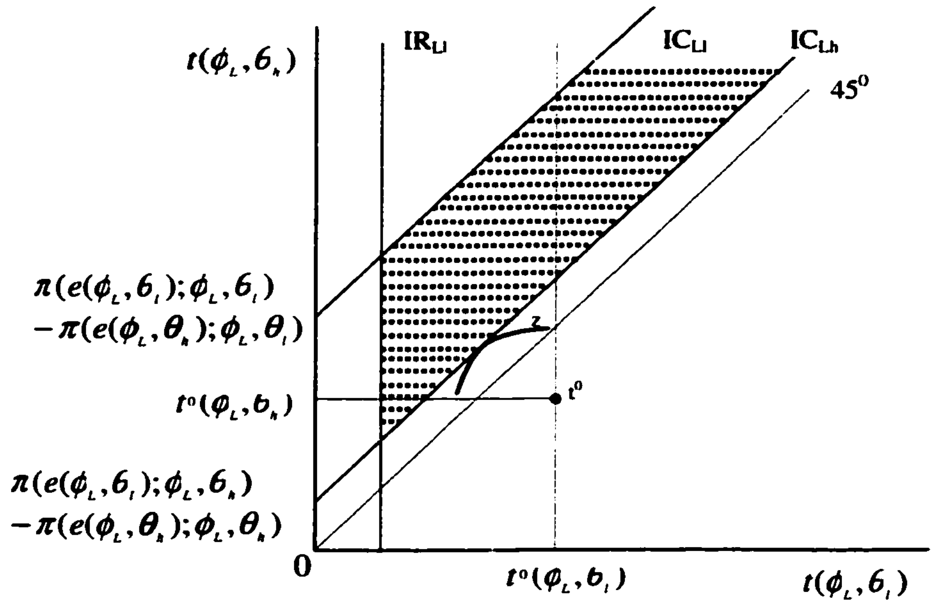

Figure 2: Transfer optimization for given $e(\phi_L,\theta_l)$ and $e(\phi_L,\theta_l)$ ---small farmers

CHAPTER 5.
GENERAL CONCLUSIONS

This dissertation includes three chapters that investigate the design of environmental incentives under dynamic behavior, asymmetric information and dual policy goals. First, I explore the incentive design for policies to sequester carbon in agricultural soils. Until policy mechanisms that appropriately incorporate the potentially temporary nature of sinks are developed, it is unlikely that agricultural sequestration will gain widespread acceptance. I introduce and discuss three such mechanisms, the PAYG System, the VLC System, and the CAA System. These mechanisms could be implemented in the context of either a private trading market or a government program (such as green payments), although we explain them in the context of a well-functioning external carbon market that determines the price of carbon abatement.

Common to all three systems are the issues of effective monitoring and enforcement, agreement on a baseline for measurement, and potential leakage (i.e. substitution of emissions from one location to another). Despite these concerns, there is ample reason to be optimistic that effective market mechanisms or government programs can be devised to include agricultural soils in an effective greenhouse gas policy.

The next chapter discusses the design of alternative bankable permit regimes. I find that the efficiency property of a bankable permit regime depends on the structure of uncertainty as well as the structure of benefit and damage functions of a pollutant. The more firms' abatement costs vary over time (i.e., have large variances), the more efficient a bankable permit regime tends to be. Moreover, negative correlations among shocks over time favor the use of bankable permits, because there is potential for cost or benefit smoothing over time. Whether allowing banking is welfare improving largely depends on the slopes of the marginal benefit and marginal damage curves. Our results have very important implications for policy makers. It may be valuable to allow banking in some tradable emission permit programs. Whether intertemporal trading should be allowed depends on the particular pollutants—both how they affect firms' profits and the social damages.

Further work on this topic may prove fruitful. If the regulator can adjust its rules when shocks are revealed more efficient designs should be possible. It may also be useful to investigate how firms' investment decisions are affected by a bankable permit system with non-unitary intertemporal trading ratios. Finally the exploration of multiplicative, instead of additive uncertainty, in the benefit functions may also yield additional insights.

In chapter four, I investigate the design of green payments in the presence of dual policy goals. Ideally, green payments encourage more environmental services from more conservation efficient farmers and, at the same time, provides income support for those who really need it, i.e., low income farmers. I find that, under incomplete information and the infeasibility of directly targeting transfers to low income farmers, the extent to which green payments can realize these dual goals depends on the correlation between a farmers' income and her conservation efficiency. Empirical work that provides evidence on the magnitude of this correlation could provide further insight into the efficacy of green payments.