ESSAYS ON THE ECONOMICS OF LAND USE AND WATER QUALITY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jordan Frederick Suter

January 2008

© 2008 Jordan Frederick Suter

ESSAYS ON THE ECONOMICS OF LAND USE AND WATER QUALITY

Jordan Frederick Suter, Ph. D.

Cornell University 2008

This dissertation is a collection of three essays that address the economic and environmental effects of land use spillovers. The first essay explores mechanisms for controlling nonpoint source water pollution, given that water quality in a watershed is a function of spillovers from agricultural production. The second essay considers the selection of an optimal wildlife corridor, on the basis that parcel values are a function of nearby land use. The third essay estimates the effect of spillovers from open space land on residential property prices. Taken together, the essays address enduring environmental economics issues through the application of contemporary economic tools.

The first essay proposes a mechanism for addressing ambient pollution through a background threat of regulation which induces nonpoint source polluters to voluntarily reduce emissions. Specifically, the severity of the threatened tax policy is endogenous to voluntary stage outcomes. Beyond showing the mechanism's theoretical properties, the essay highlights a set of economics experiments in which participants are faced with a voluntary-threat policy.

The second essay introduces a model for determining an optimal wildlife corridor that, in a departure from previous corridor research, accounts for *both* the costs and benefits of individual parcels. By combining parcel-level costs and benefits for land areas connecting three ecosystems in the Northern Rockies, the economic tradeoffs inherent in corridor design are empirically examined. By varying the granularity of the available parcels, the study explores issues related to model

specificity and computational complexity. A heuristic is also proposed that offers considerable computational advantages and approximates the optimal corridor.

The third essay applies hedonic pricing methods to estimate marginal implicit prices for open space land and other landscape amenities across distinct urban, suburban and rural submarkets in Rochester, NY, rather than the standard approach which assumes spatial uniformity. Results, controlling for spatial autocorrelation and the endogeneity of nearby open space, suggest that public open space is most highly valued in urbanized areas and that private open space is relatively valuable in rural areas. Additionally, proximity to concentrated animal feeding operations and other locational disamenitites have a significantly negative impact on home prices.

BIOGRAPHICAL SKETCH

Jordan Suter was born and raised in Webster, New York. He graduated from Vanderbilt University in May of 1999 with Honors in Economics and a second major in Political Science, as well as a minor in Environmental Science. After graduation, Jordan worked for two years as a Research Associate at the Law and Economics Consulting Group (LECG) in Washington, DC. He enrolled at Cornell University in the fall of 2001 and completed a M.Sc. in Applied Economics in 2003. In the fall of 2004, Jordan spent one semester as a visiting lecturer at Wells College in Aurora, NY, where he taught environmental economics. After graduation, he will be starting as an Assistant Professor in the Department of Economics at Oberlin College in the fall of 2007. In his free time, Jordan enjoys spending time with family and friends and engaging in a wide array of outdoor sports.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the incredible community that I have been surrounded with during my time in Ithaca. That community includes the many faculty members at Cornell that have fostered my learning and growth, my family, which has provided enduring support throughout this trying process, and the great set of friends both in Ithaca and elsewhere that have offered a helpful ear and many essential diversions.

While my education has been fostered by input from a broad range of faculty members throughout the last six years, my advisor Greg Poe has been the most consistent and influential source of guidance. From working with me on the set of essays that comprise this dissertation, to serving as the chairperson of my Master's committee, to coathoring several journal articles, Greg has provided me with the academic tools and confidence that have been essential to my success. He has also shown me the importance of the appropriate framing and presentation of academic research. I have been extremely fortunate in the degree to which Greg has supported my future in academia. This has manifested itself in his support for me while I taught for a semester at Wells College, despite the fact that it took time away from working as his research assistant. He has also been extremely generous in authorship decisions and has spent considerable time tailoring reference letters to specific academic positions.

The other three members of this dissertation committee have also contributed immensely to the document that follows. Jon Conrad has been essential in helping to spur my interest in wildlife corridor design and for helping to direct the line of research that has been followed. The work of Carla Gomes has also been critical in regards to formalizing the corridor model and in bringing her programming expertise to bear on a very challenging problem. It has been a lot of fun to work so closely with two faculty members from relatively different backgrounds on a project that has required both creativity and dedication. Finally, thanks to Dick Boisvert, who has really helped to improve the presentation of each of the essays and has directed my energies towards highlighting the policy relevance of this work. His effort in thoroughly reading and editing each of the essays is greatly appreciated and has brought the collection of essays to a higher level.

Outside of my committee, I would also like to acknowledge the support of Kathleen Segerson and Christian Vossler. The have both helped in the conceptualization of the first essay and have also provided some very helpful editorial suggestions. Nelson Bills, David Kay and Carolina Magri also deserve recognition for their considerable help in the data collection and processing for the third essay. They have also helped to provide guidance and feedback throughout the analysis.

The Cornell faculty has provided outstanding technical and educational guidance, but it has been my family and friends that have provided the essential love and support throughout the degree process. Having family so close by was one of my primary motivations for coming to Cornell six years ago and it is a decision that I am immensely thankful for. While it may not be possible to quantify the importance that a home cooked meal and a hug from the family provide, I honestly believe that my degree would not have been completed without them.

I would also like to thank all of the friends that have made Ithaca such a great place to be. In particular, thanks to my girlfriend Meike, who has been an amazing source of support and guidance through the fun and exciting times as well as the challenging times. I look forward to the fun and challenges that life has in store for us in the future. Last but not least, thanks to my dog Rio for reminding me that getting outside is sometimes more important than sitting in front of the computer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE:
Voluntary-Threat Mechanisms to Reduce Ambient Water Pollution

Abstract

This essay focuses on voluntary-threat policies that provide incentives to producers to reduce agricultural nonpoint source water pollution. The policies allow a group of polluters to meet a pollution target voluntarily, under the threat of a mandatory tax policy that is implemented only in the case of voluntary noncompliance. In a departure from the voluntary-threat mechanism proposed by Segerson and Wu (*Journal of Environmental Economics and Management*, 2006), in which the threatened tax mechanism is determined exogenously, a threatened tax mechanism is introduced where tax liabilities are conditional on the level of noncompliance in the voluntary setting. This "endogenous" voluntary-threat mechanism generates voluntary, least-cost compliance as part of a unique subgame perfect Nash equilibrium, without the need for retroactive taxation. Laboratory economics experiments are used to investigate both the Segerson and Wu mechanism and the voluntary-threat mechanism introduced in this essay. Results indicate that the voluntary-threat policies do not generally result in ambient standards being achieved voluntarily unless communication occurs among participants or the threatened tax policy is made severe.

1.1 Introduction

Improvements in surface water quality since the passage of the Federal Clean Water Act Amendments of 1972 have come primarily as a result of emissions reductions from point sources, such as wastewater treatment plants and factories. While opportunities for further point source reductions remain, nonpoint source pollution presently represents the greatest share of surface water impairment in the United States (Ribaudo 2003). Agricultural production, which occurs on approximately 60% of nonfederal land in the US (NRI 2002), is the largest contributor to nonpoint source water pollution and the leading source of water quality impairments among the rivers and lakes surveyed in the 2000 *National Water Quality Inventory* (US EPA 2002).

Given the influential role of nonpoint sources, economic theorists have devised a number of mandatory approaches designed to reduce surface water pollution stemming from agricultural production. These approaches can be roughly broken into performance-based policies, which base regulation on measurable outcomes, and design-based policies, which are predicated on input decisions (Ribaudo 1999). Since nonpoint source emissions are characterized as prohibitively costly to monitor on a firm-by-firm basis, performance-based policies have been directed toward ambient environmental conditions. Beginning with the seminal work of Segerson (1988), numerous mandatory approaches have been proposed that provide incentives to nonpoint sources based on measured ambient pollution levels (e.g., Xepapadeas 1991; Cabe and Herriges 1992; Hansen 1998; Horan et al. 1998; Karp 2004).

One main criticism of mandatory approaches, in particular policies that involve taxing nonpoint polluters based on ambient pollution, is political feasibility. Policy makers have historically addressed nonpoint source pollution almost exclusively through voluntary measures.[1] While purely voluntary programs have been widely accepted by agricultural producers, there is little evidence that they have delivered outcomes, in terms of improved water quality, that would warrant declaring them a success (Shortle, Abler and Ribaudo 2001).

In an effort to wed the political attractiveness of a voluntary policy with the theoretical effectiveness of an appropriately designed mandatory policy, Segerson and Wu (2006) develop a policy that uses voluntary and mandatory programs as complementary instruments. The proposed policy allows firms in a watershed to meet an ambient pollution standard voluntarily. As long as the ambient standard is achieved, no regulatory action takes place. If, however, the standard is not met voluntarily, then a mandatory instrument is put in place, under which all firms face an ambient-based pollution tax. The threatened tax policy is structured in such a way that profit maximizing firms are induced to meet the ambient pollution standard voluntarily.

This voluntary-threat policy has some clear advantages over a strictly mandatory or strictly voluntary approach. From a producer's standpoint the policy is attractive because it allows for flexibility in meeting pollution standards without explicit regulation or taxation. From the regulator's standpoint the policy's attractiveness comes from avoiding the potentially large costs associated with administering the tax and incurring the information costs necessary to appropriately set the tax rate. Finally, the instrument is desirable from a social planner's perspective, as it offers the potential to address the nonpoint source pollution problem cost effectively.

---

[1] Common voluntary policies include land retirement programs, such as the Conservation Reserve Program, as well as working land programs, that provide incentives to agricultural landowners for developing best management practices (BMPs) and implementing pollution prevention and control measures. Annual federal expenditures for voluntary conservation programs are projected to be nearly $5 billion by 2011 (ERS 2002).

A shortcoming of the Segerson and Wu mechanism is that there is a theoretical coordination problem in the voluntary policy stage, resulting from the existence of multiple Nash equilibria, including an equilibrium where none of the firms engage in pollution control. To address the theoretical issues related to the existence of multiple Nash equilibria, an endogenous tax mechanism is introduced where expected tax payments are an increasing function of the degree of noncompliance in the voluntary stage. This mechanism can be designed to induce optimal abatement in the voluntary stage as a unique equilibrium.

Due to the novelty of a voluntary-threat policy, empirical program evaluation using naturally occurring data is not possible since no such policy is presently being implemented.[2] The social gains from a pollution standard being achieved voluntarily at least cost, together with the theoretic potential for costly suboptimal equilibria, imply that the experimental economics laboratory is an important intermediate testing ground for gaining a perspective on how the proposed policies might work in practice.

In recent years a burgeoning experimental economics literature has complemented theoretical work on nonpoint source pollution, through laboratory tests of many of the proposed ambient-based mandatory policies (Spraggon 2002, 2004; Alpizar et al. 2004; Poe et al. 2004, Cochard et al. 2005; Vossler et al. 2006; Suter et al. forthcoming). The results from these experimental studies show that a subset of the proposed policies, including a tax policy similar to the threatened policy of Segerson and Wu, engender outcomes that are highly efficient. However, none of these studies have tested a voluntary-threat policy.

[2] Compliance mechanisms require farmers to undertake conservation measures to be eligible for some federal aid programs. For example, farmers who fail to reduce soil erosion on highly erodible land may be ineligible for some federal benefits (USDA 2004). While this is similar to a voluntary-threat policy, the threat is based more on input decisions than on the meeting a collective performance standard.

This essay introduces an endogenous voluntary-threat policy and compares experimental outcomes from the policy to outcomes under the voluntary-threat policy proposed by Segerson and Wu. The next section provides the theoretical background for the policy introduced by Segerson and Wu and derives the theoretical properties of the new endogenous policy. Section 1.3 details the experimental design and Section 1.4 presents and analyzes experimental results. Section 1.5 concludes with a summary of the findings along with a discussion of their policy relevance.

1.2 Theoretical Background

The economic model follows that of Segerson and Wu closely. In particular, it is assumed that abatement and firm characteristics can each be represented by a scalar and that the policy goal is one of meeting an ambient water quality standard on average, such that stochastic factors (e.g., weather) can presumably be suppressed.

There are n firms, indexed by i, in a given watershed. Let a_i denote the firm's abatement decision and $\boldsymbol{a} = (a_1, a_2, ..., a_n)$ the vector of abatement decisions. $C_i = C(a_i, \theta_i)$ is the abatement cost function, where θ_i is an index that represents characteristics specific to the firm. The vector of characteristics of all firms in the watershed is $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$. It is assumed that the cost function is strictly convex, with $\partial C / \partial a_i > 0, \partial^2 C / \partial a_i^2 > 0$ and $C(0, \theta_i) = 0$. Ambient pollution at a monitoring point, denoted by x, is a function of the abatement decisions and firm characteristics of all firms, i.e., $x = x(a_1, \ldots a_n; \theta_1, \ldots, \theta_n)$, with $\partial x / \partial a_i < 0$ and $\partial^2 x / \partial a_i^2 \geq 0$. Given that abatement is costly, in the absence of any policy intervention it is expected that in equilibrium a_i=0.

A social planner is interested in meeting an exogenously determined water quality standard, which is denoted by x^s. The standard could be based on a Total Maximum Daily Load (TMDL) requirement or simply be a product of political bargaining. The social planner's problem and corresponding Lagrangian, assuming an interior solution, can then be written as

$$\underset{\boldsymbol{a}}{Min}\sum_{i=1}^{n} C(a_i,\theta_i) \quad s.t. \;\; x(a_1,\ldots,a_n;\theta_1,\ldots,\theta_n) \le x^s, \;\; \boldsymbol{a} \ge 0 \tag{1}$$

$$L = -\sum_{i=1}^{n} C(a_i,\theta_i) + \lambda\left(x^s - x(a_1,\ldots,a_n;\theta_1,\ldots,\theta_n)\right). \tag{2}$$

The first-order conditions from equation (2) are solved with each firm in the watershed making abatement decision $a_i = a_i^*$, so that the vector of optimal abatement decisions can be denoted by $\boldsymbol{a} = \boldsymbol{a}^*$. The strict monotonicity and convexity of the firm cost functions imply that $\lambda^* = -\dfrac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*}$ for all *i*, where λ can be interpreted as the marginal cost to the group of polluters of decreasing the ambient standard by one unit. Since $\lambda^* > 0$, the constraint is binding and therefore, at the optimum, ambient pollution is exactly equal to the standard.

The following subsections detail the theoretical basis for three policies that seek to induce polluters in a watershed to achieve the ambient pollution standard at least cost. The first two policies, a pure ambient tax policy and a voluntary policy with a threat of an exogenously determined tax, are similar to those described by Segerson and Wu and the reader is directed to the proofs provided in that article. The primary difference between the theory presented here and that in Segerson and Wu is that here the threatened tax policy is put in place for *K* rounds rather than in perpetuity, as assumed in Segerson and Wu. A rigorous analysis of the third policy, a voluntary policy with a threat of an endogenously determined tax, is provided to establish the policy's theoretical properties.

Pure Ambient Tax Policy

The social planner is charged with meeting the ambient standard at least cost through the use of a policy that charges all firms in the watershed a marginal tax, τ, on units of ambient pollution above a specific tax threshold, $\bar{x}$. Setting $\tau = \lambda^*$ and $\bar{x} = x^s$, the cost minimization problem for firm *i*, where the superscript *t* indicates abatement under the tax, is

$$\underset{a_i^t}{Min} \quad C\left(a_i^t, \theta_i\right) + max\left(0, \tau \cdot \left(x\left(a_1, \ldots a_n; \theta_i, \ldots \theta_n\right) - x^s\right)\right). \tag{3}$$

Under the tax policy $\boldsymbol{a}^t = \boldsymbol{a}^*$ is a Nash equilibrium, where a_i^* is the optimal abatement level for firm *i* as defined previously. The intuition for this result is as follows. Suppose the vector of abatement decisions by the other $n - 1$ firms is denoted by $\boldsymbol{a}_{-i}^t = \boldsymbol{a}_{-i}^* = \left(a_1^*, \ldots, a_{i-1}^*, a_{i+1}^*, \ldots, a_n^*\right)$. Ambient pollution will be equal to or less than the standard if abatement by firm *i* is $a_i^t \geq a_i^*$. Since there is no marginal benefit to reducing ambient pollution below the standard, $a_i^t > a_i^*$ can never be a best response. Firm *i* will therefore choose $a_i^t = a_i^*$ and achieve the standard with equality or $a_i^t < a_i^*$ and pay the ambient tax.

When $a_i^t < a_i^*$, the marginal benefit of abatement by firm *i*, in terms of tax payments avoided, is $-\tau \cdot \partial x / \partial a_i^t$, where $\partial x / \partial a_i^t < 0$. Substituting, $\tau = -\dfrac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*}$ the marginal benefit from abatement becomes $\dfrac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*} \cdot \partial x / \partial a_i^t$. For $a_i^t < a_i^*$ the strict convexity of the abatement cost function implies the marginal cost of abatement by firm *i* is $\partial C / \partial a_i^t < \partial C / \partial a_i^*$. The concavity of the pollution function implies $\dfrac{\partial x / \partial a_i^t}{\partial x / \partial a_i^*} \geq 1$ and therefore $\partial C / \partial a_i^t < \dfrac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*} \cdot \partial x / \partial a_i^t$ for $a_i^t < a_i^*$. Since the marginal cost of abatement is greater than the marginal benefit to the firm, $a_i^t < a_i^*$ cannot be a best response. Thus, firm *i* will choose $a_i^t = a_i^*$, at the point where the marginal cost of abatement is equal to the marginal benefit. Assuming similar rational behavior across all firms, $\boldsymbol{a}^t = \boldsymbol{a}^*$ is a NE.

Additionally, $\boldsymbol{a}^t = \boldsymbol{a}^*$ is a *unique* NE, since abating more than a_i^* incurs marginal abatement costs in excess of the marginal benefit to an individual firm, independent of the abatement decisions of other firms in the watershed, i.e., $\partial C / \partial a_i^t > \dfrac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*} \cdot \partial x / \partial a_i^t$. This implies that $a_i^t > a_i^*$ can never be optimal and given that none of the *n* firms choose $a_i^t > a_i^*$, no firm will rationally choose $a_i^t < a_i^*$, since this would result in pollution in excess of the ambient standard. It is illustrated above that firms are strictly worse off by choosing $a_i^t < a_i^*$ and incurring the tax.

An important feature of the tax policy is that the tax threshold can be replaced with any $\bar{x} \le x^s$, such that the unique NE $\boldsymbol{a}^t = \boldsymbol{a}^*$ is maintained. Therefore $\bar{x}$ is a choice variable and setting $\bar{x}$ below x^s has the effect of increasing tax payments, while $\boldsymbol{a}^t = \boldsymbol{a}^*$ remains a unique equilibrium. The cost to each firm of one period of the tax policy can therefore be defined as

$$C(a_i^*, \theta_i) + \tau(x^s - \bar{x}). \tag{4}$$

The incentives for firms under the pure ambient tax policy are illustrated through the best response functions in Figures 1.1 and 1.2 for the simplest case involving a watershed comprised of two firms with linear pollution functions. In the example, the firms are heterogeneous with respect to their abatement costs and each has a maximal level of abatement indicated by the respective dotted lines. In both figures, firms face an ambient tax rate of $\tau = -\frac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*}$. In Figure 1.1, where $\bar{x} = x^s$, the best response for each firm is to abate just to the point where the ambient standard is achieved when abatement effort by the other firm is high. When abatement effort by the other firm is equal to or less than a_j^*, however, firm *i*'s best response is constant at a_i^*.

When the tax threshold is reduced to $\bar{x} < x^s$, each firm's best response is to abate so that ambient pollution is equal to the tax threshold as long as abatement effort by the other firm is high enough. However, as abatement effort by the other firm is reduced, the best response for each firm is again to choose a_i^*. In each case, the best response functions for firms one and two intersect at only one point, indicated in the figures by the large dot, and illustrate the unique NE whereby costs are minimized and ambient pollution is equal to the standard. This latter result is essential to the design of the voluntary-threat mechanisms.

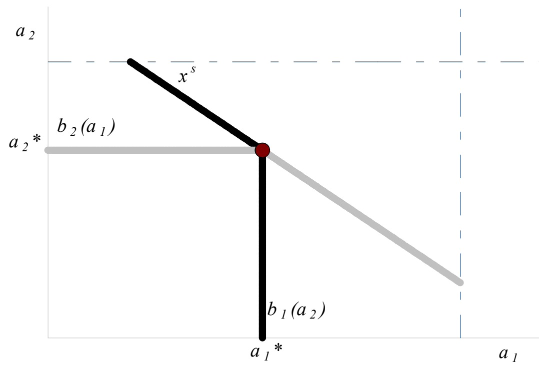


Figure 1.1 Pure Ambient Tax: $\bar{x} = x^s$

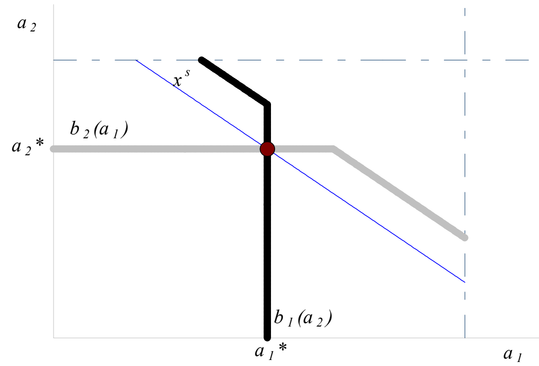


Figure 1.2 Pure Ambient Tax: $\bar{x} < x^s$

Exogenous Voluntary-Threat Policy

If the policy maker allows firms to meet the pollution standard voluntarily, and without incentives, it is expected that firms expend zero abatement effort, since abatement is expensive. However, consider a case where the policy maker allows firms to respond to the pollution standard voluntary, but includes a threat of a mandatory tax policy if the standard is not achieved. Specifically, if ambient pollution is above the standard in the voluntary setting, then the tax policy described above, with $\bar{x} < x^s$, is put into place for $K < \infty$ periods. The finite *K*-period tax penalty is a variation on Segerson and Wu, who assume that the tax policy is imposed in perpetuity. The finite nature of the tax penalty eliminates the possibility for other NE that could occur in an infinitely repeated game. Moreover, considering the finite case is important for purposes of experimental testing as it allows for ending and re-starting the game in experimental sessions where a violation occurs, akin to a situation where the regulator gives firms a second chance to comply voluntarily.

The incentive to meet the standard voluntarily is provided by employing a tax threshold that is less than the ambient standard, which makes the tax payments – even under optimal abatement in the tax stage game – strictly positive for each firm. This is labeled the "exogenous" voluntary-threat policy as the tax threshold, $\bar{x}$, is exogenous in the sense that it does not depend on behavior in the voluntary stage.

Using the superscript v to indicate outcomes in the voluntary stage, the amount of voluntary abatement chosen by firm i is denoted a_i^v. The amount of voluntary abatement necessary by firm i to ensure that the ambient standard is exactly met, given the abatement activities of the other firms in the watershed, is further defined as $a_i^{v(s)}$ (i.e., $x\left(a_i^{v(s)}, \boldsymbol{a}_{-i}^v; \theta_i, \boldsymbol{\theta}_{-i}\right) = x^s$). In every period of the tax stage, as described in the previous section, firms will choose $a_i^t = a_i^*$ as part of a unique NE when $\tau = -\frac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*}$. If the costs imposed by the tax policy are sufficiently high, firms will choose $\boldsymbol{a}^v = \boldsymbol{a}^*$ in the voluntary stage game as part of a subgame perfect NE (SPNE).

As shown in Segerson and Wu, it must be the case that firm i will optimally choose either $a_i^v = a_i^{v(s)}$ or $a_i^v = 0$ in the voluntary stage. It is never optimal for a firm to choose $a_i^v > a_i^{v(s)}$ since this causes ambient pollution to be strictly less than the standard and firm i would be better off choosing $a_i^v = a_i^{v(s)}$. It is also never optimal for firm i to choose $0 < a_i^v < a_i^{v(s)}$. If $a_i^v < a_i^{v(s)}$ then the ambient standard will not be met and the tax policy will be imposed. Since abatement is costly, firm i would have been better off choosing zero abatement.

In the voluntary period the firm therefore has a choice between abating so that the ambient standard is achieved or not abating at all and paying the tax over the next K periods. Assuming a discount factor $0 < \delta \leq 1$, the cost of voluntary abatement sufficient to meet the standard across K+1 periods is given by $\sum_{k=0}^{K} \delta^k C\left(a_i^{v(s)}, \boldsymbol{a}_{-i}^v; \theta_i, \boldsymbol{\theta}_{-i}\right)$. The cost of abating zero in the voluntary period and facing K periods of the tax policy is given by $\sum_{k=1}^{K} \delta^k \left(C\left(a_i^*, \theta_i\right) + \tau\left(x^s - \bar{x}\right)\right)$. Therefore the firm will abate voluntarily, and the standard will be achieved, if

$$\sum_{k=0}^{K} \delta^k C\left(a_i^{v(s)}, \boldsymbol{a}_{-i}^v; \theta_i, \boldsymbol{\theta}_{-i}\right) \leq \sum_{k=1}^{K} \delta^k \left(C\left(a_i^*, \theta_i\right) + \tau\left(x^s - \bar{x}\right)\right). \tag{5}$$

When $\bar{x} = x^s$, the expected liabilities are zero under the tax policy. In this case, no firm will ever choose $a_i^v \geq a_i^*$, since $C(a_i^v, \theta_i) \geq C(a_i^*, \theta_i) > C(0) = 0$. If it is not optimal for any firm to choose $a_i^v \geq a_i^*$, then the standard will not be achieved voluntarily and each firm is strictly better off by choosing $a_i^v = 0$.

When $\bar{x}$ is sufficiently below x^s, there is a SPNE in which firms optimally choose $\boldsymbol{a}^v = \boldsymbol{a}^*$ in the voluntary stage and $\boldsymbol{a}^t = \boldsymbol{a}^*$ in the tax stage. The intuition for this result is as follows. Suppose $\boldsymbol{a}_{-i}^v = \boldsymbol{a}_{-i}^*$ so that $a_i^{v(s)} = a_i^*$ for firm *i*. Recall that firm *i* will either choose $a_i^v = a_i^{v(s)}$ or $a_i^v = 0$. Choosing $a_i^v = 0$ will result in the standard not being met and the imposition of the tax policy. Therefore, if $C\left(a_i^*, \theta_i\right) \leq \sum_{k=1}^{K} \delta^k \tau \left(x^s - \bar{x}\right)$, then firm *i* will optimally choose $a_i^v = a_i^{v(s)} = a_i^*$. Note that for $\boldsymbol{a}^v = \boldsymbol{a}^*$ to be part of a SPNE, $\bar{x}$ must be chosen so that $C\left(a_i^*, \theta_i\right) \leq \sum_{k=1}^{K} \delta^k \tau \left(x^s - \bar{x}\right)$ for each firm in the watershed. In words, the discounted penalty for not achieving the standard must be at least as great as the cost of abating to a_i^*.

Under the exogenous voluntary-threat policy there will also exist SPNE whereby the ambient standard is achieved but not at least cost, unless $C\left(a_i^*, \theta_i\right) = \sum_{k=1}^{K} \delta^k \tau \left(x^s - \bar{x}\right)$ for all firms. If not, firms with $C\left(a_i^*, \theta_i\right) < \sum_{k=1}^{K} \delta^k \tau \left(x^s - \bar{x}\right)$ will strictly prefer the voluntary policy to the tax policy and thus have an incentive to abate more than a_i^*. For the condition to hold with equality for all firms requires identical total abatement costs at the optimum across all firms (i.e., $C\left(a_i^*, \theta_i\right) = C\left(a_j^*, \theta_j\right) \forall i, j \in n$) and that $\bar{x} = x^s - \left(C\left(a_i^*, \theta_i\right) \Big/ \sum_{k=1}^{K} \delta^k \tau \right)$.

When $a_i^{v(s)} \neq a_i^*$ for at least one of firm in the watershed, the aggregate cost of meeting the ambient standard is not minimized. As $\bar{x}$ diverges from x^s the range of firm-level profit maximizing voluntary abatement levels expand and the potential for free riding increases. This implies a tradeoff in the choice of the tax threshold. Setting $\bar{x}$ low relative to x^s generates a more severe penalty for the group, if the standard is not achieved voluntarily. However, it opens the door to a greater range of disparity between cost-minimizing and realized abatement choices.

In addition to multiple SPNE where the ambient standard is achieved voluntarily, there also is a SPNE whereby all firms choose zero abatement in the voluntary period. If $\boldsymbol{a}^{v}_{-i} = \mathbf{0}$, firm i will choose to abate zero units since abating to the point where the ambient standard is met is excessively costly or not feasible.[3] In past experimental analyses of ambient-based policies with a zero abatement NE, in addition to the Pareto optimal NE, groups achieved significantly lower levels of social efficiency than under the ambient policies that did not have a zero abatement NE (Spraggon 2002, Vossler et al. 2006). Unfortunately, in the case of the exogenous voluntary-threat policy, the choice of $\bar{x}$ alone cannot eliminate the zero abatement SPNE.

Figures 1.3 and 1.4 illustrate that the best response functions for firms 1 and 2 will always cross at multiple points. The overlapping best response functions reflect the multiple NE whereby the ambient standard is achieved, but not necessarily at least cost. This overlapping section increases as the exogenous threshold in the threatened tax policy is reduced, as in Figure 1.4, indicating a wider range of equilibria. In addition, the best response functions under the exogenous voluntary-threat policy always intersect at the origin, which implies that both firms choosing zero abatement cannot be eliminated as an equilibrium strategy.

[3] There is a potential that the best response for firm i is to meet the standard voluntarily even if all other firms chose zero abatement (i.e., $x\left(a_i^v, \boldsymbol{0}; \theta_i, \boldsymbol{\theta}_{-i}\right) = x^s$). However, this would require that the standard be relatively close to the baseline level of ambient pollution and is therefore not of particular interest.

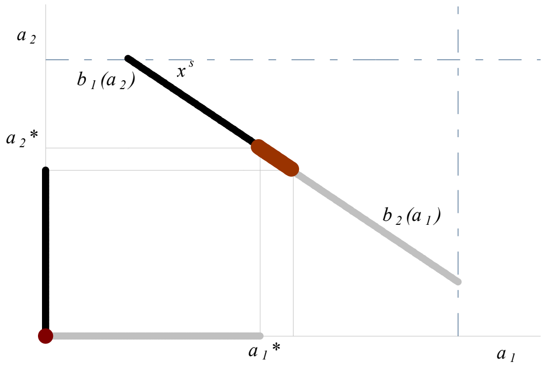


Figure 1.3 Exogenous V-T: High $\bar{x}$

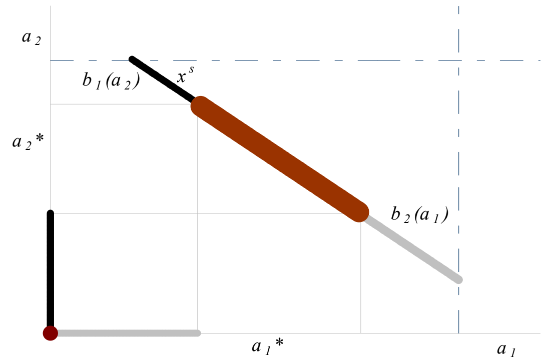


Figure 1.4 Exogenous V-T: Low $\bar{x}$

Endogenous Voluntary-Threat Policy

To eliminate suboptimal equilibria in the voluntary stage, Segerson and Wu suggest the threat of a retroactive tax policy. Under this policy, if the ambient standard is exceeded in the voluntarily stage, firms pay taxes for the violation in the voluntary stage, in addition to facing the tax policy in future periods. The tax paid on ambient pollution in excess of the standard in the voluntary stage is collected prior to the first round of the tax stage. While this eliminates the zero abatement equilibrium in the voluntary stage, it could negate much of the political attractiveness associated with the voluntary policy. The voluntary policy with a retroactive tax distinguishes itself from a pure tax policy only in the sense that rather than being collected at the end of the period, taxes in the voluntary stage are collected at the beginning of the next period. Nevertheless, it remains attractive from the standpoint that the expenditures necessary to optimally implement the tax mechanism could be avoided.

The policy instrument introduced here relies on an endogenously determined tax threshold. That is, the threshold in the tax stage is determined by the level of noncompliance in the voluntary stage. Therefore under voluntary noncompliance, this instrument makes the amount of future tax bills conditional on voluntary period behavior. This implies, for example, that even if all other firms undertake zero abatement in the voluntary period firm *i* has an incentive to abate as it reduces its own future tax payments. Formally, if the ambient standard is exceeded in the voluntary stage, the tax payment due in each period of the tax stage is defined as

$$\text{Tax Payment} = \tau\left(x^t - \tilde{x}\right) \;\; \textit{where}\; \tilde{x} = x^s - \varphi\left(x^v\text{-}x^s\right) \;\textit{and}\;\; \varphi > 0\,. \tag{6}$$

In equation (6), x^v denotes realized ambient pollution in the voluntary stage and the scale factor, φ, is freely chosen by the regulator. Increasing φ lowers the tax threshold for all levels of $x^v > x^s$ and therefore increases the severity of the threatened tax policy. The crux of the mechanism is that the tax threshold decreases as the level of pollution in excess of the standard in the voluntary stage increases, thus making the consequent tax policy more costly. The firm's tax payment in each period of the tax stage can be written as $\tau\left(x^t - x^s + \varphi\left(x^v - x^s\right)\right)$ for pollution levels greater than $\tilde{x}$, zero otherwise. From this representation, it is apparent that in the tax stage of the endogenous mechanism, firms pay a tax based on the pollution level in that period (similar to the pure ambient tax) as well as a penalty for noncompliance in the voluntary stage. Note that the exogenous-threat policy is a special case of the endogenous-threat policy where φ=0 and $\bar{x} = x^s$.

It has been shown above that in the tax stage any threshold $\bar{x} \leq x^s$ will induce a unique NE $\boldsymbol{a}^t = \boldsymbol{a}^*$, which implies that the standard is met (i.e., $x^s = x^t$) at least cost. Substituting this equality into equation (6), simplifying and multiplying it by the discount rate yields the tax penalty over K rounds from voluntary noncompliance under the endogenous threat mechanism

$$\sum_{k=1}^{K} \delta^k \tau\varphi\left(x^v - x^s\right). \tag{7}$$

In the voluntary stage, each firm compares the cost of abatement against the present value of future tax payments. However, the cost of the tax is now a function of firms' voluntary abatement decisions. The consequence is the elimination of suboptimal equilibria. The equilibria generated by the endogenous tax threat are stated in Propositions 1, 2 and 3 below along with a discussion of the intuition for each. Formal proofs of the propositions can be found in the appendix.

Proposition 1: *If* $\tau = \lambda^*$ *then* $\{\boldsymbol{a}^{v}, \boldsymbol{a}^{t}\} = \{\boldsymbol{a}^*, \boldsymbol{a}^*\}$ *is a SPNE if and only if* $\varphi \geq \left(\sum_{k=1}^{K} \delta^k\right)^{-1}$.

The intuition for Proposition 1 is based on firm *i* comparing the marginal costs and benefits of abatement in the voluntary setting. Suppose the other *n*-1 firms in the watershed choose the abatement vector $\boldsymbol{a}_{-i}^{v} = \boldsymbol{a}_{-i}^{*}$ such that pollution will be equal to or less than the standard if $a_i^v \geq a_i^*$. As in the previous cases, there is no benefit to reducing ambient pollution below the standard, and thus $a_i^v > a_i^*$ is not a best response. In such a setting, firm *i* will therefore choose $a_i^v = a_i^*$ so that the standard is achieved with equality or choose $a_i^v < a_i^*$ and face the *K*-period tax policy.

The optimal level of abatement for firm *i*, $a_i^v \leq a_i^*$, is again determined by comparing its marginal cost of abatement, $\partial C / \partial a_i^v$, to the marginal benefit of reducing its future tax burden, $-\varphi \sum_{k=1}^{K} \delta^k \tau \cdot \partial x / a_i^v$ (the entire expression is positive given that $\partial x / \partial a_i < 0$). Substituting $\tau = \lambda^* = -\frac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*}$ and rearranging, firm *i* will optimally abate to $\partial C / \partial a_i^v = \varphi \sum_{k=1}^{K} \delta^k \frac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*} \cdot \partial x / a_i^v$, where the marginal cost of abatement equals the marginal benefit. The marginal incentives of the endogenous voluntary threat policy are identical to the pure ambient tax policy when $\varphi = \left(\sum_{k=1}^{K} \delta^k\right)^{-1}$ and firm *i* will optimally choose $a_i^v = a_i^*$. When $\varphi > \left(\sum_{k=1}^{K} \delta^k\right)^{-1}$, the marginal benefit of reducing future tax payments is strictly greater than the marginal cost of abatement at $a_i^v = a_i^*$. However, since there is no benefit to reducing ambient pollution below the ambient standard, in terms of reduced future tax payments, firm *i* will optimally choose $a_i^v = a_i^*$. Thus, as long as $\varphi \geq \left(\sum_{k=1}^{K} \delta^k\right)^{-1}$, $\boldsymbol{a}^v = \boldsymbol{a}^*$ is part of a SPNE in the voluntary stage.

If $\varphi < \left(\sum_{k=1}^{K} \delta^k\right)^{-1}$ and $\boldsymbol{a}_{-i}^{v} = \boldsymbol{a}_{-i}^{*}$, the marginal cost of abatement at $a_i^v = a_i^*$ is strictly greater than the marginal benefit of abatement, i.e., $\partial C / \partial a_i^* > \varphi \sum_{k=1}^{K} \delta^k \frac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*} \partial x / a_i^*$.

Thus, firm *i* will optimally choose $a_i^v < a_i^*$ and $\boldsymbol{a}^v = \boldsymbol{a}^*$ can therefore not be part of a SPNE.

Proposition 2: *If* $\tau = \lambda^*$ *then* $\{\boldsymbol{a}^v, \boldsymbol{a}^t\} = \{\boldsymbol{a}^*, \boldsymbol{a}^*\}$ *is a unique SPNE if and only if* $\varphi = \left(\sum_{k=1}^{K} \delta^k\right)^{-1}$.

It was shown above that when $\varphi = \left(\sum_{k=1}^{K} \delta^k\right)^{-1}$, $\boldsymbol{a}^v = \boldsymbol{a}^*$ is part of a SPNE. The intuition for why $\boldsymbol{a}^v = \boldsymbol{a}^*$ is part of a *unique* SPNE is identical to that of the pure ambient tax. The marginal cost of abatement is greater than the marginal benefit of abatement, in terms of reductions in the tax burden, for all levels of abatement above a_i^*, i.e, $\partial C / \partial a_i^v > \dfrac{\partial C / \partial a_i^*}{\partial x / a_i^*} \cdot \partial x / \partial a_i^v$. Therefore none of the *n* firms will ever choose $a_i^v > a_i^*$, regardless of the abatement decisions made by other firms in the watershed. Given that other firms will never rationally over abate, it is not in a firm's best interest to abate less than a_i^*, since, as explained in Proposition 1, the marginal benefit from reduced tax payments is greater than the marginal cost of abatement for $a_i^v < a_i^*$.

When $\varphi > \left(\sum_{k=1}^{K} \delta^k\right)^{-1}$ then it is in firm *i*'s best interest to abate more than a_i^* when $a_i^{v(s)} > a_i^*$. In this case, the marginal cost of abatement at $a_i^v = a_i^*$ is less than the marginal benefit, i.e., $\partial C / \partial a_i^* < \varphi \sum_{k=1}^{K} \delta^k \dfrac{\partial C / \partial a_i^*}{\partial x / \partial a_i^*} \partial x / a_i^*$, since $\varphi \sum_{k=1}^{K} \delta^k > 1$. When one firm chooses to overabate to achieve the standard, other firms in the watershed will not deviate from their strategies and thus other SPNE are possible. It follows that is $\varphi = \left(\sum_{k=1}^{K} \delta^k\right)^{-1}$ necessary and sufficient to induce a *unique* SPNE.

Proposition 3: *If* $\tau = \lambda^*$ *then* $\{\boldsymbol{a}^v, \boldsymbol{a}^t\} = \{\boldsymbol{0}, \boldsymbol{a}^*\}$ *is never a SPNE when* $\varphi > 0$.

The intuition for Proposition 3 follows from the fact that, as long as the scale factor, φ, is greater than zero, firm *i* can potentially reduce its future tax burden through abatement. Even if every other firm in the watershed chooses not to abate, firm *i* is better off by engaging in a positive level of abatement. Since the cost of abatement tends to zero as the level of abatement approaches zero, there will always be a point at which the marginal benefit of additional abatement exceeds the costs.

The firm-level best response functions under the endogenous voluntary-threat policy are illustrated graphically in Figures 1.5 and 1.6. When $\varphi = \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$, as in Figure 1.5, the best response functions for the two firms' are identical to those under the pure ambient tax policy with $\bar{x} = x^s$. The unique NE in the voluntary stage is evidenced by the single intersection of the firms' best response functions. Figure 1.6 illustrates the effect of increasing the scale factor. The result is an overlap in the best response functions for firms one and two, which trace out the points under which the ambient standard is achieved. Importantly, the best response functions do not touch the origin, which implies that firms will never choose zero abatement in equilibrium.

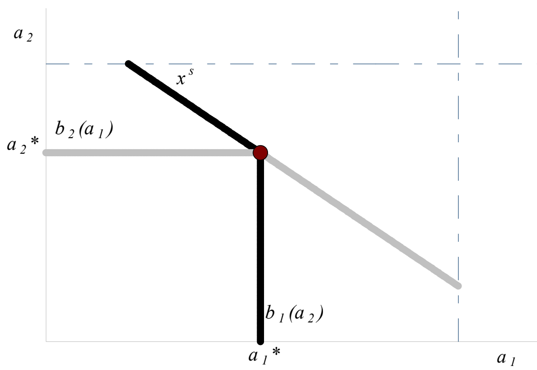


Figure 1.5 Endogenous V-T: $\varphi = \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$

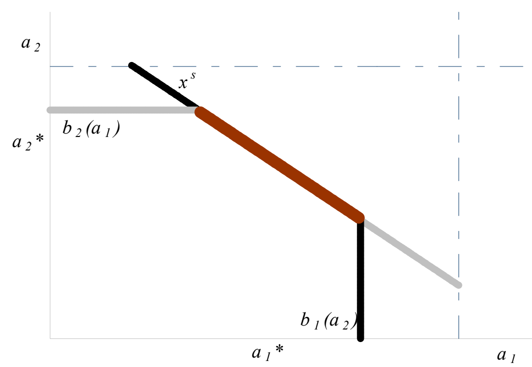


Figure 1.6 Endogenous V-T: $\varphi > \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$

The information requirements to appropriately parameterize the policies, such that voluntary least-cost abatement is a SPNE, are also an important consideration. In each of the policies considered above, the regulator needs to know λ^* so as to optimally define the tax rate, τ, for the tax policy. In the exogenous voluntary-threat policy, the regulator needs to set the threshold to ensure that the costs imposed by the threatened tax policy are large enough so that every firm in the watershed is better off abating voluntarily. This requires that the regulator spend the resources necessary to determine the abatement costs for all firms in the watershed. In comparison, to optimally design the endogenous voluntary-threat policy, the regulator need only have information regarding the discount rate, in addition to knowing λ^*. Further, even if the scale factor is set too low, firms will still voluntarily engage in a positive level of abatement effort. Under the exogenous voluntary-threat policy, setting the threshold too low results in firms choosing the zero abatement strategy as part of a unique equilibrium.

1.3 Experimental Design

To test the relative performance of the voluntary-threat policies, an economics experiment was conducted at the Cornell Lab for Experimental Economics and Decision Research. Participants were Cornell undergraduate students that had taken at least one class in economics and the majority had participated in at least one prior, but unrelated, economics experiment. The experimental sessions lasted approximately one hour and participants earned experimental tokens during each decision round, which were exchanged for dollars at the end of the session at the known rate of 70,000 tokens per $1US. Across all treatments, 264 participants took part in the experiment and individual earnings averaged $20.

Overall, there are eleven separate experimental treatments, with four groups of six participants each in a given treatment. The experiment consists of 23 decision rounds[4] and the rounds are split up into Part A (rounds 1-5) and Part B (rounds 6-23). Preceding Parts A and B, participants read through a set of written instructions, provided in the Appendix, and view a Powerpoint presentation given by the experiment administrator. In Part A there is no ambient-based policy in place, in order to establish a regulation-free baseline. In Part B, the exogenous voluntary-threat, endogenous voluntary-threat or pure ambient tax policy is instituted.

[4] The end round was determined randomly for each session, however each group completed at least 23 rounds. Participants were told that the end round would be randomly determined and the decision making spreadsheet contained 30 decision rounds.

In each treatment, all participants face the identical abatement cost function $C(a_i,\theta) = \phi a_i^{\alpha}$. As the term abatement implies reducing emissions relative to some initial condition, the participants' decision is instead framed as one of choosing a level of emissions. Specifically, participant emissions, r_i, are related to abatement through the function $r_i(a_i) = \gamma - a_i$, where γ represents the baseline level of emissions. Abatement is related to ambient pollution in a given round through the linear function $x = \sum_{i=1}^{6}(\gamma - a_i)$.

Each participant is provided an *"Emissions Decision Sheet"* that lists the "firm earnings" associated with all possible levels of emissions. To give real-world relevance to the experimental parameters, experiment earnings under a zero abatement strategy are calibrated to the net income of a medium sized dairy farm in New York State, operating with a herd size of 200 cows (Dairy Farm Management Summary 1999-2003). Table 1.1 lists the assumed functional forms and experiment parameters, which conform to the underlying assumptions of the theoretical model.

Table 1.1 Experimental Parameters

Description	Functional Form	Parameter Values
Abatement Cost Function	$\phi(a_i)^{\alpha}$	ϕ = 13; α = 3
Firm Earnings	$Y = Y^0 - \phi(a_i)^{\alpha}$	Y^0 = 75,000
Firm Level Emissions	$r_i = \gamma - a_i$	γ = 20
Ambient Pollution	$x = \sum_{i=1}^{n}(\gamma - a_i)$	n = 6

In Part A of all treatments, each participant makes their emissions decision in an environment where ambient pollution does not influence their earnings in any way. As such, the expectation is that each participant will choose an emissions level of 20, which maximizes their "firm earnings" in each of the five Part A rounds. Therefore in

Part A the expected ambient pollution level, given that each of the six participants chooses an emissions level of 20, is 120.

In Part B, there is an ambient-based policy designed to induce a 40% reduction in ambient pollution levels from an unconstrained pollution level of 120 to an ambient standard, x^s, of 72. The reduction mirrors the 40% nutrient reduction goals called for in the original Chesapeake Bay Agreement (CBP 2005). Reaching the ambient standard of 72 at least cost requires each participant to reduce their emissions to 12 units, from the unconstrained optimum of 20. Under each of the policy instruments, choosing an emissions level of 12 is a NE for all participants. A description of the 11 treatments can be found in Table 1.2.

Table 1.2 Treatment Summary

Treatment	Policy	Communication
1	Pure Ambient Tax $\bar{x} = 72$	No
2	Pure Ambient Tax $\bar{x} = 66$	No
3	Pure Ambient Tax $\bar{x} = 50$	No
4	Exogenous Voluntary-Threat: $\bar{x} = 66$	No
5	Exogenous Voluntary-Threat: $\bar{x} = 50$	No
6	Exogenous Voluntary-Threat: $\bar{x} = 0$	No
7	Exogenous Voluntary-Threat: $\bar{x} = 50$	Yes
8	Endogenous Voluntary-Threat: $\varphi = 1/3$	No
9	Endogenous Voluntary-Threat: $\varphi = 1$	No
10	Endogenous Voluntary-Threat: $\varphi = 3$	No
11	Endogenous Voluntary-Threat: $\varphi = 1$	Yes

Treatments 1-3 test the pure ambient tax policy where, in each round, participants pay a marginal tax of 2,500 tokens on every unit of ambient pollution above the threshold, $\bar{x}$. The tax payment in a given round is subtracted from the participant's "*Firm Earnings*" for that round. To aid decision making, a "*Tax Calculation Sheet*" is provided to participants that includes the tax payment corresponding to levels of ambient pollution.[5] Given that the marginal cost of

[5] Examples of the "*Tax Calculation Sheet*" and the "*Emissions Decision Sheet*" are included in the appendix.

reducing emissions to lower than 12 is greater than 2,500 tokens and the marginal cost of reducing an additional unit of emissions up to 12 is less than 2,500, optimal emissions for each firm is exactly 12 units such that the ambient standard of 72 is exactly met.

In Treatment 1 the tax threshold equals the ambient standard of 72, which replicates the tax policy in the experimental studies of Spraggon (2002), Poe et al. (2004), Cochard et al. (2004) and Suter et al. (forthcoming). When τ =2,500, emitting exactly 12 units is a unique NE for any tax threshold at or below 72. However, when the tax threshold is lower than 72 the *group* can maximize its payoff when participants emit fewer than 12 units. While collusive outcomes, in the absence of explicit group communication, are not seen in recent experimental results when total pollution was a stochastic function of total emissions (Suter et al. forthcoming), it is an open empirical question whether participants behave in a collusive manner in the non-stochastic environment presented in this study. Accordingly, treatments 2 and 3 reduce the tax threshold to 66 and 50, respectively.

In Part B of Treatments 4-7 the exogenous voluntary-threat policy are tested, with the ambient standard equal to 72. Each group begins Part B in the voluntary policy and as long as ambient pollution for the group remains at or below the ambient standard of 72, then the voluntary policy remains. If ambient pollution is greater than 72 under the voluntary policy, then the threatened tax policy is implemented in the next round and persists for three rounds (K=3) before groups again have a chance to meet the standard voluntarily. Three tax policy rounds allow for multiple restarts of the voluntary scenario while retaining the crux of a threat where participants pay a penalty over time for not meeting the standard, as suggested by Segerson and Wu. If the tax policy is implemented, each participant pays the marginal tax, τ, of 2,500 tokens for every unit of ambient pollution above the tax threshold. As in the pure ambient tax policy, optimal firm-level emissions are 12 units. Given the short time frame over which the decision rounds occur, a discount factor, δ, equal to 1 is assumed.

In Treatment 4 the threatened tax policy has an exogenous tax threshold of 66, the minimum symmetric threshold that theoretically provides the necessary incentives for voluntarily abatement. In Treatment 5 the threshold is 50, providing stronger than theoretically necessary incentives for voluntary abatement, but also increasing the range of SPNE. Treatment 6 provides the most draconian threat by reducing the threshold to 0. Varying the exogenous tax threshold provides insight into the tradeoff between a tax threshold that is relatively close to the ambient standard and a lower tax threshold, which increases the incentive to abate voluntarily but also increases the potential for meeting the voluntary standard at higher than minimum cost. In each of the exogenous voluntary-threat treatments there is also a SPNE whereby each participant chooses to emit 20 units (i.e., zero abatement) in the voluntary setting.

Meeting the ambient standard voluntarily at least cost requires a great deal of coordination by the participants, since everyone must emit exactly 12 units. Without communication, this coordination is likely to be difficult. Thus, in Treatment 7, the exogenous voluntary-threat policy is tested with a threshold of 50 while allowing groups to engage in costless, nonbinding communication, (referred to in the experimental economics literature as "cheap talk"). Groups are allowed up to five minutes of cheap talk before rounds 6, 11, 16 and 21, enhancing the potential to coordinate abatement strategies. In the cheap talk sessions, they can discuss any aspect of the experiment, but are not allowed to make threats or arrange for side payments.

Treatments 8-11 test the endogenous voluntary-threat policy. Similar to the exogenous voluntary-threat treatments, as long as groups achieve the ambient standard of 72, the voluntary policy remains in place. If ambient pollution exceeds the standard of 72 in a voluntary round then a tax policy is implemented for the following 3 rounds, where groups again pay tax rate τ for every unit of ambient pollution above the tax threshold. The tax threshold is determined by the degree of noncompliance in the voluntary setting. In Treatment 8, $\varphi = 1/K = 1/3$, so that for every unit of ambient pollution above the standard in the voluntary round, the tax threshold is reduced by 1/3 of a unit below 72 in the tax policy rounds. With $\varphi = 1/3$, under the assumption $\delta = 1$, each firm chooses to emit 12 units as part of a *unique* SPNE and thus the ambient standard is achieved at least cost. In Treatments 9 and 10 the endogenous threat is made more severe by increasing the scale parameter, φ, to 1 and 3 respectively. Again, as the severity of the threat rises, through increases in the scale parameter, the incentives for voluntary abatement increase, but additional SPNE are introduced, under which the ambient standard is achieved but not at least cost. Note that the severity of the threatened tax policy can also be increased by increasing the number of tax rounds, K, but this is not explored in the set of experiments presented here. In Treatment 12 the endogenous voluntary-threat with $\varphi = 1$ is tested with the participants allowed the chance to engage in cheap-talk. In each of the endogenous voluntary-threat treatments, the zero abatement equilibrium is eliminated.

In comparing decision making under the exogenous and endogenous voluntary-threat policies, the focus is on outcomes from treatments that are the most comparable. The exogenous voluntary-threat with threshold equal to 66 is most comparable to the endogenous voluntary-threat with the scale parameter set at 1/3. These treatments can be thought of as the minimum threat required to induce voluntary compliance and also the treatments that theoretically induce the most narrow range of SPNE. Similarly Treatment 6, with an exogenous threshold of zero is most comparable to Treatment 10, with scale parameter equal to 3, as these treatments represent the most severe threats. The zero threshold can be thought of as the maximum exogenous threat that could be administered, although theoretically the threat could be increased by introducing a negative exogenous threshold. There is not necessarily a corresponding upper bound on the endogenous scale parameter and the severity could be increased with higher scale parameters.[6] In the middle of the threat severity spectrum, the results under the exogenous voluntary-threat with threshold of 50 in Treatment 5 and the endogenous voluntary-threat with scale parameter of 1 in Treatment 9 are compared. Again, there is not a direct correspondence between the exogenous and endogenous threats in these treatments, but they are the most comparable given that they provide a mid-level point in the range of threats tested here.

1.4 Experimental Results

In this section, the outcomes from the 11 experimental treatments are analyzed to arrive at four primary results. For Result 1, an econometric model of participant emissions is used to evaluate decision making under the pure tax policy treatments. Drawing on outcomes from the voluntary stage of the voluntary-threat policies, Result 2 assesses the relationship between the severity of the tax threat and the frequency with which groups achieve the ambient standard voluntarily. For Result 3, the econometric model of participant emissions is again used to evaluate decision making in both the voluntary and tax rounds of the voluntary-threat treatments. Finally, for Result 4 treatment-level social efficiency measures are used to evaluate the relative economic performance of each policy.

Before proceeding to the primary results, it should be noted that in Part A of the experiment, where no policy mechanism is in place, participants choose zero abatement. The econometric procedures are detailed below, but the results presented in

[6] One could argue that φ=1.5 is the upper bound, given that under zero voluntary abatement the threshold in the tax policy rounds is equal to zero.

Tables 1.3 and 1.5 clearly show that in each of the treatments, participants choose emissions decisions that are not significantly different from 20, which corresponds to zero abatement. This can be taken as evidence that participants understand the decision-making framework and respond by choosing the level of emissions that maximizes their private earnings.

Since observations of participant-level decisions come from a hierarchical data generating structure, where groups are nested within treatments, participants are nested within groups and each participant makes decisions over a series of rounds, individual decisions cannot be treated as independent observations. It is reasonable to presume that participant decisions are correlated across rounds and with the decisions of other group members. To compare participant emissions decisions across treatments a mixed-effects model is estimated, with the general structure

$$r_{igr} = X\beta + \varepsilon_{igr} \ , \tag{8}$$

where r_{igr} is the emissions decision by participant i in group g and treatment r; X is known as the design matrix, which is a matrix of 1's and 0's that represent the form of the fixed treatment effects; and β is a vector of estimable fixed treatment-effect coefficients. The model error is represented by ε_{igr}.

The mixed model uses participant-level emission decisions as the dependent variable and includes treatment-specific fixed effects. The treatment effects are allowed to vary across Part A of the experiment and two aggregate round groupings, rounds 6-14 and 15-23, of Part B. In addition, the treatment effects are allowed to vary for decisions made under the voluntary and tax policies of Part B. It is assumed that $\varepsilon_{igr} = \alpha_g + u_{\mathrm{igr}}$, where a random effect α_g is included to account for within group correlation and u_{igr} is further assumed to be autocorrelated and heteroscedastic. In particular, the error variance is allowed to differ by treatment, and the model errors are assumed to follow an AR(1) process, with the autocorrelation coefficient, ρ, also allowed to vary across treatments. The mixed effects model essentially calculates mean participant emissions in each treatment scenario and then calculates standard errors analogous to an analysis of variance (ANOVA) framework that accounts for the fact that observed participant decisions are not independent. The model is estimated using the SAS "proc mixed" command and the program is included in the appendix. The goodness of fit is measured with the Akaike Information Criterion (AIC), which evaluates the degree to which the estimated model agrees with the observed structure, relative to the number of model parameters. Although one model is estimated, results are presented in two separate tables to facilitate the flow of the discussion. Coefficient results from the first three treatments are included in Table 1.3, while the remaining coefficient results are presented in Table 1.5.

Table 1.3 Mean Participant Emissions Model: Ambient Tax Results

Dependent Variable: Individual Participant Emissions (r_{igr}) AIC: 22,402
Number of Observations: 4,968

Treatment		Part A	Policy Scenario	Part B Rounds 6-14	Part B Rounds 15-23	σ^2_r	ρ
1	$\bar{x} = 72$	19.71 (0.48)	Pure Tax	13.36* (0.44)	13.73* (0.44)	8.22	0.64
2	$\bar{x} = 66$	19.45 (0.49)	Pure Tax	12.19 (0.43)	11.93 (0.43)	9.27	0.53
3	$\bar{x} = 50$	19.91 (0.53)	Pure Tax	10.18* (0.48)	10.22* (0.48)	10.62	0.69
			Estimated Group-Level Variance			0.32	

* Indicates parameter that is significantly different from the NE of 12 at the 5% level.
Note: Results in this table are from the Pure Ambient Tax treatments only although the model is estimated with data from all treatments.

Result 1: Under the pure ambient tax policy, participant emissions are not significantly different from the cost-minimizing NE of 12 when the threshold is 66, but participants tend to underabate when the threshold is 72 and overabate when the threshold is 50.

Although the focus of this essay is on the performance of voluntary-threat policies, analyzing participant emissions in the pure tax policy treatments is important as a way of gauging the incentives generated by the ambient tax. Each of the voluntary-threat policies require that the NE strategy is chosen in the tax stage of the game in order for the threat to be salient. Emissions under the pure ambient tax, provided in Table 1.3, most closely approximate the cost-minimizing NE when the threshold is equal to 66. It is therefore reasonable to assume that for some thresholds lower than the ambient standard, as required by the voluntary-threat, the tax induces the necessary incentives. As the threshold is reduced further, to a level of 50, average participant emissions are lower than the NE in both early and late rounds. Thus, participants seem to be engaging in some tacit collusion by abating less than is privately optimal when the threshold is reduced or are simply not making optimal decisions at the margin. This may serve to "water down" the threat of the tax, since group tax liabilities are lower when emissions are less than the NE. When the threshold is equal to the ambient standard of 72 in the pure ambient tax policy, participants exceed the NE of 12 in both early rounds (13.4) and late rounds (13.7). This result corroborates results in Spraggon (2002), who calculates mean emissions of 31.8 when the NE is 25, and Poe et al. (2004) who calculate emissions of 5.9 when the NE is 5. The level of emissions across the three ambient tax treatments are in contrast, however, to the findings of Suter et al. (forthcoming), where the tax threshold is found not to have a significant influence on average decision making. The lack of correspondence between the two studies is likely a result of the fact that Suter et al. consider the case where pollution is a stochastic function of group emissions, making group coordination more difficult. For example, participants may believe that individual deviations from group maximizing pollution levels in the stochastic setting are less transparent.

In Part B of the voluntary-threat treatments, groups had the opportunity to meet the ambient pollution standard voluntary and avoid the ambient tax.

Table 1.4 Summary of Voluntary Compliance

Treatment			Percent of Groups that Achieve Standard Voluntarily at Least Once	Percent of Voluntary Rounds Where Standard Achieved Among Groups That Achieve Standard
Exogenous Threat	4	$\bar{x} = 66$	25%	66.7%
	5	$\bar{x} = 50$	50%	75.0%
	6	$\bar{x} = 0$	100%	91.5%
	7	$\bar{x} = 50$ [Com]	100%	98.3%
Endogenous Threat	8	$\varphi = 1/3$	50%	33.3%
	9	$\varphi = 1$	25%	93.3%
	10	$\varphi = 3$	100%	75.6%
	11	$\varphi = 1$ [Com]	100%	100.0%

Result 2: Voluntary compliance with the ambient standard can only be expected if the threat of the tax policy is severe or if participants are able to communicate.

Table 1.4 shows that in Treatments 4, 5, 8 and 9, where the threat is theoretically sufficient but not practically severe, no more than 50% of the groups are able to achieve the standard voluntarily over the course of the session. When the tax threat is made relatively severe in Treatments 6 and 10, however, all four groups under both the exogenous and endogenous policy achieve the ambient standard voluntarily at least once. In these treatments, groups achieve the ambient standard in 92% of the voluntary rounds under the exogenous threat and 76% of the time when the threat is endogenous. These results clearly show that groups are responsive to the level of the threat and that only when they are confronted with a relatively consequential threat can voluntary compliance be reasonably expected. Even with the severe threats studied here, however, some groups still fail to voluntarily achieve the standard in every round.

In the treatments where communication is allowed, voluntary compliance is assured in both the exogenous and endogenous voluntary-threat policies. Across these two treatments, groups achieve the ambient standard in all but one voluntary round.[7] Communication allows participants the chance to coordinate their decisions and share information regarding payouts of various strategies. Under both the exogenous and endogenous policies, exactly achieving the standard voluntarily maximizes individual earnings (i.e., all individuals choose the SPNE of 12) and it also maximizes group earnings. Therefore communication provides a strong motive for groups to achieve the standard since each participant and the group as a whole are best off when the standard is achieved at least cost. This is in contrast to observed behavior in experiments that involve subsidy payments (e.g., Spraggon et al. 2002; Vossler et al. 2006), where communication results in overabatement, since the group profit maximizing level of emissions is lower than socially optimal levels.

Given that many groups in the voluntary-threat policy do not achieve the standard voluntarily, it is important to measure both the degree of noncompliance in the voluntary stage and emissions decisions in the consequent tax stage. Comparing participant emissions decisions across all of the policy settings allows for a more complete understanding of how specific policy parameters influence decision-making.

[7] The one observed instance of voluntary noncompliance was a result of a participant that accidentally typed in the wrong emissions decision number, as revealed in the cheap talk period immediately following.

Table 1.5 Mean Participant Emissions Model: Voluntary-Threat Results

Dependent Variable: Individual Participant Emissions (r_{igr}) AIC: 22,402
Number of Observations: 4,968

Treatment		Part A	Policy Scenario	Part B Rounds 6-14	Part B Rounds 15-23	σ^2_r	ρ
Exogenous Threat	4 $\bar{x} = 66$	19.39 (0.40)	Voluntary	13.85* (0.41)	17.34* (0.44)	5.56	0.34
			Tax	10.74* (0.39)	11.53 (0.36)		
	5 $\bar{x} = 50$	19.95 (0.41)	Voluntary	12.58 (0.41)	13.73* (0.42)	6.29	0.37
			Tax	9.19* (0.43)	9.40* (0.39)		
	6 $\bar{x} = 0$	19.55 (0.39)	Voluntary	11.73 (0.36)	11.99 (0.36)	4.63	0.42
			Tax	9.38* (0.67)	7.25* (0.48)		
Endogenous Threat	8 φ=1/3	19.73 (0.50)	Voluntary	12.82 (0.48)	13.75* (0.50)	8.73	0.66
			Tax	12.43 (0.46)	12.17 (0.45)		
	9 $\varphi = 1$	19.65 (0.47)	Voluntary	12.61 (0.46)	13.40* (0.47)	7.93	0.59
			Tax	11.95 (0.45)	11.29 (0.44)		
	10 $\varphi = 3$	20.02 (0.42)	Voluntary	11.91 (0.40)	12.13 (0.40)	6.23	0.47
			Tax	9.90* (0.50)	10.06* (0.42)		
			Estimated Group-Level Variance			0.32	

* Indicates parameter that is significantly different from the NE of 12 at the 5% level.
Note: Results in this table are from the Voluntary-Threat treatments only, although the model is estimated with data from all treatments.

Result 3: Voluntary-threat policies with a severe threat induce emissions decisions that are closer to the NE of 12 in the voluntary period, but further below the NE if the tax is implemented. Overall, the variability in mean emissions decisions is lower with the endogenous as opposed the exogenous voluntary-threat policy.

Inspection of Table 1.5 reveals that in the early rounds of the voluntary setting, participant emissions decisions are only significantly different from the cost-minimizing NE of 12 in the exogenous voluntary-threat treatment with the threshold equal to 66. In late rounds, however, emissions are significantly greater than 12 in the voluntary stage of all but the severe voluntary-threat treatments. Although average emissions are greater than the NE in the endogenous voluntary-threat treatments with φ = 1/3 and φ =1, they are lower than the emissions in the exogenous voluntary-threat treatments with $\bar{x}$ = 66 and $\bar{x}$ = 50 respectively. It should be noted that outcomes for treatments 4 and 8, where communication is allowed, are not included in the model results, as there is not sufficient variation among participants (i.e., in Part B participants chose the SPNE of 12 in essentially every round).

Among the groups that do not achieve the ambient standard voluntarily, there is a clear positive relationship between participant-level emissions in the tax stage of the exogenous voluntary-threat treatments and the tax threshold. When $\bar{x}$ = 50 and $\bar{x}$ = 0, emissions decisions are significantly less than the NE of 12 in both early and late rounds. The emissions decisions under the tax stage of the exogenous voluntary-threat policy where communication is not allowed reflect closely the results that observed in the pure ambient tax policy. Observed emissions are not significantly different between the exogenous voluntary-threat and pure ambient tax policies with thresholds of 66 and 50 respectively.

In the tax stage game of the endogenous voluntary-threat, the same correspondence between the severity of the threatened tax and the degree of overabatement does not exist. Only in the early rounds of the endogenous voluntary-threat with φ =3 are emissions significantly different from the NE of 12. There are two plausible explanations for this result. First, since the realized tax threshold in the endogenous voluntary-threat tends to be closer to the ambient standard of 72 than under the exogenous voluntary-threat treatments, groups facing the endogenous policy have less incentive for tacit collusion. Second, since the threshold is determined after each noncompliant voluntary round, participants have to adopt a strategy for the particular realization of each threshold. This makes a strategy of tacit collusion more challenging, since the emissions decision must be continuously readjusted. The endogenous voluntary-threat policy therefore results in emissions that are closer to the cost-minimizing NE, relative to the exogenous voluntary-threat, over the range of policy parameterizations that are investigated here. This result is consistent across both the voluntary and tax policy settings and indicates less overall variability in observed outcomes under the endogenous voluntary-threat policy.

Social Efficiency Measures

Estimating individual emissions decisions provides insight on how the various policy scenarios influence decision making, but they do not provide information useful for comparing the policy in terms of social outcomes. For example, if a policy approximates the cost-minimizing emissions level for an average participant, this does not mean that costs have been minimized if there is a significant degree of variation around the mean decision. While the objective of the policies tested here is to minimize the abatement costs of achieving the ambient standard, cost effectiveness is challenging to measure empirically. For example, there is no standard way to compare a treatment where group emissions exceed the ambient standard in a large percentage of rounds, but average abatement costs are low, to a treatment where the standard is consistently achieved, but abatement costs are relatively high.

Rather than generating an ad hoc measure of cost-effectiveness, the social outcomes of the treatments are compared based on a measure of social efficiency. The efficiency measure for each treatment assumes a damage function that is linear in total emissions with a slope of 2,500. The choice of damage function does not have a significant impact on the relative efficiencies between the treatments, and the linear damage function is assumed so as to be consistent with previous experimental analyses (e.g., Spraggon 2002; Poe et al. 2004; Vossler et al. 2006; Suter et al. forthcoming).

The aggregate efficiency measure used here is derived from the measure introduced by Spraggon (2002). The social surplus in a given round is determined by summing the pre-tax earnings of each of the six firms (the social benefit) less the social damage, determined by the aggregate emissions in that round. The observed surplus in round *t* by group *g*, S_{gt}, is then measured against the surplus in the zero abatement scenario, $\mathrm{S}_{\mathrm{zero}}$, and the maximum surplus possible, $\mathrm{S}_{\max}$, to give a measure of efficiency according to the formula

$$Social\ \ Efficiency_{gt} = \frac{S_{actual} - S_{zero}}{S_{\max} - S_{zero}}. \tag{12}$$

The social efficiency measure given in (12) is further partitioned into measures of emissions efficiency (EE) and allocative efficiency (AE) similar to those introduced in Suter et al. (forthcoming). The emissions efficiency measures the degree to which group emissions deviate from the social surplus maximizing level of group emissions, which in each treatment is 72. The allocative efficiency provides a measure of the variance in emissions decisions across participants. Given that all six firms are identical, the cost of achieving a given level of group emissions is minimized when all six participants choose the same level of emissions. Thus, allocative efficiency is higher when participants choose make similar emissions decisions. Formally, the emissions and allocative efficiency measures multiplicatively determine social efficiency according to

$$Social\ \ Efficiency_{gt} = EE_{gt} * AE_{gt} = \frac{S_{emissions} - S_{zero}}{S_{\max} - S_{zero}} * \frac{S_{actual} - S_{zero}}{S_{emissions} - S_{zero}}, \tag{13}$$

where $S_{emissions}$ is the social surplus given the level of group emissions in round *t*, and assuming that each participant contributes to the group emissions level equally. It should be noted that since emissions decisions can only be made in discrete increments, there is a small level of inefficiency included in the AE measure that results whenever group emissions are not divisible by six. In Suter et al., this third component of observed efficiency levels is defined as structural efficiency. The structural efficiency is not included here as the measure is never less than 99.8% for any of the treatments, given the relatively large range of available participant emission decisions.

Table 1.6 Social Efficiency Results

Treatment			Round Grouping	Social Efficiency	Emissions Efficiency	Allocative Efficiency
Pure Ambient Tax	1	$\bar{x} = 72$	6-14	82.68 (5.2)	92.58 (3.3)	88.85 (3.6)
			15-23	80.06 (5.2)	91.28 (3.3)	87.81 (3.6)
	2	$\bar{x} = 66$	6-14	79.59 (4.5)	95.82 (1.2)	83.41 (4.2)
			15-23	77.86 (4.5)	96.43 (1.2)	79.61 (4.2)
	3	$\bar{x} = 50$	6-14	63.52 (5.1)	87.89 (3.1)	74.62 (8.3)
			15-23	64.89 (5.1)	88.33 (3.1)	74.34 (8.3)
Exogenous Threat	4	$\bar{x} = 66$	6-14	81.59 (3.4)	90.48 (3.6)	87.51 (2.0)
			15-23	78.36 (3.4)	84.57 (3.6)	91.92 (2.0)
	5	$\bar{x} = 50$	6-14	79.89 (6.2)	90.44 (4.6)	86.67 (4.3)
			15-23	74.26 (6.2)	85.51 (4.6)	87.35 (4.3)
	6	$\bar{x} = 0$	6-14	88.72 (9.6)	96.46 (6.9)	91.68 (13.6)
			15-23	75.13 (9.6)	87.41 (6.9)	92.86 (13.6)
Endogenous Threat	7	$\varphi = 1/3$	6-14	79.77 (3.0)	96.28 (1.6)	81.60 (3.1)
			15-23	73.64 (3.0)	95.67 (1.6)	78.35 (3.1)
	8	$\varphi = 1$	6-14	82.24 (5.3)	98.16 (1.0)	85.10 (4.5)
			15-23	84.37 (5.3)	96.19 (1.0)	84.98 (4.5)
	9	$\varphi = 3$	6-14	84.83 (7.6)	93.61 (6.5)	87.83 (4.3)
			15-23	76.34 (7.6)	85.72 (6.5)	86.82 (4.3)

Note: Numbers in parentheses are standard errors.

Efficiency measures for each treatment in early and late round groupings are compared using the mixed modeling procedure described for participant-level emissions, except that the level of observation is the group rather than the individual. The model error is again assumed to follow an AR(1) process and the correlation coefficient and error variance are treatment specific. The total efficiency, emissions efficiency and allocative efficiencies results are reported for each treatment in Table 1.6.

Result 4: There are no systematic differences in aggregate social efficiency across the three policy mechanisms, but the endogenous threat policy tends to have higher emissions efficiency and lower allocative efficiency relative to the exogenous policy.

As illustrated in Table 1.6, the aggregate measure of social efficiency does not vary systematically based on policy type. All of the policy treatments have social efficiency levels that are not significantly different from 80%, with the exception of the pure ambient tax policy with a threshold of 50, which has social efficiency that is significantly less than 80%. In general, efficiency levels tend to diminish slightly in later rounds of the experiment. This is not a particularly attractive result for policy, as it suggests that groups tend to diverge from the social optimum as their experience with the mechanisms increase.

Among the voluntary-threat policies, the variability in social efficiency levels tends to be highest when the threat is made more severe, as evidenced by the differences in standard errors. These differences in variability stems from the fact that the severe threats tend to increase the frequency with which groups achieve the ambient standard voluntarily, but for the groups that do not achieve the standard, social efficiency tends to be lower as emissions are significantly less than the standard. Another important policy result is the fact that emissions efficiency tends to be higher and allocative efficiency tends to be lower in the endogenous voluntary-threat, in comparison to the exogenous voluntary-threat. On average, group emissions tend to be closer to the standard in the endogenous policy, but there is increased variance in participant emissions. Thus, if policy makers are primarily concerned with approximating ambient pollution targets, the endogenous voluntary threat policy is preferable to the exogenous policy. If, however, policy makers are concerned with equity, then the exogenous voluntary-threat is attractive, given the higher levels of allocative efficiency.

One final piece of evidence in regards to the relative merits of voluntary-threat policies comes from observed participant earnings. Clearly, when communication is allowed participants earn more money per round than in treatments where communication is not allowed. The average earnings per round in Part B when communication is allowed are approximately \$1.04 for the endogenous and exogenous policies. When communication is not allowed, earnings under the exogenous policy with threshold equal to 66 remain high at \$0.98, since the taxes levied in the tax policy rounds have a relatively minor impact on overall earnings. When the threatened tax threshold is reduced to zero, average earnings are reduced to \$0.74 per round, with earnings in groups that did not achieve the standard voluntarily significantly lower than the average. Interestingly, average earnings are higher in the endogenous policy with $\varphi=3$ than when $\varphi=1/3$ (\$0.92 and \$0.88 respectively). Thus, making the endogenous threat more severe does not necessarily result in a reduction in earnings, since ambient pollution in the voluntary policy is generally at or only slightly above the standard in each of the four groups. Therefore, the realized threshold is relatively close to the standard and the taxes collected are not extreme.

1.5 Conclusion

This essay provides empirical evidence, through controlled laboratory economics experiments, of the effectiveness of two nonpoint pollution control policies that theoretically induce voluntary compliance with a water quality standard through the background threat of an ambient tax. The first policy stems from Segerson and Wu (2006), and includes a threatened tax mechanism that is exogenous, in the respect that tax liabilities do not depend on the extent of voluntary noncompliance. The second policy, introduced in this essay, diverges from Segerson and Wu's mechanism in that the tax threshold in the threatened tax mechanism is conditional on decisions in the voluntary period. This endogenous threat mechanism theoretically eliminates undesirable, suboptimal SPNE that arise in the exogenous threat policy.

Experimental evidence suggests that making threats more severe, through lowering the threatened exogenous tax threshold or increasing the scale parameter in the endogenous voluntary-threat policy, increases the probability that groups achieve the ambient standard voluntarily. In addition, allowing participants the ability to coordinate their efforts through verbal communication leads to voluntary compliance in nearly 100% of decision periods, with essentially no individual-level deviations from the NE strategy. This positive result re-enforces the basic message of Poe et al. (2004) and Vossler et al. (2006): communication itself can be an important policy tool that can be used to greatly improve the ability of ambient policies to achieve water quality goals.

In groups where communication is not allowed, groups do not voluntarily achieve ambient pollution standards with great regularity unless the threatened tax policy is made severe. When the tax policy is put into place, participant emissions tend to deviate significantly from the cost-minimizing NE when the tax threshold is significantly below the ambient standard. This result is statistically identical to the decisions observed under the pure ambient tax policy with tax thresholds below the ambient standard. In addition, reductions in observed social efficiency levels between early and late rounds of the experiment is worrisome, as it suggests that participants learn to distrust or otherwise become frustrated with their fellow group members rather than learn to avoid the penalty provided by the tax policy. Finally, the range of earnings outcomes is wider under the exogenous rather than the endogenous voluntary-threat policy. This result arises from the penalty structure of the endogenous voluntary-threat policy, where small minor group deviations from the ambient standard result in correspondingly minor penalties in the subsequent tax policy.

Taken together, the experimental results suggest that the voluntary-threat policies are unlikely to solve ambient pollution problems unless polluter groups regularly engage in communication or the threatened tax policy is severe. While groups of agricultural producers may indeed communicate in small watersheds, communication sufficient to achieve ambient standards seems unlikely without external encouragement. In addition, the severe threats required to induce voluntary abatement in the absence of communication may serve to reduce the overall attractiveness of the voluntary-threat policies amongst agricultural producers. In addition, while the voluntary-threat policies offer firms the opportunity to achieve ambient standards without explicit regulation, other considerations may influence the policies' overall attractiveness. From the agricultural firm's perspective, the benefit of a policy that allows for voluntary compliance may be outweighed by the challenge of having to continuously alter production practices based on the type of regulatory mechanism that is in place. For example, if firms adopt drastically different strategies based on whether the voluntary or the tax policy is in place, then the management costs associated with changing cropping or abatement practices my reduce the financial benefits of the voluntary-threat policy. In addition, the attractiveness of the voluntary-threat mechanisms, from the policy maker's perspective, will be determined by the cost savings of not having to implement a tax mechanism in cases where groups achieve the standard voluntarily. When the threatened tax policy is severe, all groups are able to achieve the standard voluntarily for at least a portion of the experiment and in these cases the implementation of the tax policy is avoided.

These conclusions regarding the relative effectiveness of voluntary-threat policies could be bolstered by future research on variations of the policies tested here. Specifically, a better understanding of the tradeoffs between changing the number of threatened tax policy rounds (K) and increasing the severity of the tax in each of the tax policy rounds could lead to some important conclusions regarding the cost-effectiveness of each. In addition, the policy relevance of this research could be improved by investigating situations where participants are heterogeneous in terms of their baseline emissions and the structure of their abatement costs. Finally, explicitly incorporating uncertainty into the voluntary-threat mechanism is an important future step towards improving the policy's real-world relevance.

APPENDIX

Proof of Proposition 1: It has already been shown that in the K tax periods the strategy $\boldsymbol{a}^t = \boldsymbol{a}^*$ is a unique NE. The proof of Proposition 1 starts by showing that, when $\varphi \geq \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$ and $\boldsymbol{a}_{-i}^v = \boldsymbol{a}_{-i}^*$, firm i's best response is to choose $a_i^v = a_i^*$. In the second part of the proof it is shown that, when $\varphi < \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$ and $\boldsymbol{a}_{-i}^v = \boldsymbol{a}_{-i}^*$, it is not a best response for firm i to choose $a_i^v = a_i^*$.

When $\boldsymbol{a}_{-i}^v = \boldsymbol{a}_{-i}^*$, the standard will be achieved exactly if firm i chooses $a_i^v = a_i^*$. In this case the tax policy will not be imposed and the cost to the firm over K+1 periods will be $\sum_{k=0}^{K} \delta^k C\left(a_i^*, \theta_i\right)$. If firm i chooses $a_i^v > a_i^*$, ambient pollution will be below the standard and the tax policy will again not be imposed. However, the cost of choosing $a_i^v > a_i^*$ is greater than the cost of choosing $a_i^v = a_i^*$, which implies that this is not a best response. This result does not depend on the choice of φ.

If firm i chooses $a_i^v < a_i^*$, ambient pollution will exceed the standard and the tax policy will be put into place. The cost of the tax policy will depend on the firm's voluntary abatement decision. Firm *i's* optimal choice of voluntary abatement at or below a_i^* can be represented by the minimization problem

$$\underset{a_i^v}{Min} \quad C\left(a_i^v, \theta_i\right) + \sum_{k=1}^{K} \delta^k \left[C\left(a_i^*, \theta_i\right) + \tau\varphi\left(x\left(a_i^v, \boldsymbol{a}_{-i}^*; \theta_i, \boldsymbol{\theta}_{-i}\right) - x^s \right) \right] \quad \text{s.t. } a_i^v \leq a_i^*. \tag{a.1}$$

The associated Kuhn-Tucker (K-T) conditions are

$$\partial C \big/ \partial a_i^v + \tau\varphi \sum_{k=1}^{K} \delta^k \, \partial x \big/ \partial a_i^v + \mu = 0 \tag{a.1a}$$

$$\mu\left(a_i^* - a_i^v\right) = 0 \tag{a.1b}$$

$$\mu \geq 0. \tag{a.1c}$$

Two possible solutions must be considered; (1) $a_i^v = a_i^*$ with $\mu \geq 0$ and (2) $a_i^v < a_i^*$ with $\mu = 0$. Recall that $\tau = -\dfrac{\partial C / \partial a_i^*}{\partial x \big/ \partial a_i^*}$ so that condition (a.1a) implies $\dfrac{\partial C \big/ \partial a_i^v + \mu}{\partial x \big/ \partial a_i^v} = \varphi \sum_{k=1}^{K} \delta^k \dfrac{\partial C \big/ \partial a_i^*}{\partial x \big/ a_i^*}$. Clearly solution (1) will hold when $\varphi \geq \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$, since $\mu \geq 0$ can be chosen so that $1 + \dfrac{\mu}{\partial C \big/ \partial a_i^*} = \varphi \sum_{k=1}^{K} \delta^k$. Solution (2), however, can never hold.

Suppose otherwise, then it must be the case that $\frac{\partial C / \partial a_i^v}{\partial x / \partial a_i^v} = \varphi \sum_{k=1}^{K} \delta^k \frac{\partial C / \partial a_i^*}{\partial x / a_i^*}$. Given the strict convexity of the abatement cost function and the concavity of the pollution function, $\frac{\partial C / \partial a_i^v}{\partial x / \partial a_i^v} < \frac{\partial C / \partial a_i^*}{\partial x / a_i^*}$ for all $a_i^v < a_i^*$. Therefore when $\varphi \geq \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$, it follows that the K-T conditions do not hold, since $\frac{\partial C / \partial a_i^v}{\partial x / \partial a_i^v} < \varphi \sum_{k=1}^{K} \delta^k \frac{\partial C / \partial a_i^*}{\partial x / a_i^*}$, for $a_i^v < a_i^*$. Firm *i* minimizes costs given that the tax policy will be put in place by choosing $a_i^v = a_i^*$. This, however, represents the situation whereby the standard is achieved. The costs of choosing $a_i^v < a_i^*$ are therefore always greater than the cost of $a_i^v = a_i^*$, thus when $\varphi \geq \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$ and $\boldsymbol{a}_{-i}^v = \boldsymbol{a}_{-i}^*$, the unique best response for firm *i* is $a_i^v = a_i^*$.

When $\varphi < \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$, then $\frac{\partial C / \partial a_i^* + \mu}{\partial x / \partial a_i^*} > \varphi \sum_{k=1}^{K} \delta^k \frac{\partial C / \partial a_i^*}{\partial x / a_i^*}$ for all $\mu \geq 0$. Thus the K-T conditions (a.1a)-(a.1c) are solved only with $a_i^v < a_i^*$ and $\mu = 0$. The cost of choosing $a_i^v < a_i^*$ is strictly less than the cost of choosing $a_i^v = a_i^*$ and avoiding the tax policy. To see this, define $\hat{a}_i^v$ to be the level of voluntary abatement that solves the K-T condition (a.1a), i.e., $-\frac{\partial C / \partial \hat{a}_i^v}{\partial x / \partial \hat{a}_i^v} = \tau\varphi \sum_{k=1}^{K} \delta^k$. The cost to firm *i* of choosing $\hat{a}_i^v$ and facing the tax policy is then given by $C(\hat{a}_i^v, \theta_i) + \sum_{k=1}^{K} \delta^k \left[C(a_i^*, \theta_i) + \tau\varphi\left(x(\hat{a}_i^v, \boldsymbol{a}_{-i}^*; \theta_i, \boldsymbol{\theta}_{-i}) - x^s \right) \right]$. It was shown earlier that the cost of $a_i^v = a_i^*$ is $\sum_{k=0}^{K} \delta^k C(a_i^*, \theta_i)$. Thus the cost of $a_i^v = \hat{a}_i^v < a_i^*$ is lower than $a_i^v = a_i^*$ if $C(\hat{a}_i^v, \theta_i) + \sum_{k=1}^{K} \delta^k \tau\varphi\left(x(\hat{a}_i^v, \boldsymbol{a}_{-i}^*; \theta_i, \boldsymbol{\theta}_{-i}) - x^s \right) < C(a_i^*, \theta_i)$. Substituting $-\frac{\partial C / \partial \hat{a}_i^v}{\partial x / \partial \hat{a}_i^v} = \tau\varphi \sum_{k=1}^{K} \delta^k$, rearranging terms and dividing each side by $a_i^* - \hat{a}_i^v$, the inequality becomes

$$\left[\frac{\left(x(a_i^*, \boldsymbol{a}_{-i}^*; \theta_i, \boldsymbol{\theta}_{-i}) - x(\hat{a}_i^v, \boldsymbol{a}_{-i}^*; \theta_i, \boldsymbol{\theta}_{-i}) \right)}{a_i^* - \hat{a}_i^v} \right] \Big/ \partial x / \partial a_i^v < \left[\frac{\left(C(a_i^*, \theta_i) - C(\hat{a}_i^v, \theta_i) \right)}{a_i^* - \hat{a}_i^v} \right] \Big/ \partial C / \partial a_i^v,$$

which must hold because of the assumed curvature of the cost and pollution functions. Therefore $a_i^v = a_i^*$ is not a best response for firm *i* when $\varphi < \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$, and $\varphi \geq \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$ is necessary and sufficient to induce optimal compliance as part of a SPNE in the voluntary stage game. ■

Proof of Proposition 2: It has been shown that $\boldsymbol{a}^v = \boldsymbol{a}^*$ is part of a SPNE only when $\varphi \geq \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$. Next, it is shown that, when $\varphi = \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$, $\boldsymbol{a}^v = \boldsymbol{a}^*$ is the only possible

SPNE and, when $\varphi > \left(\sum_{k=1}^{K}\delta^k\right)^{-1}$, other SPNE are possible. Suppose that $\boldsymbol{a}_{-i}^v \leq \boldsymbol{a}_{-i}^*$, such that at least one firm is abating less than the socially optimal amount and firm *i* must overabate in order to meet the ambient standard. Let ε > 0 denote the amount of overabatement needed by firm *i* to ensure that the standard is met, i.e., $a_i^{v(s)} = a_i^* + \varepsilon$. Firm *i* would never choose $a_i^v > a_i^{v(s)}$ since this implies higher abatement costs without a reduction in tax burden. If firm *i* chooses $a_i^v < a_i^{v(s)}$ the tax policy will be imposed. The optimal choice of voluntary abatement given that the tax policy will be imposed is determined by the cost minimization problem

$$\underset{a_i^v}{Min}\ C(a_i^v,\theta_i)+\sum_{k=1}^{K}\delta^k\left[C(a_i^*,\theta_i)+\tau\varphi\left(x(a_i^v,\boldsymbol{a}_{-i}^v;\theta_i,\boldsymbol{\theta}_{-i})-x(a_i^*+\varepsilon,\boldsymbol{a}_{-i}^v;\theta_i,\boldsymbol{\theta}_{-i})\right)\right]\ \ s.t.\ \ a_i^v \leq a_i^*+\varepsilon \quad \text{(a.2)}$$

With corresponding K-T conditions

$$\partial C/\partial a_i^v + \tau\varphi\sum_{k=1}^{K}\delta^k\,\partial x/\partial a_i^v + \mu = 0 \quad \text{(a.2a)}$$

$$\mu\left(a_i^* + \varepsilon - a_i^v\right) = 0 \quad \text{(a.2b)}$$

$$\mu \geq 0. \quad \text{(a.2c)}$$

Two solutions must again be considered; (1) $a_i^v = a_i^* + \varepsilon$ with $\mu \geq 0$ and (2) $a_i^v < a_i^* + \varepsilon$ with $\mu = 0$. Substituting $\tau = -\frac{\partial C/\partial a_i^*}{\partial x/\partial a_i^*}$ condition (a.2a) implies $\frac{\partial C/\partial a_i^v + \mu}{\partial x/\partial a_i^v} = \varphi\sum_{k=1}^{K}\delta^k\frac{\partial C/\partial a_i^*}{\partial x/a_i^*}$. The only possible solution, when $\varphi = \left(\sum_{k=1}^{K}\delta^k\right)^{-1}$, has $a_i^v = a_i^*$ and $\mu = 0$. There can never be a solution with $a_i^v = a_i^* + \varepsilon$ and $\mu \geq 0$, since this would require that $\frac{\partial C/\partial a_i^v}{\partial x/\partial a_i^v} \leq \frac{\partial C/\partial a_i^*}{\partial x/a_i^*}$ for $a_i^v > a_i^*$, which is not possible given the curvature of the cost and emissions functions. Thus there cannot be a SPNE where one firm chooses $a_i^v < a_i^*$ and another firm chooses $a_i^v > a_i^*$. Setting $\varphi = \left(\sum_{k=1}^{K}\delta^k\right)^{-1}$ is therefore sufficient to induce $\boldsymbol{a}^v = \boldsymbol{a}^*$ as part of a unique SPNE.

When $\varphi > \left(\sum_{k=1}^{K}\delta^k\right)^{-1}$ then $\frac{\partial C/\partial a_i^*}{\partial x/\partial a_i^*} < \varphi\sum_{k=1}^{K}\delta^k\frac{\partial C/\partial a_i^*}{\partial x/a_i^*}$ and K-T condition (a.2a) will not be satisfied with $a_i^v = a_i^*$. Firm *i* will therefore not optimally choose $a_i^v = a_i^*$, when $a_i^v > a_i^{v(s)}$ and $\varphi > \left(\sum_{k=1}^{K}\delta^k\right)^{-1}$. To satisfy condition (a.2a), firm *i* must choose an abatement level $a_i^v > a_i^*$. Further, for all $\varphi > \left(\sum_{k=1}^{K}\delta^k\right)^{-1}$, there is some $\varepsilon > 0$ for which firm *i* will choose $a_i^v = a_i^{v(s)} = a_i^* + \varepsilon$ so that the ambient standard is achieved. When the standard is achieved with equality, none of the other n-1 firms in the watershed have an incentive to deviate from their strategies. The marginal benefit to firm j, in terms of cost savings, of reducing abatement is $\partial C / \partial a_j^v$ while the marginal cost of the increased tax payments is given by $\varphi \sum_{k=1}^{K} \delta^k \frac{\partial C / \partial a_j^*}{\partial x / \partial a_j^*} \cdot \partial x / a_j^v$. As long as $\varphi > \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$ then $\partial C / \partial a_j^v < \varphi \sum_{k=1}^{K} \delta^k \frac{\partial C / \partial a_j^*}{\partial x / \partial a_j^*} \cdot \partial x / a_j^v$ for all $a_j^v \le a_j^*$. Therefore none of the n-1 firms will optimally deviate from their strategy and there can be a SPNE where at least one firm in the watershed chooses $a_j^v < a_j^*$ and another firm chooses $a_i^v > a_i^*$. Setting $\varphi = \left(\sum_{k=1}^{K} \delta^k \right)^{-1}$ is thus necessary and sufficient for $\boldsymbol{a}^{\text{v}} = \boldsymbol{a}^*$ to be a *unique* SPNE.■

Proof of Proposition 3: Suppose $\boldsymbol{a}_{-i}^{\boldsymbol{v}} = 0$. Firm i would then make its abatement decision based on the minimization problem in equation (a.2), and condition (a.2a) implies $\partial C / \partial a_i^v + \mu = -\varphi \sum_{k=1}^{K} \delta^k \tau \cdot \partial x / \partial a_i^v$. Since $\tau > 0$ and $\partial x / \partial a_i^v < 0$ it follows that either $\mu > 0$, $\partial C / \partial a_i^v > 0$ or both. Since $\mu > 0$ requires $a_i^v > 0$ from (a.2b) and $\partial C / \partial a_i^v > 0$ also requires $a_i^v > 0$, firm i will optimally choose a positive level of abatement even when all other firms choose a strategy of zero abatement. Therefore the zero abatement equilibrium is eliminated with the endogenous voluntary-threat policy when $\varphi > 0$. ■

Experiment Instructions

Introduction

This experiment is a study of individual and group decision-making. If you follow these instructions carefully and make informed decisions you will earn money. The money you earn will be paid to you, in cash, at the end of the experiment. A research foundation has provided the funding for this study.

You will be in a group consisting of six players. Each player assumes the role of a different firm. Think of your firm and the five other firms as being located near a common water resource.

Your firm yields earnings through its operations. Your firm's operations also generate emissions, which affect the water quality of the common water resource. The combined emissions from each of the six firms located near the water resource determine the level of **Total Pollution** in each round. Pollution affects the well-being of water resource users. For example, high pollution levels affect the health of fish, causing losses to fisherman.

The experiment is broken up into many decision "rounds". There are two parts to the experiment. Part A of the experiment consists of the first 5 rounds, whereas Part B includes the remaining rounds. You will be given additional instructions after Part A is completed.

In each round you must make an **Emissions Decision**. In general, the lower your firm's emissions, the lower the level of earnings for your firm. You have been provided a sheet titled *Emissions Decision Sheet* that lists the level of **Firm Earnings** associated with various levels of emissions generated by your firm. Firm Earnings are denominated in "tokens", which will be exchanged for cash at the end of the experiment according to the exchange rate listed on the *Emissions Decision Sheet*. In addition to Firm Earnings, you are also given 5,000 tokens of **General Earnings** in each round.

All six firms in your group are <u>identical</u> in every way, meaning that all of the firms face the same relationship between earnings and emissions.

A round of the experiment is complete when all six players have made their emissions decisions. The computer will then report the **Total Pollution** for that round. It will also calculate your ***Round Earnings*** by summing up **Firm Earnings** + **General Earnings**. Pollution does not affect your earnings whatsoever in Part A of the experiment. Below we explain how to make decisions using your computer.

USING THE COMPUTER

In each round, your task is to make an **Emissions Decision**. The Emissions Decision that you type in must be a whole number that is in the range listed on your computer screen. When you type in an emissions decision and hit the enter key, the corresponding **Firm Earnings** amount will appear on your screen. You can verify that the firm earnings amount that appears is identical to that provided on the *Emissions Decisions Sheet*.

When you are satisfied with your Emissions Decision, you must then click the <**SUBMIT>** button for that round. Once you have clicked the <**SUBMIT>** button, it is no longer possible to change your decision.

After all six players have clicked the submit button, the experiment moderator will instruct you to click the <**RECEIVE>** button. After clicking the <**RECEIVE>** button, the cells indicating the **Total Pollution** and ***Round Earnings*** will be filled in. The Round Earnings cell simply sums up your Firm Earnings plus General Earnings. Recall that pollution does not affect your earnings in Part A of the experiment.

As the experiment progresses, the total number of tokens you have earned will be calculated in the **Total Tokens** box located in the lower right portion of the spreadsheet. The **Total Earnings ($)** box displays the amount of money you have earned, in U.S. dollars, after the tokens have been exchanged.

INSTRUCTIONS FOR PART B

Please click the ***Go on to Part B*** button located underneath the Total Earnings ($) cell. In Part B of the experiment you will continue to make an **Emissions Decision**. A key difference, however, is that in order to protect the water resource a **Voluntary Policy** has been put in place, whereby the firms in your group have been asked to reduce **Total Pollution** to 72 units. As long as Total Pollution for the group is reduced to the target of 72 or less, the Voluntary Policy will continue to apply. If, however, the group fails to reduce Total Pollution to at or below 72 in a round, a regulator will put in place a **Tax Policy**. The Tax Policy will begin in the round following the failure to meet the target voluntarily and will last for three rounds. The Tax Payment that will be due in each round in which the Tax Policy is in place will depend on Total Pollution in that round. Specifically, under the Tax Policy you, and everyone else in your group, must make the following **Tax Payment** in each round:

If Total Pollution is equal to or less than **50**: **Tax Payment** = 0

If Total Pollution is greater than **50**: **Tax Payment** = 2,500 * **(Total Pollution** – 50)

That is, under the Tax Policy, <u>each</u> player pays 2,500 tokens for every unit of pollution above 50 units. It is important to remember that Total Pollution is based on the emissions decisions of <u>everyone</u> in your group, not just your own. The *Tax Calculation Sheet* that has been provided to you indicates the **Tax Payment** corresponding to levels of Total Pollution under the Tax Policy.

As an example, suppose that in Round 6 Total Pollution for the group is 70. Since this means the target of 72 has been met voluntarily, no tax will be imposed and the Voluntary Policy will remain in effect for Round 7. However, in Round 7 suppose Total Pollution for the group is 80, implying that the target has not been met voluntarily in this round. No Tax Payment is due for Round 7, but in Round 8 the Tax Policy will be imposed and will stay in effect for rounds 8, 9 and 10. In these rounds, each firm will be charged a Tax Payment determined by the above formula. For example, if Total Pollution for the group is 70 you will pay a tax of 2,500 * (70 – 50) = 50,000. Thus, when the Tax Policy is in place each firm will make a Tax Payment even when the target of 72 is met.

Once your group has been subject to the Tax Policy for 3 rounds (after Round 10 in the above example), the Tax Policy will be removed and in the next round (Round 11) the Voluntary Policy will be put back in place. Firms will then be given another chance to meet the target voluntarily. Note that, if the target is met voluntarily in all rounds (starting from Round 6), the Tax Policy will never be imposed. For each round, the top cell indicates whether the Voluntary Policy or Tax Policy is in place. After everyone makes his or her emissions decision, Total Pollution will be calculated as before. If the Voluntary Policy is in place, then the Tax Payment for that round will be zero. Otherwise, the Tax Payment will be calculated using the formula above. The Tax Payment, if any, will be deducted from your earnings so that ***Round Earnings*** = **Firm Earnings – Tax Payment + General Earnings**.

Understanding How the Experiment Works

Although the computer does all of the calculations, it is important to us that you understand how the calculations work. Below we have provided two tables and we ask that you fill in the blank cells in each. Someone will look over your calculations shortly and provide any assistance that you may need.

The first table provides two examples of hypothetical Total Pollution amounts for Rounds 6-10. Please indicate whether the Voluntary or Tax Policy applies for the remaining rounds. If the Tax Policy applies, use the *Tax Calculation Sheet* to determine the amount of the Tax Payment.

		Total Pollution	**Tax** or **Voluntary?**	**Tax Payment** (if any)
Example 1	Round 6	90	Voluntary	
	Round 7	67		
	Round 8	80		
	Round 9	107		
	Round 10	31		
Example 2	Round 6	68	Voluntary	
	Round 7	72		
	Round 8	53		
	Round 9	92		
	Round 10	69		

Example 3

Suppose that the Tax Policy applies in Example 3. In the table below, first make an Emissions Decision and then make a guess at the combined emissions from the other five firms. Note that there is no right or wrong answers for these two items. Next, fill in the remaining empty fields of the table using the *Emissions Decision Sheet* and *Tax Calculation Sheet* as references.

Emissions Decision *(you choose)*	
Firm Earnings *(from Emissions Decision Sheet)*	
General Earnings	**5,000**
Combined emissions from the other 5 firms *(you choose)*	
Total Pollution	
Tax Payment *(from Tax Calculation Sheet)*	
Round Earnings	

Sample Emissions Decision Sheet

Emissions Decision Sheet	
Exchange Rate: $1 = 70,000 Tokens	
Emissions Decision	**Firm Earnings**
22	74,896
21	74,987
20	75,000
19	74,987
18	74,896
17	74,648
16	74,167
15	73,372
14	72,188
13	70,534
12	68,333
11	65,508
10	61,979
9	57,669
8	52,500
7	46,393
6	39,271
5	31,055
4	21,667
3	11,029

Sample Tax Calculation Sheet for Tax Threshold = 50

Tax Calculation Sheet					
Total Pollution	**Tax**	**Total Pollution**	**Tax**	**Total Pollution**	**Tax**
31	0	65	37,500	99	122,500
32	0	66	40,000	100	125,000
33	0	67	42,500	101	127,500
34	0	68	45,000	102	130,000
35	0	69	47,500	103	132,500
36	0	70	50,000	104	135,000
37	0	71	52,500	105	137,500
38	0	72	55,000	106	140,000
39	0	73	57,500	107	142,500
40	0	74	60,000	108	145,000
41	0	75	62,500	109	147,500
42	0	76	65,000	110	150,000
43	0	77	67,500	111	152,500
44	0	78	70,000	112	155,000
45	0	79	72,500	113	157,500
46	0	80	75,000	114	160,000
47	0	81	77,500	115	162,500
48	0	82	80,000	116	165,000
49	0	83	82,500	117	167,500
50	0	84	85,000	118	170,000
51	2,500	85	87,500	119	172,500
52	5,000	86	90,000	120	175,000
53	7,500	87	92,500	121	177,500
54	10,000	88	95,000	122	180,000
55	12,500	89	97,500	123	182,500
56	15,000	90	100,000	124	185,000
57	17,500	91	102,500	125	187,500
58	20,000	92	105,000	126	190,000
59	22,500	93	107,500	127	192,500
60	25,000	94	110,000	128	195,000
61	27,500	95	112,500	129	197,500
62	30,000	96	115,000	130	200,000
63	32,500	97	117,500	131	202,500
64	35,000	98	120,000	132	205,000

Sample SAS Program for Participant Decision Model

```
/*Imports datafile from excel database*/
Proc Import out=sw_full
            DATAFILE= "U:\User5\jfs24\Segerson Wu\Import SW.xls"
       DBMS=EXCEL REPLACE;
     SHEET="'Import SW$'";
     GETNAMES=YES;
     MIXED=NO;
     SCANTEXT=YES;
     USEDATE=YES;
     SCANTIME=YES;
run;

/*Cleans original dataset and adds period variable*/
data clean;
      set sw_full;
      if round < 6 then period=1;
      if 5 < round < 15 then period=2;
      if 14 < round < 24 then period=3;
      if period="" then delete;
      voluntary=aptax;
      keep treatment group subject period voluntary emissions surplus
      round;
run;

/*Makes necessary calculations for preparing data*/
data easy_mix;
      set clean;
      if period=1 then parta=1;
      else parta=0;
      if group < 3 then ID=subject;
      else ID=subject + 12;
      ID=24*treatment+ID;
      if treatment=8 then delete;
      if treatment=13 then delete;
      GroupID=(treatment-1)*4 + group;
run;

/*Mixed model*/
proc mixed data=easy_mix /*NOitprint*/;
      class treatment voluntary GroupID ID period round;
      model emissions= treatment period voluntary
      treatment*period*voluntary;
      lsmeans treatment*period*voluntary / ;
      repeated round / subject =ID type=ar(1) group=treatment r;
      random groupID ;
run;
```

REFERENCES

Alpízar, Francisco, Till Requate, and Albert Schram. 2004. Collective versus Random Fining: An Experimental Study on Controlling Ambient Pollution. *Environmental and Resource Economics* **29**: 231-252.

Cabe, Richard, Joseph A. Herriges. 1992 The Regulation of Non-Point Sources of Pollution Under Imperfect and Asymmetric Information. . *Journal of Environmental Economics and Management* **32** 134-146.

Chesapeake Bay Program. Updated 2005. Overview of the Bay Program. Available: http://www.chesapeakebay.net/overview.htm.

Cochard, Francois, Marc Willinger, and Anastasios Xepapadeas**.** 2005. Efficiency of Nonpoint Source Pollution Instruments: An Experimental Study. 30: 393-422 *Environmental and Resource Economics* **29**: 231-252.

Dairy Farm Management: Business Summary New York State. 1999-2003. Department of Applied Economics and Management. Cornell University. Ithaca, NY.

Economic Research Service. Updated 2002. U. S. Department of Agriculture. ERS Analysis: Conservation Programs. Available: http://www.ers.usda.gov/Features/FarmBill/analysis/conservationoverview.htm

Hansen, Lars G. 1998. A Damage Based Tax Mechanism for Regulation of Non-Point Emissions. *Environmental and Resource Economics* **12**: 99–112.

Horan, Richard D., James S. Shortle, and David G. Abler. 1998. Ambient Taxes when Polluters have Multiple Choices. *Journal of Environmental Economics and Management* **36**: 186-199.

Karp, Larry. 2005. Nonpoint Source Pollution Taxes and Excessive Tax Burden. *Environmental and Resource Economics* **32**: 229–251.

National Resources Inventory. 2002. 2002 Annual NRI. United States Department of Agriculture. Natural Resources Conservation Service. Available: http://www.nrcs.usda.gov/technical/land/nri02/

Poe, Gregory L., William D. Schulze, Kathleen Segerson, Jordan F. Suter and Christian A. Vossler. 2004. Exploring the Performance of Ambient-Based Policy Instruments when Non-Point Source Polluters can Cooperate. *American Journal of Agricultural Economics* **86**: 1203-1210.

Ribaudo, Marc. 2003. Agricultural Resources and Environmental Indicators, 2003. Chapter 2.3 Water Quality Impacts of Agriculture. *Agricultural Handbook AH722*. United States Department of Agriculture. Economic Research Service.

Ribaudo, Marc. 2004. Policy Explorations and Implications for Nonpoint Source Pollution Control: Discussion. *American Journal of Agricultural Economics* **86**: 1220-1221.

Segerson, Kathleen. 1988. Uncertainty and Incentives for Non-Point Source Pollution. *Journal of Environmental Economics and Management* **15**: 87-98.

Segerson, Kathleen, and T.J. Miceli. 1998. Voluntary Environmental Agreements: Good News for Environmental Quality? *Journal of Environmental Economics and Management* **36:** 109–130.

Segerson, Kathleen, J. Wu. 2006. Nonpoint Source Pollution Control: Introducing First Best Outcomes Through the Use of Threats. *Journal of Environmental Economics and Management* **51**: 165-184.

Shortle, J., D. Abler, and M. Ribaudo. 2001. Agriculture and Water Quality: The Issues. In J. Shortle and D. Abler (Eds.) *Environmental Policies for Agricultural Pollution*. Walingford, UK: CAB International.

Spraggon, John. 2002. Exogenous Standarding Instruments as a Solution to Group Moral Hazards. *Journal of Public Economics* **84**: 427-456.

Spraggon, John. 2004. Testing Ambient Pollution Instruments with Heterogeneous Agents. *Journal of Environmental Economics and Management* **48**: 837-856.

Suter, Jordan F., Christian A. Vossler, Gregory L. Poe and Kathleen Segerson. Forthcoming. Experiments on Damage-Based Ambient Taxes for Nonpoint Source Polluters. *American Journal of Agricultural Economics*.

U.S. Environmental Protection Agency. 2002. *National Water Quality Inventory 2000 Report*. Office of Water, Washington, DC.

Vossler, Christian A., Gregory L. Poe, William D. Schulze, and Kathleen Segerson. 2006. Communication and Incentive Mechanisms Based on Group Performance: An Experimental Study of Nonpoint Pollution Control. *Economic Inquiry* **44**: 599-613.

Xepapadeas, Anastasios. 1991. Environmental Policy under Imperfect Information: Incentives and Moral Hazard. *Journal of Environmental Economics and Management* **20**: 113-126.

CHAPTER TWO:
Incorporating Economic and Ecological Information into the Optimal Design of a Wildlife Corridor for the Northern Rockies

Abstract

In an attempt to address the negative ecological impacts of habitat fragmentation, wildlife corridors have been proposed as a way to connect areas of biological significance. In this essay, a model to maximize the amount of suitable wildlife habitat in a fully connected parcel network linking core habitat areas subject to a budget constraint is introduced. The standard economic framework of maximizing benefits subject to a budget constraint that is employed is a divergence from other recently proposed models that focus only on minimizing the cost of a single parcel-wide corridor. While the budget constrained optimization model that is introduced is intuitively appealing, it presents substantial computational challenges above determining the cost-minimizing corridor. The optimization model is applied to the design of a wildlife corridor for grizzly bears in the U.S. Northern Rockies and is shown to drastically increase the total habitat suitability of the corridor over parcel selection based on cost minimization alone. In addition, an efficiency frontier is generated that illustrates the relative tradeoffs between corridor cost and habitat suitability. Finally, for cases where optimization is computationally impractical, a heuristic is suggested that closely approximates the optimally selected corridor.

2.1 Introduction

In many parts of the world, land development has resulted in a reduction and fragmentation of natural habitat, leading to increased rates of species decline and extinction. To combat the negative consequences of anthropogenic habitat fragmentation, the procurement of biologically valuable conservation land has been promoted by biologists as a way to ensure species viability. A large number of models for optimally selecting land parcels for conservation, formally referred to as the reserve site selection problem (RSSP), have been proposed in the conservation biology literature. These models select parcels to ensure that all targeted species in a given region are protected, as in the Set Covering Problem (SCP) (e.g., Margules, Nicholls and Pressey 1988; Underhill 1994), or they select a constrained number of parcels that maximize species richness, as in the Maximal Covering Problem (MCP) (e.g., Church, Stoms and Davis 1996; Camm et al. 1996).

A number of subsequent studies have added to the conservation biology literature by incorporating economic variables into the RSSP. These studies seek to procure conservation parcels, given a budget constraint, that maximize the number of species protected (e.g., Ando et al. 1998; Polasky et al. 2001; Costello and Polasky 2004) or maximize the environmental benefits of the sites selected (e.g., Ferraro 2003; Messer 2005; Newburn, Berck and Merenlender 2006). The results of these economic-based studies show that incorporating spatially heterogeneous financial costs into reserve site selection models leads to a substantially different set of priority parcels than standard SCP or MCP models that ignore parcel costs. Moreover, the parcels selected based on budget constrained optimization obtain considerably greater environmental benefits for the same conservation budget than traditional site selection models (Balmford, Gaston and Rodrigues 2000; Naidoo et al. 2006).

In recent years, biologists and economists have recognized that a parcel's spatial location relative to other protected parcels is also an essential attribute to consider in reserve site selection. Reflecting this, a variety of models that seek to increase the degree of spatial coherence in the set of parcels selected for conservation have been developed (Williams, ReVelle and Levin 2005 provide a thorough review). One primary way in which spatial attributes have been incorporated into site selection models is through the optimal selection of a connected reserve network, which is referred to here as a *wildlife corridor*.[1] The focus on developing models for the design of optimal wildlife corridors has come as biologists have highlighted the environmental imperative of connecting core areas of biological significance (Noss 1987).

Beginning with the work of Sessions (1992), several models have been developed that attempt to optimally select a spatially connected set of parcels (e.g., Williams 1998, 2002; Williams and Snyder 2005; Cerdeira et al. 2005; Onal and Briers 2006; Fuller et al. 2006). The design of an optimal corridor is, in its essence, a standard economic problem where the conservation planner is attempting to select the most ecologically beneficial corridor given the conservation funding available. Previous models of optimal corridor design, however, have not considered the case of budget constrained optimization. In fact, with the exception of Sessions (1992) and Williams (1998), spatially heterogeneous parcel cost has been ignored altogether. The formulation of the corridor design problem that is presented in this essay is the most relevant to a conservation planner, operating with limited funds with which to secure conservation land. The budget constrained optimization model, however, introduces some additional computational challenges and design elements into reserve selection, over previous corridor selection models that have sought to minimize the number of sites selected such that a specific number of species are preserved (Onal and Briers 2006; Fuller 2006; and Cerdeira 2006) or minimize the amount of unsuitable habitat in the corridor (Williams 1998, 2002; Williams and Snyder 2005).

---

[1] Wildlife corridors are also referred to more or less interchangeably as conservation, habitat, and movement corridors.

In this essay, a spatially explicit model that seeks the optimal construction of a wildlife corridor between multiple areas of biological significance is proposed. The model is then applied to the design of a wildlife corridor for grizzly bears connecting the Yellowstone, Salmon-Selway and Northern Continental Divide Ecosystems in Idaho, Wyoming and Montana. The results from the budget constrained optimization model are then used to define an efficiency frontier that highlights the tradeoffs between corridor cost and overall suitable habitat included in the corridor.

In the next section, the optimization model is motivated by highlighting the implementation of corridor projects in various parts of the world and reviewing the literature on optimal corridor design. In section 2.3, specific corridor design problem statements are introduced and in section 2.4 the programming model is described. The application of a wildlife corridor for grizzly bears in the United States Northern Rocky Mountain region is outlined in section 2.5 and the data sources used in the analysis are described in 2.6. Results of the corridor optimization in the Northern Rockies are provided in section 2.7. In section 2.8, a heuristic is suggested that is computationally more practical than optimization when a large number of parcels are available for selection. The concluding section describes implications for policy and future directions.

2.2 Review of Corridor Implementation and Literature

Properly implemented wildlife corridors provide numerous ecological benefits by returning the landscape to its natural connected state. By allowing species the ability to migrate between core areas of biological significance, corridors increase gene flow and reduce rates of inbreeding, thereby improving species fitness and survival (Schmitt and Seitz 2002). Corridors also allow for greater mobility (Andreassen et al. 1996), thus allowing the potential for species to escape predation and respond to stochastic events such as fire. Additionally, corridors allow species to respond more easily to long term climatic changes (McEuen 1993).

Responding to the ecological benefits of connected ecosystems, a wide range of corridor projects have been proposed or are currently being implemented. The projects range from local scale projects, such as the Quimper Wildlife Corridor, which provides a 3.5 mile greenbelt in Jefferson County, WA, to much wider scale projects like the 'Yellowstone to Yukon' initiative, which seeks to implement a viable corridor stretching from Yellowstone National Park in northwestern Wyoming to the Yukon region of western Canada. Corridor projects are currently being planned or implemented by governments and private organizations across the world. In Europe, for example, numerous countries have initiated wildlife corridor projects, such as the National Ecological Network in the Netherlands. Near the India-Bangladesh border, the Siju-Rewak Corridor currently connects elephant populations in the Siju Wildlife Sanctuary and Rewak Reserve Forests. Similarly, the proposed Selous-Niassa Wildlife Protection Corridor Project in Africa would link game reserves in Tanzania and Mozambique to form Africa's largest protected area. In northern Brazil, the Amapa Biodiversity Corridor connects 11 million hectares of some of the most pristine remaining areas of the Amazon Rainforest. In eastern Australia, a proposed 2,800 km corridor would link existing reserves in a corridor project dubbed the 'Alps to Artherton'. This is by no means an exhaustive list of corridor projects currently being implemented around the world, but it is meant to illustrate the policy relevance of wildlife corridor design on several continents.

Despite the increasing number of corridors being implemented around the world, and several studies documenting the positive ecological benefits of existing corridors (e.g., Tewksbury et al. 2002; Haddad et al. 2003; Dixon et al. 2006), models for the optimal selection of corridor parcels have received comparatively little attention. The problem of optimal corridor design was first posed by Sessions (1992), who models the selection of a hypothetical corridor as a network Steiner tree (NST) problem. The hypothetical formulation employed by Sessions involves a landscape composed of a set of available parcels to connect a subset of critical parcels. The cost of each parcel is defined as the opportunity cost of not harvesting the parcel's timber, which is assumed to be known and the model objective is to connect the critical parcels with the least-cost set of available parcels. Noting that arriving at a solution may not possible in polynomial time for a large set of parcels, Sessions uses a shortest path heuristic to select parcels that minimize the cost of connecting the critical parcels.

Williams (1998) also models the optimal selection of a hypothetical corridor as a NST problem, and addresses the dual objectives of minimizing corridor cost and minimizing the amount of unsuitable area included in the corridor. Using linear integer programming, Williams generates an efficiency frontier by varying the weights placed on each of the two objectives. In other words, he finds a set of points where one model objective (e.g., cost) cannot be made lower without increasing the other objective (e.g., unsuitable habitat). The efficiency frontier that is generated allows for a comparison of the tradeoffs between corridor cost and habitat unsuitability.

In subsequent work, Williams modifies his original model to consider cases where there are no predefined reserves and the planner is simply trying to form a connected reserve (Williams 2002; Williams and Snyder 2005). In Williams (2002) he considers a relaxation of the contiguity requirement by incorporating a separate contiguity parameter that can be adjusted to control the overall degree of connectivity in the parcels selected. In Williams and Snyder (2005), the authors take up the special case of percolating clusters, where the corridor is selected so as to connect one end of the landscape to the other (i.e., from north to south).

The studies by Sessions and Williams are groundbreaking in the formulations of the corridor problem that they introduce. Their models, however, only allow each parcel to be connected to two other parcels in the corridor. Considering only a one parcel-wide corridor, however, rules out the possibility of a corridor being "thicker" (i.e., multiple parcels wide) for at least some portion of the path. This would be beneficial, for example, if there is an agglomeration of high quality and low cost habitat in some portion of the corridor that could be cost-effectively incorporated into the reserve system. In addition, the authors do not extend their research to the study of an applied corridor instance, making it difficult to determine how the models perform in practice. Finally, although the problem Williams poses in his 1998 article is novel in that it incorporates both the financial and environmental attributes of each parcel, including measures of unsuitable habitat in the objective function results in some perverse incentives. For example, suppose that two parcels have identical cost, cover the same linear distance between two reserves, and have the same percentage of suitable to unsuitable habitat. The only difference between the two parcels is that parcel A is wider, and therefore covers more area, than parcel B. The optimization model would tend to prefer parcel A over parcel B, all else being equal, since parcel B has more aggregate unsuitable area. This perverse incentive can be remedied by maximizing suitable area, as is done in this study, rather than minimizing unsuitable area.

Recent articles by Cerdeira et al. (2005), Önal and Briers (2006) and Fuller et al. (2006) introduce models of optimal corridor design and apply them to specific study areas. Cerdeira et al. (2005) formulate a linear integer programming approach to solve a fully connected set covering problem and apply their model to the case of 496 uniform and contiguous parcels in the county of Hertfordshire, UK. They find that a minimum of 22 contiguous sites are needed to optimally cover the 45 species of butterflies in the study area. A heuristic method that they develop in the paper selects 23 sites for conservation, which the authors take as evidence that their heuristic performs well in comparison to exact methods. Önal and Briers (2006) also formulate a fully connected set covering problem as a linear integer program. They apply their model to 121 bird species dispersed over 391 parcels in Berkshire County, UK and show that the model is too complex to be solved. They then outline a procedure that involves solving the problem at a more aggregate scale and then selecting the minimum set of small disaggregate sites within the aggregate solution that cover all 121 species. This procedure is not guaranteed to find the optimal solution, since the minimum disaggregate number of sites could occur outside of the first stage, aggregate solution. The algorithm performs more favorably, however, than a heuristic procedure that is an extension of the greedy algorithm, where parcels are selected sequentially based on their contribution to the number of remaining unpreserved species. Finally, Fuller et al. (2006) apply a three stage algorithm to select a connected conservation network in central Mexico. They begin by selecting sites for conservation based on the habitat requirements of 99 species. They then define a set of paths that link the conservation areas with parcels containing suitable habitat. Finally, in the third stage, the paths that have the smallest area and impact on human populations are selected to form the connected reserve network.

The model presented in the next section diverges from previous corridor design studies in four important ways. First, the problem is modeled as one of finding the set of corridor parcels that maximize habitat suitability, subject to a budget constraint. This is a change from previous studies that have modeled the problem as one of minimizing some aspect of parcel cost, either in terms of number of parcels, financial cost, or cost to wildlife traversing the corridor. This is the first corridor model to explicitly include a budget constraint; something that likely improves the relevance of the model for conservation planners, who generally operate in an environment with limited budgets. The second primary divergence from the studies reviewed above is that the model presented here does not limit the selected corridor to being only one parcel wide. This is important because it means that if the budget allotted for the corridor is higher than the minimum cost corridor, then the benefits of the corridor can be improved either by selecting a new route, or by simply making the corridor wider so that it cost-effectively includes adjacent parcels. This reformulation, while intuitively appealing, presents additional computational challenges, as illustrated by the model results. The third contribution of this study is that it incorporates both estimated parcel costs and habitat suitability measures from a naturally occurring landscape. Williams (1998), the only other study to consider both parcel costs and habitat benefits relies on empirical results from a purely hypothetical instance. Finally, by changing the granularity of the parcels available for selection, a greater understanding of the relationship between computational complexity and the number of parcels in the landscape is gained.

2.3 Connection Subgraph Problem

The corridor model that is presented here assumes a landscape that is divided up into a set of contiguous, non-overlapping parcels. Utilizing terminology from graph theory, the landscape is represented by a graph (*G*) made up of vertices (parcels) and edges (parcel adjacencies) so that $G = G(V,E)$. A subset of the vertices in the graph are predefined as terminal vertices (reserves), $T \subseteq V$. Next, it is assumed that associated with each vertex is nonnegative cost, *c*, representing the amount necessary to secure the vertex for inclusion in the corridor, and a nonnegative utility, *u*, which represents the environmental benefit (i.e., habitat suitability) of the vertex. Finally, it is assumed that the conservation planner has a finite budget constraint, *B*, and a desired level of aggregate utility, *U*. The Connection Subgraph Problem requires finding a subgraph *H* of *G* such that (1) *H* is fully connected (2) $T \subseteq V(H)$, i.e., the subgraph includes all terminal vertices (3) $\sum_{v \in V(H)} c(v) \leq B$, i.e., the subgraph has aggregate cost no greater then the available budget and (4) $\sum_{v \in V(H)} u(v) \geq U$, i.e., the subgraph has aggregate utility of at least the desired level.

The last two conditions can then be relaxed to obtain three separate optimization problems of interest to the conservation planner.

(1) Budget Constrained Utility Optimization

(2) Utility Constrained Cost Minimization

(3) Unconstrained Cost Minimization

Note that unconstrained utility optimization could also be added to the set of problems above, but since it is assumed that each vertex has nonnegative utility this would simply entail a subgraph that is identical to the graph itself, i.e., every parcel in the landscape is acquired.

By comparing the connection subgraph problem to the network Steiner tree problem, it can be shown that the connection subgraph problem is NP-complete (Conrad et al. 2007). NP-completeness is a term used in computational complexity theory to define a problem where it is "easy" to verify that a particular solution satisfies the constraints[2] (i.e., the reserves are connected and the utility and budget constraint is met), but it is potentially not possible to prove that a particular feasible solution is an optimum. Proving optimality may not be possible, because the time necessary to solve the problem increases exponentially as the number of vertices increase. It is possible that the optimality of the connection subgraph problem cannot be proven in a realistic timeframe, if there are a large number of available parcels in the landscape.

The differences in terms of parcel selection and computational complexity of cost constrained utility optimization, as opposed to the unconstrained cost minimization, are illustrated in the hypothetical 3x3 parcel map presented in Figure 2.1, where parcels C and G are to be connected with a contiguous corridor. In this simple example, corridor costs are minimized with the selection of parcels B, E, and H as shown in panel I. With this selection, the cost is 7 units and the utility of the parcels selected is 5. Now suppose that the conservation planner has available a budget of 10 units. Rather than simply selecting the least cost path, the planner would now be interested in finding the corridor that yields the highest utility, with a cost of no more than 10 units. Panel II, shows that for a budget of 10 units, the planner maximizes utility by selecting E, F, H, and I for a total utility of 9. If the conservation planner's budget is further increased to 11 units, as in panel III, the optimal selection of parcels is A, B, D, with a corresponding aggregate utility of 10. It is not surprising that considering only parcel costs in panel I results in a

[2] Computationally speaking, the term "easy" in this case refers to a feasibility condition that can be checked in polynomial time.

very different set of selected parcels from that in panels II and III, where both parcel cost and utility are considered. What is unique about the constrained corridor optimization problem is that a marginal change in the available budget can result in the selection of mutually exclusive sets of parcels, as illustrated in panels II and III. Given the constraint that all of the selected parcels must be connected, the model outcomes can change drastically as budget levels are varied, which is different from typical reserve site selection models where marginal changes in budget levels generally only influence the selection of a small subset of the available parcels.

I. Minimum Cost Corridor Cost = 7, Habitat = 5 | II. Optimal Corridor B=10 Cost = 10, Habitat = 9 | III. Optimal Corridor B=11 Cost = 11, Habitat = 10

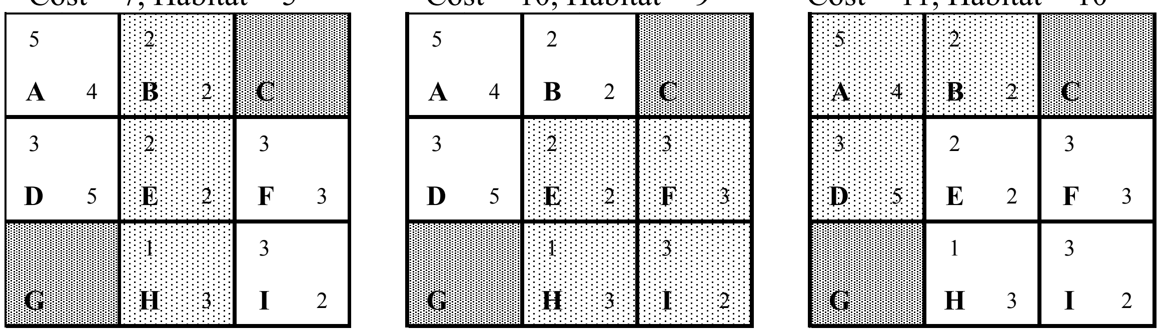


Figure 2.1 Hypothetical Corridor Optimization

Note: Parcel labels are provided in the lower left, costs are in the lower right and utilities are in the upper left.

Figure 1 also illustrates the computational challenges of the budget constrained utility maximization problem. If the objective is to find a least cost path, as has been done in all previous studies, only six possible paths in the 3x3 parcel grid need to be considered. The optimal selection will never include paths that are more than one parcel wide, as this can only add to the cost of the corridor. For the case of constrained utility maximization, however, the set of potentially optimal corridors jumps from six to thirty. Thus, even in this small hypothetical case, the challenge of maximizing utility given a budget constraint is considerably greater than simply finding the single-parcel-wide least cost path. The computation complexity of the problem is analyzed more rigorously in

Conrad et al. (2007) and the challenges of reaching an optimal solution for the Northern Rockies corridor are dealt with later in the essay.

2.4 Mixed Integer Linear Programming Model

To solve the connection subgraph problem, a Mixed Integer Linear Programming Model is formulated where the binary variable x_i represents each vertex $i \in V$ and indicates whether i is included in the connected subgraph. The budget constrained utility maximization problem can be written

$$Max \sum_{i \in V} u_i x_i \tag{1}$$

$$s.t. \sum_{i \in V} c_i x_i \leq B \tag{2}$$

$$x_i \in \{0,1\} \quad \forall \ i \in V . \tag{3}$$

To ensure that connectivity is achieved, four additional constraints are included by applying a particular network flow model. In the model, each edge is represented by a nonnegative variable y_{ij}, which reflects the amount of flow from vertex i to vertex j. Flow that is identical in volume to the n vertices in the graph is "injected" from an external vertex x_0 into one of the terminal vertices. This constraint is formalized in equation (4) below. Further, the constraint provided in equation (5) ensures that only flow that is injected into the terminal parcel is utilized by the network.

$$x_0 + y_{0t} = n \tag{4}$$

$$y_{0t} = \sum x_i \tag{5}$$

Next, each of the vertices that are included in the subgraph retains one unit of flow. This implies that the flow from vertex i to vertex j must be less than the total amount of flow injected into the system,

$$y_{ij} < nx_j , \ \forall \{i, j\} \in E . \tag{6}$$

The conservation of flow in the network requires that the sum of all flow entering a vertex must be identical to the amount remaining at the vertex plus the amount of flow that leaves the vertex. This constraint is formalized as

$$\sum_{i:\{i,j\} \in E} y_{ij} = x_j + \sum_{i:\{i,j\} \in E} y_{ij} , \ \forall \ j \in V. \tag{7}$$

Finally, to ensure that all of the terminal vertices are included in the subgraph, each of the terminal vertices is forced to retain one unit of flow,

$$x_t = 1, \ \forall \ t \in T. \tag{8}$$

The constrained utility optimization problem above can be transformed into its dual, utility constrained cost minimization problem by essentially swapping (1) and (2). The utility constrained cost minimization problem is written as

$$Min \sum_{i \in V} c_i x_i \tag{9}$$

$$s.t. \sum_{i \in V} u_i x_i \geq U\ , \tag{10}$$

with the four connectivity constraints remaining the same. The unconstrained cost minimization problem, the so called "least-cost path", can be obtained by eliminating constraint (9).

The formulation presented here allows for the possibility that the corridor can be more than one parcel wide. This is a favorable attribute of the model, given that the overall utility of the parcels selected can be increased by widening the corridor or by incorporating paths to areas of high quality habitat. It may be, however, that the conservation planner wishes to eliminate the possibility of having peninsulas in the network, which could represent dead ends to wildlife in the corridor. While this option is not explored empirically in this essay, in practice peninsulas could be reduced[3] through the institution of an additional constraint, which requires that every vertex receiving flow must output flow to at least one other vertex that is different from the input vertex. Formally, the constraint is

$$y_{ji} > x_i + y_{ij}\, , \forall i, j \in H \neq T. \tag{11}$$

2.5 Wildlife Corridor Application

The U.S. Northern Rocky Mountain Region is unparalleled in the continental U.S. in terms of resident wildlife species. The region is home to significant populations of

[3] It is still possible with this constraint for there to exist a multiple parcel wide peninsula, however one parcel wide peninsulas would be eliminated.

grizzly bears, mountain lions, gray wolves, bighorn sheep, elk, moose and bison. In addition, the region contains three of the largest wild, undeveloped areas in the continental U.S; The Greater Yellowstone, Salmon-Selway[4] and Northern Continental Divide Ecosystems together comprise a land area of approximately 80,000 square miles, larger than the combined size of the states of New York, Massachusetts, New Hampshire and Vermont. The 25% population growth rate of the mountain West, however, was the highest of any region of the United States in the 1990's. Moreover, many rural counties in the region gained population at higher rates than the urban counties (Hansen et al. 2002). While many people are attracted to the region because of the abundant natural amenities, the sprawling development that has resulted is leading to an increasingly fragmented landscape. As such, the three large undeveloped areas are becoming isolated biological islands, separated by development.

Development and habitat fragmentation in the Northern Rockies has led to a situation where populations of grizzly bears, which were once abundant across the region, now live almost exclusively in the Yellowstone and Northern Continental Divide Ecosystems. While the number of grizzlies in the Yellowstone Ecosystem, estimated between 400-600, has been stable enough to warrant their recent removal from the endangered species list (US FWS 2007), grizzly populations outside of Yellowstone remain federally protected under the Endangered Species Act. It is clear that for sustained populations of grizzlies to return to the Salmon-Selway Ecosystem, and for the general viability of grizzly populations across the Northern Rockies, wildlife corridors must be instituted that enable movement between the core ecosystems. Further, the grizzly bear is referred to as an "umbrella species", meaning that its survival improves the persistence of a wide range of other species living in the region (Walker and Craighead 1997). Therefore a viable corridor connecting the Greater Yellowstone,

[4] The Salmon-Selway Ecosystem is also referred to as the Bitterroot Ecosystem.

Salmon-Selway and Northern Continental Divide Ecosystems would not only sustain populations of grizzlies, but would also improve the overall biodiversity of the region.

Given the large geographic size of the region and the heterogeneity in terms of both habitat quality and land costs, the design of a corridor connecting the core wildlife areas is challenging. The overall viability of a proposed corridor will undoubtedly be determined by the sites that are selected to connect the landscape. Given the limited conservation funding available, a corridor that is overly expensive will make the project a budgetary impossibility. Moreover, a corridor that incorporates sites with limited environmental quality will fail to provide the habitat necessary for wildlife populations to move freely between the core areas and will therefore not fulfill the conservation objectives. By incorporating both habitat suitability and land acquisition costs, the model presented in this essay allows for the best chance of corridor implementation. The data that are used to measure the relative benefits and costs of acquiring land to form the corridor are explained in the next section.

2.6 Description of Data Sources

The study area for this analysis is comprised of 64 counties in Idaho and western Montana, located in the U.S. Northern Rockies region. At the aggregate level, the parcels that are considered for inclusion in the corridor are the 64 counties themselves. While securing an entire county to be included in the reserve may seem infeasible, the county-level analysis provides an illustrative example of a case where the optimization problem is relatively simple from a computational perspective. The county level model allows us to identify general corridor areas that contain low cost, suitable habitat, similar to Ando et al. (1998). The county model also provides a means of comparing the results of an aggregate model with relatively few sites, to more granular models with greater numbers of parcels. A map of the study area is included below as Figure 2.2.

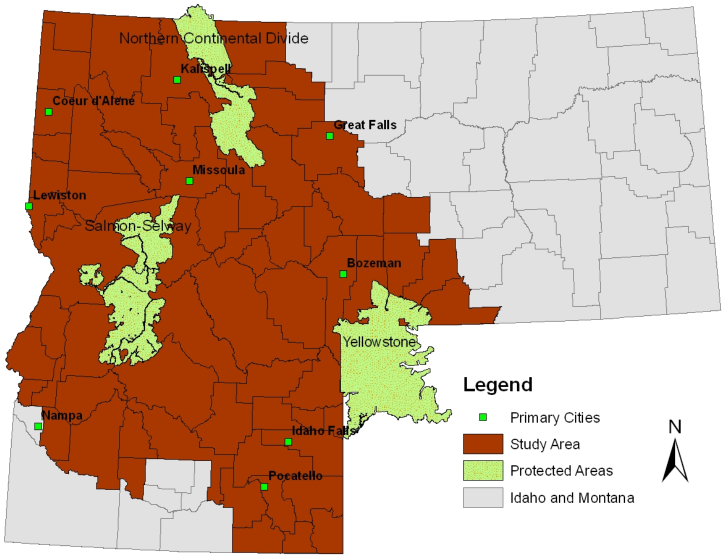


Figure 2.2 U.S. Northern Rockies Study Area

To investigate the impact of increasing the granularity of the available parcels, the study area is further segmented into continuous sets of square grid cells. The largest grid cells are 60km on each side and segment the study area into 118 parcels. The parcel size is then incrementally reduced to square grids with sides of 50km, 40km, 25km, 10km and 5km. With the most granular grid size of 5km, the study area is segmented into 12,788 cells. Given the relatively large range of an adult grizzly (the home range of an adult female grizzly bear is approximately 125 square km), grid sizes smaller than 5km are unlikely to be suitable for grizzly bear movement (Mace and Waller 1997). Increasing the granularity of the grid cells allows for much more precision in defining parcel habitat suitability and acquisition costs and it also increases the number of parcels in the landscape. Given the greater number of parcels available for the corridor, increasing the granularity also increases the complexity of the optimization problem. Thus, comparing results across the continuum of cell sizes allows for an investigation into the tradeoffs inherent in the granularity of the model that allows for increased specificity at the cost of greater computational complexity.[5] In addition, increasing the granularity of the parcels

[5] In addition to square grid cells, a grid composed or hexagonal parcels is also considered. The hexagonal grid results in significantly lower costs than a square grid, given that a corridor moving in a diagonal

is equivalent to increasing the scope of a study area with fixed parcel sizes. Therefore increasing the parcel granularity provides insight into the scales at which corridor optimization is possible.

Grizzly bear habitat suitability data, developed and provided by the Craighead Environmental Research Institute (CERI), is used to measure the utility of each parcel. These data spatially define habitat that is considered to be suitable for grizzlies. The suitable habitat is measured on a 30 meter grid and each grid cell is given a score from 2 to 4, with 4 being the highest quality habitat. The habitat suitability data are then aggregated to the larger grid and county levels used in the analysis by summing the habitat scores within each parcel boundary. This method of aggregation implicitly assumes, for example, that a cell with a habitat suitability value of 4 is twice as beneficial as a habitat suitability value of 2.

The estimate of parcel cost is calculated in three steps. First, spatial data on land stewardship, available for the states of Montana and Idaho from the GAP Analysis project (USGS 1999), are used to classify privately and publicly owned land in the study area. Next the amount of private land acreage within each parcel is calculated. The private land acreage is then multiplied by the county specific average value of farm real estate per acre, available from the USDA's Census of Agriculture (2002). For grid cells with land acreage in multiple counties, the county specific real estate value per acre is multiplied by the amount of private acreage in each county and then summed. Using the value of farm real estate is a proxy for the cost of all private land, as it is reflects, although not perfectly, the opportunity costs faced by private land owners. Ando et al. (1998) similarly use county level average farm real estate value in their reserve selection model.

---

direction contains fewer cells than comparably sized square grid cells. These results are available in the appendix.

In delineating the cost of each parcel, the assumption is that land already in the public domain is freely available for inclusion in the corridor. One could, however, imagine incorporating the opportunity cost of lost timber or mining contracts as proxies for the cost of acquiring public land as in Polasky et al. (2001) and Sessions (1992). Costs of incorporating public land in the corridor are not included in the present analysis as there is insufficient data with which to accurately predict the heterogeneity in lost resource profitability associated with each parcel. In addition, it is possible that some limited resource extraction could occur on land included in the corridor. A depiction of the spatial distribution of parcel costs and parcel utilities at the 10km grid level is included as Figure 2.3.

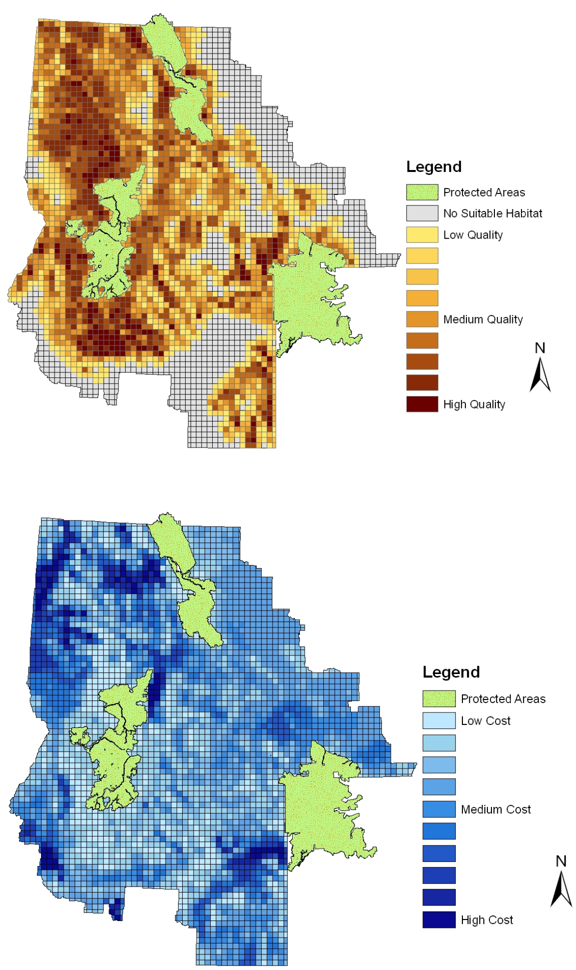


Figure 2.3 Habitat Suitability and Parcel Cost for 10km Grid Parcels

By calculating the cost of each parcel based on the real estate value of its privately owned acreage, the assumption is that the parcels included in the corridor will be acquired with fee-simple purchases. For large projects, such as a corridor connecting the three large ecosystems in the Northern Rockies, the funds necessary to purchase a viable corridor outright will be large. Yet the cost estimates should be put into perspective by comparison to the significant amount of both public and private funding currently being spent on land conservation. In 2006, 133 separate ballot initiatives across the U.S. approved $6.7 billion in public funds for the procurement of conservation land (LTA 2007). This funding is in addition to the efforts of private land trusts at the local, state and national levels, who conserved 37 million acres in 2005 (LTA 2005) and other federal conservation programs such as the Conservation Reserve Program (CRP), which had annual expenditures exceeding $1.8 billion in 2006 (USDA 2006).

It should also be noted that parcels may not necessary need to be purchased outright in order to be included in the corridor, as easements and other voluntary agreements may be sufficient to maintain habitat. This voluntary type of arrangement is being used, for example, in the 'Alps to Artherton' project in Australia, where the Australian government is seeking agreements with private land owners to abstain from certain land use practices in exchange for annual payments. In the U.S., an initiative such as the USDA's Forest Legacy Program (FLP) could provide the means for funding voluntary easements on private forest land. The FLP works in partnership with states to acquire conservation easements that place restrictions on the property rights of private landowners owning forested acreage. Annual expenditures under the FLP have exceeded $55 million in each of the last four years (USDA 2007).

While securing voluntary agreements for habitat protection may be a more viable strategy for cost-effectively targeting parcels to include in the corridor, there is insufficient data on the incentives necessary to secure such voluntary arrangements. The real estate value can therefore be thought of more as an upper-bound on a parcel's cost[6], noting that the potential for voluntary habitat protection could significantly reduce the funds necessary to acquire the corridor. Future research on the incentives necessary for voluntary habitat protection could provide useful information for conservation planners.

One additional consideration in terms of the overall cost of the corridor is the transaction and management costs associated with securing property rights and maintaining the selected parcels. Researchers have identified transaction and management costs as being an important consideration in reserve design (e.g., Naidoo et al. 2006; Newburn, Berck and Merenlender 2006), yet these costs are rarely included in optimal conservation models. One notable exception is Groeneveld (2005) who looks at the theoretical implications of varying transactions costs on the number of sites included in a reserve. In the present analysis, the influence of transaction costs on corridor design for the 5km grid parcels is investigated by finding the cost minimizing corridor both with and without transaction costs. Transaction costs are likely to play a more significant role when the cell granularity is small, as the transaction cost represents a greater proportion of the overall cost of the parcel and the number of potential paths is large. A fixed $5,000 transaction cost is added to each parcel that is included in the corridor, which would cover legal fees, signage and other fees associated with defining a parcel as part of the corridor. The actual transaction and maintenance cost of a particular parcel is likely to be variable, but $5,000 is chosen as an approximation that is in line with reported transaction costs for conservation lands in New York State (LTA 2004).

Beyond defining the costs and utilities of parcel acquisition, it is also necessary to define the parcel adjacencies for all of the parcels in the study area. The adjacencies for both the aggregate-county and square grid parcels are defined based on shared borders/edges. For the grid parcels this implies that interior parcels are adjacent to exactly

---

[6] This of course assumes that landowners are willing to sell land for inclusion in the corridor, which is likely not be valid in many circumstances.

four other parcels. This is referred to as a “rook” pattern of adjacency, which differentiates itself from a “queen” adjacency pattern where adjacency is defined based on shared edges and corners.[7]

Finally, it should be noted that there are regions within the study area that may represent either natural or man-made barriers to grizzly bear movement. For example, grizzlies may not be able to cross the Mission Mountain Range in Northwestern Montana or parts of the Clark Fork of the Columbia and other large rivers that flow through the study area. Man-made obstructions such as Interstate 90 or highly urbanized areas near major cities may also be impenetrable. As such, the land use planner is advised to ground truth the optimization results to ensure that these barrier areas are not included in a proposed corridor.

2.7 Model Results

This section reports the optimization results for two variations of the connection subgraph problem described in section 2.3 for the county-level parcels and the separate grid parcel granularities. First, the results from the unconstrained cost minimization model are reported along with an explanation of the algorithm that is employed. Second, the results and algorithms used for the budget constrained utility maximization model are described. The unconstrained minimum cost corridor problem corresponds to the minimum Steiner tree problem. In the case where the number of reserves is bounded, the minimum Steiner tree problem can be computed in polynomial time (Promel and Steger 2002). For the case considered here with three reserves, the algorithm that is implemented runs in time roughly equivalent to n^3, where n is the number of parcels. The algorithm first computes the shortest path between each parcel and all other parcels in the study area, which generates what is referred to as an all-pairs shortest path (APSP) matrix (Corman et al. 2001). In this case, the path length is measured strictly in terms of the cost

---

[7] Williams and Snyder (2005) refer to rook adjacency as the nearest-neighbor rule and queen adjacency as the next-nearest neighbor rule.

of each parcel. Next, the algorithm determines the parcel that minimizes the distance between that parcel and the three reserves, which is referred to the center point. In the case of three reserves, it is guaranteed that there will be exactly one center point.[8] The minimum cost corridor is then determined based on the shortest path between the chosen center point and the three reserves. The minimum cost corridor results for each granularity are reported in Table 2.1.

After computing the unconstrained minimum cost corridor, the budget constrained utility maximization corridor is determined for budgets greater than the minimum cost. For parcel granularities down to 50km, the optimal budget constrained corridor can be computed using standard, off-the-shelf CPLEX optimization software using the MIP formulation described in section 2.4. For parcel granularities smaller than 50km, a preprocessing step is executed using the all-pairs shortest path matrix generated in the minimum cost solution. Specifically, if the minimum cost of connecting a given parcel to its two closest reserves exceeds the budget, then that parcel is "pruned" from the set of available parcels. This preprocessing step allows for the calculation of the optimal corridor for the 40km parcel granularities. Unfortunately, for parcels granularities smaller than 40km, optimal corridors cannot be determined even with this preprocessing step. For the smaller parcel granularities, a heuristic method can be implemented based on the minimum cost corridor. This heuristic is explained in greater detail in section 2.8.

For the county, 60km, 50km and 40km parcel granularities, the utility maximizing corridor for a budget that is 10% greater than the cost minimum is provided in Table 2.2. For the 50km and 40km grids, the effect of varying the size of the budget on the parcels selected is evaluated in greater detail. To this end an efficiency frontier is generated illustrating the tradeoffs between parcel cost and the amount of suitable habitat in the corridor.

---

[8] When there are more than three reserves, the algorithm finds all k-2 centerpoints through a Dreyfus-Wagner algorithm (Corman et al. 2001)

Cost Minimizing Corridor

Table 2.1 Cost Minimization Results

Parcel size	Number of parcels	Parcels selected	Corridor Cost (thousand)	Total HSI (thousand)	Acres Preserved (thousand)	% Private	Cost per acre ($)
County	64	5	1,904,355	7,038	9,649	27.2%	197.4
60km	118	11	1,657,740	7,188	8,234	27.1%	201.3
50km	167	12	1,329,090	5,902	6,777	30.7%	196.1
40km	239	16	891,052	5,807	5,409	13.6%	164.7
25km	570	23	449,430	3,743	3,408	12.5%	131.9
10km	3,296	120	99,341	3,679	4,096	1.9%	24.3
5km	12,788	265	10,865	2,147	1,637	0.5%	6.6
5km†	12,788	196	11,824	1,576	1,210	0.7%	9.8

† Includes a $5,000 transaction cost per parcel selected.

The number of parcels available for acquisition ranges from 64, in the case of the county-level grid, to 12,788, for the 5km grid. When the parcels available for acquisition are the counties themselves, the minimum cost corridor has a price tag of over $1.9 billion. Not surprisingly, purchasing all of the private land in five counties is extremely expensive. As the sizes of the available parcels are reduced, the corresponding cost of the cheapest corridor diminishes considerably. At the 5km granularity, the cost of the corridor is slightly less than $11 million. A portion of the decrease in cost is a result of the fact that less overall land area is being purchased; the 5km cost minimizing corridor covers over 1.6 million acres, while the county-level minimum cost corridor covers nearly 10 million acres. The difference in cost between the county corridor and 5km corridor cannot be explained by differences in preserved acreage alone, however. Increasing the parcel granularity allows for greater specificity, so that the corridor is better able to select low cost areas, which are composed primarily, and in some cases exclusively, of zero cost national forest land. As evidence of this, the cost per acre in the county grid is $197, while the cost per acre of the 5km grid is only $6. Moreover, the optimal parcels for the larger parcel sizes (county, 60km, 50km) are comprised of approximately 30% private land. For the 5km parcels, the amount of private land included in the parcel drops to below 1%. The corridor cost could be further reduced by evaluating parcels smaller than 5km, yet there are tradeoffs in terms of the minimum corridor width necessary for a wide-ranging species such as the grizzly bear. It should be noted that higher percentages of public land also tends to increase the amount of suitable habitat per acre as national forest land generally has high habitat suitability.

Changing the parcel granularity not only influences the cost of the parcels selected and the complexity of the problem, but it also influences the general path that the corridor follows. For the county level, 60km and 50km parcel maps, the minimum cost corridor essentially forms the shape of an upside-down T, where the parcels selected are concentrated in the region in the middle of the three ecosystems. When the parcel size is reduced to 40km and below, the minimum cost corridor traces a path connecting the three reserves that resembles the shape of a C, with the Salmon-Selway Ecosystem connecting directly to the Northern Continental Divide Ecosystem via a parcel path in the northwestern portion of the study area. By increasing the parcel granularity, the model avoids higher priced areas in southwestern Montana and instead chooses a slightly longer corridor that incorporates more national forest land. Thus, influencing the parcel granularity not only influences the estimated cost of the cheapest corridor, but it also has a significant influence on the general path that the corridor follows across the landscape. Maps of the optimally selected minimum cost corridors for each of the parcel sizes are included as Figure 2.4.

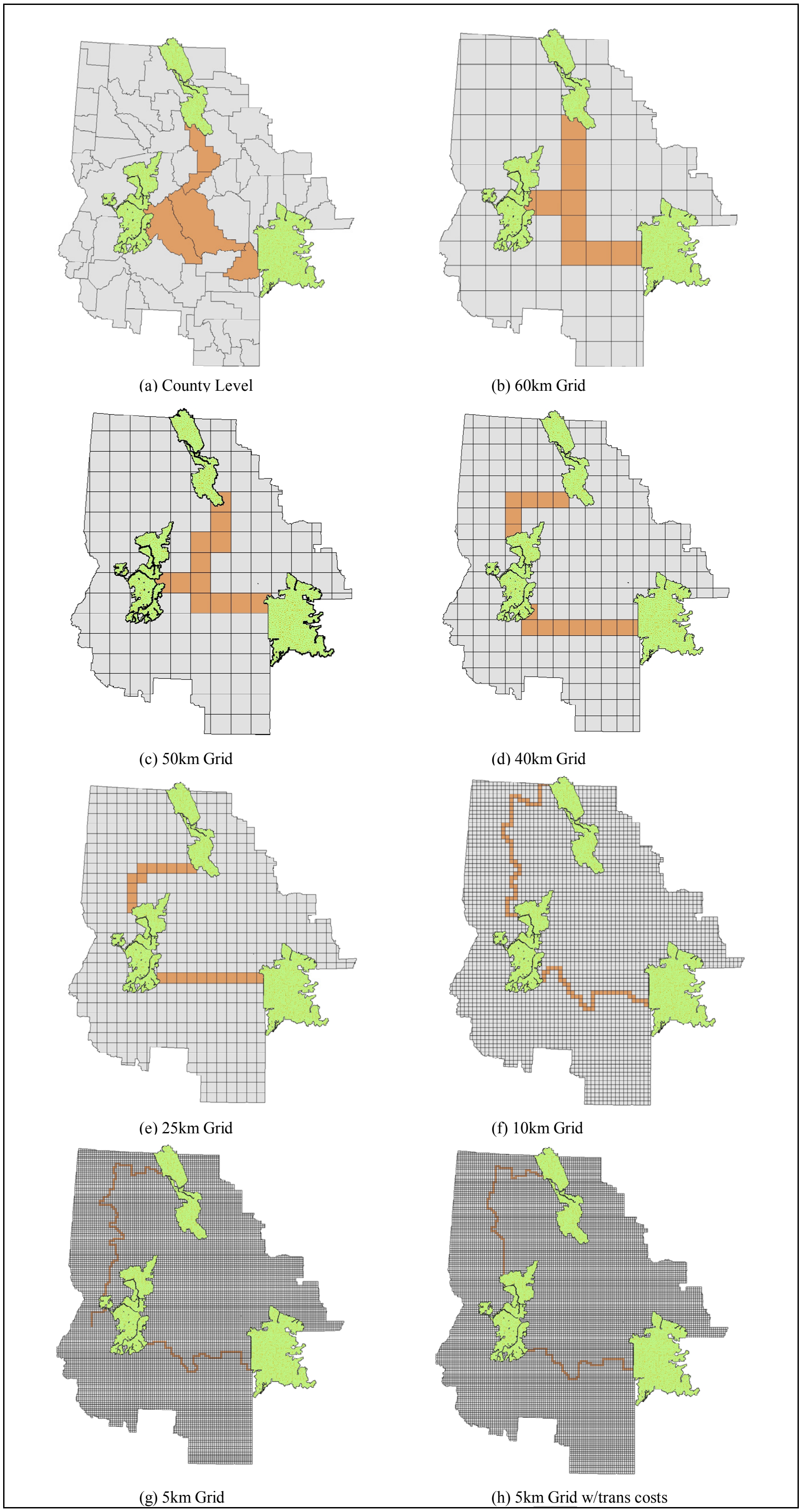


Figure 2.4 Unconstrained Cost Minimum Corridor for Each Granularity

The addition a $5,000 transaction cost per parcel at the 5km level reduces the number of parcels selected from 265 to 196. When transaction costs are considered, each parcel adds incrementally to the overall cost of the corridor. Thus, the minimum cost corridor tends to select parcels that provide more of a direct link between the reserve sites, rather than following a slightly longer path that includes more zero cost, national forest parcels. This difference is illustrated in panels (g) and (h) of Figure 2.4, which show the chosen 5km corridor both with and without transaction costs. The most noticeable difference between the two corridors is the portion of the corridor connecting the Salmon-Selway to the Northern Continental Divide Ecosystem. With the inclusion of the transaction costs, the parcels selected link directly to the northern portion of the Salmon-Selway, rather than the longer parcel selected without transaction costs that connects to the western edge of the ecosystem. With transaction costs the model also does not select the zero cost parcels that form a peninsula starting from the western edge of the Salmon-Selway. Thus, incorporating transactions costs has a significant influence on both the number and shape of the resulting corridor that is selected and represents an important consideration for land use planners.

Cost Constrained Utility Maximizing Corridor

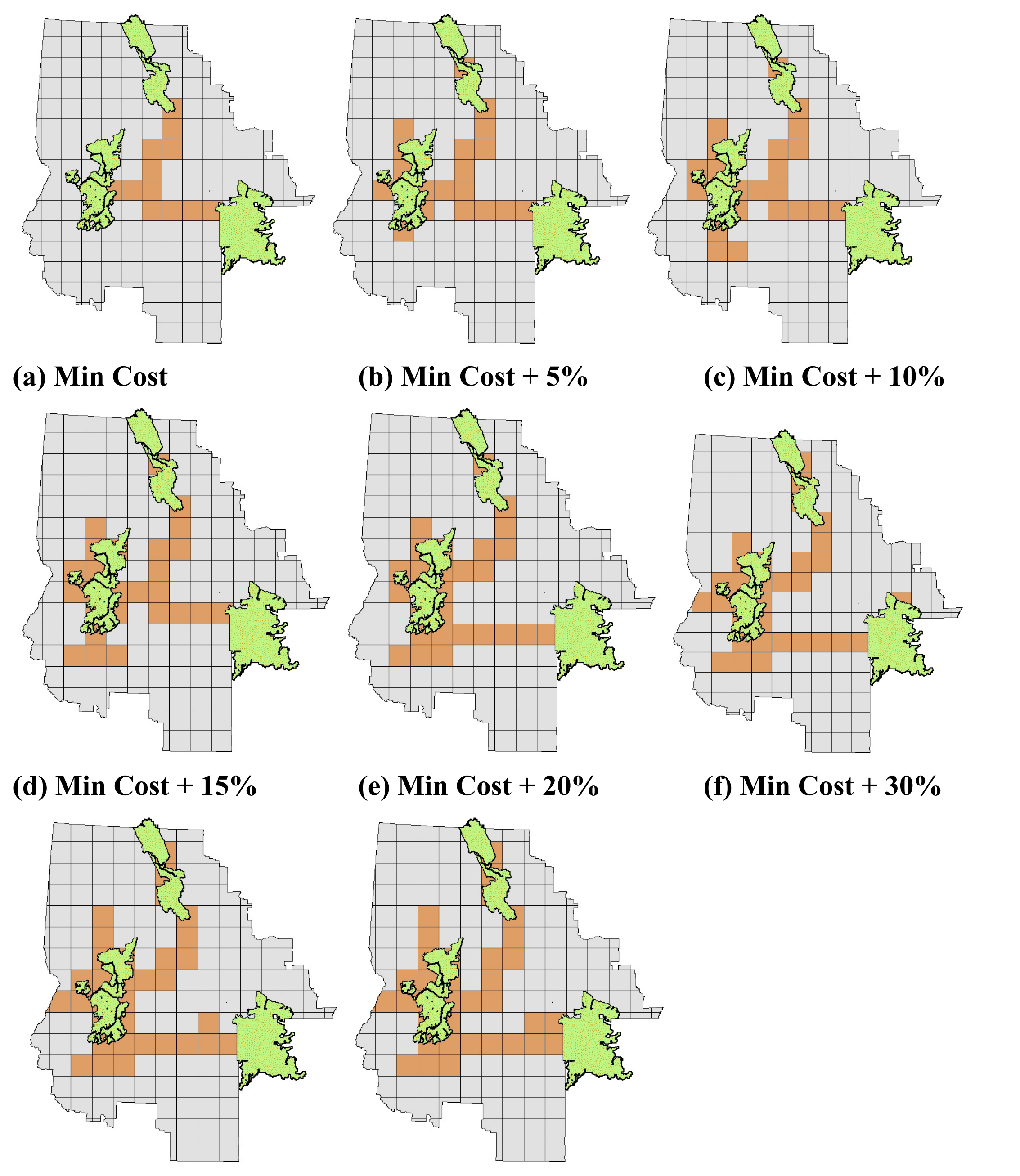
While determining the minimum cost corridor connecting core areas of biological significance is important for land use planners in determining the financial feasibility of a wildlife corridor, selecting a corridor based on cost alone is likely to yield outcomes that leave out relatively low cost parcels with high quality habitat. If a land use planner has a budget that is larger than the minimum cost corridor, then she would ideally determine the corridor that maximizes the amount of suitable habitat given the budget that is available. Thus, after determining the minimum cost corridor, the corridor that maximizes the amount of suitable habitat for a budget that is 10% greater than the minimum cost corridor is selected. Unfortunately, at parcel granularities smaller than 40km, it is not possible to prove the optimality of the cost constrained utility maximizing corridor. In other words, the program can find a feasible solution that connects the reserves and meets the budget constraint, but proving that a particular feasible solution is the optimal solution requires excessive computational time. It is therefore apparent that cost constrained optimization is a significantly more computationally complex problem than corridor cost minimization. At very high budget levels, it is possible to prove optimality with a greater number of parcels, but these high budget cases are not particularly realistic. In the most extreme case, where the planner has funds available to purchase all of the land in the study area, the optimal corridor would be composed of all of the parcels.

Table 2.2 Budget Constrained Utility Maximization Results

Parcel size	Number of parcels	Parcels selected	Corridor Cost (million)	Total HSI (thousand)	Acres Preserved (thousand)	% Private	Cost per acre
County	64	5	1,904	7,038	9,649	27.2%	197.3
60km	118	20	1,821	14,240	14,209	32.1%	128.2
50km	167	22	1,461	12,188	11,303	19.4%	129.3
40km	239	23	999	11,832	9,932	8.4%	100.6
25km	570	-	-	-	-	-	-
10km	3,296	-	-	-	-	-	-
5km	12,788	-	-	-	-	-	-

Note: Budget is set 10% higher than the cost minimum solution.

A summary of the parcels selected for the utility maximizing corridors, given a budget that is 10% higher than the cost minimum, is provided in Table 2.2 for the county-level, 60km, 50km and 40km grids. At the county level, increasing the budget by 10% does not change the parcels selected in the optimal corridor. For this coarse parcel size, the budget increase is not enough to motivate the selection of a different set of counties. At the 60km level, increasing the budget by 10% results in the optimal selection of 20 parcels as opposed to the 11 parcels selected in the cost minimization model. When the grid size is further reduced to 50km, the number of selected parcels jumps from 12 to 22, while at the 40km level the number of selected parcels goes from 15 to 23. In each case, budget constrained optimization results in a significant increase in aggregate habitat suitability, for example at the 50km level, suitable habitat increases from 5,902,000 in the cost minimization model to 12,187,572. It should be noted that the minimum cost 50km grid obtains the maximum habitat suitability possible for a budget equal to the minimum cost. Therefore the increase in aggregate habitat suitability at the higher budget level is strictly a result of the increase in budget.



(a) Min Cost **(b) Min Cost + 5%** **(c) Min Cost + 10%**

(d) Min Cost + 15% **(e) Min Cost + 20%** **(f) Min Cost + 30%**

(e) Min Cost + 40% **(f) Min Cost + 50%**

Figure 2.5 Cost Constrained Utility Maximization of 50km Parcels

A visual depiction of the optimal corridors for the 50km parcels under varying budgets is presented in Figure 2.5. The shape of the 50km corridor changes considerably under budget constrained maximization. Rather than forming the upside-down T as in the cost minimization solution, the budget constrained optimization solution looks more like a C, with the inclusion of a number of buffer parcels around the Salmon-Selway Ecosystem that are both low cost and contain a high degree of suitable habitat.

Efficiency Frontier

The results of the cost minimizing corridor and the utility maximizing corridor with a 10% greater budget allude to the tradeoffs that exist between corridor cost and overall habitat suitability measures. The conservation planner is likely to be interested both in reducing the overall cost of the corridor and in incorporating as much suitable habitat as possible, but may not have an exact idea of the relative tradeoffs between these two objectives. If the planner does have a specific target for the cost of the corridor and a specific target for the overall amount of habitat in the corridor, then goal programming techniques, first introduced by Charnes, Cooper and Ferguson (1955), could be used to solve the optimal corridor parcels. If the planner is interested in minimizing the costs of the corridor and maximizing the amount of suitable habitat, but does not have specific targets for either, then an efficiency frontier would provide a useful basis for decision making. The efficiency frontier represents the locus of non-inferior parcel combinations, where in order to improve one of the objectives, the other objective must be made worse off. In this case to increase the amount of suitable habitat in the corridor for a point on the efficiency frontier, the cost of the corridor must increase.

The weighting method and the constraint method are two possible techniques for generating the efficiency frontier (Willis and Perlack 1980). The weighting method, used in the context of corridor design in Williams (1988), incorporates both cost and habitat objectives in the objective function and multiplies each of the objectives by a weighting vector that sums to one. Thus, if all of the weight is placed on the cost objective, then the model finds the minimum cost corridor and if all of the weight is placed on the habitat objective then the model, for the case of maximizing suitable habitat, would select all of the available parcels. The efficiency locus is then derived by systematically changing the weights between 0 and 1.

The constraint method for delineating the efficiency frontier, which is pursued in this essay, requires first solving for the minimum cost corridor and then solving the cost constrained utility maximization problem for various budget levels greater than the minimum cost corridor.[9] While both the weighting method and the constraint method can be used to define the efficiency frontier, the constraint method is preferable in this case because it is able to determine all of the noninferior solutions. In contrast, when the objective space is nonconvex (e.g., when the decision variable is zero-one), the weighting method is not able to find all noninferior solutions (Willis and Perlack 1980). Since it is minimizing a linear combination of the two objectives, the weighting method finds only the convex hull of noninferior solutions and ignores solutions in the interior of this hull. This problem is referred to as the "duality gap" (Williams 1998). The constraint method employed here is also preferable to the weighting method because it allows planners to explicitly compare the habitat benefits of specific budget levels. For example, if the planner has a choice between three specific budgets for a project, then he or she can solve the corridor optimization problem subject to each of the three budget constraints and compare the resulting parcel outcomes. With the weighting method, the weights would have to be adjusted in such a way so as to approximate each of the three budget levels and, because of the duality gap, an exact solution for each of the budgets may not be found. A second reason for not using the weighting method in the present analysis is that it is not possible to weight two objectives with different signs. In this case, we are

[9] The identical efficiency frontier could also be generated by minimizing the cost of the corridor and systematically changing the constraint on the overall level of suitable habitat.

seeking to minimize cost and maximize habitat suitability and thus the weighting method would require an alternative specification.

The efficiency frontier for the Northern Rockies corridor is illustrated by using the constraint method on the 50km. Beginning with the minimum cost corridor for each, the budget constraint is systematically increased. The numeric results of the budget constrained maximization are presented in Table 2.3. In addition, the graphical efficiency frontier is provided for the 50km grid in Figure 2.6.

The results show that aggregate habitat suitability measures increase nonlinearly with increases in the conservation budget. For the case of the 50km grid, increases in the budget up to 10% above the minimum cost corridor increase overall habitat suitability at an approximate rate of 48 units for every $1 increase in the budget. At budget levels between 50% and 150% greater than the minimum cost corridor, however, the rate of increase in habitat suitability associated with increases in the budget is less than 6 units for every 1 dollar increase in the budget.

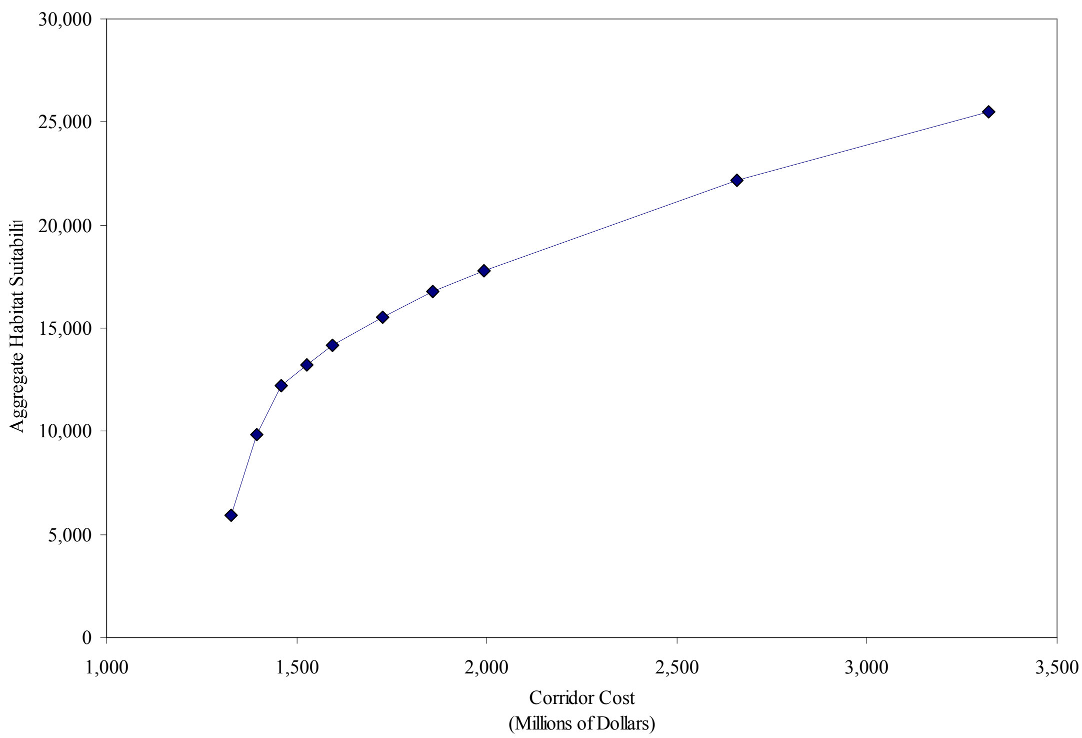


Figure 2.6 Corridor Efficiency Frontier for 50km grid

Table 2.3 reveals several noteworthy trends as the conservation budget is increased above the minimum cost. First, similar to the overall HSI, the number of acres preserved increases at a decreasing rate with changes in the budget. For budget increases above the cost minimum between 0 and 10%, approximately 0.033 additional acres are preserved for each additional dollar in the case of the 50km grid. For budget increases between 50 and 150%, the number of additional acres preserved per dollar falls to 0.005. In addition, the percentage of private land that is included in the corridor follows a U shaped trajectory as the budget is increased. For marginal increases in the budget above the cost minimum, the optimal corridor incorporates higher quantities of public land. As parcels with high percentages of public land are exhausted at higher budget levels, the optimal corridor adds additional parcels with greater percentages of private land. The overall percentages of private land are reflected in the cost per acre, which also decreases for initial budgetary increases and then grows as the low cost parcels are exhausted.

Table 2.3 Budget Constrained Utility Maximization for 50km Parcels

Budget (million)	Cost (million)	Total HSI (thousand)	Acres Preserved (thousand)	Percent Private	Cost per Acre ($)	HSI per Acre
-	1,329	5,902	6,777	30.7%	196.1	0.87
1,396	1,394	9,842	9,608	22.2%	145.1	1.02
1,462	1,461	12,188	11,303	19.4%	129.3	1.08
1,528	1,526	13,220	12,176	18.5%	125.3	1.09
1,595	1,594	14,145	12,874	15.5%	123.8	1.10
1,728	1,727	15,533	14,131	15.7%	122.2	1.10
1,861	1,857	16,777	15,119	15.2%	122.8	1.11
1,994	1,992	17,811	16,239	16.1%	122.7	1.10
2,658	2,658	22,151	20,105	16.2%	132.2	1.10
3,323	3,321	25,500	23,298	16.5%	142.5	1.09

The HSI per acre preserved increases dramatically for initial budgetary increases, but then plateaus and finally decreases slightly for the higher budget levels. The plateau in the case of the 50km grid occurs at approximately 1.10 units per acre. This again illustrates the fact that greater granularity allows for greater targeting of low cost, high benefit parcels. Taken together, the results indicate significant marginal benefits in terms of greater numbers of acres preserved, higher habitat suitability per acre and lower costs per acre for marginal in increases in the conservation budget above the cost minimum. These marginal benefits are reduced at higher budget levels, as the number of low cost and high benefit parcels becomes scarce.

While the marginal benefits of budget increases are high in close proximity to the cost minimum, the marginal cost in terms of computational complexity is great. As the budget is increased above the cost minimum, the time necessary to prove utility maximization, in this example, decreases. Thus it may be possible to prove an optimal solution when the allotted budget is high, but as the budget becomes increasingly constrained, proving optimality may not be possible. The optimization results therefore illustrate the significant tradeoffs between computational complexity and the amount of suitable habitat in the corridor. Given that land use planners are likely to have budgetary constraints that restrict their spending to levels that approximate the cost-minimizing corridor, these tradeoffs represent a significant policy dilemma. At low budget levels the overall utility of the corridor can be drastically increased by selecting parcels optimally, yet proving optimality represents a considerable and, in the case of a large set of available parcels, potentially impossible challenge.

2.8 Minimum Cost Extension Heuristic

Proving that a particular corridor is optimal given a budget constraint is not possible in a reasonable amount of time for the case of parcel granularities smaller than 40km. Conservation planners, however, need not be completely without guidance for selecting corridor parcels when the number of available parcels is large. This section describes a heuristic that can be applied, which generates a feasible corridor that approximates, and in some cases is equivalent to, the optimal corridor. Results from the application of the heuristic to the 40km parcels are illustrated in detail to show the degree to which the heuristic results approximate the optimal results.

The heuristic procedure is performed using the minimum cost corridor as a baseline. The parcels selected for the minimum cost corridor are then treated as if they themselves are reserves, guaranteeing an initial connected path. Next, the optimal extension of the minimum cost corridor is calculated using the optimization procedure described above, where parcels that are not feasible given the budget constraint are pruned and the additional parcels are optimally selected from the remaining set using CPLEX.

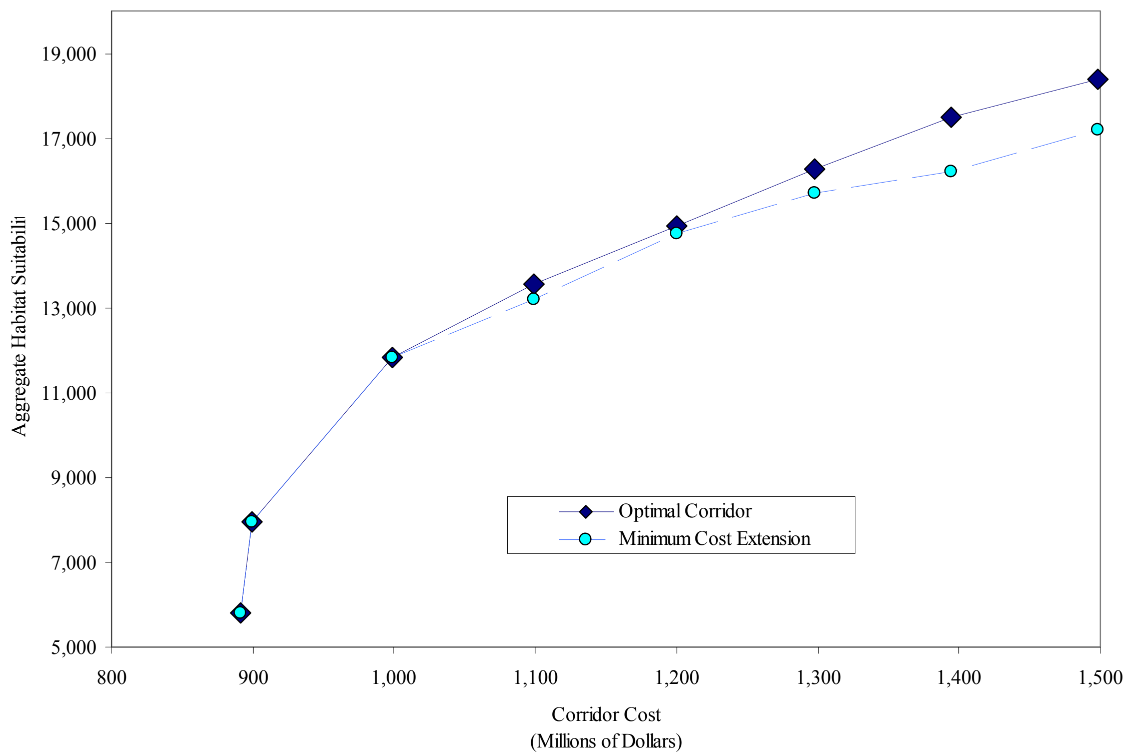


Figure 2.7 Comparison of Optimal and Minimum Cost Extension 40km Grid Efficiency Frontier

The comparative results depicted in Figure 2.7 reveal that the minimum cost extension heuristic closely approximates the optimal corridor for the 40km parcels. Indeed, for budgets up to approximately 25% more than the cost minimizing corridor, the minimum cost extension heuristic selects the identical set of parcels as the optimization algorithm. At higher budget levels the correspondence between the heuristic and the optimum is not exact, but the difference is relatively minor. For budget levels up to 70% more than the minimum cost corridor, the difference in habitat suitability between the optimally selected and heuristically selected corridor is never more than seven percent. The results suggest that even in cases where optimization is not possible, the minimum cost extension heuristic will allow conservation planners to capture the majority of the habitat benefits for a given budget, especially for budget levels that are close to the cost minimum. Importantly, the minimum cost extension heuristic is significantly less computationally intensive and provides feasible solutions for granularities at least as small as the 5km grid, in a reasonable timeframe. The close correspondence between the minimum cost extension heuristic and the optimal solution is facilitated by the correlation between parcel utilities and costs in the Northern Rockies study area. It is an open empirical question whether the same relationship will hold for a general case, where corridor costs and utilities are not correlated.

2.9 Conclusion

As human populations grow and increasing amounts of habitat are lost to development, maintaining habitat connectivity will play a major role in fostering biological diversity. The design of wildlife corridors that connect key areas of biological significance is in its essence a classic economic problem that involves selecting the most suitable corridor habitat given a particular conservation budget. The case of an optimal design for a wildlife corridor connecting the Northern Continental Divide, Salmon-Selway and Yellowstone Ecosystems in the U.S. Northern Rockies is considered in this essay using both heterogeneous parcel costs and utilities. To gain a better understanding of how parcel shape and quantity influence computational complexity, as well as the aggregate costs and benefits of the selected parcels, optimization is conducted over a range of parcel granularities. The results indicate that greater parcel granularity allows for more specificity in targeting low cost parcels in the cost minimization problem. Moreover, as the granularity of the parcels change, the cost minimizing corridor is likely to follow considerably different paths, reflecting the tradeoff between parcel cost and benefit as well as the parcel's location in the landscape. The results also provide evidence that determining the connected set of parcels that minimize corridor cost is computationally easier than proving that a particular set of parcels maximize the amount of suitable corridor habitat for a given budget level. In the study area that is evaluated in this essay it is not possible to prove optimality for budget levels near the cost minimum, when the number of parcels is greater than 240.

For small scale problems, budget constrained maximization allows conservation planners to optimally utilize the funds allotted for corridor acquisition. The efficiency frontier, created by applying the constraint method on the 50km parcel grid, illustrates the tradeoffs that exist between corridor cost and overall habitat suitability. Budgets in excess of the cost minimum corridor have the potential to provide considerably higher levels of habitat suitability, though the marginal benefit of budgetary increases is concave. This implies that the greatest potential for an optimally selected corridor occurs for budget levels that are close to the cost minimum. Unfortunately, this budget range is also the most challenging in terms of computational complexity. For conservation planners that operate with relatively small budgets, in areas with a large number of available parcels, optimization may not be practical. In this case, evidence is provided that suggests that a heuristic, which finds the optimal extension of the minimum cost corridor in a practical timeframe, closely approximates the optimal solution. Thus, conservation planners operating with budgets greater than the minimum cost corridor, in areas with large numbers of available parcels, can still select parcels such that aggregate habitat suitability closely approximates optimal levels. Future corridor research comparing the cost effectiveness of heuristics over a variety of parcel costs and utilities will be useful to land use planners as corridor projects continue to be proposed for increasingly large landscapes.

APPENDIX

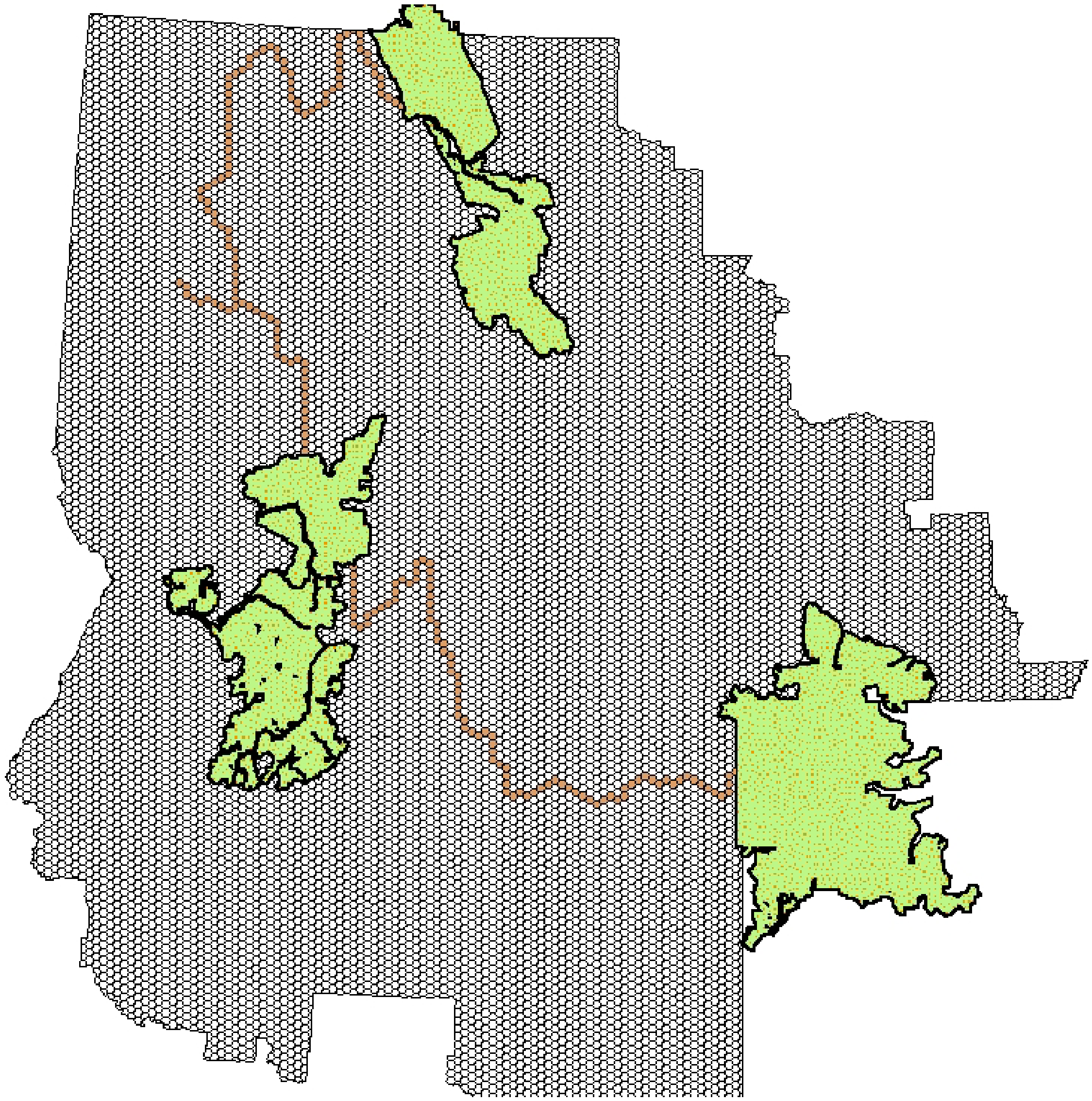


Figure A.2.1 Minimum Cost Hexagonal Grid No Transaction Costs

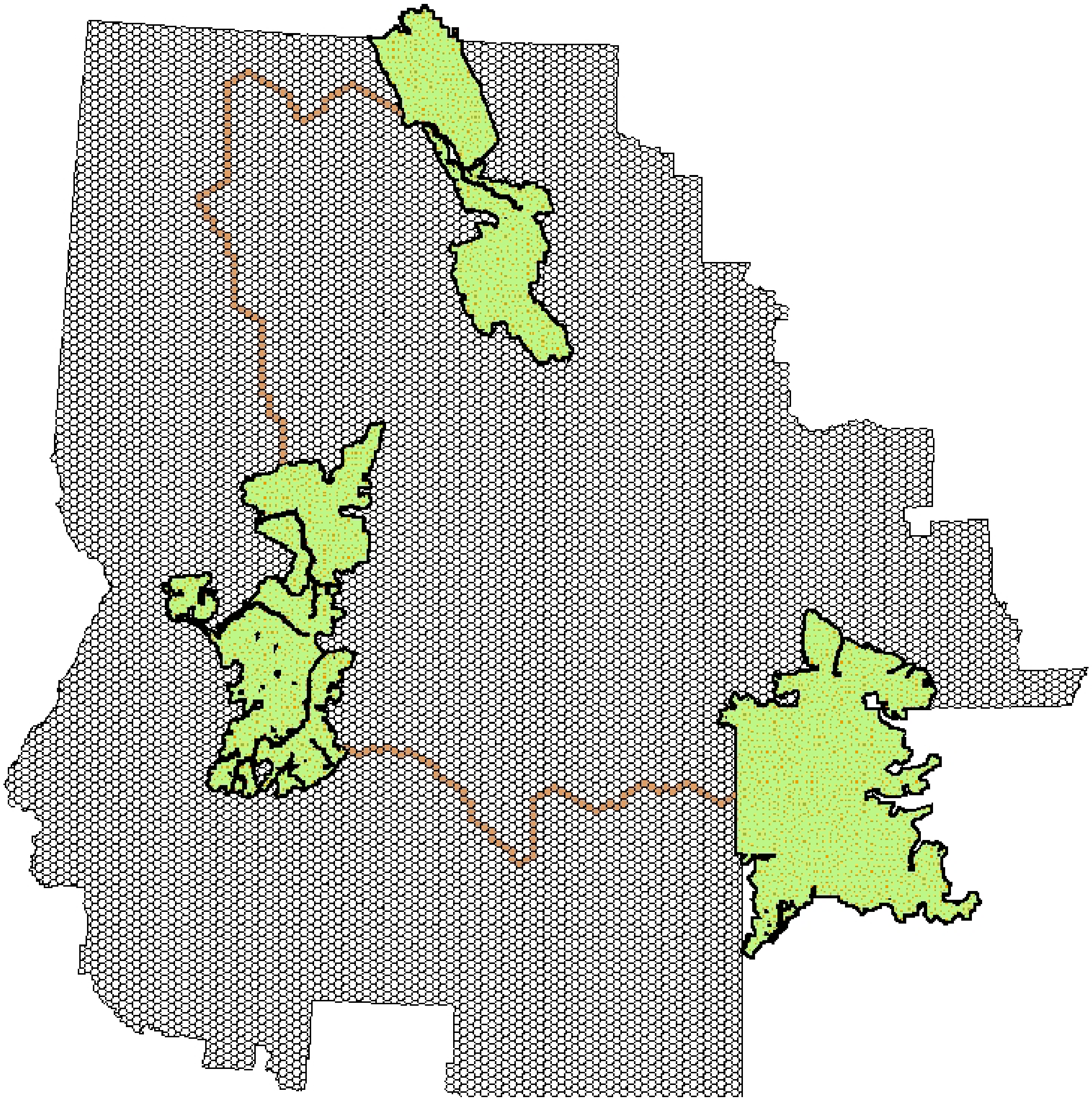


Figure A.2.2 Minimum Cost Hexagonal Grid with Transaction Costs

REFERENCES

Ando, A., J. Camm, S. Polasky, and A. Solow. 1998. "Species distributions, land values, and efficient conservation." Science. 279: 2126-2128.

Andreassen H.P., S.Halle and R. A. Ims. 1996. "Optimal width of movement corridors for root voles: not too narrow and not too wide." Journal of Applied Ecology. 33: 63-70.

Balmford A., K.J. Gaston, and S.L. Rodrigues. 2000. "Integrating Costs of Conservation into International Priority Setting." *Conservation Biology*. 14(3): 597-605.

Charnes, A., W.W. Cooper, and R.O. Ferguson. 1955. "Optimal Estimation of Executive Compensation by Linear Programming." *Management Science*. 1: 138-151.

Camm, J. D., S. Polasky, A. Solow, and B. Csuti. 1996. "A note on optimal algorithms for reserve site selection." Biological Conservation. 78: 353-355.

Cerdeira, J.O., Gaston, K.J., Pinto, L.S. 2005. "Connectivity in priority area selection for conservation." Environmental Modeling and Assessment. 10: 183–192.

Church, R. L., D. M. Stoms, and F. W. Davis. 1996. "Reserve selection as a maximal covering location problem." Biological Conservation. 76: 105-112.

Conrad, J.C, C.P. Gomes, W.V. Hoeve, A. Sabharwal, and J.F. Suter. 2007. "Connections in Networks: Hardness of Feasibility versus Optimality." Proceedings of the 4th International Conference on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems. Brussels, Belgium. May 2007.

Corman, T.H., C.E. Leiserson, R.L. Rivest, and C.L. Stein. 2001. *Introduction to Algorithms*. MIT Press.

Dixon, J.D., M.K. Oli, M.C. Wooten, T.H. Eason, J.W. Mccown, and D. Paetkau. "Effectiveness of a regional corridor in connecting two Florida Black Bear populations." Conservation Biology. 20(1): 155-162.

Ferraro, P.J., 2003. "Conservation contracting in heterogeneous landscapes: An application to watershed protection with threshold constraints." Agricultural and Resource Economics Review. 32: 53–64.

Fuller, T., M. Munguia, M. Mayfield, V. Sanchez-Cordero, S. Sarkar. 2006. "Incorporating connectivity into conservation planning: A multi-criteria case study from central Mexico." Biological Conservation I. 33: 131-142.

Groeneveld, R. 2005. "Economic considerations in the optimal size and number of reserve sites." Ecological Economics. 52: 219-228.

Haddad, N. M., David R. B., A. Cunningham, B. J. Danielson, D. J. Levey, S. Sargent and Tim Spira. 2003. "Corridor use by diverse taxa." Ecology. 84(3): 609–615.

Hansen, A., R. Rasker, B. Maxwell, J. J. Rotella, J. D. Johnson, A. W. Parmenter, U. Langer, W. B. Cohen, R. L. Lawrence, and M. P. V. Kraska. 2002. "Ecological Causes and Consequences of Demographic Change in the New West." Bioscience. 52(2): 151–162.

Land Trust Alliance 2004. "New York State Conservation Partnership Program 2003 Awards." Available: http://www.lta.org/yourregion/ny/nyscpp_2003_awards.htm.

Land Trust Alliance. 2005. "2005 National Land Trust Census." Available: http://www.lta.org/census/.

Land Trust Alliance. 2006. "Private Land Conservation in U.S. Soars." Available: http://www.lta.org/newsroom/pr_113006.htm.

Mace, R. D., J. S. Waller. 1997. "Spatial and Temporal Interaction of Male and Female Grizzly Bears in Northwestern Montana." Journal of Wildlife Management. 61(1) 39-52.

McEuen, A. 1993. "The Wildlife Corridor Controversy: A Review." Endangered Species Update 10: 1-6.

Margules, C. R., Nicholls, A. O., and Pressey, R. L. 1988. "Selecting networks of reserves to maximize biological diversity." Biological Conservation. 43: 63-76.

Messer, K. D. 2006. "The conservation benefits of cost-effective land acquisition: A case study in Maryland." Journal of Environmental Management. 79: 305-315.

Naidoo, R., A. Balmford, P.J. Ferraro, S. Polasky, T.H. Ricketts, and M. Rouget. 2006. "Integrating economic costs into conservation planning." Trends in Ecology & Evolution. 21(12): 681-687.

Newburn, D.A., P. Berck, A.M. Merenlender. 2006. "Habitat and open space at risk: targeting strategies for land conservation." American Journal of Agricultural Economics. 88(1): 28-42.

Noss, R.F. 1987. "Corridors in real landscapes: a reply to Simberloff and Cox." Conservation Biology. 1(2): 159-164.

Önal, Hayri, Robert A. Briers. 2006. “Optimal Selection of a Connected Reserve Network.” Operations Research. 54(2). 379–388.

Polasky, S., Camm, J.D., Garber-Yonts, B. 2001. “Selecting Biological Reserves Cost-effectively: An application to terrestrial vertebrate conservation in Oregon.” Land Economics 77: 68–78.

Promel, H.G. and A. Steger. 2002. *The Steiner Tree Problem: A Tour Through Graphs, Algorithms, and Complexity*. Vieweg.

Schmitt T. and A. Seitz. 2002. “Influence of habitat fragmentation on the genetic structure of Polyommatus coridon (Lepidoptera: Lycaenidae): implications for conservation.” Biological Conservation. 107: 291–297.

Sessions, J. 1992. “Solving for Network Connections as a Steiner Network Problem.” Forest Science. 38(1): 203-207.

Simberloff, D., J.A. Farr, J. Cox, and D.W. Mehlman. 1992. “Movement Corridors: Conservation Bargains or Poor Invesments?” Conservation Biology 6: 493-504.

Tewksbury, J. J., D. J. Levey, N. M. Haddad, S. Sargent, J. L. Orrock, A. Weldon, B. J. Danielson, J. Brinkerhoff, E. I. Damschen, and P. Townsend. “Corridors affect plants, animals, and their interactions in fragmented landscapes.” PNAS. 99(20): 12923-12926.

Underhill, L. G. 1994. “Optimal and suboptimal reserve selection algorithms.” Biological Conservation. 70: 85–87.

United States Fish and Wildlife Service. 2007. “Successful recovery efforts bring Yellowstone grizzly bears off the endangered list.” Available: http://www.fws.gov/news/NewsReleases/

United States Department of Agriculture. 2002. Census of Agriculture.

U.S. Department of Agriculture. 2006. “Conservation Reserve Program Monthly Summary.” Available: http://www.fsa.usda.gov/Internet/FSA_File/dec2006.pdf.

U.S. Department of Agriculture. 2007. “Forest Legacy Program: Funded and Completed Projects.” Available: http://www.fs.fed.us/spf/coop/programs/loa/flp_projects.shtml

United States Geologic Survey. 1999. Gap analysis- Land stewardship coverage. Available: http://gapanalysis.nbii.gov/

Urban, Dean and Timothy Keith. 2001. "Landscape Connectivity: A Graph-Theoretic Perspective." Ecology. 82(5): 1205–1218.

Walker, Richard, and Lance Craighead. 1997. "Least-Cost-Path Corridor Analysis Analyzing Wildlife Movement Corridors in Montana Using GIS." ESRI Users Conference.

Williams, J.C. 1998. "Delineating Protected Wildlife Corridors with Multi-objective Programming." Environmental Modeling and Assessment 3: 77–86.

Williams, J.C. 2002. "A Zero-One Programming Model for Contiguous Land Acquisition." Geographical Analysis. 34(4): 330-348.

Williams, J.C. 2005. "Restoring habitat corridors in fragmented landscapes using optimization and percolation methods." Environmental Modeling and Assessment 10: 239–250.

Williams, J.C. and C. S. ReVelle. 1997. "Applying mathematical programming to reserve selection." Environmental Modeling and Assessment. 2: 167-175.

Williams, J.C., C.S. ReVelle, and S.A. Levin. 2005. "Spatial attributes and reserve design models: A review." Environmental Modeling and Assessment. 10: 163-168.

Willis, C.E. and R.E. Perlack. 1980. "A Comparison of Generating Techniques and Goal Programming for Public Investment, Multiple Objective Decision Making." *American Journal of Agricultural Economics*. 62(1): 66-74.

CHAPTER THREE:
Assessing Differences in Implicit Prices for Open Space Across Urban, Suburban and Rural Areas: A Hedonic Pricing Study of Upstate New York

Abstract

This essay investigates differences in implicit prices for open space across different densities of development through estimation of separate hedonic equations for urban, suburban and rural residential areas surrounding Rochester, NY. Separate measures for public, agricultural, private forested and developed open space are included to measure heterogeneity in open space values within each of the submarkets. The hedonic model controls for spatial error autocorrelation through a sampling technique that eliminates nearest neighbors and corrects for endogeneity in surrounding open space through instrumental variable regression. Estimation results indicate that implicit prices for open space differ across the study area, with public open space valued more highly in urban settings and private open space generally valued more in rural areas. In addition, the results suggest that significant disamenities are associated with rural residential property in close proximity to Concentrated Animal Feeding Operations (CAFOs) and wastewater treatment facilities. Similar disamenities exist for residential property in urban areas in close proximity to waste transfer facilities and high traffic roadways.

3.1 Introduction

Open space land provides a vector of public amenities, from scenic views and recreational opportunities to serving as a buffer from land development. In recent years, significant financial resources have been devoted toward the preservation of threatened agricultural and forested open space land. According to the Land Trust Alliance's 'National Land Trust Census,' conducted in 2006, over 36 million acres of land are currently being conserved by private land trusts at local, state and national levels. Moreover, in 2006 alone, \$6.7 billion in public funds were approved as part of 133 separate ballot measures across the U.S. (LTA 2006) for the public protection of open space land. The economic motivation for these expenditures is bolstered by a growing number of hedonic studies of residential property sales, which illustrate that home buyers place a premium on proximity to open space land (e.g., Irwin 2002; Geoghegan, Lynch and Bucholtz 2003; Ready and Abdalla 2005; Anderson and West 2006).

Despite the extensive set of research in this area, the degree to which marginal implicit prices for specific types of open space evolve across the landscape has not been thoroughly investigated. Palmquist (1992) shows that to estimate the benefits (costs) that accrue from a localized externality, such as nearby open space, "it is only necessary to estimate the hedonic equation for a relatively homogeneous neighborhood." (p. 61) Given that the hedonic equation represents the locus of intersection between the supply and demand for specific attributes, estimating the hedonic equation across a heterogeneous study area makes benefit estimation difficult, as supply and demand are likely to be different in diverse locations (Michaels and Smith 1990). Yet in nearly all previous open space studies, a single hedonic equation is estimated for an entire, likely heterogeneous, urban area (e.g., Lutzenhizer and Netusil 2001; Irwin and Bockstael 2001; Paterson and Boyle 2002). Demonstrating that implicit prices for open space differ across a given study area is therefore critical to understanding differences in the benefits of nearby open space. From the social planner's perspective, determining the relative values for open space across a heterogeneous landscape may therefore be important in the optimal use of open space preservation funds, as there are significant differences in terms of land costs and the number of local residents affected by open space in urban and rural locations.

In this essay, data from a ten county "commute-shed" surrounding Rochester, New York are used to estimate marginal values for open space over a gradient of housing densities. Open space in each of the submarkets is further disaggregated into publicly owned open space, private open space that is forested, agricultural open space, and developed open space (e.g., golf courses). Additionally, proximity to Concentrated Animal Feeding Operations (CAFOs), an open space related disamenity (Palmquist, Roka and Vukina 1997; Ready and Abdalla 2005; Herriges, Secchi and Babcock 2005), and proximity to other locational disamenities, such as high traffic roads, waste management and wastewater treatment facilities, are included in the hedonic model.

The estimation results clearly show that nearby open space has a significantly positive impact on home price in all submarkets. Marginal implicit prices for different types of open space, however, are highly variable and differ significantly, although not necessarily systematically, between the three submarkets. In addition, disamenities resulting from proximity to CAFOs, wastewater treatment facilities, waste transfer stations and high traffic roads are significant, though results vary substantially by submarket.

The next section describes the hedonic method for valuing housing attributes. Section 3.3 highlights the results of several recent studies that provide the motivation for this essay and explores some of the issues related to the estimation of the hedonic price function brought forth in the open space valuation literature. Section 3.4 details the specification of the hedonic equation, section 3.5 describes study area and section 3.6 describes the data sources and variables used in the analysis. The estimated hedonic function parameters are presented in section 3.7 followed by a discussion of the policy implications and a review of some of the lessons learned from this analysis in section 3.8. The final section offers some concluding thoughts.

3.2 Review of Hedonic Theory

The hedonic pricing model is premised on the Lancasterian idea that the sales price of an individual housing unit is a function of the characteristics specific to that unit. The seminal paper by Rosen (1974) is consistently referenced in the literature as the first to show that the hedonic price function traces the loci of equilibria that arise from perfectly competitive home "suppliers" seeking to maximize their profit and utility maximizing consumers seeking to maximize the utility that they derive from a home's characteristics. In equilibrium, the price of residential property in a particular market can be written as a function of attributes specific to nearby open space (*O*), locational amenities and disamenities (*L*), and property specific features (*S*),

$$P = P(O, L, S). \tag{1}$$

Residential consumer *i* is assumed to have utility over housing attributes as well as a composite commodity (*X*), representing all other goods. In addition, consumer *i* faces a budget constraint, given his income, y^i. Formally, the utility function and budget constraint for consumer *i*, assuming that he purchases exactly one housing unit are

$$U^i = U(O, L, S, X) \tag{2}$$

$$y^i = P(O, L, S) + X. \tag{3}$$

Taking home prices as exogenous, the first-order conditions associated with consumer *i* maximizing utility subject to his budget constraint are such that, in the example of neighborhood open space,

$$\frac{\partial P}{\partial O} = \frac{\left(\partial U^i / \partial O\right)}{\left(\partial U^i / \partial X\right)} \tag{4}$$

In words, the consumer will purchase a residential property such that marginal implicit price of open space ($\frac{\partial P}{\partial O}$) is exactly equal to his marginal rate of substitution between open space land and money. Therefore in equilibrium, the marginal implicit price for each housing attribute provides a measure of an individual's marginal willingness to pay for that attribute (Taylor 2003).

Using the prices of transacted residential property and the characteristics associated with the property, it is possible to econometrically derive the hedonic price function for a given area. The econometric issues associated with estimating the hedonic function are described in greater detail in the next section. Once the hedonic price function has been estimated, the marginal implicit price for a particular attribute can be obtained by differentiating the hedonic equation with respect to that particular attribute.

The implicit prices obtained from the hedonic equation provide a measure of the average marginal willingness to pay of households in the study area for a marginal change in a given attribute. To derive the entire marginal willingness to pay function for an attribute in a single market requires a second stage estimation procedure to be performed, where the estimated marginal implicit prices are regressed on attribute quantities as well as a set of potential demand shifting demographic attributes. The second stage analysis is rarely executed, however, because of the econometric challenges associated with estimating a demand function that relies on estimated implicit prices. The implicit prices are estimated by the hedonic equation, which generally contains a similar set of variables included in the demand function. This leads to a classic identification problem, where it is not possible to differentiate between the effects of demand shifting variables and the actual slope of the demand function (Freeman 2003).

To obtain measures of welfare benefits from discrete changes in the quantity of a particular attribute across a market, it is generally necessary to estimate the entire marginal willingness to pay function. This would allow one, for example, to integrate the marginal willingness to pay function over the change in the attribute. Fortunately, in the case of a localized externality, Palmquist (1992) shows that estimation of marginal implicit prices is enough to estimate the welfare benefits associated with marginal changes in an attribute, assuming negligible costs associated with relocation. As long as the externality affects a relatively small number of homes, a marginal change in the externality does not appreciably alter the supply of the attribute in the market. Thus the marginal implicit prices accurately represent the benefits (costs) associated with marginal changes in a localized externality such as open space. Over a heterogeneous study area, marginal implicit prices are likely to vary, given the allocation of neighborhood features such as housing density. Palmquist therefore recommends estimation of hedonic equations involving localized externalities across only homogeneous areas, to get a true measure of the benefits associated with open space in those areas.

3.3 Review of Relevant Literature

A significant and growing set of studies use hedonic methods to estimate values for open space land. A recent paper by McConnell and Walls (2005) highlights approximately 40 hedonic studies conducted between 1967 and 2003 that include some measure of open space as an explanatory variable for residential property price. In this section, a subset of recent hedonic studies that incorporate open space measures are described. Across this set of studies, there is much variation in estimated marginal implicit prices for open space land. While some of the differences in estimated marginal implicit prices are undoubtedly due to differences in how open space variables are defined and measured, differences in the characteristics of the respective study areas are also likely to contribute to the variability. The following review of the literature concentrates primarily on recent hedonic studies that have estimated the implicit prices of different open space types within a specified neighborhood of each observed home sale. The studies by Cheshire and Sheppard (1995) and Geoghegan, Lynch and Bucholtz (2003) specify separate hedonic equations for housing markets defined based on political boundaries, while the remaining studies estimate one equation for the entire study area.

Cheshire and Sheppard (1995) were among the first to evaluate hedonic values for differentiated open space types. The authors estimate the separate impacts of publicly accessible open space and private open space land on residential home prices in two English towns. Estimation results indicate that publicly accessible open space has a significantly more positive hedonic value than private open space in the town of Darlington, but that the opposite is true for the town of Reading. They attribute this difference to the relative scarcity of the open space types between the two towns, consistent with diminishing marginal utility for open space land. For example, Reading, which has a relatively high percentage of publicly accessible open space, has a correspondingly lower marginal value for additional public land.

Irwin and Bockstael (2001) utilize residential property sales data in four Maryland counties to evaluate the implicit prices for private open space, public open space and land held in conservation easements. Correcting for identification issues associated with endogeneity and spatial autocorrelation, Irwin and Bockstael find that privately held open space, public open space and land with conservation easements all have a positive impact on home price. The marginal implicit price for privately held land is estimated to be roughly twice that of public land and land under conservation easements. Following a similar estimation procedure for the same four counties in Maryland, Irwin (2002) incorporates a more heterogeneous set of open space types that includes public open space, cropland, pasture, private forested open space and land held in conservation easements. The study finds that publicly owned open space and land under conservation easements have a significantly greater effect on residential prices than pasture and cropland. In addition, she provides evidence that privately-owned forest land has a significantly negative impact on residential home prices relative to agricultural land. In both of these studies, the open space variables are measured as the percentage of the particular type of open space within 400 meters of the residential sales point.

A set of studies by Geoghegan, Wainger and Bockstael (1997), Geoghegan (2002) and Geoghegan, Lynch and Bucholtz (2003) also look at hedonic values for open space in Maryland. One of the primary interests of the Geoghegan, Wainger and Bockstael (1997) study is determining whether "…the value of open space might be higher in urban areas relative to rural areas, where there is a scarcity of open space." (p. 261) To investigate differences in open space value across the landscape, they estimate a varying parameters model, which allows model parameters to vary according to distance from Washington, DC. Their results indicate that the value of open space within 100 meters of a house is not significantly different from zero. The value of open space within 1.0 km of a house is negative in close proximity to the center of Washington, DC and then increases slightly as the distance from the city increases. Although Geoghagan et al.'s interest in measuring how the value of neighborhood attributes vary with distance to the city center is novel, their results are difficult to interpret. First, despite the fact that they incorporate interesting landscape variables into their model, such as indices measuring the diversity and fragmentation in nearby land covers, they include only one measure of open space. Second, the fact that open space has a negative value at close proximity to the city and then increases as distance from the city center increases is puzzling and goes against their hypothesis.

The Geoghegan (2002) and Geoghegan, Lynch and Bucholtz (2003) studies differentiate between permanent and developable open space within close proximity to each home sale point. Model results in Geoghegan (2002), from one Maryland county, show that hedonic values for permanent open space are significantly positive, while the value of developable open space is not significantly different from zero. Geoghegan, Lynch and Bucholtz (2003) present a similar analysis, but estimate separate hedonic equations for three Maryland counties and address issues of spatial autocorrelation and endogeneity. Permanent open space within close proximity (100 meters) is found to have a significantly positive hedonic value in only one of the three counties, while developable open space had an insignificant hedonic value in all three counties.

A recent study by Ready and Abdalla (2005) in Bucks County, PA incorporates open space variables and looks more specifically at spillovers from agricultural open space by measuring the influence of proximity to animal agriculture operations and other locational disamenities on housing prices. The authors differentiate between open space in pasture or crops, public open space, private forested, eased land and vacant land. They find that all open space types within 400 meters, except for vacant land, have a significantly positive impact on housing price, while only public open space and eased land have a significant impact between 400 and 1,600 meters. The study also finds a significantly negative relation between proximity to livestock facilities and house price. These findings are similar to a study by Herriges, Secchi and Babcock (2005), who find significant disamenities resulting from hog operations in Iowa. While including livestock facilities in an analysis of open space values may seem unjustified, in many areas agricultural open space is defined by the predominance of animal agriculture. Failure to include the potential disamenities resulting from livestock facilities thus has the potential to bias measures of agricultural open space.

A study by Lutzenhiser and Netusil (2001) in Portland, Oregon incorporates measures of neighboring land in urban parks, natural area parks, cemeteries and golf courses. The authors find that all of the open space types, with the exception of cemeteries, have positive impacts on home price, with natural area parks and golf courses having the highest marginal value. They also show that the influence of golf courses on home price decreases rapidly with distance, whereas the influence of parks tends to decrease more gradually. A more recent study by Netusil (2005) also utilizes residential home sales data from Portland and incorporates a similar set of open space types as well as measures of the tree canopy and environmental zoning. In this study, golf courses and quantity of tree canopy are the only open space types found to have significantly positive marginal implicit prices, but it is unclear if this result arises from the inclusion of additional variables or simply because it is based on a different sample of housing transactions.

Paterson and Boyle (2002) model the amount of visible open space, in addition to overall quantities of open space within one kilometer of each housing transaction in their suburban Connecticut study area. The estimation results indicate that aggregate amounts of forested open space and agricultural open space within one kilometer does not have a significant impact on home price, although visible forest land has a negative implicit price. While visibility is likely an important attribute of open space amenities, the results from this study suggest that excluding visibility considerations is unlikely to bias estimates of open space benefits from models that exclude visibility measures.

The studies cited above measure the influence of nearby open space land by calculating the amount of open space within a specific radius of each residential transaction, thus simultaneously capturing elements of both scale and proximity. A recent study by Anderson and West (2006) captures the impact of nearby open space slightly differently, by including independent variables that measure the distance from the residential property to neighborhood parks, regional parks, cemeteries, golf courses and open water. In addition, they interact the distance measure with the size of the nearest open space type. They also include interactions between distance to open space and housing density, distance to the CBD, income and other demographic measures. The estimation results indicate that proximity to regional parks and open water is highly valued and that, in general, marginal implicit prices for open space are higher in wealthier, more densely populated areas closer to the CBD.

The studies reviewed here account for spatial externalities resulting from a wide range of open space types across a diverse set of study areas. This essay builds on previous studies in defining the open space types that are included in the hedonic model. Following the majority of recent analyses, the amount of each open space type is measured within a specific radius of each residential transaction. In addition, hedonic values for the various open space types are estimated across urban, suburban and rural areas, instead of treating the effect of open space as fixed across the entire study area. This reflects the intuition of Geoghegan, Wainger and Bockstael (1997) as well as Anderson and West (2006), who allow open space variables to vary with distance from the central business district. It also draws on the analyses of Cheshire and Sheppard (1995) as well as Geoghegan, Lynch and Bockstael (2003), where separate equations are estimated for specific regions to evaluate differences in open space variables. Finally, estimating separate equations for urban, suburban and rural municipalities is in line with Palmquist's (1992) recommendation that localized externalities are most appropriately measured within relatively homogeneous areas.

3.4 Model Specification

This section describes four model specification issues that have arisen in past hedonic studies and how they are treated in this study. While the hedonic price function allows for an analytically simple way to use revealed preferences to infer implicit prices for the characteristics of residential property, there are numerous empirical estimation issues that arise. This section outlines specification issues related to the choice of functional form, market extent, and dealing with endogeneity and spatial autocorrelation. To begin with, a functional form that relates a home's price to its characteristics must be assumed. Given that the hedonic function represents the relationship between home buyers and sellers, economic theory does not provide much guidance on the choice of functional form (Taylor 2003). Authors have often used goodness of fit criteria to justify their choice of a particular functional form (e.g., Archarya and Bennett 2001; Irwin 2002), or estimate a flexible functional form model such as the Box-Cox. Cropper, Deck and McConnell (1988), however, show that the quadratic Box-Cox model measures implicit prices with greater error than either the linear or semi-log functions in the presence of omitted variables. Further, specifying a linear dependent variable implies that the marginal price of one attribute is independent of the levels of other attributes (Freeman 2003). This seems unlikely for the case of open space variables, where the implicit price of public open space, for example, is likely dependent on the amount of other types of open space surrounding the home. In this study a semi-log functional form is employed in all three markets. Keeping the estimation equation consistent across submarkets facilitates the comparison of individual coefficient estimates.

The choice of market extent is also an empirical issue that must be resolved in estimation of the hedonic price function. According to Freeman (1993), a housing market should be segmented if the structure of demand or supply is different across segments and if purchasers in one segment do not participate “significantly” in other segments. While these criteria appear to be rather restrictive, Bourassa et al. (2003) note that housing submarkets have been defined in the literature based on geographic and political boundaries, socio-economic and environmental characteristics, physical characteristics of the dwellings as well as by the opinion of real estate agents. Indeed, Freeman suggests that no matter how submarkets are defined, the estimates from each submarket will accurately reflect the implicit prices in that submarket.

In this study, the descriptive statistics provided in Table 3.1 illustrate that the structure of the residential housing supply between three residential submarkets, defined based on commuting patterns, is indeed different. For example, the mean property size in urban areas is 0.3 acres, while in rural areas it is 0.8 acres. Further, the amount of aggregate open space within 400 meters varies from 65.5 acres in urban areas to 101.1 in rural areas. No assumption is made regarding the structure of demand across segments, as there is no reason to believe that demand differs systematically across the study area. Whether purchasers in one segment participate “significantly” in other segments is also an open empirical question. Certainly there is some cross-segment participation, given that there are no barriers influencing mobility between markets. The reason for segmenting markets in this study is motivated by the importance of evaluating differences in marginal implicit prices across a heterogeneous area, rather than specifying unique demand systems for each submarket. The desire for obtaining accurate implicit prices estimates in a heterogeneous market follows the motivation of Michaels and Smith (1990), who estimate separate hedonic equations in four submarkets in Boston, MA to measure differences in the influence of hazardous waste sites on property prices.

In addition to the choice of functional form and market delineation, issues also arise in regards to the exogeneity of model variables. Specifically, Irwin and

Bockstael (2001), Irwin (2002), Geoghegan, Lynch and Bucholtz (2003), Ready and Abdalla (2005) have found that the amount of open space in a particular location is a function of nearby residential property prices. This is a result of land development over time in which open space areas with favorable attributes for residential use are developed. The quantity of open space surrounding a residential property at time *t* is therefore a function of residential prices, which reflect the favorable locational attributes of the area. This endogeneity problem results in measures of nearby private open space that are correlated with the model error term and thus biased parameter estimates.

Model misspecification can also occur as a result of model errors that are correlated across space. Errors can be spatially correlated if a spatially heterogeneous variable that explains variation in housing prices is left out of the model. This type of spatial relationship is often referred to as a spatial error or spatial autoregressive disturbance model and typically results in unbiased, but inefficient coefficient estimates (Anselin 2002). Additionally, spatial autocorrelation can be attributable to substantive correlation in the prices of nearby housing units. This can occur, for example, if individuals prefer homes that are surrounded by high priced homes simply because of the fact that the homes have high prices. Substantive spatial autocorrelation leads to biased parameter estimates (Anselin 2002). Kimm, Phipps and Anselin (2003) describe ways of treating autocorrelation in a hedonic study of air pollution in Seoul, South Korea, through both implementation of a spatial lag model, which includes an independent variable that measures the weighted price of nearby homes, or a spatial error model, which corrects standard errors based on the weighted errors of nearby housing sales. Numerous other hedonic studies have estimated spatial error and spatial lag models (e.g., Basu and Thibodeau 1998; Patterson and Boyle 2002; Brasington and Hite 2005).

When endogenous variables are present in the model, however, traditional spatial econometric approaches are not sufficient. Irwin and Bockstael (2001) show that in the presence of endogenous open space variables, spatial autocorrelation will result in biased coefficient estimates. This is attributable to the fact that the open space variables are correlated with the error terms, which are themselves correlated across space. Further, Irwin (2002) states that standard spatial models are not capable eliminating the spatial autocorrelation, since "it is not possible to separate the endogenous spatial spillover effects from the spatial error, given that they are both positively correlated."(p.474). She therefore accounts for spatial error correlation by taking a random sample of housing transactions so that no two transactions in the dataset are within a certain 100 meter neighborhood of one another. In this study, a sampling method similar to that introduced by Irwin is used to control for spatial autocorrelation and is described in section 3.7.

3.5 Study Area and Hedonic Model

The extent of the study area is defined by the "commute-shed" of Rochester, NY, which is composed of 83 municipalities in which at least 5% of the workforce commute into the City of Rochester for employment. These 83 municipalities are distributed across parts of 10 counties. The commuting patterns of the workforce are derived from the 2000 U.S. Census for each municipality. Given that the primary focus of this essay is on open space amenities, including agricultural land, housing transactions inside the City of Rochester are excluded.

To address differences in the implicit prices of open space and other housing attributes, the study area is divided into urban, suburban and rural submarkets. The Census defines land as being either urban or rural, but they do not have a strict definition for suburban areas. As such, the urban, suburban and rural submarkets are defined based on the percentage of the workforce that commute to Rochester for employment. The urban submarket is defined as municipalities in which at least 40% of the workforce commute to Rochester, while the suburban submarket includes municipalities in which between 25 and 40% commute to Rochester. According to these definitions, there are 4 urban and 15 suburban municipalities. The remaining 63 municipalities have fewer than 25% of the workforce commuting to Rochester and are classified as rural. The names of the municipalities included in the urban, suburban and rural submarkets are provided in appendix Table A.3.1 and a map of the submarkets is provided in Figure A.3.2.

Since data from downtown Rochester are excluded, the three submarkets can alternatively be thought of as representing inner ring suburbs, outer ring suburbs and the peripheral rural areas. Given that the Rochester market could potentially be divided into submarkets based on a number of criteria, two alternative models are estimated to test the robustness of the results from the primary model, defined based on commuting patterns. The first alternative model is estimated with submarkets determined by census definitions of urban areas, rural areas and urban clusters. The second alternative model is estimated based on political boundaries, with the submarkets defined based on whether the transaction occurred inside or outside of Monroe County; the county in which Rochester is located. The results of these models are provided in the appendix and discussed in section 3.7.

Using the natural log of the sales price of each residential transaction as the dependent variable, the estimated hedonic equation is given by

$$\ln(P_{mi}) = \alpha_m + \beta_m O_{mi} + \lambda_m L_{mi} + \delta_m S_{mi} + \varepsilon_{\mathrm{mi}}, \tag{5}$$

where O_{mi}, L_{mi} and S_{mi} are vectors of variables representing neighborhood open space, location and property specific characteristics respectively for housing unit *i* in submarket *m* and α_m, β_m, λ_m and δ_m are vectors of parameters to be estimated for each submarket. Additionally, the possibility that the open space measures are endogenously determined (i.e., $E(O_{mi}, \varepsilon_{mi}) \neq 0$) is considered as well as the possibility that model errors are correlated across space (i.e., $E(\varepsilon_{mi}, \varepsilon_{mj}) \neq 0$ for all i,j). The selection and calculation of the variables in each of the attribute vectors is described in the next section.

3.6 Data Description

The hedonic model presented above is estimated using spatially explicit data derived from a wide variety of sources. This section describes the data used and the calculation of all of the variables included in the hedonic model. Data on the sales price, property specific characteristics and geocoded location of each residential transaction are from the New York State Office of Real Property Services (NYS ORPS). The open space and locational variables are then linked to the transaction data with a geographic information system (GIS). The open space variables are derived from a GIS layer of the 2001 National Land Cover Dataset (NLCD). Data on the location of CAFOs, wastewater treatment facilities, transfer stations, and high traffic roads are from the NYS Department of Environmental Conservation (DEC), while data on the location of entrances to the NYS Thruway are made available by the NYS Thruway Authority. Data on neighborhood demographic characteristics come from the 2000 U.S. Census, school quality measures come from the NYS Department of Education and school tax rates are from the Office of State Comptroller. The derivation of each of the model variables is described in greater detail below.

The transactions data are provided by NYS ORPS for the years 1998 through 2006. To ensure that the dataset is comprised exclusively of arms-length transactions of detached residential housing, only residential property that sold for at least $10,000, with a lot size between 0.1 and 5.0 acres and at least 600 square feet of living area are included. These criteria closely follow that used by Ready and Abdalla (2005). The resulting dataset is composed of 66,609 observed arms-length transactions of detached residential housing units. The urban submarket contains 34% (22,654) of the total observations, while the suburban and rural submarkets comprise 39% (26,196) and 27% (17,759) of observations in the dataset, respectively. A map depicting the spatial distribution of the transactions included in the dataset is included in the appendix as Figure A.3.1.

Property Specific Variables

The set of property specific variables are taken directly from the ORPS transactions database. Although the primary objective of this study lies in determining the marginal implicit prices of neighborhood open space characteristics, it is essential to control for variation in property specific characteristics of the housing units. To this end, six property variables are included in the hedonic equation. The overall condition of the home measures the quality of the residential structure and is defined based on a five point scale from poor to excellent. The square feet of living area and site acreage are included as independent variables that control for the size of the structure and property respectively. The age of the home is also included as an independent variable, with 2006 serving as the baseline year. Finally, the number of bathrooms and a variable defining whether the home has central air conditioning are included to measure household-level amenities. Property specific variables, similar to the set included in this study, are essential control variables. The year of sale, relative to the 2006 base year, is included as an independent variable to control for house price trends over time.

Neighborhood Open Space Variables

The data on land cover in open space are derived from the 2001 NLCD, which was developed through the efforts of the Multi-Resolution Land Characteristics (MRLC) Consortium, a group comprised of various federal agencies. The NLCD classifies land cover for the conterminous 48 states in the U.S. using Landsat 7 satellite imagery. Land use is assigned based on contiguous 30M grid cells according to 29 separate classifications. The land covers corresponding to open water, developed open space, forest, pasture/hay and cultivated crops were then extracted for the entire study area. The developed open space classification is composed primarily of large-lot residential property, golf courses and land planted for recreation or aesthetic purposes. The NLCD coverage was merged with spatial data on the location of publicly owned land, available from the New York State Office of Cyber Security and Critical Infrastructure Coordination. The publicly owned land is comprised of both municipal parks, which tend to be smaller parcels used primarily for recreation, as well as state and county parks, which are generally larger than municipal parks and offer a greater range of recreational activities in addition to wildlife habitat.

The neighborhood open space variables measure the quantity of land within close proximity to each residential sale that is in open space. The amount of proximate open space land is measured as the acreage in each open space type within a radius of 1600 meters from each observed residential sale. To investigate preferences for the spatial distribution of open space amenities, the amount of open space located within 400 meters of the home and between 400 and 1600 meters are estimated separately. The 400 meter distance essentially measures open space that is in the immediate vicinity of the residence that might influence the viewshed or be encountered on a short walk. The 400-1600 meter range measures open space that is contained within a short drive from the house or that might be encountered on a bike ride or longer walk. The intuitive expectation is that open space within 400 meters has a larger impact on home price than open space quantities at the larger 400 to 1600 meter scale. The open space scales of 400 and 1600 meters match the methodology used in Ready and Abdalla (2005). Irwin (2002) evaluates land use only within 400 meters, while Geoghegan, Lynch and Bucholtz (2003) evaluate open space within both a 100 meter and 1600 meter radius. Finally, Netusil (2005) looks at open space amenities within 200 feet of the home, between 200 feet and ¼ mile and between ¼ and ½ mile.

To calculate the open space acreage for each transaction, the number of 30 meter grid cells of each open space type, within 400 meters and within 1600 was calculated for each residential transaction using the ARCGIS software and then converted to acres (each 30 meter cell = 0.22 acres). The percentage of each open space type between 400 and 1600 meters is calculated by subtracting the acreage within 400 meters from the acreage within 1600 meters. The coefficients on the open space variables in the hedonic model can be interpreted as the predicted percent change in home price associated with changing one acre of land area within either 400 meters or between 400 and 1600 meters of a home from a developed use to a particular type of open space (or conversely the percent loss in home price associated with the open space being developed).

The acreage in open water within 400 meters and between 400 and 1600 meters of each home are also included as independent variables. Given Rochester's location next to Lake Ontario, as well as the fact that a portion of the Finger Lakes region lies in the study area, access to water is an important variable in describing residential price differentials. The open space variables that are included in the model have all been shown to significantly influence property prices in previous hedonic analyses.[1] Agricultural open space is found to positively influence property prices in Ready and Abdalla (2005) and negatively influence property prices in Smith, Poulos and Kim (2002). Private forest land is shown to have a positive influence in Tyrväinen and Miettinen (2000), but a negative effect in Irwin (2002). Developed open space, such as large lot residential and golf courses has a positive and significant influence on

[1] This is by no means an exhaustive list, but is included to give the reader a flavor for the number of hedonic studies that have found significant implicit prices for open space land.

residential property in studies by Smith, Poulos and Kim (2002) as well as Do and Grudsnitski (1995). Smaller municipal parks have a strong positive impact in Espey and Owusu-Edusei (2001), but a negative impact in Lutzenhiser and Netusil (2001), while larger parks are shown to have the highest impact relative to other open space types in Anderson and West (2006) and Irwin (2002). Finally, proximity to open water is shown to have a strongly positive impact in Earnhart (2001) and Mahan, Polasky, and Adams (2000).

Location Specific Variables

In addition to the property specific characteristics and the relative acreage in various types of open space surrounding a residential unit, the price of a home is influenced by its specific location in the landscape. A set of variables that capture the influence of nearby amenities and disamenities are therefore included in the hedonic equation. Although agricultural open space is hypothesized to have a positive effect on home price, there are some aspects of agricultural production that may generate negative external effects. Over half of all agricultural receipts in New York State are generated from dairy farming (NYS Ag and Markets 2002). While the number of dairy farms across New York has been steadily decreasing, the number of animals housed on the average farm has increased (NASS 2002). As the remaining dairy farm operations have increased in size, the potential for negative impacts on nearby residential property has also risen (Palmquist, Roka and Vukina 1997).

To capture the influence of large scale animal agriculture, a proximity measure to all 98 farms[2] in our study area that qualify as Concentrated Animal Feeding Operations (CAFOs) is calculated. The geographic location of each CAFO coincides with the center of the farmstead, as determined by the NYS Department of Environmental Conservation (DEC). According to the DEC classification system, the

---

[2] The Finger Lakes Racetrack, which is identified as a CAFO, is not included in the analysis.

study area contains 60 medium-size CAFOs (200-599 mature dairy cows) and 38 large-size CAFOs (at least 600 mature dairy cows). Dairy farms account for the overwhelming majority of the CAFOs in the dataset, though some sheep and chicken farms are included in the total. The impact of large and medium sized CAFOs is measured using an index developed in Ready and Abdalla (2005) that allows for the measurement of the influence of multiple nearby CAFOs, weighted by their proximity. For each housing unit, the proximity index is measured as

$$C_i = \sum_{j=1}^{J} \left(Dist_{ij}^{-1} - K^{-1} \right) \quad , \tag{6}$$

where J is the number of each type of CAFO within K meters of housing unit i and K = 1,600 meters. The 1,600 meter extent of influence is based on Ready and Abdalla's finding that the influence of nearby disamenities does not extend past 1,600 meters and also corresponds to the 1600 meter extent of measured nearby open space. Summing across all of the CAFOs within 1,600 meters allows for measuring the influence of multiple sites, weighted by their distance from the housing unit. In addition to Ready and Abdalla, separate studies by Herriges, Secchi and Babcock (2005) as well as Palmquist, Roka and Vukina (1997) have shown significantly negative spillovers associated with residential proximity to hog operations in Iowa and North Carolina respectively.

Other potential disamenities included in the hedonic model are proximity to waste management and wastewater treatment facilities as well as proximity to high traffic roadways. Data on the spatial distribution of all three of these disamenities are made available by the NYS DEC. Negative implicit prices for proximity to waste management facilities have been found in Hite et al. (2001) and Ready and Abdalla (2005), among others. For the proximity to waste management facilities variable, the proximity index is calculated for landfills and waste transfer stations. Although a transfer station, defined as a waste management facility where solid waste is received for the purpose of subsequent transfer to another facility, likely has less of an impact on housing price than a landfill, there are too few landfills in the study area (a total of 3) to differentiate between the two. It is hypothesized that that the types of sites, smells and sounds associated with both landfills and transfer stations make living nearby unattractive. Overall there are 45 landfill and transfer stations within the study area whose proximity influence nearby residential sites. In addition, 47 wastewater treatment facilities that have been issued a State Pollution Discharge Elimination System (SPDES) permit are contained within the study area. Wastewater treatment plant proximity is included in the study by Ready and Abdalla, but was not found to significantly impact property values.

High traffic roadways may also be seen as a disamenity to individuals who live nearby. Although major roadways are essential for commuting to work and to commercial areas, living very close to a high traffic road entails enduring the noise and exhaust associated with heavy vehicular traffic. An extensive review of hedonic studies measuring the noise and pollution impacts of high traffic roads on residential property can be found in Bateman et al. (2001). Previous studies have found that the disamenity impact of high traffic roads is typically limited to a relatively small geographic scale. For example, Palmquist (1992) finds that the impact of high traffic roads extends for approximately 1,000 feet. Therefore the upper distance bound on the high traffic road variable used in this analysis is 400 meters. The high traffic roadways included in this study were chosen based on the NYS DEC major roadways designation. They include all of the interstates that pass through Rochester in addition to several major state routes. Smith, Poulos and Kim (2002) as well as Palmquist (1992) find significant negative externalities associated with living in very close proximity to major highways. The proximity index is again used as a measure of the impact of living near high traffic roads. A map illustrating the distribution of all of the disamenities described above is provided in the appendix as Figure A.3.3.

The proximity of residential property to the central business district and to major transportation routes should positively impact a home's price. The distance from each observed housing transaction to the center of Rochester and to the closest of the six NYS Thruway entrances in the study area are included as independent variables. Based on results from myriad hedonic studies that include distance to central cities (e.g., Brasington and Hite 2006; Anderson and West 2006), and the original insights of von Thunen, property prices are expected to decrease with distance from the CBD, holding all else constant. The distance to a Thruway entrance applies to all of the homes in the dataset, whereas the proximity index for high traffic roads mentioned previously is limited to a 400 meter buffer. While living in close proximity to a high traffic road is likely to be seen as a nuisance, a location that is relatively close to a Thruway entrance is hypothesized to have a positive influence on home price (e.g., Geoghegan 1997).

The demographic characteristics of other households in the neighborhood also have been found to influence a home's price (e.g., Cheshire and Sheppard 1995; Irwin 2002). Variables measuring median income and the percentage of the population that is white are included using data from the 2000 US Census at the Census tract level. Based on previous studies, it is expected that both median income and the percentage of the population in the Census tract that identify themselves as white will have positive influences on home price.

The quality of local public schools has also been shown to have a positive impact on residential prices (e.g., Braisington 1999; Ready and Abdalla 2005; Clapp, Nanda, and Ross 2005). In this study, two proxy measures of school quality are included, which were provided by the NYS Department of Education. The school performance index is determined based on the results of statewide tests administered to 4th grade students in language arts and mathematics in 2003. A measure of the student to teacher ratio at the elementary school level is also included as an independent variable. This variable is derived by dividing the number of enrolled students in each elementary school by the number of full and part time faculty. The school quality measures for each residential sale are assigned by merging the location of the housing unit with a contiguous map of elementary school boundaries, assembled from data obtained from individual school districts. This level of specificity represents a marked improvement from school quality measures defined at the district level (e.g., Brasington 1999; Clapp, Nanda, and Ross 2005). Brasington's study on the influence of school quality on residential property prices finds that both school performance measures and student-teacher ratio are capitalized into home prices in the majority of the six Ohio housing markets that he evaluates. It is hypothesized that both school quality measures have a positive influence on home price, but comparing the two provides important insights, in that the student to teacher ratio represents an input to education while test score performance measures a school quality output. While high quality schools are expected to have a positive influence on housing prices, the school budgets are financed in large part from local property taxes. Both Brasington (1999) and Klapp, Nanda, and Ross (2007) find that school taxes are capitalized into home prices. To control for heterogeneity in school taxes across the study area, a measure of the average school tax rate adjusted for property value and provided by the Office of State Comptroller, is included as an independent variable.

Two final location variables included in the hedonic model are whether or not the residential parcel is located in a village and whether the residence has access to public sewer systems. Many studies include the distance to nearby rural towns centers or villages (e.g., Michael, Boyle and Bouchard 2000; Herriges, Secchi and Babcock

2005) as an explanatory variable, but whether or not the residence is located inside or outside of a village boundary is likely to have an impact on the services received (e.g, refuse collection, road maintenance) and taxes. Villages often provide additional public services that are not available to residents living outside of their boundaries. These services, however, are again financed by local taxes. Moreover, a home situated in a rural village often has a different set of surrounding amenities than a home outside such boundaries. A dummy variable indicating whether the home is located in a village or city[3] is included as an independent variable to capture this effect. The availability of a connection to the public sewer system is clearly an advantage over the potentially costly maintenance involved with private septic systems (e.g., Ready and Abdalla 2005). As such, it is expected that the availability of a public sewer connection will increase home price, holding all other factors constant.

Descriptive Statistics

Descriptive statistics for all of the variables included in the hedonic equations for each of the three primary submarkets are included as Table 3.1. It is clear from the descriptive statistics that considerable differences exist between the urban, rural and suburban housing markets. Most notably, percentages of agricultural land and private open space increase significantly between urban to rural submarkets, while developed open space and public open space display the opposite trend. In addition, homes in the suburban submarket tend to be newer, larger, in better condition and more expensive than homes in either the rural or urban submarkets. It should be noted that differences in average housing characteristics alone do not necessitate estimating parameters that vary by submarket. Indeed, heterogeneity across the study area is essential to accurately estimating the hedonic price function. It is the systematic variation that

[3] The cities of Batavia and Canandaigua are located within the study area.

exists across urban, suburban and rural areas, however, that points to the importance of separating the effects of the covariates based on submarket.

Table 3.1 Descriptive Statistics for Model Variables

		Full Market		Rural		Suburban		Urban	
Variable	**Units**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Observations	#	66,609		17,759		26,196		22,654	
Sale Price	Dollars	131,151	78,959	116,370	85,283	154,471	82,328	115,771	61,286
ln(Sale Price)	ln(Dollars)	11.67	0.46	11.52	0.52	11.85	0.43	11.58	0.38
Neighborhood: 0 - 400 Meters (130 acres)									
Agriculture	Acres	25.1	31.8	49.1	35.2	26.3	29.6	5.0	11.8
Developed OS	Acres	42.9	29.1	23.1	18.1	55.1	32.4	44.2	23.3
Private Forest	Acres	17.8	18.1	22.0	19.6	20.0	17.9	11.9	15.5
Municipal Park	Acres	1.0	4.3	0.7	2.8	1.6	5.7	0.5	3.3
State/County Park	Acres	1.1	6.3	0.6	4.8	0.7	5.3	1.8	8.1
Water	Acres	2.6	11.4	5.7	17.0	1.1	6.2	2.0	10.2
Neighborhood: 400 - 1600 Meters (1,900 acres)									
Agriculture	Acres	443.3	426.8	853.8	374.0	459.4	361.4	102.8	162.8
Developed OS	Acres	474.5	308.0	202.5	128.5	594.0	336.7	549.6	233.8
Private Forest	Acres	286.3	199.9	354.9	225.6	321.7	186.4	191.5	152.5
Municipal Park	Acres	21.4	31.3	6.2	13.5	30.5	35.9	22.8	31.2
State/County Park	Acres	36.1	97.4	14.3	64.2	21.7	81.5	69.8	123.5
Water	Acres	57.8	184.0	90.0	242.9	31.1	120.5	63.4	186.7
Location Based									
Medium Cafo	Index	0.00002	0.00041	0.00008	0.00079	0.00000	0.00004	-	-
Large Cafo	Index	0.00001	0.00013	0.00002	0.00025	-	-	-	-
Wastewater Plant	Index	0.00005	0.00031	0.00015	0.00053	0.00002	0.00021	0.00001	0.00006
Transfer Station	Index	0.00001	0.00050	0.00010	0.00070	0.00001	0.00040	0.00001	0.00040
High Traffic Road	Index	0.00052	0.00275	0.00079	0.00374	0.00031	0.00192	0.00054	0.00262
Dist. Roch. CBD	Kilometers	19.31	14.24	39.40	11.79	15.31	4.36	8.19	2.54
Dist. NYS Thruway	Kilometers	16.58	9.07	18.99	13.19	13.79	7.79	17.92	4.47
Median HH Income	Dollars (000)	55.0	15.1	46.1	9.0	65.1	16.0	50.4	10.4
White Population	%	0.93	0.06	0.95	0.05	0.93	0.05	0.92	0.06
Public Sewer	Binary	0.79	0.41	0.59	0.49	0.78	0.42	0.96	0.19
Village	Binary	0.14	0.34	0.35	0.48	0.10	0.31	0.003	0.06
Pupil-Teacher Ratio	Ratio	13.4	1.8	12.6	1.7	13.9	1.7	13.5	1.6
School Performance	Test Score	180	9	177	7	182	9	181	8
School Tax Rate	Dollars	19.26	1.77	18.54	2.39	19.00	1.09	20.14	1.43
Structural Characteristics									
Sale Year	Year	9.11	2.56	9.14	2.60	9.09	2.55	9.12	2.55
Overall Condition	Scale (1-5)	3.05	0.337	3.06	0.155	3.07	3.18	3.02	0.203
Age	Years	44	32	56	44	33	26	46	22
Living Area	Sq Feet	1,694	578	1,626	576	1,840	600	1,579	514
Property Size	Acres	0.55	0.69	0.84	1.00	0.58	0.61	0.30	0.28
Number Bathrooms	#	1.7	0.7	1.6	0.6	1.9	0.7	1.6	0.6
Central AC	Binary	0.34	0.474	0.15	0.362	0.427	0.495	0.386	0.487

3.7 Model Estimation Results

This section presents the estimation results of the hedonic equation parameters, and the discussion proceeds according to the framework presented in Table 3.2. To begin, the hedonic parameters are estimated through ordinary least squares (OLS) regression of the full 66,609 observation dataset. The full dataset is then segmented into rural, suburban and urban submarkets and the separate hedonic equations are estimated using OLS in each of the three submarkets. The results of the full dataset OLS model for both the aggregate market area and each of the three submarkets are presented in Table 3.3, followed by a discussion of the estimated parameter results.

Table 3.2 Hedonic Estimation Framework

	OLS	**2SLS**
Full Dataset	Table 3.3	Table A.3.3
Sampled Dataset	Table A.3.4	Table 3.4

Next, the sampling procedure used to control for spatial autocorrelation and the two-stage least squares (2SLS) estimation method used to correct for endogeneity in the private open space variables are explained. The estimation results from the sampled 2SLS model are provided in Table 3.4, followed by a discussion of the implications from this model. For completeness, the results of the 2SLS estimation on the full sample and the OLS estimation on the sampled dataset are included in the appendix as Tables A.3.3 and A.3.4 respectively. Comparisons across all four sets of estimation results provide insights into the consistency of the estimation results across estimation techniques. Keeping with the objectives of this study, attention is focused primarily on the variables that relate to open space proximity and locational disamenities as well as the differences across submarkets.

Table 3.3 OLS Estimation Results from Full Dataset

	Aggregate (66,609)		Rural (17,759)		Suburban (26,196)		Urban (22,654)	
	$R^2 = 0.773$		$R^2 = 0.716$		$R^2 = 0.795$		$R^2 = 0.798$	
Variable	**Coefficient**	**Error**	**Coefficient**	**Error**	**Coefficient**	**Error**	**Coefficient**	**Error**
Intercept	9.36 *	0.03	9.36 *	0.09	9.79 *	0.05	9.41 *	0.04
Neighborhood: 0 - 400 Meters								
Municipal Park	-0.00046 **	0.00021	-0.00022	0.00079	-0.00043 ***	0.00023	-0.00032	0.00036
State/County Park	0.00128 *	0.00017	0.00019	0.00056	0.00131 *	0.00030	0.00110 *	0.00018
Agriculture	0.00183 *	0.00007	0.00209 *	0.00016	0.00150 *	0.00011	0.00133 *	0.00018
Developed OS	0.00160 *	0.00007	0.00238 *	0.00023	0.00092 *	0.00010	0.00182 *	0.00009
Private Forest	0.00213 *	0.00009	0.00237 *	0.00023	0.00119 *	0.00014	0.00208 *	0.00013
Water	0.00938 *	0.00015	0.01085 *	0.00033	0.00604 *	0.00032	0.00706 *	0.00021
Neighborhood: 400 - 1600 Met								
Municipal Park	0.000082 **	0.000031	-0.000260	0.000179	0.000025	0.000041	0.000094 **	0.000042
State/County Park	0.000002	0.000012	0.000024	0.000044	0.000081 *	0.000020	0.000181 *	0.000015
Agriculture	-0.000060 *	0.000006	-0.000130 *	0.000015	-0.000050 *	0.000012	0.000127 *	0.000017
Developed OS	0.000079 *	0.000008	0.000106 **	0.000036	0.000084 *	0.000013	0.000247 *	0.000011
Private Forest	-0.000020 ***	0.000009	-0.000130 *	0.000023	-0.000030 **	0.000015	0.000057 *	0.000014
Water	-0.000050 *	0.000010	-0.000040	0.000025	0.000008	0.000020	-0.000070 *	0.000013
Location Based								
CAFO	-11.56 *	1.92	-10.01 *	2.42	-	-	-	-
Wastewater Tmt	-12.10 *	2.87	-18.04 *	4.14	15.73 **	5.82	108.96 *	19.61
Transfer Station	-4.06 **	1.67	2.34	2.87	-7.49 **	2.76	-10.05 **	3.05
High Traffic Road	-1.09 *	0.32	-0.97 ***	0.56	-0.70	0.64	-2.89 *	0.46
Dist. CBD	-0.0035 *	0.0001	-0.0026 *	0.0003	-0.0041 *	0.0005	-0.0171 *	0.0016
Dist. NYS Thru	-0.0043 *	0.0001	-0.0063 *	0.0003	-0.0005 **	0.0002	0.0010	0.0006
Median Income	0.0026 *	0.0001	0.0029 *	0.0004	0.0021 *	0.0001	0.0032 *	0.0002
Percent White	0.328 *	0.018	0.251 *	0.045	0.218 *	0.038	0.068 **	0.023
Public Sewer	0.0098 **	0.0030	0.0206 **	0.0063	-0.0061	0.0040	0.0091	0.0070
Pupil-Teacher	0.0046 *	0.0006	0.0047 **	0.0016	0.0290 *	0.0012	-0.0111 *	0.0010
School Perf.	0.00233 *	0.00012	0.00314 *	0.00038	0.00233 *	0.00017	0.00160 *	0.00020
School Tax Rate	-0.0116 *	0.0006	-0.0189 *	0.0011	-0.0316 *	0.0015	0.0100 *	0.0015
Village	0.0813 *	0.0034	0.0412 *	0.0072	0.0838 *	0.0051	-0.0453 **	0.0206
Property Characteristics								
Sale Year	0.034 *	0.000	0.034 *	0.001	0.035 *	0.000	0.034 *	0.000
Overall Cond.	0.171 *	0.003	0.208 *	0.005	0.103 *	0.004	0.159 *	0.006
Age	-0.0026 *	0.0000	-0.0023 *	0.0001	-0.0029 *	0.0001	-0.0026 *	0.0001
Living Area	0.00034 *	0.00000	0.00033 *	0.00000	0.00034 *	0.00000	0.00035 *	0.00000
Property Size	0.0502 *	0.0016	0.0522 *	0.0026	0.0650 *	0.0025	0.0743 *	0.0049
Number Bath	0.1077 *	0.0021	0.1073 *	0.0047	0.0933 *	0.0029	0.0962 *	0.0030
Central AC	0.0667 *	0.0020	0.0790 *	0.0065	0.0653 *	0.0028	0.0530 *	0.0025

Note: *,**,*** Indicate significance at the 0.01, 0.05, 0.10 levels respectively.

OLS Results from Full Dataset

The OLS estimation results from the full dataset, presented in Table 3.3, reveal that proximity to open space is highly valued. With the exception of municipal parks, and county/state parks in rural areas, all of the open space variables within 400 meters have positive and significant implicit prices. Between 400 and 1600 meters all open space types, except for municipal parks, have a significantly smaller impact on residential prices than open space within 400 meters. Interpreting the coefficients on open space within 400-1600 meters is challenging, however, as it is not clear how to reconcile an open space measure that has a positive impact on home price within 400 meters and negative impact at the 400 to 1600 meter scale. It may be that home buyers have different preferences for open space acreage that is accessible to their home, but not necessarily adjacent to their property. While open space in the 400-1600 meter range is included in the model as a control variable, the primary focus of the discussion will be on the findings from open space coefficients within 400 meters of the home.

Proximity to open water is clearly the most highly valued open space in all three submarkets within a 400 meter radius, but the relative values of the remaining open space types are highly variable. In particular state/county parks have significantly higher implicit prices in suburban and urban areas than in rural areas, possibly reflecting preferences for accessible open space in the more urbanized neighborhoods. Conversely, private open space appears to be more highly valued in rural areas relative to urban areas. The coefficients on agricultural, developed open space and forested open space in rural areas are not statistically different and are higher than private open space types in either the suburban or urban markets.

Proximity to concentrated animal feeding operations has a significantly negative impact on home price. In appendix Table A.3.2, the hedonic model is estimated with separate measures for medium and large CAFOs. Parameter results from this model suggest that medium CAFOs have a significantly negative impact on home price, while the proximity of large CAFOs is negative, but not significantly different from zero. This result is similar to the findings in Herriges, Secchi and Babcock (2003), who hypothesize that larger animal agriculture facilities tend to be newer and better managed, potentially because they are also more heavily regulated, and therefore have less impact on nearby residential property. In this case, however, the finding that large CAFOs do not have a significant effect on home price may also be a result of the relatively small number of homes that are located within 1600 meters of a large CAFO (448; 0.6%) compared to a medium CAFO (1,188; 1.8%). In addition to the fact that there are fewer large than medium CAFOs, it could also be that fewer homes are affected by large CAFOs because they tend to account for a larger land area. Thus the location of the barn housing the animals is less likely to be within 1600 meters of residential property in the case of large CAFOs. A Wald test of the equality of medium and large CAFO coefficients, however, reveals no statistical difference between the influence of medium and large CAFOs (Wald Statistic = 0.21 which is less than the critical chi-squared value of 3.84 for the 0.05 level of significance) and thus in the model results presented in the text all CAFOs are aggregated into one variable.

The proximity indices representing wastewater treatment facilities, transfer stations and high traffic roads are significantly negative in aggregate, but reveal substantive differences across the three markets. Wastewater treatment facilities have a significantly negative impact in rural areas, while the impact is actually found to be significantly positive in suburban and rural areas. This result is similar to the findings of Ready and Abdalla (2005) who estimate a significantly positive influence of wastewater treatment plants in their OLS hedonic model. In the case of the Rochester area, a likely reason for the positive implicit price is the fact that wastewater treatment plants in the urban and suburban municipalities tend to be located along Lake Ontario, where waterfront homes tend to have high prices, while in rural areas they tend to be located along smaller, linear waterways. Proximity to transfer stations and high traffic roads follow the opposite trend, with the most negative localized impacts occurring in the more urban municipalities.

The census derived demographic statistics are relatively consistent across market areas and reveal positive implicit prices for homes located in high income, white neighborhoods. The indicator variables for access to public sewers and location in a village setting are both significantly positive in rural areas, suggesting a premium on local services, and insignificant and negative respectively in urban areas. Given that the overwhelming majority of homes in urban areas (96%) have access to public sewer, the insignificance of this variable is not surprising. The public school related variables also reveal some significant differences between markets. Although performance measures based on test scores are significantly positive across the three markets, the pupil-teacher ratio ranges from significantly negative in urban areas to positive in rural and suburban areas. One would expect, based on intuition and the results of Brasington (1999), that pupil-teacher ratio would have a negative implicit price (i.e., having fewer students per teacher is a benefit), but this is not the case in rural and suburban areas. This unexpected result could be a function of the fact that the pupil-teacher ratio measures an input to school quality that may be adjusted based on performance (i.e., the number of teachers is increased in lower performing schools in the rural and suburban areas). The variable measuring adjusted school tax rates follows a similar pattern, as increases in school tax rates have a negative influence on home price in rural and suburban areas and a positive impact in urban areas. The property specific parameters are significant in the predicted direction and generally stable across all markets. The exception to this is the finding that urban residents have higher implicit prices for lot size and interior square footage relative to rural residents. This result is likely driven by the constraints on property size in urban settings that are not are less restrictive in rural areas. Finally, the estimation results across the three markets reveal that the hedonic equation does a notably better job predicting variation in residential property prices in urban and suburban municipalities (R-sq = 0.798 and 0.795 respectively) compared to rural areas (R-sq = 0.716). This is not surprising given the fact that the rural transactions are distributed across a much larger geographic area and therefore rural prices are likely a function of a greater number of unobservable, location specific factors.

Overall, the estimation results suggest significant variation across the three housing markets that validate the estimation of three separate equations. A Chow test reveals that, in aggregate, the coefficient estimates from the three hedonic equations are significantly different.[4] A Tiao-Goldberger test is also performed, which tests the stability of individual coefficient estimates across the three submarkets. The test amounts to comparing the squared deviation of each regression coefficient from the weighted average coefficient across the three models, relative to the sum of squared errors. An application of the test to the estimation of hedonic functions across submarkets is provided in Michaels and Smith (1990). The results of the test are provided in appendix Table A.3.12 and show that all but 2 of the 33 coefficients (municipal parks within 400 meters and the year of sale) vary significantly across the three submarkets. This is strong evidence that the hedonic functions are indeed different across the three markets.

[4] The F-statistic is 150.7, which is significant at the 0.01 level

Two Stage Least Squares Estimation on Sampled Dataset

Previous studies have found quantities of private open space to be endogenous to residential property prices and have found evidence of spatial autocorrelation in model errors (e.g., Irwin and Bockstael 2001; Irwin 2002; Geoghegan, Lynch and Bucholtz 2003). A joint Hausman test reveals that endogeneity in the private open space measures is present in the Rochester study area as well.[5] Further, econometric tests reveal significant spatial error autocorrelation.[6] This subsection describes how endogeneity and spatial autocorrelation are controlled for in the hedonic estimation. These explanations are followed by a presentation and discussion of the 2SLS results on the sampled dataset.

To control for endogeneity, a two stage least squares (2SLS) estimation procedure is undertaken, with instruments provided for the private open space variables; developed open space, agricultural land and private forested open space. In separate first stage regressions, the amount of each of the three private open measures within 1600 meters of each home are included as the dependent variable in an OLS regression, with all of the instruments, described below, and all of the exogenous variables serving as independent variables. In the second step, the predicted open space values are inserted into the hedonic equation, which is estimated with a second OLS regression. Given that the hedonic equation includes open space measures within 400 meters and between 400 and 1600 meters, the predicted amount of open space within 1600 meters from the first stage equation is multiplied by the percentage of the

---

[5] The test statistics of 168.5 for the full dataset and 13.08 for the sampled dataset are greater than the critical value of 11.3, which represents a significance level of 0.01 of a chi-squared distribution with 3 degrees of freedom. This test is performed in two steps. First, residuals are obtained from first stage regression of each of the private open space variables on the set of instruments and all of the exogenous variables. Second, the residuals from each of the first stage regressions are inserted into the hedonic equation as independent variables. The test statistic then measures if the coefficients on the residuals are jointly significantly different from zero.

[6] The Lagrange Multiplier statistic is 13,828 which is distributed chi-square and significant at the 0.01 level. This statistic is obtained using Matlab's spatial toolbox. In addition, a spatial error model was calculated and the spatial parameter (lambda) was positive and highly significant.

open space variable that lies within 400 and the percentage of open space between 400 and 1600 meters. Thus, expanding equation (5), the hedonic equation is estimated as

$$\begin{aligned}\ln(P) = \alpha + \hat{A}_{1600} \cdot \left(\beta_{a4} \cdot a_4 \cdot + \beta_{a4-16} \cdot a_{4\_16}\right) + \hat{D}_{1600} \cdot \left(\beta_{d4} \cdot d_4 + \beta_{d4-16} \cdot d_{4\_16}\right) + \\ \hat{F}_{1600} \cdot \left(\beta_{f4} \cdot f_4 + \beta_{f4-16} \cdot f_{4\_16}\right) + M_{1600} \cdot \left(\beta_{m4} \cdot m_4 + \beta_{m4-16} \cdot m_{4\_16}\right) + \\ C_{1600} \cdot \left(\beta_{c4} \cdot c_4 + \beta_{c4-16} \cdot c_{4\_16}\right) + W_{1600} \cdot \left(\beta_{w4} \cdot w_4 + \beta_{w4-16} \cdot w_{4\_16}\right) + \\ \lambda\ L + \delta\ S + \varepsilon,\end{aligned} \tag{7}$$

where the upper case letters $\hat{A}_{1600}$, $\hat{D}_{1600}$, $\hat{F}_{1600}$, M_{1600}, C_{1600}, W_{1600} represent the total acreage within 1600 meters that is in agriculture, developed open space, forested, municipal park, county/state park, or open water respectively and the lower case letters represent the percentage of each land use type that is within 400 meters (e.g., a_4) and the percentage between 400 and 1600 meters (e.g., $a_{4\text{-}16}$). The "hats" on the *A*, *D,* and *F*, variables indicate that the predicted values of agricultural, developed open space and forest land from the first stage equation are used instead of the actual values. The implicit assumption is that the overall amount of each private open space type within 1600 meters is endogenous, but that the relative amount of open space within 400 meters and between 400 and 1600 meters is exogenous.

The instruments used in the first stage regressions are chosen based on the hypothesis that they are correlated with the amount of open space surrounding a home, but do not influence the sales price of the home. Specifically, variables are included that influence the direct costs and opportunity costs of development. The first instrumental variable is whether the residential parcel is located in an agricultural district. In New York State, agricultural districts are implemented to "forestall the conversion of farmland to non-agricultural uses," by allowing agricultural landowners to receive preferential tax treatment and protections against local zoning laws and nuisance lawsuits (NYS Ag and Markets 2007). Being in an agricultural district increases the opportunity costs of taking a parcel out of agricultural production and therefore it is expected that homes inside of agricultural districts have higher amounts of surrounding open space.

The next three instruments are all derived from a GIS coverage of soil characteristics developed by the USDA Natural Resource Conservation Service (NRCS), and include measures of building suitability, septic suitability and soil productivity. To calculate these variables, two steps were taken. First a 1600 meter lattice grid was overlaid onto the soil map[7] and an area weighted average was calculated for each of the three suitability measures. Next, the grid averages were linked to the residential transactions based on the 1600 meter grid cell in which the transaction is located.

The four instruments associated with soil characteristics and opportunity costs of conversion are augmented with two instrumental variables that measure proximity to transportation and other infrastructure, which tend to reduce overall construction costs. The variables included are the natural log of distance from Rochester and the natural log of distance to Thruway exists. Irwin (2002) similarly includes the log of distance to various metro areas as instruments and Geoghegan, Lynch and Bucholtz (2003) include the log of distance to the nearest transportation node.

The F-statistics from first stage regressions, where the three endogenous variables are regressed on the six instruments, fall comfortably above the threshold of 10 recommended by Staiger and Stock (1997). The first stage regression results from the sampled dataset, where the three private open space types are regressed on the instruments and all of the exogenous variables, are also included in the appendix in Table A.3.11.

---

[7] The Soil Survey Geographic (SSURGO) Database is used as the soil map in all counties with the exception of Livingston, Wayne and Yates counties, where SSURGO data are not available and the comparable, but slightly less refined State Soil Geographic (STATSGO) Database is used.

In addition to econometric issues associated with endogenous independent variables, spatial error autocorrelation is found to be statistically significant based on a Lagrange Multiplier test ($p < 0.01$). To control for the autocorrelation, the data are sampled so that no residential location in the final dataset has a neighbor within 100 meters. This sampled dataset is constructed by picking one sale at random from the set of 66,609 residential locations. Next, a second home is chosen at random from the remaining pool of 66,608 observations and if this home is not within 100 meters of the first home selected, then it is included in the final dataset. If it is within 100 meters of the previously chosen observation then it is thrown out. The sampling proceeds in this fashion until the entire dataset has been evaluated. In the resulting "no nearest neighbor" dataset that is used in the final set of regressions, there are 22,721 total observations.

It should be noted that the sample generated from this procedure is not unique. The transaction chosen first determines in part the inclusion of the next transaction chosen and so on. The probability that an individual residential sale is included in the sampled is, however, not random. For example, a sales point that does not have any neighbors within 100 meters in the full dataset will always be included in the reduced dataset. This also hints at a potential concern with standard hedonic models that estimate equations over heterogeneous neighborhoods. Clearly, if no sampling or weighting is undertaken, then model results will reflect the preferences of residential consumers in densely populated areas more so than rural buyers. Sampling alone does not solve this dilemma, but sampling in conjunction with the separate estimation of equations in rural, suburban and urban markets allows for a better understanding of the differences in implicit prices across different densities.

To test the robustness of the hedonic parameter estimates to particular sample draws, the sampling technique is repeated ten times. The ten samples are then used to generate the minimum and maximum coefficient estimate from each of the 2SLS model variables. The results of this robustness test from the aggregate market and each of the three submarkets and are presented in appendix Table A.3.8 and show that the estimated coefficients are generally robust to the particular sample draw. Among the coefficients that are significantly different from zero in the sampling draw featured in Table 3.4, both the minimum and maximum parameter estimate from the 10 draws are of the same sign.

The hedonic equation coefficients estimated with OLS on the sampled dataset are included in the appendix as Table A.3.4. The primary effect of the sampling procedure is an increase in the standard errors of the estimated coefficients. Across the 32 variables in the aggregate market model, the standard error in the sampled dataset is an average of 1.83 times higher than in the full dataset. Higher standard errors are a natural result of reducing the number of observations in the estimated equation. Theoretically, decreasing the number of observations from 66,609 to 22,712 should result in a $\sqrt{66609}/\sqrt{22712} = 1.71$ increase in the standard error. The estimated standard errors from the sampled dataset are therefore in line with the theoretical expectations. Further, only one of the coefficients that is significantly different from zero in the full dataset OLS model switches signs in the sampled OLS model. Given that 128 coefficients are estimated in total, the fact that only one of them switches sign and significance (open water acreage between 400 and 1600 meters) is strong evidence that coefficients are robust between the model estimated with the full dataset and the sampled dataset.

Table 3.4 2SLS Estimation Results from Sampled Dataset

	Aggregate (22,721) $R^2 = 0.7509$		Rural (8,507) $R^2 = 0.6739$		Suburban (9,050) $R^2 = 0.7652$		Urban (5,164) $R^2 = 0.7657$	
Variable	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**
Intercept	9.29 *	0.07	8.40 *	0.30	10.01 *	0.12	9.63 *	0.11
Neighborhood: 0 - 400 Meters								
Municipal Park	0.00170 **	0.00078	0.00225	0.00179	0.00210 **	0.00066	0.00167	0.00102
State/County Park	0.00332 *	0.00082	0.00132	0.00135	0.00381 *	0.00078	0.00261 *	0.00059
Agriculture	0.00609 *	0.00138	0.00366 **	0.00144	0.00503 *	0.00089	0.00776 *	0.00220
Developed OS	0.00453 *	0.00116	0.00413 **	0.00204	0.00394 *	0.00085	0.00210 *	0.00056
Private Forest	0.00539 *	0.00115	0.00472 **	0.00180	0.00503 *	0.00092	0.00079	0.00054
Water	0.01127 *	0.00087	0.01229 *	0.00159	0.00781 *	0.00078	0.00838 *	0.00068
Neighborhood: 400 - 1600 Meters								
Municipal Park	-0.000320 **	0.000110	0.000722 ***	0.000398	0.000067	0.000132	0.000372 **	0.000128
State/County Park	0.000008	0.000044	0.000445 *	0.000107	0.000222 *	0.000055	-0.000020	0.000068
Agriculture	-0.000240 **	0.000085	0.000256 **	0.000088	0.000045	0.000083	-0.000430 **	0.000191
Developed OS	0.000082	0.000060	0.001582 *	0.000363	0.000333 *	0.000076	0.000050	0.000075
Private Forest	-0.000170 **	0.000062	0.000254 **	0.000104	0.000143	0.000110	0.000342 *	0.000064
Water	-0.000020	0.000046	0.000375 *	0.000085	0.000292 *	0.000074	-0.000140 ***	0.000072
Location Based								
CAFO	-6.79 **	3.29	-14.63 *	4.43	-	-	-	-
Wastewater Tmt	1.08	6.68	-23.28 **	7.89	43.69 **	13.92	49.44	38.33
Transfer Station	-1.95	2.54	3.69	3.71	9.33	10.23	-9.06 ***	4.71
High Traffic Road	-0.64	0.62	-0.59	0.94	1.87	1.37	-3.03 **	1.04
Dist. CBD	-0.0041 *	0.0002	-0.0005	0.0008	-0.0057 **	0.0018	-0.0100	0.0130
Dist. NYS Thru	-0.0043 *	0.0003	-0.0088 *	0.0007	-0.0034 *	0.0008	-0.0030	0.0029
Median Income	0.0013 **	0.0004	0.0026 *	0.0007	0.0008 **	0.0003	0.0042 *	0.0005
Percent White	0.244 *	0.040	-0.016	0.080	-0.137	0.100	0.155 **	0.063
Public Sewer	0.0355 *	0.0078	0.0202	0.0153	0.0115	0.0110	0.0629 **	0.0193
Pupil-Teacher	0.0113 *	0.0014	0.0026	0.0045	0.0278 *	0.0026	-0.0138 *	0.0033
School Perf.	0.00197 *	0.00025	0.00363 *	0.00078	0.00011	0.00047	0.00189 *	0.00057
School Tax Rate	-0.0129 *	0.0011	-0.0105 *	0.0023	-0.0305 *	0.0034	0.0016	0.0048
Village	0.1324 *	0.0199	0.0296 ***	0.0172	0.1763 *	0.0193	-0.1203 **	0.0438
Property Characteristics								
Sale Year	0.034 *	0.001	0.034 *	0.001	0.036 *	0.001	0.032 *	0.001
Overall Cond.	0.188 *	0.005	0.218 *	0.007	0.122 *	0.008	0.121 *	0.013
Age	-0.0023 *	0.0001	-0.0021 *	0.0001	-0.0029 *	0.0001	-0.0027 *	0.0002
Living Area	0.00032 *	0.00000	0.00030 *	0.00001	0.00032 *	0.00001	0.00036 *	0.00001
Property Size	0.0468 *	0.0026	0.0536 *	0.0041	0.0621 *	0.0038	0.0501 *	0.0096
Number Bath	0.1085 *	0.0043	0.1030 *	0.0079	0.0911 *	0.0058	0.1030 *	0.0069
Central AC	0.0663 *	0.0045	0.0599 *	0.0144	0.0687 *	0.0056	0.0518 *	0.0060

Note: *,**,*** Indicate significance at the 0.01, 0.05, 0.10 levels respectively.

The coefficient estimates of the hedonic model from the 2SLS estimation and the sampled dataset are included as Table 3.4. Overall, the open space coefficient estimates are considerably higher under the sampled, 2SLS model than under the full OLS model. Inspection of the 2SLS model results from the full dataset reveals that the higher coefficient estimates are a result of correcting for endogeneity rather than sampling. This suggests that the OLS model generates downwardly biased open space estimates, which is intuitively reasonable. If a given area is endowed with locational amenities, for example close proximity to Rochester and Lake Ontario, then it will command a higher price than other areas all else being equal. Given the higher prices that consumers are willing to pay, this land is more likely to be developed. Thus, holding all else constant, high priced residential neighborhoods will have less open space. This relationship runs contrary to the expectation that higher amounts of open space are correlated with higher property values. By estimating the amount of open space in a given area in the first stage equation and then incorporating these predictions in the second stage equation, the result should be higher estimates of marginal implicit prices for open space.

The relative implicit prices for open space between the three markets remain generally unchanged. Public open space, especially larger state and county parks, have a higher value in urban and suburban areas as opposed to the rural market. Additionally, marginal implicit prices for all of the private open space types remain relatively high in the rural areas. Interestingly, between the full dataset OLS model in Table 3.3 and the sampled 2SLS model in Table 3.4, the relative implicit prices of agricultural land are reversed. In the 2SLS model, agricultural land is most highly valued in urban areas relative to rural areas, while the opposite is true in the OLS model. This suggests that endogeneity problems associated with agricultural open space are most acute in urban settings, where the majority of agricultural land has been developed (only 4% of the land area within 400 meters of a home in the urban area is classified as agricultural).

The coefficient estimates from the sampled 2SLS model indicate that in rural areas, proximity to CAFOs and wastewater treatment facilities has a significantly negative effect. In urban areas, transfer stations and high traffic roads have a significantly negative impact on residential property prices. These results are consistent with the results of the OLS model with the full dataset. Among the remaining location based variables, only the coefficient on the percentage of white residents in the census tract changes sign between the sampled 2SLS model and the full OLS model, though the coefficient is only significantly different from zero in the full OLS model. None of the property specific variables change sign between the two models.

Alternatively Defined Submarkets

To investigate the robustness of the estimation results to alternative submarket specifications, ancillary hedonic functions are estimated with the submarkets defined based on census boundaries and based on county boundaries. For the census based model, separate submarkets are determined based on the Rochester urban area, urban clusters and rural areas inside the general study area. These submarket definitions are defined by the U.S. Census and differ from the urban, rural and suburban submarkets analyzed above. The primary difference is that the urban clusters are composed of high population census blocks located in rural areas, for example village and town centers.

The hedonic estimates of the open space coefficients are generally similar between the primary urban, suburban and rural submarkets and the alternative census defined submarkets. Measures of public open space are significantly positive in the urban submarket under both specifications. A primary difference, however, is that the private open space coefficients are considerably higher in the urban cluster submarket as opposed to the rural submarket. Thus, the significantly positive open space coefficients in the rural submarket in the primary model may be driven in large part by the homes located in more densely populated villages and town centers. Additionally, the coefficients on the locational disamenities are similar across the census defined and primary submarkets, except for the fact that CAFOs in rural areas are not significantly different from zero, though still negative.

The county defined submarkets are segmented into Monroe County, which is the county that contains the City of Rochester, and homes in the study area located in the nine counties surrounding Rochester. The results from the county defined submarkets closely mirror those from the primary model. Specifically, public open space is significantly positive among homes in Monroe County, but not significantly different from zero in the surrounding counties. Further, the private open space measures corresponding to developed open space and private forests are similarly highly significant in both submarkets. The locational disamenities are also very similar between the county and primary submarket definitions. CAFOs and wastewater treatment facilities are significant disamenities in the rural counties, while high traffic roads are a significant inside Monroe County.

Overall, the majority of the hedonic coefficients are intuitively comparable across the three alternative sets of submarkets. Given that each set of submarkets is defined according to unique criteria, it is not surprising that differences exist in the estimated coefficients across the three sets. Where coefficients differ, the results can generally be explained based on the specific submarket boundary definition. Thus, the results suggest that the appropriate submarket definition is important in making inferences based on estimated implicit prices. Since submarkets can be defined in myriad ways, it is important for the researcher to understand that the choice of market boundaries will have an impact on the estimated hedonic coefficients. Of course, given the fact that 32 coefficients are estimated in each equation, differences in estimated coefficients are not surprising. The fact that the relative implicit prices are generally consistent across the different submarket deliniations is important evidence, however, of the systematic trends that occur within the market as a whole.

Predicted Marginal Values

The estimated open space coefficients in Table 3.4 indicate the predicted percentage change in home price resulting from a marginal change in open space. This subsection uses the coefficient estimates from Table 3.4 and the mean values of each variable to estimate the marginal values of open space and locational disamenities in dollar terms. Using the coefficient values to predict marginal values for open space and other localized externalities relies on the assumption that the change in open space under consideration is small relative to the entire market for open space. Therefore, a one acre change in each open space type is used to obtain the marginal value. If the change in open space being considered is large relative to the market area, then the housing market would adjust to reflect the new distribution of open space and relative prices would be altered. In the analysis that follows, the marginal value of open space is measured as the cost of a one acre loss in a given open space type.

The marginal values are calculated by multiplying the vector of estimated coefficients times the mean values of each of the model variables. The antilog of the sum of the products is the predicted mean home price. Next, one acre is subtracted from the mean acreage of a given open space type and the new predicted home price is obtained. The difference between the predicted home price based on the mean value of each variable and the predicted price with one less open space acre indicates the marginal value, or in this case the marginal cost, of losing an acre of open space to development. The estimated marginal values as well as the percentage of homes that have at least one acre of a given open space within either 400 meters or between 400 and 1600 meters are reported in Table 3.5. The marginal cost of each of the four locational disamenities (CAFO, wastewater treatment facility, transfer station, high traffic road) are determined in a similar fashion. The cost of adding one disamenity at distances of 400 meters and 1000 meters are estimated separately and the results are reported in Table 3.5, along with the percentage of homes within a proximity of 400 meters and 1000 meters of each disamenity. Since the proximity index for high traffic roads is measured only up to a distance of 400 meters, the impact of high traffic roads is estimated for a distance of 200 meters only.

Table 3.5 Marginal Cost of 1 Acre Open Space Loss or Addition of a Disamenity

	Aggregate		**Rural**		**Suburban**		**Urban**	
Total Observations	22,721		8,507		9,050		5,164	
	% Obs	Marg. Value	% Obs	Marg. Value	% Obs	Marg. Value	% Obs	Marg. Value
Neighborhood: 0 - 400 Meters								
Municipal Park	8.2%	-201.79	6.1%	-	11.8%	-301.56	5.5%	-
State/County Park	5.2%	-394.10	3.5%	-	3.9%	-548.42	10.1%	-295.91
Agriculture	84.6%	-720.98	97.1%	-364.05	89.4%	-722.12	55.6%	-878.19
Developed OS	99.9%	-537.33	99.9%	-411.48	100.0%	-566.59	100.0%	-238.74
Private Forest	94.0%	-638.38	97.9%	-469.90	96.3%	-722.51	83.3%	-
Neighborhood: 400 - 1600 Meters								
Municipal Park	44.1%	37.81	23.6%	-72.16	56.6%	-	56.2%	-42.26
State/County Park	22.2%	-	15.1%	-44.24	17.2%	-32.06	42.9%	-
Agriculture	98.9%	29.10	100.0%	-25.51	100.0%	-	95.2%	48.43
Developed OS	100.0%	-	100.0%	-157.52	100.0%	-48.22	100.0%	-
Private Forest	100.0%	19.76	100.0%	-25.25	100.0%	-	99.8%	-38.84
Located 400 Meters								
CAFO	0.5%	-1,504.02	1.3%	-2,697.37	-	-	-	-
Wastewater Tmt	0.6%	-	1.2%	-4,259.08	0.2%	12,289.00	-	-
Transfer Station	0.3%	-	0.5%	-	0.1%	-	0.3%	-1,915.27
High Traffic Road*	4.8%	-	6.0%	-	3.1%	-	5.9%	-859.98
Located 1000 Meters								
CAFO	3.5%	-302.34	8.2%	-545.41	-	-	-	-
Wastewater Tmt	8.9%	-	17.3%	-866.75	4.0%	2,378.00	3.7%	-
Transfer Station	1.8%	-	3.5%	-	0.8%	-	0.7%	-385.66

Note: Marginal values are calculated only for variables that are significantly different from zero in Table 3.4.

* The impact of high traffic roads is measured at 200 meters, since the maximum distance of influence is 400 meters.

Inspection of Table 3.5 reveals that, based on the results from the aggregate market, the marginal cost of converting one acre of land from open space to some developed use varies from \$201 to \$720 per residence within 400 meters. The cost of open space conversion is consistently higher in the suburban submarket than in the market as a whole. This result arises from higher home prices along with relatively high coefficients on the open space variables in the suburban hedonic equation. The highest cost for a one acre loss of open space is \$828 for agricultural land in urban areas. One could argue that this result arises from the relative scarcity of agricultural open space in urban settings, where agriculture makes up less than 4% of the land surrounding an average home. In rural areas, private open space within 400 meters that is lost to development has a marginal cost of between \$364 for agricultural land to \$470 for private forest land. Not surprisingly, the marginal cost of open space loss is considerably smaller in the 400-1600 meter range, reaching a maximum of \$157 for losses in developed open space in rural areas.

The marginal cost of the introduction of a locational disamenity is also highly variable both across markets and across disamenity types. In rural areas, having a CAFO located at a distance of 400 meters is predicted to depress the price of an average home by nearly \$2,700. The effect of CAFOs in other submarkets is not estimated, as CAFOs are found almost exclusively in rural areas. A wastewater treatment facility located at a distance of 400 meters is predicted to depress home price by more than \$4,000 in the rural submarket, while proximity to wastewater treatment facilities is actually positively correlated with the price of homes in suburban areas. This latter result, as mentioned previously, likely stems from wastewater treatment plants being located in desirable lakefront areas. Waste management facilities and high traffic roads have negative impacts on average home prices of \$1,915 and \$960 respectively in urban areas, but no significant impact in rural and suburban submarkets. Locational disamenities located 1000 meters from a residence have a roughly five times smaller impact on home price compared to the same disamenity located 400 meters from the home. This result implies that the influence of the disamenities considered in this analysis on home price diminish at an increasing rate with increases in distance.

3.8 Policy Implications and Lessons Learned

In this section, the policy implications of the estimated hedonic functions are discussed, followed by a brief summary of the lessons that have been learned in undertaking this analysis. The estimated hedonic equations provide predictions of how marginal changes in open space, location based characteristics, and property specific attributes influence home prices. These estimated marginal implicit prices represent a measure of the marginal willingness to pay of homebuyers in each market submarket for specific housing attributes, such as open space. Utilizing the mean values of housing characteristics and the estimated coefficients from the hedonic function, the marginal cost of open space loss per household is also calculated. This analysis could be taken a step further, by multiplying the marginal cost per home by the number of homes affected in various neighborhoods. This would then provide an estimate of the social costs of lost open space acreage in a particular location.

Care must be taken, however, in interpreting all of these results. First, the predicted marginal costs are based on changes in marginal open space quantities. If large open space areas are lost to development in a region, the supply curve for open space in a given area would shift and a new hedonic relationship would apply. Second, the social cost of changes in open space ultimately depends on what the open space land is converted to. For example, if a farm is subdivided into five acre lots for an upscale residential development, the resulting impact on nearby home prices will undoubtedly be very different from the case where the same farm is developed into a parking lot for a large retail store. Third, the social value of open space derived from a hedonic model is based solely on changes in property prices. Ceteris paribus, this necessarily implies that open space in more densely populated areas is of greater value than open space in sparsely populated areas with very little development. Just because a particular open space type does not influence property prices, however, does not mean that it has no value for society. For example, in this study there does not appear to be a preference among homebuyers for residential units in close proximity to municipal parks. This does not mean, however, that a municipal park is not valued by the general public. While the noise and congestion spillovers of a municipal park may not be desirable for adjacent properties, others in the larger community may benefit greatly from the recreational opportunities made available by the park. Indeed, there are a wide range of open space benefits that are not capitalized into home prices, such as the ecological benefits of large contiguous areas of natural habitat for wildlife. Finally, although the hedonic analysis illustrates costs of marginal losses of open space land, it does not consider the costs of open space preservation. These costs are likely to be highly variable across the study area and would need to be considered in the design of an optimal open space preservation program.

Numerous lessons have been learned over the course of conducting this study, which may be helpful to future researchers undertaking a similar analysis. The first lesson that became readily apparent is that large amounts of spatially explicit data can be both a blessing and a curse. While one could include literally hundreds of variables related to residential location and property characteristics, it is important to focus on variables of particular interest to the researcher and variables that serve as essential controls. Adding variables can make coefficient interpretation more difficult, similar to the attempts in this analysis to incorporate open space variables at two different ranges. This analysis employs data from a large, 10 county region and incorporates over 60,000 observations and close to 40 variables. Assuming that the data represent one distinct housing market seems overly restrictive; yet defining submarkets is also not completely straightforward and could be accomplished in myriad ways. While one might be tempted to incorporate a large number of submarkets, there are costs to segmenting markets. This first cost is obvious; more submarkets mean more equations, which inevitably mean more coefficients to interpret or more possibilities that at least some of the variables will have counterintuitive coefficient estimates. The second cost of delineating a plethora of housing markets is that this greatly reduces intra-market variation and decreases the number of observations in each market, both of which lead to higher standard errors on model coefficients.

Other considerations for future researchers revolve around estimation of the hedonic equation. To begin, it is important to have solid theoretical arguments for believing that a given variable is endogeneous. The test for endogeneity of a given variable involves comparing the coefficients of the hedonic equation, estimated with OLS, to the same equation estimated with 2SLS. Therefore, one must have an appropriate instrument just to test for an endogenous regressor. Thus it is essential to have a theoretical basis for presuming that a give variable is endogenous, to avoid an exhaustive search for instruments. Also, when estimating the hedonic equation for each submarket with 2SLS, the first stage regression is run separately in each submarket. This flexibility should allow for more accurate predictions of open space in each submarket, however, it also introduces another element of variation that may lead to greater variability in the estimated results. For example, it may not be possible to tell the degree to which observed results vary as a result of differences in the first stage prediction equation as opposed to the second stage hedonic estimation.

The sampling routine employed in this paper also requires some rather large assumptions. It is not immediately obvious what the optimal sampling distance should be. Irwin (2002) tested several sampling distances and found that while the majority of variables were robust to sampling at greater distances (i.e., so that no two observations are within 200, 400, and 600 meters respectively), some variables did in fact change sign. She therefore used the 100 meter subset as the basis of her discussion on the basis that it provided the most robust estimates. Sampling at greater distances necessarily results in fewer overall observations and a greater percentage of observations from less densely populated areas. Future researchers should understand these tradeoffs and investigate the empirical implications of different sampling schemes.

3.9 Conclusion

In conclusion, the estimation results presented in this paper illustrate the significant differences that exist in the marginal implicit prices for attributes of residential property across urban, suburban and rural areas near Rochester, NY. According to the estimated hedonic price equations, large public parks have a relatively high implicit price in urban and suburban areas, while private open spaces such as agricultural and private forested land is of consistently high value in rural locations. In addition, proximity to Concentrated Animal Feeding Operations (CAFOs) and wastewater treatment facilities have significantly negative impacts on housing prices in rural areas, while in urban areas proximity to high traffic roads and waste transfer stations significantly diminish residential property prices. Overall, the results do not support the assumption that open space always has the highest impact on home prices in urban as opposed to rural areas. The impact of open space on residential property values is clearly dependent both on the location and the type of open space being considered. Failure to account for these differences across open space types and market areas may lead policy makers to infer incorrect values of specific types of open space, if considering only aggregate data.

To validate the robustness of the results, several additional steps should be undertaken in future work. First, allowing parameter estimates to vary across a more incremental gradient of population densities than what is presented here would allow for a comparison of the value of obtaining potentially more accurate estimates of implicit prices in each of the submarket delineations, against the loss in efficiency caused by lower variation within each submarket. Second, future research will extend this analysis to included housing markets in Binghamton and Albany, NY, which will allow for comparisons of implicit prices both within and between housing markets in Upstate New York.

APPENDIX

A.3.1 Names of Municipalities Included in Dataset

Market	Number of Transactions	Municipality Names				
Urban	22,654	Brighton	Gates	Greece	Irondequoit	
		Chili	Clarkson	East Roc.	Hamlin	Henrietta
Suburban	26,196	Mendon	Ogden	Parma	Penfield	Perinton
		Pittsford	Riga	Rush	Webster	Wheatland
		Albion	Avon	Barre	Batavia	Bergen
		Bethany	Bristol	Byron	Caledonia	Canadice
		Canandaigua	Carlton	Clarendon	Conesus	Dansville
		East Bloomfield	Elba	Farmington	Gaines	Genesee Falls
		Geneseo	Gorham	Grove	Groveland	Hopewell
		Huron	Kendall	Le Roy	Leicester	Lima
Rural	17,759	Livonia	Macedon	Manchester	Marion	Middlesex
		Mount Morris	Murray	Naples	North Dansville	Nunda
		Oakfield	Ontario	Ossian	Palmyra	Pavilion
		Perry	Portage	Richmond	Sodus	South Bristol
		Sparta	Springwater	Stafford	Sweden	Victor
		Walworth	Wayland	West Bloomfield	West Sparta	Williamson
		Yates	York			

Table A.3.2 OLS Estimation Results with Medium and Large CAFOs

	Aggregate (66,609) $R^2 = 0.773$	
Variable	Coefficient	Error
Intercept	9.36 *	0.03
Neighborhood: 0 - 400 Meters		
Municipal Park	-0.00046 **	0.00021
State/County Park	0.00128 *	0.00017
Agriculture	0.00183 *	0.00007
Developed OS	0.00160 *	0.00007
Private Forest	0.00213 *	0.00009
Water	0.00938 *	0.00015
Neighborhood: 400 - 1600 Meters		
Municipal Park	0.000082 **	0.000031
State/County Park	0.000002	0.000012
Agriculture	-0.000060 *	0.000006
Developed OS	0.000079 *	0.000008
Private Forest	-0.000020 ***	0.000009
Water	-0.000050 *	0.000010
Location Based		
Medium CAFO	-12.01 *	2.15
Large CAFO	-8.62	6.78
Wastewater Tmt	-12.12 *	2.87
Transfer Station	-4.06 **	1.67
High Traffic Road	-1.09 *	0.32
Dist. CBD	-0.0035 *	0.0001
Dist. NYS Thru	-0.0043 *	0.0001
Median Income	0.0026 *	0.0001
Percent White	0.327 *	0.018
Public Sewer	0.0098 **	0.0030
Pupil-Teacher	0.0046 *	0.0006
School Perf.	0.00233 *	0.00012
School Tax Rate	-0.0116 *	0.0006
Village	0.0813 *	0.0034
Property Characteristics		
Sale Year	0.034 *	0.000
Overall Cond.	0.171 *	0.003
Age	-0.0026 *	0.0000
Living Area	0.00034 *	0.00000
Property Size	0.0502 *	0.0016
Number Bath	0.1077 *	0.0021
Central AC	0.0667 *	0.0020

A.3.3 2SLS Estimation Results from Full Dataset

	Aggregate (66,609)		Rural (17,759)		Suburban (26,196)		Urban (22,654)	
	$R^2 = 0.756$		$R^2 = 0.687$		$R^2 = 0.777$		$R^2 = 0.746$	
Variable	**Coefficient**	**Error**	**Coefficient**	**Error**	**Coefficient**	**Error**	**Coefficient**	**Error**
Intercept	9.23 *	0.03	8.69 *	0.16	10.21 *	0.06	9.55 *	0.06
Neighborhood: 0 - 400 Meters								
Municipal Park	0.00112 *	0.00033	0.00296 **	0.00110	0.00160 *	0.00032	0.00240 *	0.00053
State/County Park	0.00301 *	0.00034	0.00252 **	0.00092	0.00414 *	0.00041	0.00256 *	0.00030
Agriculture	0.00605 *	0.00065	0.00500 *	0.00089	0.00495 *	0.00042	0.01206 *	0.00156
Developed OS	0.00467 *	0.00045	0.00577 *	0.00120	0.00419 *	0.00039	0.00227 *	0.00025
Private Forest	0.00500 *	0.00043	0.00650 *	0.00109	0.00482 *	0.00042	0.00049 ***	0.00028
Water	0.01098 *	0.00033	0.01415 *	0.00097	0.00749 *	0.00039	0.00872 *	0.00037
Neighborhood: 400 - 1600 Meters								
Municipal Park	-0.000320 *	0.000049	0.000236	0.000266	0.000069	0.000082	0.000412 *	0.000056
State/County Park	-0.000020	0.000022	0.000054	0.000073	0.000144 *	0.000032	0.000013	0.000027
Agriculture	-0.000320 *	0.000043	-0.000090	0.000061	-0.000040	0.000050	-0.000470 *	0.000095
Developed OS	0.000021	0.000029	0.000978 *	0.000151	0.000205 *	0.000047	0.000077 **	0.000029
Private Forest	-0.000220 *	0.000025	-0.000190 **	0.000068	0.000029	0.000073	0.000353 *	0.000038
Water	-0.000120 *	0.000020	0.000006	0.000055	0.000187 *	0.000045	-0.000150 *	0.000028
Location Based								
CAFO	-6.89 **	2.14	-10.58 *	2.84	-	-	-	-
Wastewater Tmt	7.36 ***	3.77	-22.20 *	4.89	39.83 *	6.85	130.13 *	22.64
Transfer Station	-1.71	1.75	9.26 **	3.46	-3.61	2.92	-12.85 *	3.39
High Traffic Road	-0.82 **	0.33	-0.21	0.62	1.19 ***	0.70	-2.38 *	0.53
Dist. CBD	-0.0025 *	0.0001	-0.0004	0.0005	-0.0048 *	0.0012	-0.0226 *	0.0059
Dist. NYS Thruway	-0.0039 *	0.0002	-0.0078 *	0.0004	-0.0018 *	0.0005	-0.0028 **	0.0014
Median Income	0.0015 *	0.0002	0.0030 *	0.0005	0.0010 *	0.0002	0.0042 *	0.0002
Percent White	0.278 *	0.020	0.117 **	0.055	-0.244 *	0.056	0.190 *	0.029
Public Sewer	0.0161 **	0.0051	0.0030	0.0090	-0.0026	0.0064	0.0804 *	0.0113
Pupil-Teacher Ratio	0.0075 *	0.0007	-0.0061 ***	0.0035	0.0318 *	0.0014	-0.0111 *	0.0015
School Performance	0.00260 *	0.00013	0.00469 *	0.00057	0.00054 **	0.00024	0.00122 *	0.00025
School Tax Rate	-0.0103 *	0.0006	-0.0133 *	0.0017	-0.0349 *	0.0018	0.0024	0.0028
Village	0.1189 *	0.0076	0.0298 **	0.0097	0.1644 *	0.0090	-0.0632 **	0.0230
Property Characteristics								
Sale Year	0.034 *	0.000	0.035 *	0.001	0.035 *	0.001	0.033 *	0.001
Overall Condition	0.181 *	0.003	0.208 *	0.005	0.111 *	0.005	0.161 *	0.007
Age	-0.0025 *	0.0000	-0.0021 *	0.0001	-0.0030 *	0.0001	-0.0023 *	0.0001
Living Area	0.00033 *	0.00000	0.00032 *	0.00001	0.00034 *	0.00000	0.00036 *	0.00000
Property Size	0.0458 *	0.0022	0.0563 *	0.0029	0.0641 *	0.0027	0.0463 *	0.0072
Number Bathrooms	0.0993 *	0.0024	0.0971 *	0.0056	0.0851 *	0.0031	0.0953 *	0.0034
Central AC	0.0615 *	0.0022	0.0604 *	0.0081	0.0624 *	0.0029	0.0544 *	0.0028

A.3.4 OLS Estimation Results from Sampled Dataset

Variable	Aggregate (22,721) $R^2 = 0.766$ Coefficient	Error	Rural (8,507) $R^2 = 0.7088$ Coefficient	Error	Suburban (9,050) $R^2 = 0.7888$ Coefficient	Error	Urban (5,164) $R^2 = 0.7923$ Coefficient	Error
Intercept	9.43 *	0.05	9.30 *	0.13	9.64 *	0.09	9.53 *	0.08
Neighborhood: 0 - 400 Meters								
Municipal Park	-0.00040	0.00040	-0.00041	0.00134	-0.00042	0.00043	-0.00025	0.00076
State/County Park	0.00097 **	0.00031	-0.00033	0.00079	0.00049	0.00047	0.00154 *	0.00036
Agriculture	0.00153 *	0.00014	0.00144 *	0.00025	0.00109 *	0.00020	0.00146 *	0.00034
Developed OS	0.00104 *	0.00014	0.00139 *	0.00039	0.00033 ***	0.00020	0.00181 *	0.00020
Private Forest	0.00165 *	0.00017	0.00132 *	0.00034	0.00084 *	0.00025	0.00208 *	0.00027
Water	0.00875 *	0.00028	0.00952 *	0.00049	0.00544 *	0.00054	0.00734 *	0.00044
Neighborhood: 400 - 1600 Meters								
Municipal Park	-0.000030	0.000064	-0.000380	0.000271	-0.000090	0.000076	-0.000050	0.000097
State/County Park	0.000050 **	0.000023	0.000113 ***	0.000062	0.000115 *	0.000034	0.000167 *	0.000031
Agriculture	-0.000010	0.000012	-0.000050 **	0.000025	-0.000030	0.000022	0.000108 **	0.000035
Developed OS	0.000128 *	0.000016	0.000272 *	0.000058	0.000108 *	0.000024	0.000291 *	0.000026
Private Forest	0.000016	0.000016	-0.000030	0.000034	-0.000020	0.000028	0.000081 **	0.000030
Water	0.000055 **	0.000020	0.000088 **	0.000038	0.000056	0.000036	-0.000008	0.000030
Location Based								
CAFO	-11.17 *	3.32	-9.50 **	3.94	-	-	-	-
Wastewater Tmt	-10.85 **	5.19	-17.97 **	6.79	17.92	11.66	38.74	34.17
Transfer Station	-2.87	2.43	-0.27	3.24	-8.11	9.45	-6.69	4.39
High Traffic Road	-1.48 **	0.55	-1.42 ***	0.81	-0.74	1.23	-3.48 *	0.91
Dist. CBD	-0.0042 *	0.0002	-0.0026 *	0.0005	-0.0033 *	0.0008	-0.0200 *	0.0031
Dist. NYS Thru	-0.0045 *	0.0002	-0.0062 *	0.0004	-0.0010 **	0.0004	0.0025 **	0.0012
Median Income	0.0027 *	0.0002	0.0041 *	0.0006	0.0023 *	0.0002	0.0033 *	0.0004
Percent White	0.318 *	0.033	0.138 **	0.065	0.310 *	0.070	0.037	0.051
Public Sewer	0.0198 *	0.0051	0.0294 **	0.0097	0.0034	0.0069	0.0040	0.0121
Pupil-Teacher	0.0082 *	0.0011	0.0056 **	0.0023	0.0252 *	0.0023	-0.0089 *	0.0021
School Perf.	0.00177 *	0.00022	0.00289 *	0.00052	0.00245 *	0.00033	0.00206 *	0.00043
School Tax Rate	-0.0138 *	0.0010	-0.0148 *	0.0016	-0.0282 *	0.0027	0.0040	0.0031
Village	0.0702 *	0.0064	0.0145	0.0121	0.0711 *	0.0101	-0.0974 **	0.0406
Property Characteristics								
Sale Year	0.034 *	0.001	0.034 *	0.001	0.036 *	0.001	0.033 *	0.001
Overall Cond.	0.182 *	0.005	0.222 *	0.007	0.113 *	0.007	0.122 *	0.012
Age	-0.0025 *	0.0001	-0.0022 *	0.0001	-0.0029 *	0.0001	-0.0028 *	0.0001
Living Area	0.00033 *	0.00000	0.00031 *	0.00001	0.00033 *	0.00001	0.00035 *	0.00001
Property Size	0.0492 *	0.0021	0.0501 *	0.0031	0.0605 *	0.0034	0.0656 *	0.0072
Number Bath	0.1158 *	0.0037	0.1099 *	0.0068	0.1008 *	0.0053	0.1029 *	0.0063
Central AC	0.0709 *	0.0041	0.0847 *	0.0102	0.0730 *	0.0053	0.0486 *	0.0056

Table A.3.5 Census Defined Market Areas: OLS Results on Sampled Data

	Rural (8,444) $R^2 = 0.7323$		Urban Cluster (2,062) $R^2 = 0.7094$		Urban Area (12,213) $R^2 = 0.7998$	
Variable	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**
Intercept	9.26 *	0.13	9.16 *	0.34	9.44 *	0.06
Neighborhood: 0 - 400 Meters						
Municipal Park	-0.00062	0.00135	0.00053	0.00157	0.00002	0.00038
State/County Park	-0.00006	0.00064	-0.00044	0.00177	0.00103 **	0.00034
Agriculture	0.00093 **	0.00028	0.00292 *	0.00047	0.00167 *	0.00018
Developed OS	0.00050	0.00047	0.00376 *	0.00063	0.00107 *	0.00014
Private Forest	0.00124 *	0.00038	0.00313 *	0.00067	0.00170 *	0.00019
Water	0.00911 *	0.00055	0.01226 *	0.00096	0.00602 *	0.00036
Neighborhood: 400 - 1600 Meters						
Municipal Park	-0.000220	0.000208	-0.001820 **	0.000633	0.000156 **	0.000058
State/County Park	0.000222 *	0.000051	-0.000450 *	0.000129	0.000054 **	0.000025
Agriculture	0.000079 **	0.000030	-0.000340 *	0.000051	-0.000004	0.000016
Developed OS	0.000228 *	0.000057	-0.000110	0.000110	0.000160 *	0.000016
Private Forest	0.000075 **	0.000037	-0.000360 *	0.000078	0.000083 *	0.000020
Water	0.000140 *	0.000042	-0.000110	0.000092	-0.000004	0.000023
Location Based						
CAFO	-1.60	4.75	-3.95	7.62	-44.95	71.45
Wastewater Tmt	-8.48	9.13	-30.22 **	11.67	-15.87	10.54
Transfer Station	0.30	3.56	8.12	9.71	-3.29	6.19
High Traffic Road	-1.26	1.01	0.84	1.34	-1.60 **	0.76
Dist. CBD	-0.0043 *	0.0004	-0.0063 *	0.0012	-0.0037 *	0.0006
Dist. NYS Thru	-0.0050 *	0.0004	-0.0027 **	0.0009	-0.0015 *	0.0003
Median Income	0.0032 *	0.0004	0.0037 *	0.0010	0.0024 *	0.0002
Percent White	0.272 *	0.068	0.395 *	0.118	0.226 *	0.042
Public Sewer	0.0378 *	0.0099	-0.0031	0.0233	-0.0042	0.0063
Pupil-Teacher	0.0144 *	0.0023	0.0107 **	0.0047	0.0006	0.0014
School Perf.	0.00149 **	0.00051	0.00369 **	0.00147	0.00228 *	0.00024
School Tax Rate	-0.0124 *	0.0017	-0.0122 **	0.0044	-0.0009	0.0017
Village	0.0356 **	0.0152	0.0766 *	0.0229	0.0593 *	0.0089
Property Characteristics						
Sale Year	0.032 *	0.001	0.032 *	0.002	0.036 *	0.001
Overall Cond.	0.213 *	0.007	0.228 *	0.016	0.104 *	0.007
Age	-0.0023 *	0.0001	-0.0026 *	0.0002	-0.0028 *	0.0001
Living Area	0.00032 *	0.00001	0.00031 *	0.00001	0.00034 *	0.00000
Property Size	0.0503 *	0.0029	0.0559 *	0.0120	0.0622 *	0.0044
Number Bath	0.1060 *	0.0068	0.1287 *	0.0140	0.1062 *	0.0044
Central AC	0.0930 *	0.0098	0.0898 *	0.0199	0.0635 *	0.0041

A.3.6 Census Defined Market Areas: 2SLS results on sampled data

Variable	Rural (8,444) $R^2 = 0.7148$ Coefficient	S.E.	Urban Cluster (2,062) $R^2 = 0.6286$ Coefficient	S.E.	Urban Area (12,213) $R^2 = 0.7905$ Coefficient	S.E.
Intercept	8.54 *	0.32	7.26 *	0.79	9.53 *	0.06
Neighborhood: 0 - 400 Meters						
Municipal Park	-0.00019	0.00158	0.00053	0.00216	0.00156 **	0.00057
State/County Park	0.00040	0.00098	0.00182	0.00254	0.00210 *	0.00049
Agriculture	0.00143	0.00103	0.00822 *	0.00196	0.00544 *	0.00107
Developed OS	0.00071	0.00156	0.01096 *	0.00254	0.00287 *	0.00053
Private Forest	0.00255 ***	0.00134	0.00971 *	0.00219	0.00290 *	0.00050
Water	0.01040 *	0.00128	0.01722 *	0.00187	0.00676 *	0.00047
Neighborhood: 400 - 1600 Meters						
Municipal Park	0.000124	0.000238	-0.003080 **	0.001270	-0.000030	0.000082
State/County Park	0.000725 *	0.000178	-0.000700 *	0.000205	-0.000050	0.000035
Agriculture	0.000558 *	0.000157	-0.000920 *	0.000185	-0.000360 *	0.000072
Developed OS	0.001080 **	0.000337	-0.000080	0.000378	-0.000010	0.000040
Private Forest	0.000664 *	0.000185	-0.000310	0.000225	-0.000050	0.000053
Water	0.000558 *	0.000143	-0.000410 **	0.000188	-0.000110 *	0.000030
Location Based						
CAFO	-1.92	4.95	7.09	11.37	1.68	73.97
Wastewater Tmt	-3.58	10.14	-23.33 ***	13.93	-23.93 **	11.86
Transfer Station	3.27	3.81	36.58 **	16.26	-4.59	6.38
High Traffic Road	-0.41	1.09	1.53	1.54	-1.51 ***	0.79
Dist. CBD	-0.0040 *	0.0004	-0.0046 **	0.0023	0.0010	0.0014
Dist. NYS Thru	-0.0077 *	0.0009	-0.0023	0.0020	-0.0013 **	0.0005
Median Income	0.0016 **	0.0007	0.0078 *	0.0021	0.0028 *	0.0002
Percent White	0.119	0.084	0.293 ***	0.166	0.152 **	0.050
Public Sewer	0.0528 *	0.0128	-0.0883 **	0.0407	-0.0057	0.0097
Pupil-Teacher	0.0163 *	0.0025	0.0089	0.0058	-0.0033 **	0.0017
School Perf.	0.00135 **	0.00054	0.01288 *	0.00254	0.00230 *	0.00025
School Tax Rate	-0.0095 *	0.0020	-0.0090	0.0120	0.0005	0.0018
Village	0.1001 *	0.0303	0.1765 *	0.0419	0.0874 *	0.0127
Property Characteristics						
Sale Year	0.033 *	0.001	0.029 *	0.003	0.036 *	0.001
Overall Cond.	0.202 *	0.008	0.238 *	0.018	0.108 *	0.007
Age	-0.0022 *	0.0001	-0.0020 *	0.0003	-0.0027 *	0.0001
Living Area	0.00032 *	0.00001	0.00032 *	0.00002	0.00034 *	0.00000
Property Size	0.0473 *	0.0034	0.0485 *	0.0141	0.0566 *	0.0056
Number Bath	0.1051 *	0.0072	0.1163 *	0.0163	0.1043 *	0.0046
Central AC	0.0877 *	0.0108	0.0736 **	0.0248	0.0623 *	0.0042

A.3.7 County Defined Market Areas: OLS Results from Sampled Dataset

	Outside Monroe (8,294) $R^2 = 0.7097$		Monroe County (14,425) $R^2 = 0.7867$	
Variable	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**
Intercept	9.33 *	0.13	9.41 *	0.06
Neighborhood: 0 - 400 Meters				
Municipal Park	-0.00062	0.00135	-0.00002	0.00038
State/County Park	-0.00058	0.00080	0.00130 *	0.00031
Agriculture	0.00149 *	0.00026	0.00143 *	0.00016
Developed OS	0.00164 *	0.00041	0.00106 *	0.00014
Private Forest	0.00145 *	0.00035	0.00185 *	0.00018
Water	0.01014 *	0.00050	0.00610 *	0.00036
Neighborhood: 400 - 1600 Meters				
Municipal Park	-0.000520 ***	0.000275	0.000152 **	0.000060
State/County Park	0.000087	0.000065	0.000077 *	0.000023
Agriculture	-0.000080 **	0.000026	0.000022	0.000016
Developed OS	0.000208 *	0.000060	0.000171 *	0.000016
Private Forest	-0.000080 **	0.000035	0.000055 **	0.000019
Water	0.000025	0.000039	0.000017	0.000023
Location Based				
CAFO	-5.13	4.03	-30.94	32.93
Wastewater Tmt	-24.93 *	7.22	14.97	10.85
Transfer Station	8.65	7.46	-1.31	2.55
High Traffic Road	-0.09	0.86	-1.68 **	0.75
Dist. CBD	-0.0033 *	0.0005	-0.0058 *	0.0006
Dist. NYS Thru	-0.0056 *	0.0004	-0.0013 *	0.0003
Median Income	0.0042 *	0.0006	0.0027 *	0.0002
Percent White	0.186 **	0.066	0.242 *	0.042
Public Sewer	0.0229 **	0.0099	-0.0084	0.0059
Pupil-Teacher	0.0063 **	0.0024	0.0021	0.0015
School Perf.	0.00261 *	0.00053	0.00222 *	0.00024
School Tax Rate	-0.0138 *	0.0016	-0.0023	0.0017
Village	0.0235 ***	0.0125	0.0779 *	0.0088
Property Characteristics				
Sale Year	0.034 *	0.001	0.035 *	0.001
Overall Cond.	0.221 *	0.007	0.115 *	0.006
Age	-0.0022 *	0.0001	-0.0028 *	0.0001
Living Area	0.00031 *	0.00001	0.00034 *	0.00000
Property Size	0.0510 *	0.0032	0.0596 *	0.0030
Number Bath	0.1099 *	0.0069	0.1062 *	0.0042
Central AC	0.0932 *	0.0106	0.0673 *	0.0041

A.3.8 County Defined Market Areas: 2SLS Results from Sampled Dataset

	Outside Monroe (8,294) $R^2 = 0.6837$		Monroe County (14,425) $R^2 = 0.7786$	
Variable	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**
Intercept	8.53 *	0.26	9.47 *	0.06
Neighborhood: 0 - 400 Meters				
Municipal Park	0.00158	0.00168	0.00129 **	0.00053
State/County Park	0.00061	0.00122	0.00259 *	0.00049
Agriculture	0.00284 **	0.00114	0.00480 *	0.00086
Developed OS	0.00322 ***	0.00166	0.00311 *	0.00054
Private Forest	0.00394 **	0.00146	0.00365 *	0.00056
Water	0.01205 *	0.00129	0.00724 *	0.00052
Neighborhood: 400 - 1600 Meters				
Municipal Park	0.000445	0.000389	-0.000060	0.000088
State/County Park	0.000394 *	0.000106	-0.000020	0.000035
Agriculture	0.000238 **	0.000086	-0.000280 *	0.000061
Developed OS	0.001431 *	0.000317	0.000029	0.000041
Private Forest	0.000206 **	0.000101	-0.000090 ***	0.000052
Water	0.000318 *	0.000085	-0.000080 **	0.000031
Location Based				
CAFO	-11.08 **	4.46	-28.05	33.88
Wastewater Tmt	-32.09 *	8.06	25.48 **	11.92
Transfer Station	17.43 **	8.82	-1.89	2.61
High Traffic Road	0.34	0.93	-1.69 **	0.77
Dist. CBD	-0.0014 **	0.0007	-0.0023 ***	0.0012
Dist. NYS Thru	-0.0079 *	0.0007	-0.0011 **	0.0004
Median Income	0.0027 *	0.0007	0.0028 *	0.0002
Percent White	0.045	0.079	0.156 **	0.049
Public Sewer	0.0151	0.0135	-0.0110	0.0102
Pupil-Teacher	0.0033	0.0037	-0.0007	0.0016
School Perf.	0.00337 *	0.00070	0.00226 *	0.00025
School Tax Rate	-0.0097 *	0.0022	-0.0009	0.0019
Village	0.0277	0.0176	0.1178 *	0.0136
Property Characteristics				
Sale Year	0.034 *	0.001	0.035 *	0.001
Overall Cond.	0.220 *	0.007	0.120 *	0.007
Age	-0.0021 *	0.0001	-0.0027 *	0.0001
Living Area	0.00031 *	0.00001	0.00033 *	0.00000
Property Size	0.0540 *	0.0039	0.0594 *	0.0035
Number Bath	0.1043 *	0.0076	0.1037 *	0.0044
Central AC	0.0694 *	0.0139	0.0662 *	0.0041

Table A.3.9 Coefficient range from 10 sample draws: aggregate and rural

	Aggregate				Rural			
Variable	Coefficient	S.E.	Min	Max	Coefficient	S.E.	Min	Max
Intercept	9.29 *	0.07	9.15	9.34	8.40 *	0.30	8.01	8.43
Neighborhood: 0 - 400 Meters								
Municipal Park	0.00170 **	0.00078	0.00110	0.00238	0.00225	0.00179	0.00177	0.00474
State/County Park	0.00332 *	0.00082	0.00261	0.00414	0.00132	0.00135	0.00083	0.00213
Agriculture	0.00609 *	0.00138	0.00480	0.00766	0.00366 **	0.00144	0.00267	0.00496
Developed OS	0.00453 *	0.00116	0.00341	0.00578	0.00413 **	0.00204	0.00271	0.00555
Private Forest	0.00539 *	0.00115	0.00427	0.00698	0.00472 **	0.00180	0.00356	0.00635
Water	0.01127 *	0.00087	0.01090	0.01290	0.01229 *	0.00159	0.01147	0.01399
Neighborhood: 400 - 1600 Meters								
Municipal Park	-0.000320 **	0.000110	-0.000320	-0.000120	0.000722 ***	0.000398	0.000580	0.001290
State/County Park	0.000008	0.000044	-0.000040	0.000017	0.000445 *	0.000107	0.000399	0.000590
Agriculture	-0.000240 **	0.000085	-0.000330	-0.000190	0.000256 **	0.000088	0.000244	0.000358
Developed OS	0.000082	0.000060	0.000028	0.000100	0.001582 *	0.000363	0.001502	0.002000
Private Forest	-0.000170 **	0.000062	-0.000250	-0.000120	0.000254 **	0.000104	0.000211	0.000353
Water	-0.000020	0.000046	-0.000090	-0.000020	0.000375 *	0.000085	0.000327	0.000454
Location Based								
CAFO	-6.79 **	3.29	-6.86	-1.70	-14.63 *	4.43	-16.31	-10.21
Wastewater Tmt	1.08	6.68	-8.25	2.06	-23.28 **	7.89	-32.62	-23.22
Transfer Station	-1.95	2.54	-3.26	1.34	3.69	3.71	4.26	6.41
High Traffic Road	-0.64	0.62	-1.00	0.18	-0.59	0.94	-0.72	0.53
Dist. CBD	-0.0041 *	0.0002	-0.0044	-0.0041	-0.0005	0.0008	-0.0008	0.0004
Dist. NYS Thru	-0.0043 *	0.0003	-0.0044	-0.0042	-0.0088 *	0.0007	-0.0096	-0.0085
Median Income	0.0013 **	0.0004	0.0009	0.0020	0.0026 *	0.0007	0.0020	0.0031
Percent White	0.244 *	0.040	0.266	0.315	-0.016	0.080	-0.016	0.071
Public Sewer	0.0355 *	0.0078	0.0320	0.0440	0.0202	0.0153	0.0058	0.0278
Pupil-Teacher	0.0113 *	0.0014	0.0096	0.0112	0.0026	0.0045	-0.0014	0.0023
School Perf.	0.00197 *	0.00025	0.00169	0.00221	0.00363 *	0.00078	0.00359	0.00413
School Tax Rate	-0.0129 *	0.0011	-0.0133	-0.0120	-0.0105 *	0.0023	-0.0108	-0.0077
Village	0.1324 *	0.0199	0.1147	0.1524	0.0296 ***	0.0172	0.0129	0.0420
Property Characteristics								
Sale Year	0.034 *	0.001	0.033	0.035	0.034 *	0.001	0.033	0.035
Overall Cond.	0.188 *	0.005	0.183	0.190	0.218 *	0.007	0.212	0.222
Age	-0.0023 *	0.0001	-0.0023	-0.0023	-0.0021 *	0.0001	-0.0022	-0.0020
Living Area	0.00032 *	0.00000	0.00031	0.00032	0.00030 *	0.00001	0.00029	0.00031
Property Size	0.0468 *	0.0026	0.0444	0.0484	0.0536 *	0.0041	0.0530	0.0573
Number Bath	0.1085 *	0.0043	0.1041	0.1170	0.1030 *	0.0079	0.1012	0.1117
Central AC	0.0663 *	0.0045	0.0642	0.0716	0.0599 *	0.0144	0.0360	0.0691

Table A.3.10 Coefficient range from 10 sample draws: suburban and urban

	Suburban				Urban			
Variable	**Coefficient**	**S.E.**	**Min**	**Max**	**Coefficient**	**S.E.**	**Min**	**Max**
Intercept	10.01 *	0.12	9.87	10.08	9.63 *	0.11	9.44	9.61
Neighborhood: 0 - 400 Meters								
Municipal Park	0.00210 **	0.00066	0.00159	0.00236	0.00167	0.00102	0.00138	0.00391
State/County Park	0.00381 *	0.00078	0.00354	0.00448	0.00261 *	0.00059	0.00202	0.00361
Agriculture	0.00503 *	0.00089	0.00416	0.00524	0.00776 *	0.00220	0.00679	0.01326
Developed OS	0.00394 *	0.00085	0.00326	0.00435	0.00210 *	0.00056	0.00222	0.00351
Private Forest	0.00503 *	0.00092	0.00442	0.00564	0.00079	0.00054	0.00063	0.00205
Water	0.00781 *	0.00078	0.00748	0.00916	0.00838 *	0.00068	0.00834	0.01015
Neighborhood: 400 - 1600 Meters								
Municipal Park	0.000067	0.000132	0.000022	0.000157	0.000372 **	0.000128	0.000333	0.000610
State/County Park	0.000222 *	0.000055	0.000166	0.000202	-0.000020	0.000068	-0.000110	0.000005
Agriculture	0.000045	0.000083	0.000037	0.000110	-0.000430 **	0.000191	-0.000810	-0.000420
Developed OS	0.000333 *	0.000076	0.000292	0.000359	0.000050	0.000075	-0.000060	0.000096
Private Forest	0.000143	0.000110	0.000075	0.000214	0.000342 *	0.000064	0.000275	0.000429
Water	0.000292 *	0.000074	0.000213	0.000296	-0.000140 ***	0.000072	-0.000220	-0.000060
Location Based								
CAFO	-	-			-	-		
Wastewater Tmt	43.69 **	13.92	31.37	64.57	49.44	38.33	48.39	88.20
Transfer Station	9.33	10.23	-6.85	23.27	-9.06 ***	4.71	-19.60	-9.31
High Traffic Road	1.87	1.37	-0.51	1.75	-3.03 **	1.04	-3.80	-2.72
Dist. CBD	-0.0057 **	0.0018	-0.0077	-0.0053	-0.0100	0.0130	-0.0367	-0.0036
Dist. NYS Thru	-0.0034 *	0.0008	-0.0033	-0.0024	-0.0030	0.0029	-0.0058	0.0023
Median Income	0.0008 **	0.0003	0.0007	0.0012	0.0042 *	0.0005	0.0033	0.0045
Percent White	-0.137	0.100	-0.111	0.022	0.155 **	0.063	0.095	0.217
Public Sewer	0.0115	0.0110	0.0039	0.0162	0.0629 **	0.0193	0.0370	0.1039
Pupil-Teacher	0.0278 *	0.0026	0.0245	0.0299	-0.0138 *	0.0033	-0.0159	-0.0084
School Perf.	0.00011	0.00047	-0.00028	0.00048	0.00189 *	0.00057	0.00181	0.00259
School Tax Rate	-0.0305 *	0.0034	-0.0345	-0.0265	0.0016	0.0048	-0.0086	0.0049
Village	0.1763 *	0.0193	0.1560	0.1759	-0.1203 **	0.0438	-0.1175	-0.0476
Property Characteristics								
Sale Year	0.036 *	0.001	0.035	0.037	0.032 *	0.001	0.032	0.035
Overall Cond.	0.122 *	0.008	0.109	0.120	0.121 *	0.013	0.126	0.152
Age	-0.0029 *	0.0001	-0.0029	-0.0028	-0.0027 *	0.0002	-0.0028	-0.0025
Living Area	0.00032 *	0.00001	0.00032	0.00033	0.00036 *	0.00001	0.00033	0.00036
Property Size	0.0621 *	0.0038	0.0580	0.0635	0.0501 *	0.0096	0.0336	0.0542
Number Bath	0.0911 *	0.0058	0.0887	0.0964	0.1030 *	0.0069	0.1005	0.1183
Central AC	0.0687 *	0.0056	0.0691	0.0770	0.0518 *	0.0060	0.0499	0.0594

Table A.3.11 First stage regression results

	Agricultural Land		Developed Open Space		Forested Open Space	
	R^2 = 0.714		R^2 = 0.594		R^2 = 0.460	
	F-Stat of Instruments=1939		F-Stat of Instruments= 442.9		F-Stat of Instruments=861.7	
	Coefficient	**S.E.**	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**
Intercept	-7289.53 *	143.64	1530.39 *	116.59	-1087.90 *	95.17
In Ag District	208.67 *	6.54	-93.93 *	5.31	-62.93 *	4.33
Building Suitability	134.90 *	5.36	-94.85 *	4.35	-78.33 *	3.55
Septic Suitability	-56.43 *	7.70	58.84 *	6.25	8.25	5.10
Soil Productivity	8.14 *	0.23	1.93 *	0.19	-7.89 *	0.15
ln_dist_roch	726.52 *	11.45	-323.80 *	9.30	380.55 *	7.59
ln_thru_dist	37.97 *	6.77	87.67 *	5.50	-60.73 *	4.49
Municipal Park_4	-2.29 *	0.43	0.86 **	0.35	-0.56 **	0.28
State/County Park_4	-0.06	0.32	0.58 **	0.26	-0.55 **	0.21
Water_4	-0.18	0.29	0.44 ***	0.24	-0.73 *	0.19
Municipal Park_4_16	-2.39 *	0.07	1.83 *	0.05	-0.79 *	0.04
State/County Park_4_16	-0.17 *	0.02	-0.42 *	0.02	-0.13 *	0.02
Water_4_16	-0.52 *	0.02	-0.19 *	0.02	-0.18 *	0.01
Pct. Ag_4_16	-79.86 *	16.62	142.48 *	13.49	85.94 *	11.01
Pct. Dev_o_4_16	-589.99 *	48.78	565.73 *	39.59	-323.15 *	32.32
Pct. Lg. for_4_16	-59.78	42.16	392.29 *	34.22	-375.39 *	27.93
CAFO	17770.00 *	3571.85	-3146.53	2899.32	-677.98	2366.54
Wastewater Tmt	85713.00 *	5558.66	-54933.00 *	4512.03	-3478.34	3682.90
Transfer Station	10527.00 *	2602.18	-10469.00 *	2112.23	-1010.17	1724.08
High Traffic Road	-1301.64 **	588.41	7.68	477.62	-191.80	389.85
Dist. CBD	-15.84 *	0.51	8.36 *	0.42	-14.29 *	0.34
Dist. NYS Thru	-5.60 *	0.53	-7.03 *	0.43	13.83 *	0.35
Median Income	-0.67 *	0.16	4.65 *	0.13	2.38 *	0.11
Percent White	227.67 *	37.67	117.20 *	30.58	-211.72 *	24.96
Public Sewer	-224.10 *	5.13	162.53 *	4.17	-38.28 *	3.40
Pupil-Teacher	-17.90 *	1.19	-4.70 *	0.97	-5.95 *	0.79
School Perf.	3.66 *	0.24	0.31	0.20	-0.47 **	0.16
School Tax Rate	17.73 *	1.09	-12.32 *	0.89	-6.69 *	0.72
Village	-90.09 *	6.90	7.44	5.60	-79.45 *	4.57
Sale Year	2.58 *	0.69	-3.76 *	0.56	1.29 **	0.46
Overall Cond.	25.52 *	4.87	-37.11 *	3.96	15.09 *	3.23
Age	-0.39 *	0.06	0.28 *	0.05	-0.09 **	0.04
Living Area	0.01 **	0.00	0.01 *	0.00	-0.02 *	0.00
Property Size	4.27 ***	2.26	-23.09 *	1.84	25.36 *	1.50
Number Bath	-10.26 **	3.97	22.46 *	3.22	0.43	2.63
Central AC	-10.97 **	4.40	20.94 *	3.57	-2.03	2.91

A.3.12. Tiao-Goldberger F-Statistics

Variable	Full Dataset OLS	Sampled Dataset 2SLS
Intercept	18.23 *	28.46 *
Neighborhood: 0 - 400 Meters		
Municipal Park	0.06	0.09
State/County Park	2.37 **	3.91 *
Agriculture	8.76 *	39.06 *
Developed OS	146.15 *	53.50 *
Private Forest	23.64 *	70.64 *
Water	88.04 *	19.70 *
Neighborhood: 400 - 1600 Meters		
Municipal Park	2.93 **	2.15 **
State/County Park	12.43 *	14.44 *
Agriculture	64.79 *	15.37 *
Developed OS	78.77 *	26.07 *
Private Forest	33.56 *	1.58
Water	4.17 *	24.63 *
Location Based		
CAFO	-	-
Wastewater Tmt	23.67 *	9.75 *
Transfer Station	5.37 *	2.08 **
High Traffic Road	4.44 *	3.21 **
Dist. CBD	38.32 *	5.01 *
Dist. NYS Thru	175.85 *	17.79 *
Median Income	12.05 *	15.72 *
Percent White	9.56 *	3.04 **
Public Sewer	8.23 *	2.39 **
Pupil-Teacher	253.87 *	51.88 *
School Perf.	8.29 *	12.62 *
School Tax Rate	161.20 *	17.78 *
Village	23.99 *	33.23 *
Property Characteristics		
Sale Year	1.15	2.77 **
Overall Cond.	152.61 *	46.70 *
Age	31.03 *	17.90 *
Living Area	10.40 *	11.16 *
Property Size	11.43 *	1.45
Number Bath	4.40 *	1.07
Central AC	10.49 *	1.49

Note: *,**,*** Indicate significance at the 0.01, 0.05, 0.10 levels respectively.

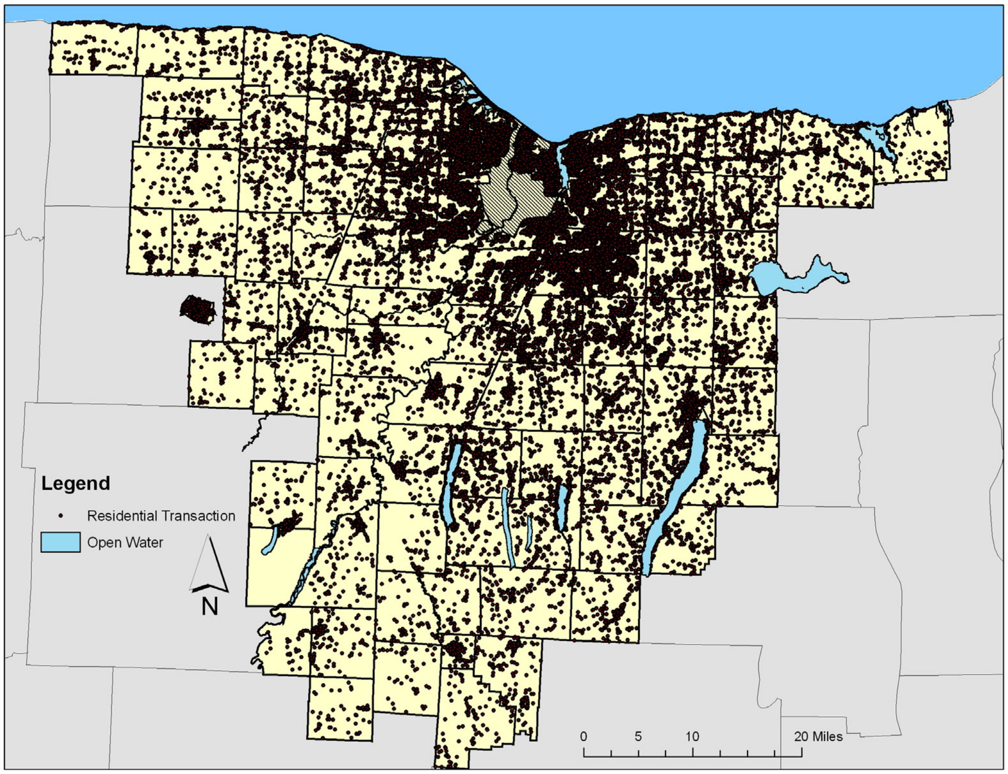


Figure A.3.1. Spatial Distribution of Residential Transactions

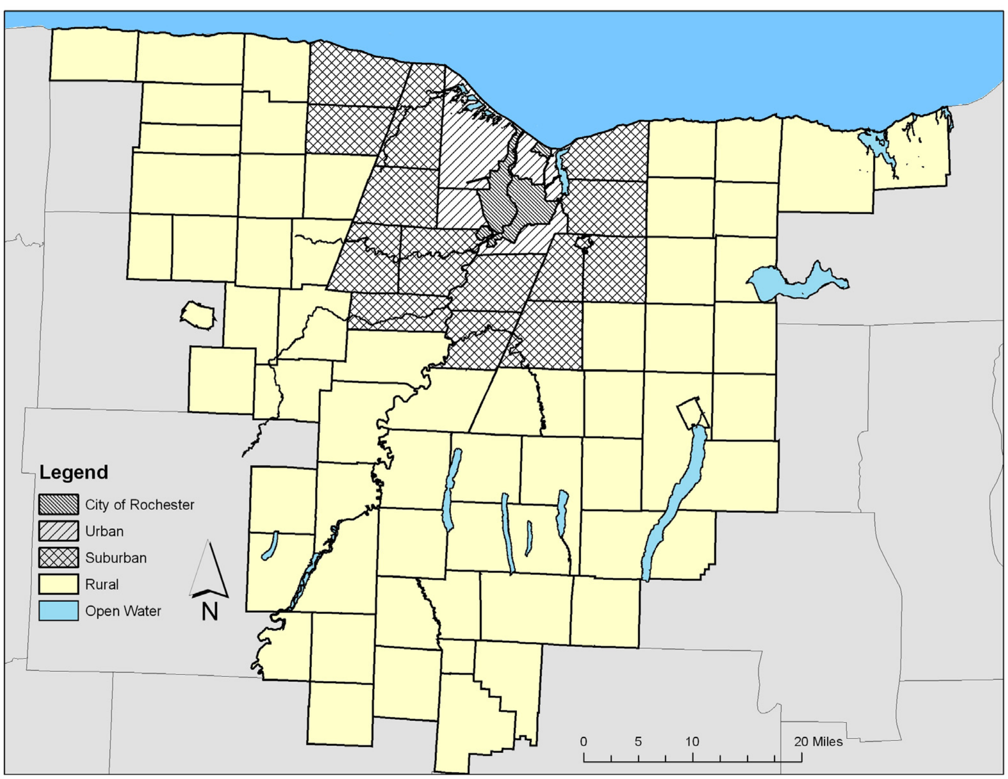


Figure A.3.2. Urban, Suburban and Rural Municipalities in the Study Area

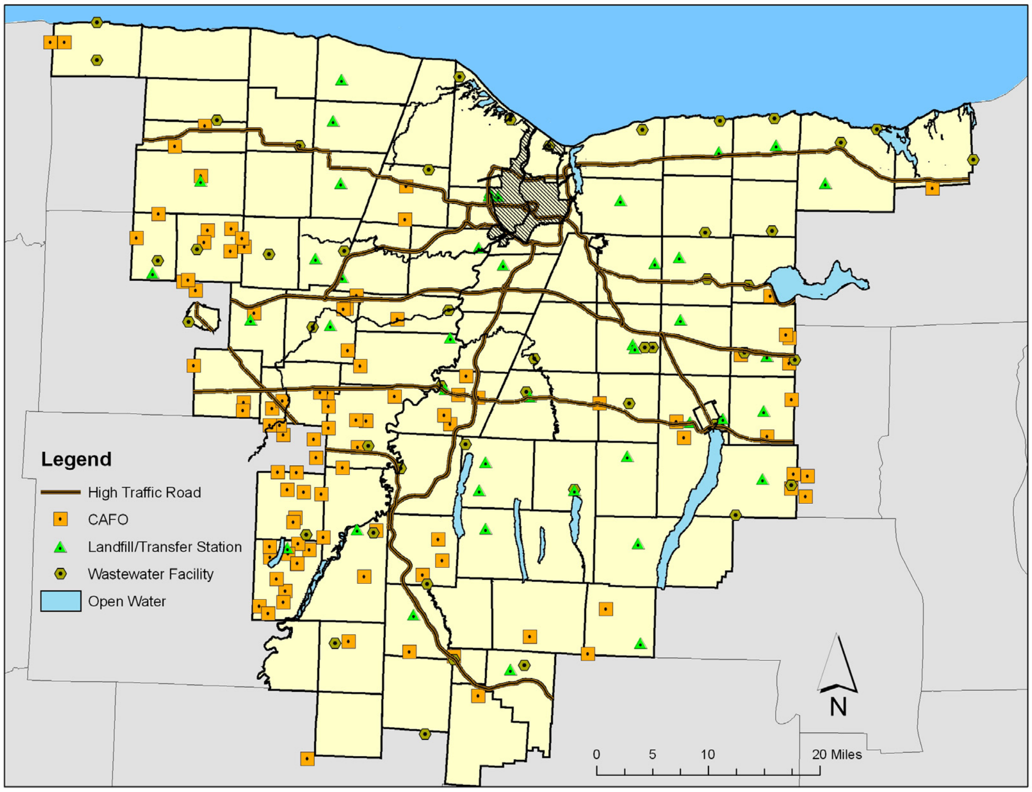


Figure A.3.3. Spatial Distribution of Potential Disamenities

REFERENCES

Anderson, S.T., and S.E. West. 2006. "Open space, residential property values and spatial context." *Regional Science and Urban Economics*. 36: 773-789.

Anselin, L. 2002. "Under the hood: Issues in the specification and interpretation of spatial regression models." *Agricultural Economics*. 17: 247-267.

Archarya, B, and L.L. Bennett. 2001. "Valuing Land Use and Open Space Patterns in Urban Watersheds." *Journal of Real Estate Finance and Economics*. 23(2/3): 221-237.

Basu, S. and Thibodeau, T.G. 1998. "Analysis of spatial autocorrelation in house prices." *Journal of Real Estate Finance and Economics*. 17: 61-86.

Batemen, I., B. Day, I. Lake, and A. Lovett. 2001. "The Effect of Road Traffic on Residential Property Values: A Literature Review and Hedonic Pricing Study." Norwich: University of East Anglia. Economic and Social Research Council. University College London.

Bourassa, S. C., M. Hoesli, and V. Peng. 2003. "Do Housing Submarkets Really Matter?" *Journal of Housing Economics*. 12: 12-28.

Brasington, D. M. 1999. "Which measures of school quality does the housing market value? Spatial and non-spatial evidence." *Journal of Real Estate Research*, 18(3): 95-413.

Brasington, D.M. and D. Hite. 2005. "Demand for Environmental Quality: a Spatial Hedonic Analysis." *Regional Science and Urban Economics*. 35(1): 57-82.

Cheshire, P and S. Sheppard. 1995. "On the Price of Land and the Value of Amenities." *Economica*. 62(246): 247-267.

Clapp, J.M., A. Nanda, and S.L. Ross. 2005. "Which School Attributes Matter? The Influence of School District Performance and Demographic Composition on Property Values," Working papers 2005-26, University of Connecticut, Department of Economic.

Cropper, M.L., L. Deck and K.E. McConnell. 1988. "On the Choice of Functional Forms for Hedonic Price Functions". *Review of Economics and Statistics.* 70: 668-675.

Do, A.Q., and G. Grudnitski 1995. Golf courses and residential house prices: an empirical examination. *Journal of Real Estate Finance and Economics*. 10 (3), 261–270.

Espey, M., and K. Owusu-Edusei. 2001. "Neighborhood Parks and Residential Property Values in Greenville, South Carolina." *Journal of Agricultural and Applied Economics*. 33(3): 487-492.

Freeman, A. M. 1993. The Measurement of Environmental and Resource Values: Theory and Methods. Resources for the Future. Washington, DC.

Geoghegan, J., L. Lynch, S.J. Bucholtz. 2003. "Are Agricultural Land Preservation Programs Self-Financing?" *Agricultural and Resource Economics Review*, 32(1): 33-45.

Geoghegan, J., L. A. Wainger, and N. Bockstael. 1997. "Spatial Landscape Indices in a Hedonic Framework. An Ecological Economics Analysis Using GIS." *Ecological Economics*, 23 (3) 251-264.

Herriges, J., S. Secchi, and B.A. Babcock. 2005. "Living with Hogs in Iowa: The Impact of Livestock Facilities on Rural Residential Property Values." *Land Economics*, 81 (4): 530-545.

Hite, D., W. Chern, F. Hitzusen, and A. Randall. 2001. "Property-Value Impacts of an Environmental Disamenity: The Case of Landfills." *Journal of Real Estate Finance and Economics*. 22(2/3): 185-202.

Irwin, E.G, and N. Bockstael. 2001. "The Problem of Identifying Land Use Spillovers: Measuring the Effects of Open Space on Residential Property Values." *American Journal of Agricultural Economics*. 83(1): 698-704.

Irwin, E.G. 2002. "The Effects of Open Space on Residential Property Values." *Land Economics*, 78 (4): 465 – 480.

Land Trust Alliance. 2006. "Private Land Conservation in U.S. Soars." Available: http://www.lta.org/newsroom/pr_113006.htm.

Lutzenhiser M., and N.R. Netusil. 2001. "The Effect of Open Space's on a Home's Sale Price." *Contemporary Economic Policy*. 19(3): 291-298.

Kim C.W., T.T. Phipps, and Anselin L. 2003. "Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach." *Journal of Environmental Economics and Management*, 45: 24-39.

Klapp, J.M, A. Nanda, and S.L. Ross. 2007. "Which School Attributes Matter? The Influence of School District Performance and Demographic Composition on Property Values." University of Connecticut. Department of Economics Working Paper Series.

Mahan, B., S. Polasky, and R. Adams. 2000. "Valuing Urban Wetlands: A Property Price Approach." *Land Economics*, 76 (1): 100-113.

McConnell, V. and M. Walls. "The Value of Open Space: Evidence from Studies of Non-Market Benefits," RFF Report. January 2005.

Michael, H.J, K.J. Boyle, and R. Bouchard. 2000. "Does the Measurement of Environmental Quality Affect Implicit Prices Estimated from Hedonic Models." *Land Economics*. 76(2): 284-298.

Michaels, R.G., and V.K. Smith. 1990. "Market Segmentation and Valuing Household Amenities: The Case of Hazardous Waste Sites." *Journal of Urban Economics*. 28(2): 223-242.

New York State Department of Agriculture and Markets. 2002. "Commissioner Announces Dairy Farm Transition Survey Results." Available: http://www.agmkt.state.ny.us/AD/release.asp?ReleaseID=1111.

New York State Department of Agriculture and Markets. 2007. "Agricultural Districts." Available: http://www.agmkt.state.ny.us/AP/agservices/agdistricts.html

Netusil, N.R. 2005. "The Effect of Environmental Zoning and Amenities on Property Values: Portland, Oregon." *Land Economics*. 81(2): 227-246.

Palmquist, R.B. 1992. "Valuing localized externalities." *Journal of Urban Economics*, 31 (1): 59–68.

Palmquist, R.B., F. Roka, and T. Vukina. 1997. "Hog Operations, Environmental Effects and Residential Property Values." *Land Economics*. 73(1): 114-124.

Paterson, R.W., K.J. Boyle. 2002. "Out of Sight Out of Mind? Using GIS to Incorporate Visibility in Hedonic Property Value Models." *Land Economics*. 78(3): 417-425.

Ready, R., C. Abdalla. 2005. "The Amenity and Disamenity Impacts of Agriculture: Estimates from a Hedonic Pricing Model." *American Journal of Agricultural Economics*, 87 (2): 314-326.

Rosen, S. 1974. "Hedonic prices and implicit markets: Product differentiation in pure competition." *Journal of Political Economy*, 82(1): 34-55.

Smith, V.K., C. Poulos, H. Kim. 2002. Treating open space as an urban amenity. Resource and Energy Economics 24 (1–2), 107–129.

Staiger, D., and J.H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica*, 65: 557–586.

Tyrväinen, L., A. Miettinen. 2000. Property prices and urban forest amenities. *Journal of Environmental Economics and Management.* 39(2): 205–223.

Taylor, L.O. "The Hedonic Method." In *A Primer on Nonmarket Valuation*. 2003. Kluwer Academic Publishers. Boston, MA.

U.S. Department of Agriculture. National Agricultural Statistics Service. 2002. 2002 Census of Agriculture - State Data. Available: http://www.nass.usda.gov/Census of_Agriculture/.