Incentive-compatible pollution control policies under asymmetric information on both risk preferences and technology

Jeffrey M. Peterson and Richard N. Boisvert

Incentive-Compatible Pollution Control Policies under Asymmetric Information on Both Risk Preferences and Technology

by

Jeffrey M. Peterson

and

Richard N. Boisvert

May 2003

Send correspondence to:
   Jeffrey M. Peterson
   Department of Agricultural Economics
   331B Waters Hall
   Kansas State University
   Manhattan, KS 66506-4011
   Phone: 785-532-4487
   Fax: 785-532-6925
   Email: jpeters@ksu.edu

Incentive-Compatible Pollution Control Policies under Asymmetric Information on Both Risk Preferences and Technology

Jeffrey M. Peterson and Richard N. Boisvert

This article proposes a method to accommodate asymmetric information on farmers' risk preferences in designing voluntary environmental policies. By incorporating stochastic efficiency rules in a mechanism design problem, the government can find incentive-compatible policies by knowing only the general class of risk preferences among farmers. The model also accounts for hidden information on technology types and input use. The method is applied empirically to simulate a pollution control program in New York. Results suggest that participation incentives would be inadequate for many risk-averse producers if the government does not account for the diversity in risk preferences.

Keywords: asymmetric information, mechanism design, nonpoint pollution, risk preferences

Incentive-Compatible Pollution Control Policies under Asymmetric Information on Both Risk Preferences and Technology

Agricultural nonpoint source pollution remains a significant policy challenge, with obstacles that stem in large part from asymmetric information. The ability to assess producers' responses to policies "... requires empirical knowledge of the production function, the environmental impacts of input use, and the risk attitudes of producers...." (Leathers and Quiggin, p. 763). The challenge is to devise policies that function effectively when the government's information on these attributes is limited.

One recently studied approach to finding such policies is based on the principles of mechanism design (e.g., Wu and Babcock, 1995, 1996; Peterson and Boisvert, 2001a,b). In cases where pollution and commodity outputs depend on input levels as well as a farmer's privately known "technology type," these authors propose voluntary policies designed to give farmers the incentive to truthfully reveal their private information. The government is aware of the different technology types but cannot (or chooses not to) match them to individual farmers. Instead, the government designs a policy "menu" consisting of (input level, government payment) pairs, and the policy problem is to design this menu so that farmers of different types self-select the appropriate policies. Feasible policies in this problem must satisfy two kinds of constraints, which ensure that farmers of each type would benefit by: (i) participating in the program and (ii) selecting the policy designed for their type over other policies.

Although promising, these proposals address only the hidden information on technology types. In practice, a farmer's incentive to truthfully reveal information on technology depends on other attributes that are also hidden from the government, such as risk preferences. Leathers and Quiggin caution that agricultural commodity and environmental policies may well have unintended consequences unless one knows something quite specific about a producer's utility

function. For example, policy responses depend on whether producers exhibit decreasing or increasing absolute risk aversion (DARA and IARA, respectively), so that changes in inputs and environmental quality can only be predicted if we know "…how many farmers are DARA? how many are IARA? how many are 'extremely' DARA? and how many are only 'slightly' DARA?" (Leathers and Quiggin, p. 673). Even then, the results are only qualitative unless a specific functional form for utility is assumed. More recently, other authors have underscored the fact that risk preferences are likely to have a significant empirical impact on input use and policy responses (Isik; Roosen and Hennessy). All these authors are silent on what can be done when obtaining such detailed information about risk preferences is impossible or prohibitively costly.

In the broader mechanism design literature, unknown preferences have been dealt with in two major ways. The first approach is to characterize the mechanism with a utility function whose form is left unspecified (e.g., Townsend and Mueller). While this method can be used to derive general properties of a mechanism, its limitation is that it cannot be applied to compute a mechanism in practice. The other approach is to assume a particular functional form for utility. Under certain conditions, this approach does lead to computable mechanisms (Rochet and Stole), allowing it to serve as the basis for empirical applications of mechanism design; naturally, the danger is that empirical results would be misleading because preferences are misspecified.

This article proposes a new method to accommodate limited information on agent preferences in policy design. The government's policy goal is for farmers' to truthfully reveal their privately known technology type when risk preferences are treated as an additional piece of hidden information. By incorporating stochastic efficiency criteria in a mechanism design problem, the government need only know the general class of preferences among agents (e.g., money-loving and risk-averse) and the support of their income distributions. If a solution to the

government's problem exists, it ensures that all agents in the chosen class will truthfully reveal their private information on technology. Although we apply this method to pollution control policies, it is equally applicable to other mechanism design problems where decision makers are faced with substantial risk, and where it is believed that many or all of them are risk averse (e.g., designing insurance or credit institutions).

By appealing to stochastic efficiency criteria, we address the specific concerns of Leathers and Quiggin and others mentioned above. Our mechanism leads to policies that "get it right" for producers in any given range of risk aversion. We are assured that the policies provide the appropriate incentives to change inputs and environmental quality in the anticipated and desired directions. It is well understood that these policies are likely to be more costly than if preferences are known. By comparing policies estimated from two limiting assumptions (one where all producers are risk neutral and a second, where all we know is that producers are risk averse), we can use our method to assess the consequences (both monetary costs and effects on environmental quality) of a policy design based on the wrong (or very specific) assumption about risk aversion. The costs of "getting it wrong" represent what the government can afford to pay to collect more specific data about risk preferences.

We also devote substantial effort to examining the conditions under which a solution will exist and how it can be determined empirically. We show that the model can be numerically simulated under a broad range of conditions, and we also derive an empirically testable necessary condition for self-selection to be possible. Further, we demonstrate that in certain cases the computational burden of the simulations can be dramatically lowered. In all cases, the stochastic efficiency approach leads to a policy problem that can ultimately be solved with linear programming methods.

In addition, we extend existing models by accounting for hidden information on input use. Past studies assumed that farmers could not "cheat" on agreed input restrictions because the government perfectly observes input use. This assumption may be valid if the input in question is discrete technology or an input such as irrigation water that is regularly metered. However, many polluting agricultural inputs are very difficult to monitor. We show that the "cheating" problem can be circumvented by setting payments that depend on crop yield. This scheme shifts the monitoring burden to crop outputs that are in many instances are already publicly observed. Such payments are in fact a logical extension of many past commodity and insurance programs in the United States, where payments were tied to government-certified yields.

We apply our model to simulate a pollution control program for New York agriculture. Our results suggest that diversity of risk preferences should not be ignored in designing such a program: if the program were designed assuming all farmers are risk neutral, payments would be too small for risk averse farmers to be willing to participate. We also find that even if payments could be conditioned on monitored input levels, government costs would not be appreciably higher if payments were conditioned on monitored output levels instead.

Conceptual Model

Consider a region where nonpoint pollution is generated from an input used for some crop (e.g., fertilizer in corn production). Different production technologies in the region (e.g., soil types) are indexed by $i \in \Theta$. Crop yield per acre for technology i obeys $y^i = y^i(x, b_i)$, where x represents a random uncontrollable input such as rainfall and b_i is the controllable polluting input. b_i, which may or may not be publicly observable, could represent a continuously variable input such as fertilizer or it could be a binary variable representing some production practice. Emissions of pollution e^i are jointly produced with output, so that $e^i = g^i(x, b_i, y^i)$. In exchange for setting b_i at

some prescribed level, the government offers payments to farmers of $s - ty^i$ per acre. Net income per acre for technology i is then

(1) $$m_i = (p_y - t)y^i(x, b_i) - p_b b_i - k + s \equiv \pi^i(x, b_i, t) + s$$

where p_y and p_b are output and input prices, respectively, k is fixed cost, and $\pi^i(x, b_i, t)$ is production profit per acre. Let the support of x be the interval $[\underline{x}, \overline{x}]$ and assume that x and b_i are defined such that $\pi^i_x \geq 0$ and $de/db_i = g^i_b + g^i_y y^i_b \geq 0$ for all i.

Farmers also differ by their risk preferences. Each farmer selects b_i by solving: $\max_{b_i} \{Eu(m_i)\}$, where E is the expectation with respect to x and $u(\cdot)$ belongs to some set Ω of continuous, real-valued utility functions. Letting $b_i(t, s)$ represent the solution to the farmer's problem, the maximized value of the objective function is denoted $Eu(m_i(t, s)) \equiv Eu(\pi^i(x, b_i(t, s), t) + s)$. If emissions are a negative externality, the pre-policy input level $b_i(0, 0)$ (and consequently e^i) exceeds the socially optimal level; suppose the government wishes to implement $b^*_i \leq b_i(0,0)$ as the target input level for technology i.[1]

Although the government is assumed to know the possible types of farmers (i.e., the elements of Θ and Ω), it may not know which farmers are of each type. It must create a policy menu that induces farmers of type i to meet the input target b^*_i through voluntary actions. Designing this menu can be viewed as a two-staged game of imperfect information, where the government chooses a set of policies for the menu in the first stage, and farmers select from these policies in the second stage (Smith and Tomasi).[2] The government must solve this game by backward induction; setting policies in the first stage requires predictions of how farmers will respond in the second stage. In our model, a policy can be described an ordered pair (t, s).

This policy formulation differs from past studies in that government payments depend on

crop yield ($-t$ is a marginal output payment). Previous studies considered only a fixed payment s in exchange for setting inputs at b_i^*, implicitly assuming that the government can perfectly monitor inputs to guarantee farmers will comply with the target input levels.[3]

The government's problem is to find a set of policies $\{(t_i, s_i) : i \in \Theta\}$ satisfying the following constraints:

(2) $$ b_i(t_i, s_i) \leq b_i^* \qquad \text{for all } i \in \Theta, u \in \Omega $$

(3) $$ Eu(m_i(t_i, s_i)) \geq Eu(m_i(0, 0)) \qquad \text{for all } i \in \Theta, u \in \Omega $$

(4) $$ Eu(m_i(t_i, s_i)) \geq Eu(m_i(t_j, s_j)) \qquad \text{for all } i, j \in \Theta, u \in \Omega $$

Equation (2) represents *self-compliance constraints*. Policies must be set so that privately optimal input use is no larger than the socially desirable level. The *participation constraints* in (3) require that post-policy expected utility is at least as large as pre-policy expected utility. The *self-selection constraints* in (4) guarantee that expected utility for type *i*'s own policy exceeds the expected utility for all other policies.

This combination of constraints is the most general case where the government has limited information on input use, technology types, and risk attitudes. Complete information in any one of these areas is a nested case where one set of constraints can be ignored. For example, if the government has complete information on technology types, then policies could be assigned to individual farmers, so that a type-*i* farmer would have only the choices of the policy (t_i, s_i) or not participating. In this case, the self-selection conditions in (4) could be ignored. The special cases and their policy consequences are discussed in more detail below.

Stochastic Efficiency Representation

The primary limitation of the formulation above is that it is difficult to implement. To compute a feasible policy, the constraints would have to be evaluated for all $u \in \Omega$, a prohibitive number of

computations in most applications where the set of admissible utility functions is extremely large. Such an enumeration can be avoided using stochastic efficiency criteria. For several specifications of Ω, the statement that $Eu(m) \geq Eu(m')$ for all $u \in \Omega$ can be equivalently expressed by a single stochastic efficiency condition on the distributions of m and m'.

An appropriate such condition in our case is that of second-degree stochastic dominance (SSD). A cumulative distribution $G(m)$ dominates $H(m')$ by SSD if and only if the area under G is nowhere more than that of H and somewhere less than the area under H:

$$\int_{-\infty}^{\tilde{m}} G(m)\,dm \leq \int_{-\infty}^{\tilde{m}} H(m')\,dm' \quad \text{for all } \tilde{m},$$ with strict inequality somewhere. Dominance by SSD is equivalent to greater expected utility for all increasing and concave utility functions (Hadar and Russel).[4] In our model, the cumulative distribution function (cdf) of income takes the form

(5) $$F_i(m; b, t, s) \equiv \Pr\{\,\pi^i(x, b, t) + s \leq m\,\}.$$

Type i farmers face an income distribution F_i that is conditioned on the values of b, t, and s.

One consequence of unknown risk preferences is that the optimal input level cannot be uniquely predicted. In an SSD setting, the candidates for an optimal input level are those that generate income distributions which are not dominated by any other distribution. These input levels comprise what is known as the second-degree stochastic efficient (SSE) set. Among the three input levels depicted in figure 1, both b and b' belong to the SSE set but b'' does not: b'' is dominated by both b and b' but neither b nor b' are dominated.[5] As shown by the dashed curves in figure 1, the value of s does not influence the SSD rankings of input levels because any $s > 0$ would shift the all the cdfs by the same distance. Therefore, group i's SSE set depends only on t; let $B_i(t) \subset \Re_+$ denote the optimal input correspondence for group i.

To illustrate the use of SSD in the policy scheme, consider two groups (i.e., $\Theta = \{1, 2\}$).

The government must choose policies (t_1, s_1) and (t_2, s_2) to implement the input targets b_1^* and b_2^*. The constraints (2) - (4), written in terms of SSD, require the policies to satisfy:

(6)
$$b_i \leq b_i^* \qquad \forall b_i \in B_i(t_i), i = 1, 2$$

(7)
$$F_i(m; b_i, t_i, s_i) \succ F_i(m; b_i^0, 0, 0) \quad \forall b_i \in B_i(t_i), \forall b_i^0 \in B_i(0), i = 1, 2$$

(8)
$$F_1(m; b_1, t_1, s_1) \succ F_1(m; b_1^2, t_2, s_2) \quad \forall b_1 \in B_1(t_1), \forall b_1^2 \in B_1(t_2)$$

(9)
$$F_2(m; b_2, t_2, s_2) \succ F_2(m; b_2^1, t_1, s_1) \quad \forall b_2 \in B_2(t_2), \forall b_2^1 \in B_2(t_1)$$

where "$\succ$" denotes dominance by SSD. The self-compliance constraints in equation (6) require each t_i to be set so that the SSE set of input levels lies entirely below b_i^*. The participation constraints in equation (7) require the post-policy distributions of income to dominate the pre-policy distributions. Similarly, the self-selection constraints in equations (8) and (9) require group i's distributions of income under the policy (t_i, s_i) to dominate those under (t_j, s_j).

When written in stochastic efficiency terms, the constraints reveal that unknown risk attitudes affect the policy in two ways. First, producers with different risk preferences will make different program enrollment decisions when presented a given policy menu. To account for this fact, the policy designer must compare the entire distributions of returns $F_i(\cdot)$ for the different menu items, rather considering only one statistic such as mean returns. Second, different producers' input responses to a particular menu item will also differ. This requires the designer to evaluate each constraint over a range of stochastic efficient input levels $b_i \in B_i(\cdot)$. In essence, feasible payments include a magnified risk premium that has two "layers," implying that ignoring risk and/or risk aversion is likely to lead to policies that are not incentive compatible.[6]

Computational Method for Finding Policies

The SSD conditions also suggest a computational procedure for finding policies. Given a random

sample of x, SSD comparisons can be made numerically using income distributions generated from the sample points (Anderson, Dillon, and Hardaker). The steps in the procedure follow.

1. Find the pre-policy SSE input sets $B_1(0)$ and $B_2(0)$. For group i, fix $t_i = 0$ and discretize the domain of input levels. Then select two input levels b_i and b_i' in the domain and compare the distribution $F_i(m; b_i, 0, 0)$ to $F_i(m; b_i', 0, 0)$ based on SSD. After repeating this comparison for all (b_i, b_i') pairs, $B_i(t_i)$ can be identified as the set of b_i's that were never dominated.

2. Find t_i sufficiently large so that (6) holds. For each i, repeat the procedure in step 1 for successively larger values of $t_i > 0$ until $B_i(t_i)$ lies entirely below b_i^*.

3. Find the restrictions imposed on s_i by the participation constraints in (7). This restriction is depicted in figure 2. $F_i(\cdot, 0, 0)$ represents a pre-policy distribution of income associated with an input level in $B_i(0)$, and $F_i(\cdot, t_i, 0)$ is an income distribution for an input level in $B_i(t_i)$ but with no acreage payment (i.e., under the policy $(t_i, 0)$). An acreage payment of $s_i > 0$ will shift the distribution to the right in a parallel fashion, as shown by the curve $F_i(\cdot, t_i, s_i)$. The participation constraint says that s_i must be large enough so that $F_i(\cdot, t_i, s_i)$ dominates $F_i(\cdot, 0, 0)$ by SSD, which implies that area A in the figure must exceed area B. By iterating over the input levels in the sets $B_i(0)$ and $B_i(t_i)$, the smallest value of s_i that satisfies (7) can be found numerically. Denoting this minimum value P_i, the participation constraints reduce to $s_i \geq P_i$.

4. Find the restrictions on s_i imposed by the self-selection constraints in (8) and (9). For each i, this requires knowledge of the cross-policy input set $B_i(t_j)$, which can be computed similar to the procedure in step 2. The self-selection condition is depicted in figure 3. Initially assuming that $s_i = s_j = 0$, the distributions $F_i(\cdot, t_i, 0)$ and $F_i(\cdot, t_j, 0)$ in the figure represent incomes for some $b_i \in B_i(t_i)$ and some $b_i^j \in B_i(t_j)$, respectively. Assuming that the policy $(t_j,$

0) is preferred to $(t_i, 0)$, as shown in the figure, s_i must be enlarged to $\tilde{s}_i$, so that $F_i(\cdot, t_i, \tilde{s}_i)$

$\succ F_i(\cdot, t_j, 0)$. Letting I_i represent the smallest value of $\tilde{s}_i$ satisfying this condition over all

admissible b_i and b_i^j's, group i's self-selection constraint (when $s_j = 0$) becomes $s_i \geq I_i$. If s_j

> 0, the income distribution under group j's policy shifts to the right by s_j units, as shown by

the dashed curve $F_i(\cdot, t_j, s_j)$. In this case the SSD condition requires s_i to be enlarged by an

extra s_j units, implying the constraint $s_i \geq I_i + s_j$.

5. Find the acreage payments s_i that meet the restrictions found in steps 3-4. The government's

   minimum cost acreage payments can be found by solving the following linear program:

   (10)   Minimize $a_1 s_1 + a_2 s_2$

   (11)   Subject to: $s_i \geq P_i,$ $i = 1, 2$

   (12) $s_i \geq I_i + s_j,$ $i = 1, 2$

where a_i is the number of acres of land in group i. This problem is depicted in figure 4. The

constraints in (11) require that s_1 is on or to the right of the vertical line at P_1 and that s_2 is

on or above the horizontal line at P_2. The constraints in (12) require s_1 to lie on or to the

right of the 45-degree line starting at I_1 and s_2 to lie on or above the 45-degree line starting

at I_2. For the situation depicted in the figure, the feasible region is the shaded area and the

objective function is minimized at point d.

  This five-step procedure reveals that t_i and s_i need not be computed simultaneously (for

instance, using a grid search) but instead can be found in a sequence of self-contained

computational tasks. Such a sequence is possible because s_i (which is determined in steps 3-5)

does not affect the SSD ranking of input levels (see figure 1), and therefore does not affect the

pre-policy input levels (step 1) or the feasible level of t_i (step 2). Although it has been outlined

for the two group case, the procedure can also be extended to higher-dimension problems.[7]

Existence of a solution in step 5 requires the feasible region to be nonempty, which will be true in general if: (i) P_1 and P_2 are finite, and (ii) $I_1 \leq -I_2$ (see figure 4). The first of these conditions holds by assumption, while the second depends on the technologies of the two groups. A necessary condition for existence can be derived as follows. There are two necessary conditions for one distribution to dominate another by SSD: neither the mean of the dominant distribution nor its lowest observation may be smaller (Anderson, Dillon, and Hardaker). For the self-selection constraints in (8) and (9), these requirements can be written: $E\pi^i(x, b_i, t_i) + s_i \geq E\pi^i(x, b_i^j, t_j) + s_j$, and $\pi^i(\underline{x}, b_i, t_i) + s_i \geq \pi^i(\underline{x}, b_i^j, t_j) + s_j$, where $b_i \in B_i(t_i)$ and $b_i^j \in B_i(t_j)$. Equivalently, s_i must equal or exceed the larger of $[E\pi^i(x, b_i^j, t_j) - E\pi^i(x, b_i, t_i)] + s_j$ and $[\pi^i(\underline{x}, b_i^j, t_j) - \pi^i(\underline{x}, b_i, t_i)] + s_j$, for all permissible b_i and b_i^j. Thus, I_i is at least as large as:

$$(13) \qquad \tilde{I}_i = \max_{\substack{b_i^j \in B_i(t_j), \\ b_i \in B_i(t_i)}} \left\{ E\pi^i(x, b_i^j, t_j) - E\pi^i(x, b_i, t_i), \pi^i(\underline{x}, b_i^j, t_j) - \pi^i(\underline{x}, b_i, t_i) \right\}$$

The necessary condition for separate self-selecting policies to exist is that $\tilde{I}_1 \leq -\tilde{I}_2$. This condition requires some measure of group 2's loss in returns (either in terms of the mean or the lower tail of the distribution) to exceed group 1's loss. That is, one technology is required to be more productive in a stochastic sense. This requirement is an instance of the more general "single-crossing property" encountered in the literature (Mas-Collel, Whinston, and Green).

The Government's Policy Alternatives

As mentioned above, the constraints in equations (6) - (9) encompass several special cases that reflect different levels of government information. The government's information set varies along three dimensions. First, the government may or may not be able to observe input levels. If observing inputs is impossible or prohibitively costly, but outputs can be observed instead, then the government can set a marginal output payment $(-t_i)$ that gives farmers the incentive to self-

comply. If inputs can be observed, the government can directly condition payments on compliance with the target input level b_i^*, in which case the self-compliance constraints (equation (6)) can be ignored and payments need not depend on output (i.e., $t_i = 0$).

Second, the government's level of information on farmers' technology types may vary. If the government does not know individual farmers' technology types, then the government can set acreage payments s_i to ensure self-selection (at least under certain conditions as shown above).[8] Assuming input monitoring and $b_1^* < b_2^*$, Peterson and Boisvert (2001b) showed that the minimum-cost self-selecting payments always occur at point d in figure 4. If the government has complete information on technology types, it could assign policies to each technology group and the self-selection constraints (equation (7)) could be ignored. The minimum cost payments would then be at point c in figure 4. These results imply that technology information is valuable to the government because it could be used to reduce payments (point c is less costly than point d). However, even if this type of information already exists, the government may avoid using it to differentiate policies across farmers due to political constraints (Chambers).

The third dimension concerns information on risk attitudes. Policies based on SSD are conservative in the sense that the government is assumed to know nothing about risk attitudes other than that farmers are risk averse—the government believes that a farmer's risk aversion coefficient, $r(m)$, could like anywhere in the range $[0, \infty)$.[9] In some cases, information on risk attitudes is more precise because $r(m)$ is known to lie in a narrower range (e.g., based on empirical studies of risk behavior). Policies can be conveniently computed in these cases by replacing SSD in the procedure above with what is now called stochastic dominance with respect to a function (Meyer, King; see footnote 4). Doing so for various assumed bounds on $r(m)$ would trace out the relationship between better knowledge of risk attitudes and government cost.

The Case of Simply Related Variables

While the SSD formulation is feasible to implement numerically as outlined above, the computations can be dramatically simplified under certain conditions. This simplification is due to the concept of *simply related random variables* (Hammond). Two random variables are simply related if their cdf's cross at most once. Each of the SSD conditions in (7) - (9) compares some random variable of the form $m = \pi^i(x, b, t) + s$ to another random variable $m' = \pi^i(x, b', t')$ $+ s'$ (e.g., for the participation constraint (7), $t = t_i$, $s = s_i$, and $t' = s' = 0$). The following result describes a sufficient condition for the cdf's of these random variables (F_i and F_i', respectively) to intersect only once, for a given combination of (b, t) and (b', t'):

RESULT 1: *If* $\delta_\pi = \pi_x^i(x,b,t) - \pi_x^i(x,b',t')$ *is positive (negative) for all x, then F_i and F_i' intersect*

at most once, and F_i intersects F_i' from above (from below) if the distributions do cross.

Proofs for this and all other results are in the appendix.

Intuitively, the simply related property follows from the one-to-one relationship between x and income: each realization of income is associated with a unique value of x, and larger incomes are associated with larger x's because $\pi_x^i > 0$. If $\delta_\pi > 0$, then a given change in x causes a larger change in m than in m', so that F_i is geometrically "flatter" and can only intersect F_i' from above. The opposite case is where $\delta_\pi < 0$, so that F_i is "steeper" than F_i'. If δ_π switches sign somewhere in the domain of x, then F_i and F_i' may intersect more than once.

Although the condition in Result 1 must be checked empirically and is not guaranteed to hold, it is not unlikely. Kramer and Pope argued that the simply related property holds for many agricultural applications. This property is actually guaranteed if inputs can be monitored (so that $t = t' = 0$) and y_{xb}^i does not change sign in the domain of x, because δ_π is positive (negative) for

all x if and only if y^i_{xb} is positive (negative). Peterson and Boisvert (2001b) showed that if y^i_{xb} is positive (negative), then b is a risk increasing (risk decreasing) input.

The advantage of simply related variables is that the SSD conditions can be very easily evaluated, because the two necessary conditions for SSD are also sufficient. Formally:

RESULT 2: *Suppose m and m' are simply related, with cdfs F_i and F_i', respectively. The sufficient conditions for F_i to dominate F_i' by SSD are: (i) $\underline{m} \geq \underline{m}'$ and (ii) $Em \geq Em'$, where $\underline{m}$ and $\underline{m}'$ are the lowest observations with positive probability.*

This result has two useful implications for solving the policy problem in practice:

RESULT 3: *Suppose that the profits at any two input levels (i.e., $\pi^i(x, b, t)$ and $\pi^i(x, b', t')$) are simply related random variables. Then the SSE set of input levels $B_i(t)$ is a closed interval of real numbers bounded by $\underline{b}(t) = \arg\max_b \pi^i(\underline{x}, b, t)$ and $\overline{b}(t) = \arg\max_b E\pi^i(x, b, t)$.*

RESULT 4: *Consider a condition of the form:*

(14) $$F_i(m; b, t, s) \succ F_i(m; b', t', s') \;\forall\; b \in B_i(t), b' \in B_i(t').$$

If the two cdfs are simply related, then condition (14) is equivalent to the requirement:

(15) $$s - s' \geq \max\left\{ E\pi^i(x, b', t') - E\pi^i(x, b, t), \;\; \pi^i(\underline{x}, b', t') - \pi^i(\underline{x}, b, t): \; b = \underline{b}(t), \overline{b}(t), \; b' = \underline{b}(t'), \overline{b}(t') \right\}$$

Result 3 allows each of the SSE sets required in steps 1 and 2 of the procedure above to be found by solving two nonlinear maximization problems. Result 4 simplifies the search for the bounds P_i and I_i in steps 3 and 4, which involve several SSD comparisons of the form in (14). Even though each SSD condition must be satisfied for *all* pairs $(b, b') \in B_i(t) \times B_i(t')$, a sufficient condition for this to hold (equation (15)) involves computations at only four points (b is either the upper or lower bound of $B_i(t)$ and b' is either the upper or lower bound of $B_i(t')$). Result 4 also implies that if I_1 and I_2 are computed using equation (15), then the condition $I_1 \leq -I_2$ is both

necessary and sufficient for self-selecting payments to exist.

Empirical Application to Nitrate Loss from New York Corn Production

The model is applied to simulate a program to reduce nitrate contamination of drinking water in central New York. Nitrates reach water supplies by leaching into aquifers and through surface runoff following storm events. Elevated levels of nitrate leaching and runoff in the region are due in large part to the use of nitrogen fertilizer on corn acreage. In the simulated program, corn producers would receive payments in return for reductions in nitrogen fertilizer use.

Empirical Model Components

Corn silage yield functions were estimated from data collected at field trials run by the Department of Soil, Crop, and Atmospheric Sciences at Cornell University. The data include 276 observations of corn silage yield (y), commercial fertilizer, manure application, and growing season rainfall (x) at several sites in New York over several crop years. Total nitrogen applied (b) was computed as the sum of fertilizer and manure nitrogen. The soils at the sites were divided into two groups (indexed by $i = 1, 2$) based on their hydrologic characteristics, which become the different technologies in the simulations.[10]

To gain efficiency, the functions were estimated in a pooled regression using a quadratic specification. The model was fit by maximum likelihood, with the parameters bounded so that the derivative in x is positive to be consistent with the theoretical model. The results are:

$$(16) \quad y = -15.12 + 0.699d_m + 25.71d_2 + 0.1001b - 0.00024b^2 + 0.000057d_2b^2 + 1.51x$$
$$\quad\quad\quad (-5.01) \quad (1.56) \quad\quad (9.38) \quad\quad (6.67) \quad\quad (-6.09) \quad\quad\quad (2.04) \quad\quad\quad (10.08)$$

$$\quad\quad -1.37d_2x - 0.0007bx, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad R^2 = 0.56,$$
$$\quad\quad (-9.59) \quad (-1.40)$$

where *t*-ratios are in parentheses, and d_m and d_2 are dummy variables for manure application and

group 2 soils, respectively. The interaction terms d_2b^2 and d_2x allow the shape of the yield function in nitrogen and rainfall to differ by group. The estimated coefficients on d_2b^2 and d_2x are both statistically different from zero, and their signs imply that group 2 has a higher marginal product of nitrogen but a smaller marginal product of rainfall. At the data means, a one-pound increase in nitrogen increases yield by 0.023 tons and 0.038 tons per acre for groups 1 and 2, respectively, while a one-inch increase in rainfall raises yield by 1.42 tons and 0.05 tons, respectively. Additional details on the data and estimation are in Peterson and Boisvert (2001b).

Profits for group i were simulated by equation (1), where the policy variable b_i is nitrogen from commercial fertilizer.[11] In these simulations, x takes on values from a sample of growing season rainfall observations at the Ithaca weather station over the 30-year period 1963-1992. The prices p_y and p_b were set at the mean of observed corn silage and nitrogen prices (in constant 1992 dollars) over the same 30 years, where corn silage prices were imputed as a corn grain equivalent. Other costs k were based on enterprise budgets from USDA and Schmit.

Nitrate emissions were defined as the sum of leaching and runoff: $e^i = e^i_R + e^i_L$. Emissions were simulated using a recursive system estimated by Boisvert, Regmi, and Schmit, of the form:

$$(17) \qquad e^i_R = e_R(b_i, \mathbf{x}, \mathbf{c}_i), \qquad e^i_L = e_L(b_i, \mathbf{x}, \mathbf{c}_i, e^i_R)$$

$\mathbf{x}$ contains four weather variables (total annual rainfall; rainfall within 14 days of planting, fertilizer, and harvest), and $\mathbf{c}_i$ denotes group i's soil characteristics (field slope, percent organic matter, soil horizon depth, and the erodibility factor "K"). The New York soils data used to estimate this model and its translog specification are described in Boisvert, Regmi, and Schmit.

Policy targets for fertilizer were computed using chance constraints (Lichtenberg and Zilberman). Probability distributions of emissions from each soil were simulated from annual weather observations in Ithaca from 1963-1992. Values of b_i^* were then found by iteratively

reducing fertilizer until nitrate emissions exceeded $e^* = 25$ pounds per acre with a probability of no more than $\alpha = 0.1$. This procedure led to estimated targets of $b_1^* = 55$ and $b_2^* = 82$.

Policy Simulations

Policies were simulated under several scenarios to study the effects of hidden information. All simulations were conducted using the five-step procedure outlined above, which could be simplified to involve only mean and lower-tail profits because the sufficient condition for simply related random variables was satisfied (Result 1).[12] Table 1 reports the fertilizer levels, profits, and production for each of the policies; payments are in table 2. The policy scenarios vary along the three dimensions in the government's information set.

The first dimension is information on fertilizer use. In the *output monitoring* scenarios, fertilizer use cannot be observed and farmers receive payments of the form $s_i - t_i y^i$. In the *input monitoring* scenarios, the government can observe fertilizer use so that the acreage payments s_i could be conditioned directly on meeting the fertilizer targets b_i^*; the marginal output payments t_i were set to zero in these cases (table 2). The effects of information on fertilizer use are well illustrated by the SSD-based policies that are assigned by soil (the upper-left block of numbers). If fertilizer use were unobservable (but output could be monitored), the expected net payments per acre, $E[s_i - t_i y^i]$, were in the range of about \$4-\$7 for group 1 and \$9-\$14 for group 2 (table 2).[13] If fertilizer could be observed, then the groups would receive acreage payments of about \$6 and \$12 per acre, respectively. Based on these results, information on input use does not appear to have a large impact on government costs.

The second dimension is soils information. One set of scenarios here is *assigned* policies by soil type, which requires the use of soils information by farm. In these cases, the self-selection constraints were ignored and acreage payments s_i were set at the participation bounds (P_i). The

other possibilities, where soils information is not used, are *self-selected* policies or a *uniform* policy for all soils. The estimated yield relationships imply that self-selecting policies do not exist, so that the government's only alternative to assigned policies is a uniform policy.[14] The uniform policies were computed by first setting the marginal output payment (or fertilizer levels for input monitoring) at the more stringent of the two individual policies, to ensure that environmental targets were met for both groups. The acreage payments were then adjusted to ensure that both groups would be willing to participate.

The value of soils information can be revealed by comparing a uniform and assigned policies. Consider first this comparison for the SSD scenarios with output monitoring (the first column of numbers in table 2). If the government obtained enough soils information to assign policies, then payments to group 2 would not be affected (they would receive about $9-$14 in either case), but the payment to group 1 could be reduced from about $7-$10 to $4-$7. This cost savings is the value of soils information to the government, and reflects an "information rent" of about $4 earned by group 1 in a uniform policy. In the input monitoring case (the second column of numbers), the implied value of information would be even larger because both groups would earn information rents, of about $14 and $19 per acre, respectively.[15]

Third, scenarios differ by information on risk attitudes. Here, one group of scenarios assumes that the government knows only that farmers are *risk averse* to varying degrees, so that policies must be computed based on SSD. The other scenarios assume that all farmers have identical preferences and are *risk neutral*. In the risk neutral cases, policies were computed by considering only mean profits in each step. The value of risk information can be found by comparing the SSD and risk neutral results. If the government initially believed farmers are risk averse to unknown degrees (the SSD case), and then learned that all farmers are risk neutral, this

information would be valuable because payments could be reduced by up to $7 per acre (table 2).

More plausibly, the government would improve its information on risk attitudes by learning that

risk aversion coefficients have some upper bound $\bar{r}$. Peterson and Boisvert (2001b) calculated

payments under input monitoring using stochastic dominance with respect to a function, with

$\bar{r} = 0.03$.[16] Payments for this bounded case were up to $5 smaller than those under SSD.

The results also illustrate the consequences of incorrect assumptions on risk attitudes.

The difference between SSD and risk neutral payments represents a risk premium that accounts

for up to about 47% of SSD payments. If the government incorrectly assumed that all farmers

were risk neutral, then payments would be too small for risk-averse farmers to be willing to

participate.[17] In the bounded model studied by Peterson and Boisvert (2001b), the estimated risk

premium was somewhat smaller (up to about $3 or 37% of payments under bounded risk

aversion). Nevertheless, non-participation in the program was a dominant strategy for risk-averse

farmers if payments were computed under risk neutrality.

Sensitivity Analysis

As mentioned above, self-selecting policies were not possible because the estimated marginal

products were quite similar across groups. At the data means, the difference in estimated

marginal products of nitrogen is 0.015 tons (30 pounds), which translates to less than 0.1% of

mean yield. The estimated differential may have been muted because of the specific soils used in

the field trials; previous agronomic evidence suggests that light and heavy soils in New York

respond to nitrogen quite differently (Peterson and Boisvert, 2001a).

To explore the effect of a larger productivity differential, a sensitivity analysis was

performed on the coefficient for $d_2 b^2$ in the yield equation, which had a point estimate of

0.000057. Table 3 presents the results for a coefficient value of 0.0001, which raises the

difference in marginal products at mean nitrogen to 0.026 tons (52 pounds). In this case, self-selected policies are possible and are shown in the bottom half of the table. The most significant difference between these and the base results for a uniform policy (table 2) is that output monitoring is substantially more costly than input monitoring. The expected net payments under output monitoring are in the range of $30 and $21 for the two groups, respectively, which are three to four times the payments under input monitoring (table 3). This suggests that the least-cost monitoring scheme is sensitive to technical parameters and must be evaluated empirically.

Policy Implications

This article demonstrates both the theoretical and empirical possibility of successfully designing a voluntary environmental program when the government's information is limited. In particular, we identified the structure of policies necessary to ensure incentive compatibility when both risk attitudes and technology are unknown. We outlined a computational procedure for finding policies that accommodates unknown risk attitudes through stochastic efficiency criteria. We also derived an empirically verifiable necessary condition for self-selection to be possible, and showed that in certain cases the stochastic efficiency comparisons can be simplified to involve a small number of computations involving lower-tail and mean income.

The model was simulated for a program that would offer government payments to New York corn producers in exchange for fertilizer reductions. The results suggest that monitoring corn yields could be substituted for the potentially costly and intrusive monitoring of fertilizer use. Although the results differ with technical parameters, expected net payments were frequently lower when they were tied to crop yields instead of fertilizer levels. Overall, average net payments were below $15 per acre, which are smaller than typical farm program payments received by New York producers in the past. Self-selection would be possible in cases where the

marginal productivity of nitrogen differs substantially across soils.

Simulated payments could be reduced if the government assigned policies to specific soils, and in this sense information on the distribution of soils across farms is valuable to the government. This type of information already exists in many states. In New York, for example, the use-value assessment program requires local officials to record each farm's acreage in each of ten soil productivity groups. Policy makers would need to weigh these cost savings against the political and other consequences of conditioning policy eligibility and benefits on a farmer's resource setting.

Policies were also simulated for both the risk averse and risk neutral cases. Pannell, Malcom, and Kingwell have argued that the insight gained by modeling risk aversion is more pronounced for discrete decisions (such as the adoption of new technology) than for continuous decisions regarding input use, since in the latter case the results are often very similar to a risk neutral model. In the New York application of our policy model, the risk neutral and risk averse results for optimal inputs and profits are indeed similar, with expected incomes that often differ by less than $1 per acre. However, with respect to the program participation decision, the incentive compatibility constraints magnify the effect of risk aversion. If payments were designed assuming risk neutrality, environmental objectives could well be sacrificed because risk averse farmers would have insufficient incentives to participate in the program voluntarily.

Footnotes

[1]The method for finding this target level is not modeled explicitly. In practice, the choice is often made through a second-best standards approach, where input targets are set to meet some predetermined emissions target (Baumol and Oates).

[2]In general, this game involves the government and all producers, so that any farmer's choice may depend strategically on the choices of all other famrers. If all policy options are available to any farmer regardless of others' choices, and the distributions of technology and risk attitudes are independent, this strategic interdependence can be ignored and the policy becomes a large number of two-player games between the government and each producer.

[3]Output-dependent payments require monitoring of yields, which are generally much easier to observe than inputs and need to be measured only once a year. In cases where neither input nor output can be observed, the alternative is to use random spot checks on target input levels and assess penalties on violators. This is the typical approach for industrial-type emissions that occur continuously or at least very frequently (e.g., Malik; Florens and Foucher). The mechanism design below could be modified to accommodate spot checks and penalties; these modifications are described in a supplement available from the authors.

[4]Formally, if $G(m)$ dominates $H(m')$ by SSD, then $Eu(m) \geq Eu(m')$ for all continuous and twice differentiable $u(\cdot)$ with $u' > 0$ and $u'' \leq 0$ (Hadar and Russell, p. 31). Other stochastic efficiency criteria exist for other specifications of the utility set Ω. For example, Meyer has discovered a set of criteria, named stochastic dominance with respect to a function, which can order distributions when the coefficient of absolute risk aversion lies in a specified range.

[5]b does not dominate b' because $F_i(m; b, t, 0)$ starts to the left of $F_i(m; b', t, 0)$, and b' does not dominate b because area A is smaller than area B.

[6]To illustrate, consider the participation constraint in equation (7). Given a t_i, s_i must be large enough so that the policy (t_i, s_i) dominates $(0, 0)$. If risk neutrality is assumed, participation could be secured by a payment of $\bar{s}_i = E\pi^i(x,b_0,0) - E\pi^i(x,b_i,0)$, where b_0 and b_i are the solutions to $\max\{E\pi^i(x, b, 0)\}$ and $\max\{E\pi^i(x, b, t_i)\}$, respectively. But if farmers are risk averse in reality, this payment will generally be insufficient because it does not include a risk premium. Based on the usual definition, the risk premium required is the value r_i such that $F_i(m;b_i,t_i,\bar{s}_i + r_i) \succ F_i(m;b_i,0,0)$. However, this risk premium is generally not large enough because the SSD condition must hold *for all* $b_i \in B_i(t_i)$ and $b_0 \in B_i(0)$.

[7]With additional groups, the computational burden increases linearly in steps 1-2, because the pre-policy input levels and the t_i's must be determined separately for each of n groups. Similarly, the participation constraints in step 3 must be found for each $i = 1, ..., n$. In step 4, there would be a total of $n!/(n-2)!$ self-selection constraints to find, one for each (i, j) pair.

[8]If self-selection is desired, the government has a commitment problem, in that producers may believe policies will be assigned to them once they reveal their type by selecting a policy. This problem could be avoided through the use of a multiple-year, binding contract, although the government would still know the farmer's type for future contract periods. We are indebted to a reviewer for making this observation.

[9]This coefficient is defined as $r(m) = -u''(m)/u'(m)$; it is positive for risk-averse individuals.

[10]Manure was credited with 3 pounds of nitrogen per ton to obtain total nitrogen applied. 52 of the 276 observations are from group 1, corresponding to Hydrologic Group A soils, while the 224 remaining observations (group 2) are from Hydrologic Group B soils. A soils are coarser and more vulnerable to leaching than B soils (Boisvert, Regmi, and Schmit).

[11]Farmers were assumed to always apply 20 tons of manure per acre to dispose of animal

waste.

[12]Because the estimated yield functions $y^i(x, b)$ are linear in x, the quantity $\delta_\pi = \pi_x^i(x,b,t) - \pi_x^i(x,b',t')$ is a constant (either positive or negative) for all (b, t) and (b', t').

[13]The expected values assume a range of values under SSD because they correspond to a range of optimal fertilizer levels (table 2).

[14]The self-selection bounds were $I_1 = -11.35$ and $I_2 = 18.59$, which violate the necessary condition for self-selection because $I_1 > -I_2$. The underlying difficulty is the similarity in marginal products of fertilizer across groups. Sensitivity results with a wider gap in marginal productivities are presented below for comparison.

[15]If fertilizer is monitored, the uniform policy would have to require $b^* = 55$ pounds per acre to meet both environmental targets (table 1). An acreage payment of $25.24 is then needed to ensure both group would be willing to participate in the program. This payment exceeds the assigned payments by about $19 and $14 for groups 1 and 2, respectively (table 2).

[16]This value was chosen as a plausible upper bound based on published estimates of risk coefficients. Such estimates must be properly calibrated because they are not invariant to the level of income (Grube). The bound of 0.03 reflects empirical evidence on risk coefficients when the amount of income at stake is in the range of $100 to $1,000 (Grube, Tauer).

[17]For our model specification, the risk neutral results are equivalent to those from assuming no uncertainty. As a reviewer points out, the quadratic production function implies that the effect of rainfall on the marginal product of fertilizer is constant, so that $\max\{E\pi^i(x, b, t)\}$ is equivalent to $\max\{\pi^i(Ex, b, t)\}$. Thus, in this case, ignoring risk as the same consequences as ignoring risk aversion; both lead to violations of incentive compatibility.

References

Anderson, J.R., J.L. Dillon, and B. Hardaker. *Agricultural Decision Analysis*. Ames IA: The Iowa State University Press, 1977.

Baumol, W.J., and W.E. Oates. *The Theory of Environmental Policy, Second Edition.* New York: Cambridge University Press, 1988.

Boisvert, R., A. Regmi, and T. Schmit. "Policy Implications of Ranking Distributions of Nitrate Runoff and Leaching from Corn Production by Region and Soil Productivity." *Journal of Production Agriculture* 10(1997):477-83.

Chambers, R.G. "On the Design of Agricultural Policy Mechanisms." *American Journal of Agricultural Economics* 74(1992):646-54.

Florens, J., and C. Foucher. "Pollution Monitoring: Optimal Design of Inspection—An Economic Analysis of the Use of Satellite Information to Deter Oil Pollution." *Journal of Environmental Economics and Management* 38(1999):81-96.

Grube, A.H. "Participation in Farm Commodity Programs: A Stochastic Dominance Analysis: Comment." *American Journal of Agricultural Economics* 68(1986):185-188.

Hadar, J., and W. Russell. "Rules for Ordering Uncertain Prospects." *American Economic Review* 59(1969):25-34.

Hammond, J.S. "Simplifying the Choice Between Uncertain Prospects Where Preference is Nonlinear." *Management Science* 20(1974):1047-72.

Isik, M. "Resource Management under Production and Output Price Uncertainty: Implications for Environmental Policy." *American Journal of Agricultural Economics* 84(2002):557-71.

Kramer, R.A., and R.D. Pope. "Participation in Farm Commodity Programs: A Stochastic Dominance Analysis: Reply" *American Journal of Agricultural Economics* 68(1986):189-90.

Leathers, H.D., and J.D. Quiggin. "Interactions between Agricultural and Resource Policy: Importance of Attitudes Toward Risk." *American Journal of Agricultural Economics* 73(1991):757-64.

Lichtenberg, E., and D. Zilberman. "Efficient Regulation of Environmental Health Risks." *Quarterly Journal of Economics* 103(1988):167-78.

Malik, A.S. "Self-Reporting and the Design of Policies for Regulating Stochastic Pollution." *Journal of Environmental Economics and Management* 24(1993):241-57.

Mas-Collel, A., M.D. Whinston, and J.R. Green. *Microeconomic Theory.* New York: Oxford University Press, 1995.

Meyer, J. "Choice Among Distributions." *Journal of Economic Theory* 14(1977):326-36.

Pannell, D.J., B. Malcolm, and R.S. Kingwell. "Are We Risking Too Much? Perspectives on Risk in Farm Modelling." *Agricultural Economics* 23(2000):69-78.

Peterson, J.M., and R.N. Boisvert. "Control of Nonpoint Source Pollution Through Voluntary Incentive-Based Policies: An Application to Nitrate Contamination in New York." *Agricultural and Resource Economics Review* 30(2001a): 127-38.

_____. "Designing Nonpoint Source Pollution Policies with Limited Information about both Risk Attitudes and Production Technology." Working Paper 2001-13, Department of Applied Economics and Management, Cornell University, 2001b.

Rochet, J., and L.A. Stole. "Nonlinear Pricing with Random Participation." *Review of Economic Studies* 69(2002):277-311.

Roosen, J., and D.A. Hennessy. "Tests for the Role of Risk Aversion on Input Use." *American Journal of* Agricultural *Economics* 85(2003):30-43.

Schmit, T.D. "The Economic Impact of Nutrient Loading Restrictions on Dairy Farm Profitability." MS thesis, Cornell University, 1994.

Smith, R.B.W., and T.D. Tomasi. "Multiple Agents, and Agricultural Nonpoint-Source Water Pollution Control Policies." *Agricultural and Resource Economics Review* 28(1999):37-43.

Tauer, L.W. "Use of Life Insurance to Fund the Farm Purchase of Heirs." *American Journal of Agricultural Economics* 67(1985):60-69.

Townsend, R.M., and R.A. Mueller. "Mechanism Design and Village Economies: From Credit to Tenancy to Cropping Groups." *Review of Economic Dynamics* 1(1998): 119-72.

U.S. Department of Agriculture, Economic Research Service. *Economic Indicators of the Farm Sector: Costs of Production—Major Field Crops & Livestock and Dairy, 1992.* ECIFS 12-3, August 1994.

Wu, J., and B. Babcock. "Optimal Design of a Voluntary Green Payment Program Under Asymmetric Information." *Journal of Agricultural and Resource Economics* 20(1995):316-27.

_____. "Contract Design for the Purchase of Environmental Goods from Agriculture." *American Journal of Agricultural Economics* 78(1996):935-945.

Appendix

Proof of Result 1

Suppose that $\delta_m = \dfrac{\partial F_i(\hat{m})}{\partial m} - \dfrac{\partial F_i'(\hat{m})}{\partial m'}$ is negative (positive) for all $\hat{m}$; i.e., F_i is everywhere flatter

(steeper) than F_i'. This implies that if the distributions cross, F_i intersects F_i' from above (below).

We will prove that $\delta_m < (>) 0$ if and only if $\delta_\pi > (<) 0$. Let F_x be the cdf of x (i.e., $F_x(a) \equiv \Pr\{x \le a\}$), and define $X(m)$ and $X'(m')$ as the inverse functions of m and m' respectively, such that $X(\pi^i(x, b, t) + s) = x$ and $X'(\pi^i(x, b', t') + s') = x$. Applying $X(\cdot)$ and $X'(\cdot)$ to both sides of the inequalities inside F_i and F_i', based on the definition in (5):

$$F_i(m) = \Pr\{x \le X(m)\} = F_x(X(m)) \quad \text{and} \quad F_i'(m') = \Pr\{x \le X'(m')\} = F_x(X'(m'))$$

At an intersection point $\hat{m}$:

$$F_i(\hat{m}) = F_i'(\hat{m}) \quad \Rightarrow \quad F_x(X(\hat{m})) = F_x(X'(\hat{m})) \quad \Rightarrow \quad X(\hat{m}) = X'(\hat{m})$$

Letting $\hat{X}$ represent the value of $X(\hat{m}) = X'(\hat{m})$, δ_m can be written in terms of F_x as follows:

$$(18) \qquad \delta_m = \frac{\partial F_x(\hat{X})}{\partial X}\frac{\partial X}{\partial m} - \frac{\partial F_x(\hat{X})}{\partial X}\frac{\partial X'}{\partial m'} = \frac{\partial F_x(\hat{X})}{\partial X}\left[\frac{\partial X}{\partial m} - \frac{\partial X'}{\partial m'}\right]$$

By the inverse function theorem, $\partial X/\partial m = 1/\pi^i_x(x, b, t)$ and $\partial X'/\partial m' = 1/\pi^i_x(x, b', t')$. Substituting

these relationships into (18) and noting that $\partial F_x/\partial X > 0$ by the definition of a cdf, $\delta_m < (>) 0$ is

equivalent to: $1/\pi^i_x(x, b, t) < (>) = 1/\pi^i_x(x, b', t')$. Rearranging, $\pi^i_x(x, b, t) > (<) \pi^i_x(x, b', t')$, which

is the desired result.

Proof of Result 2

SSD requires that $S(\tilde{m}) = \displaystyle\int_{-\infty}^{\tilde{m}} [F_i(m) - F_i'(m)]dm \le 0$ for all $\tilde{m}$, with strict inequality for some $\tilde{m}$.

We prove this condition holds if F_i and F_i' are simply related and hypotheses (i) and (ii) are met.

There are three cases to consider:

Case 1: F_i and F_i' do not intersect. Under hypotheses (i) and (ii) Result 2, F_i must lie strictly to the right of F_i' in this case. Thus, $F_i(m) < F_i'(m)$ for all m, and $S(\tilde{m}) < 0$ for all $\tilde{m}$.

Case 2. F_i and F_i' intersect at their lower tails. Here $\underline{m} = \underline{m}'$; since the distributions cannot cross a second time, F_i must lie either strictly to the right or left of F_i' for $m > \underline{m}$. Hypothesis (ii) precludes the second possibility, which implies that $F_i(m) < F_i'(m)$ for all $m > \underline{m}$. Thus, $S(\tilde{m}) = 0$ for all $\tilde{m} \leq \underline{m} = \underline{m}'$ and $S(\tilde{m}) < 0$ for all $\tilde{m} > \underline{m}$.

Case 3. F_i and F_i' intersect above their lower tails. In this case $F_i(\hat{m}) = F_i'(\hat{m})$ for some $\hat{m} > \underline{m}, \underline{m}'$. Since $\hat{m}$ can be the only intersection point, hypothesis (i) implies that $\underline{m} > \underline{m}'$. Thus, F_i must lie strictly to the right of F_i' up to $\hat{m}$ (i.e., $F_i(m) < F_i(m) \ \forall m < \hat{m}$), so that $S(\tilde{m}) < 0$ for all $\tilde{m} \leq \hat{m}$. For $\tilde{m} > \hat{m}$, $S(\tilde{m}) = S(\hat{m}) + \int_{\hat{m}}^{\tilde{m}} [F_i(m) - F_i'(m)]dm$. The second component is positive and monotonically increasing in $\tilde{m}$ since F_i lies to the left of F_i' after the intersection point (i.e., $F_i(m) > F_i'(m)$ for $m > \hat{m}$). However, hypothesis (ii) guarantees that it never becomes large enough to exceed $S(\hat{m})$ in absolute value. To see this, note that $Em \geq Em'$ means that

$$\int_{-\infty}^{\infty} m[dF_i - dF_i'] \geq 0, \text{ or, integrating by parts, } m[F_i(m) - F_i'(m)]\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} [F_i(m) - F_i'(m)]dm \geq 0.$$

Since $F_i(\underline{m}) = F_i'(\underline{m}) = 0$ and $F_i(\bar{m}) = F_i'(\bar{m}) = 1$, the first term in brackets equals zero. Therefore,

$$\lim_{\tilde{m} \to \infty} S(\tilde{m}) \leq 0.$$

Proof of Result 3

We will show that a b below or above both $\underline{b}(t)$ and $\bar{b}(t)$ is dominated by some input level but a

b between $\underline{b}(t)$ and $\overline{b}(t)$ is not. We do not know a priori whether $\underline{b}(t)$ is larger or smaller than

$\overline{b}(t)$. Suppose first that $\underline{b}(t) < \overline{b}(t)$. To begin, note that strict concavity of $\pi^i(\underline{x},b,t)$ and

$E\pi^i(x,b,t)$ in b implies that: (i) $\pi^i(\underline{x},b,t)$ is strictly increasing (decreasing) for all $b < (>) \underline{b}(t)$,

and (ii) $E\pi^i(x,b,t)$ is strictly increasing (decreasing) for all $b < (>) \overline{b}(t)$. Thus, for any $b < \underline{b}(t)$:

$$\pi^i(\underline{x},b,t) < \pi^i(\underline{x},\underline{b}(t),t) \quad \text{and} \quad E\pi^i(x,b,t) < E\pi^i(x,\underline{b}(t),t)$$

where the inequalities follow from the definition of $\underline{b}(t)$ and fact (ii), respectively. Therefore,

by result 1, $\underline{b}(t)$ dominates b by SSD. Similarly, for any $b > \overline{b}(t)$:

$$\pi^i(\underline{x},b,t) < \pi^i(\underline{x},\overline{b}(t),t) \quad \text{and} \quad E\pi^i(x,b,t) < E\pi^i(x,\overline{b}(t),t)$$

by fact (i) and the definition of $\overline{b}(t)$, implying that $\overline{b}(t)$ dominates b by SSD. Finally, consider

any two input levels b, b' such that $\underline{b}(t) \le b < b' \le \overline{b}(t)$. Neither of these input levels can

dominate the other because:

$$\pi^i(\underline{x},b,t) < \pi^i(\underline{x},b',t) \quad \text{and} \quad E\pi^i(x,b,t) > E\pi^i(x,b',t)$$

That is, one of the necessary conditions for either b or b' to dominate is violated. A parallel set

of arguments verifies that if $\underline{b}(t) > \overline{b}(t)$ then: $\overline{b}(t)$ dominates all $b < \overline{b}(t)$; $\underline{b}(t)$ dominates all b

$> \underline{b}(t)$; and for any $b < b'$ in the interval $[\overline{b}(t),\underline{b}(t)]$, neither input level dominates the other.

Proof of Result 4

By Result 3, $B(t)$ is the closed interval of real numbers bounded by $\underline{b}(t)$ and $\overline{b}(t)$. To begin, we

must establish that $\forall b \in B(t)$, $\pi^i(\underline{x},b,t)$ is bounded between $\pi^i(\underline{x},\underline{b}(t),t)$ and $\pi^i(\underline{x},\overline{b}(t),t)$, and

that $E\pi^i(x,b,t)$ is bounded between $E\pi^i(x,\underline{b}(t),t)$ and $E\pi^i(x,\overline{b}(t),t)$. By the definition of a

maximum $\pi^i(\underline{x},b,t) \le \pi^i(\underline{x},\underline{b}(t),t)$ for all $b \in B(t)$. Since $B(t)$ is a closed and bounded interval,

any $b \in B(t)$ can be written $b = \alpha \underline{b}(t) + (1-\alpha)\overline{b}(t)$ for some $\alpha \in [0, 1]$. By the concavity of

$\pi^i(\underline{x}, b, t)$ in b, $\pi^i(\underline{x}, b, t) \geq \alpha \pi^i(\underline{x}, \underline{b}(t), t) + (1-\alpha)\pi^i(\underline{x}, \overline{b}(t), t) \geq \pi^i(\underline{x}, \overline{b}(t), t)$. A parallel set of

arguments verifies that $E\pi^i(x, \underline{b}(t), t) \leq E\pi^i(x, b, t) \leq E\pi^i(x, \overline{b}(t), t)$ for all $b \in B(t)$.

Now suppose group i faces the policy alternatives (t, s) and (t', s') and that all of the

following conditions are met:

(19) $\qquad \pi^i(\underline{x}, \underline{b}(t), t) + s \geq \pi^i(\underline{x}, \underline{b}(t'), t') + s', \qquad E\pi^i(x, \underline{b}(t), t) + s \geq \pi^i(\underline{x}, \underline{b}(t'), t') + s'$

(20) $\qquad \pi^i(\underline{x}, \underline{b}(t), t) + s \geq \pi^i(\underline{x}, \overline{b}(t'), t') + s', \qquad E\pi^i(x, \underline{b}(t), t) + s \geq E\pi^i(x, \overline{b}(t'), t') + s'$

(21) $\qquad \pi^i(\underline{x}, \overline{b}(t), t) + s \geq \pi^i(\underline{x}, \underline{b}(t'), t') + s', \qquad E\pi^i(x, \overline{b}(t), t) + s \geq E\pi^i(x, \underline{b}(t'), t') + s'$

(22) $\qquad \pi^i(\underline{x}, \overline{b}(t), t) + s \geq \pi^i(\underline{x}, \overline{b}(t'), t') + s', \qquad E\pi^i(x, \overline{b}(t), t) + s \geq E\pi^i(x, \overline{b}(t'), t') + s'$

Conditions (19) - (20) and the bounds on $\pi^i(\underline{x}, b', t')$ and $E\pi^i(x, b', t')$ established above imply:

(23) $\quad \pi^i(\underline{x}, \underline{b}(t), t) + s \geq \pi^i(\underline{x}, b', t') + s', \quad E\pi^i(x, \underline{b}(t), t) + s \geq E\pi^i(x, b', t') + s' \qquad \forall b' \in B(t')$

Similarly, (21) - (22) and the bounds on profits imply:

(24) $\quad \pi^i(\underline{x}, \overline{b}(t), t) + s \geq \pi^i(\underline{x}, b', t') + s', \quad E\pi^i(x, \overline{b}(t), t) + s \geq E\pi^i(x, b', t') + s' \qquad \forall b' \in B(t')$

Finally, (23) - (24) and the bounds on $\pi^i(\underline{x}, b, t)$ and $E\pi^i(x, b, t)$ imply that:

(25) $\pi^i(\underline{x}, b, t) + s \geq \pi^i(\underline{x}, b', t') + s', \quad E\pi^i(x, b, t) + s \geq E\pi^i(x, b', t') + s' \quad \forall b \in B(t), \forall b' \in B(t')$

By Result 2, (25) is sufficient to guarantee that:

(26) $\qquad\qquad\qquad F_i(m; b, t, s) \succ F_i(m; b', t', s') \ \forall b \in B(t), \forall b' \in B(t')$

An equivalent way of expressing the conditions in (19) - (22) is that:

(27) $\qquad s - s' \geq \max \begin{bmatrix} \pi^i(\underline{x}, \underline{b}(t'), t') - \pi^i(\underline{x}, \underline{b}(t), t) & E\pi^i(x, \underline{b}(t'), t') - E\pi^i(x, \underline{b}(t), t) \\ \pi^i(\underline{x}, \overline{b}(t'), t') - \pi^i(\underline{x}, \underline{b}(t), t) & E\pi^i(x, \overline{b}(t'), t') - E\pi^i(x, \underline{b}(t), t) \\ \pi^i(\underline{x}, \underline{b}(t'), t') - \pi^i(\underline{x}, \overline{b}(t), t) & E\pi^i(x, \underline{b}(t'), t') - E\pi^i(x, \overline{b}(t), t) \\ \pi^i(\underline{x}, \overline{b}(t'), t') - \pi^i(\underline{x}, \overline{b}(t), t) & E\pi^i(x, \overline{b}(t'), t') - E\pi^i(x, \overline{b}(t), t) \end{bmatrix}$

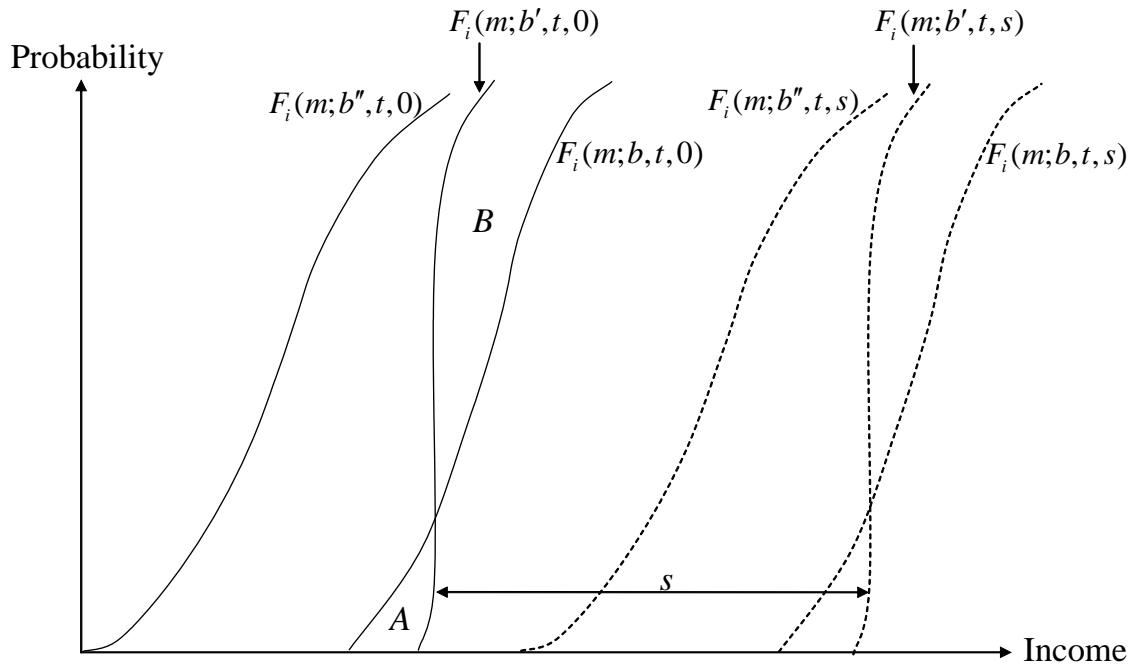That is, (27) implies (26), which is the desired result.

Figure 1. Geometry of stochastic efficient input levels

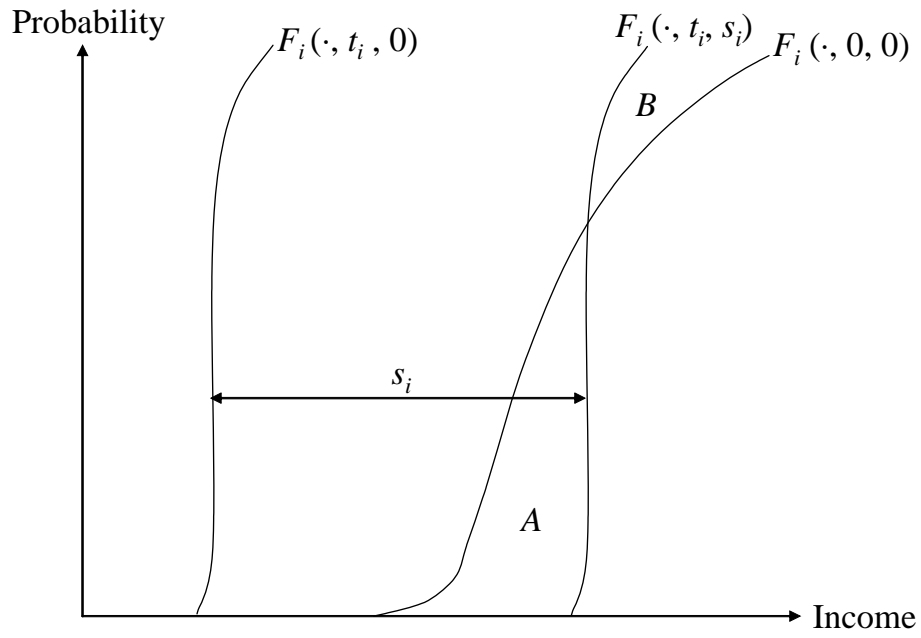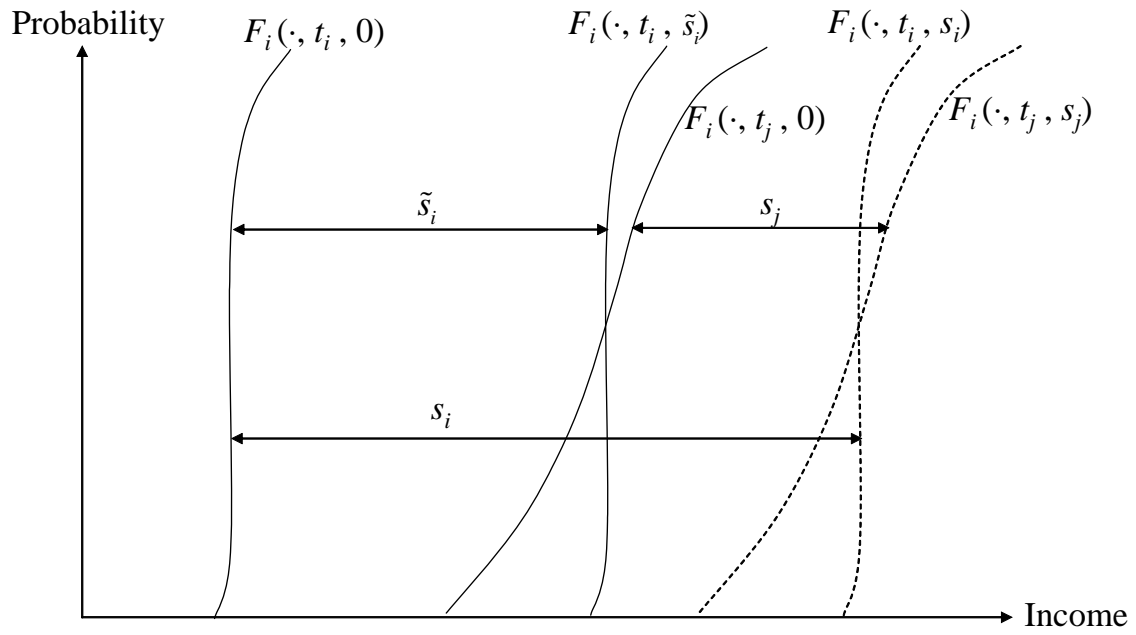**Figure 2. Geometry of the participation constraint**

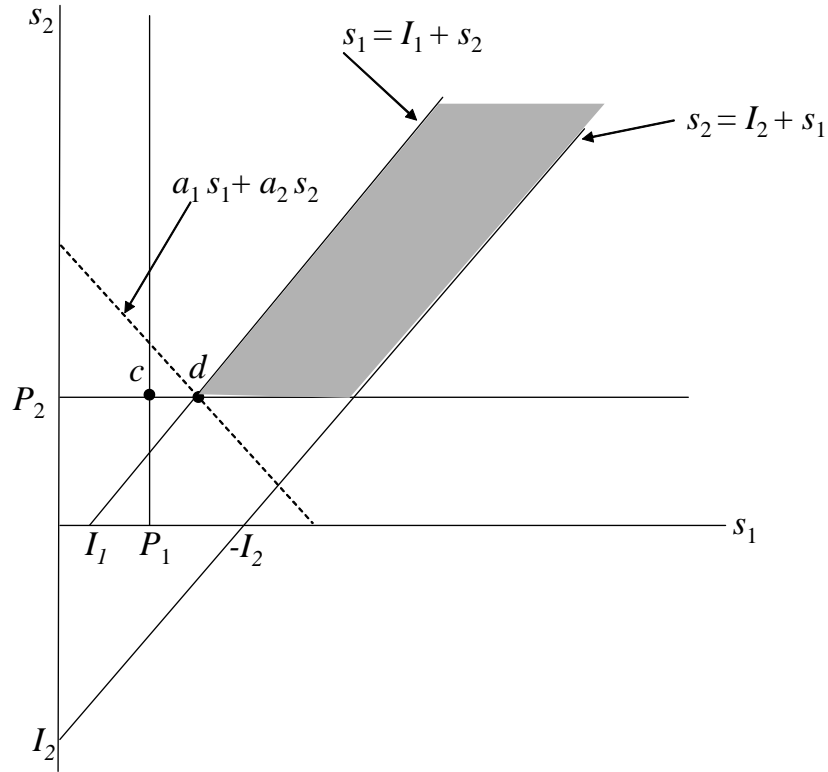**Figure 3. Geometry of the self-selection constraint**

Figure 4. Geometry of the policy design problem

Table 1. Pre- and Post-Policy Fertilizer, Income, and Production, Various Policy Scenarios

Item	Risk Aversion (SSD)			Risk Neutrality		
	Pre-Policy	Output Monitoring	Input Monitoring	Pre-Policy	Output Monitoring	Input Monitoring
Assigned Policies by Soil						
Fertlizer, group 1 (lb/acre)	83-92	46-55	55	83	55	55
Mean income, group 1 ($/acre)	225-226	226	229	226	226	226
Mean yield, group 1 (tons/acre)	23.7-23.9	22.8-23.0	23.0	23.7	23.0	23
Fertlizer, group 2 (lb/acre)	127-139	70-82	82	127	82	82
Mean income, group 2 ($/acre)	212-213	215	217	213	213	213
Mean yield, group 2 (tons/acre)	23.8-24.0	22.2-22.7	22.7	23.8	22.7	22.7
Uniform Policy						
Fertlizer, group 1 (lb/acre)	83-92	39-48	55	83	48	55
Mean income, group 1 ($/acre)	225-226	228	248	226	226	241
Mean yield, group 1 (tons/acre)	23.7-23.9	22.5-22.8	23.0	23.7	22.8	23
Fertlizer, group 2 (lb/acre)	127-139	70-82	55	127	82	55
Mean income, group 2 ($/acre)	212-213	215	219	213	213	213
Mean yield, group 2 (tons/acre)	23.8-24.0	22.2-22.7	21.6	23.8	22.7	21.6

Table 2. Mean Optimal Payments and Information Premiums

	Risk Aversion (SSD)		Risk Neutrality	
Item	Output Monitoring	Input Monitoring	Output Monitoring	Input Monitoring
Assigned Policies by Soil				
Output payment, group 1 ($/ton)	-9.86	0.00	-8.49	0.00
Acreage payment, group 1 ($/acre)	230.81	6.33	199.24	3.60
Expected net payment, group 1 ($/acre)	3.61-6.53	6.33	3.61	3.60
Output payment, group 2 ($/ton)	-10.68	0.00	-9.54	0.00
Acreage payment, group 2 ($/acre)	251.06	11.62	223.43	7.29
Expected net payment, group 2 ($/acre)	9.13-13.69	11.62	7.30	7.29
Uniform Policy				
Output payment ($/ton)	-10.68	0.00	-9.54	0.00
Acreage payment ($/acre)	251.06	25.24	223.43	18.63
Expected net payment, group 1 ($/acre)	7.29-10.77	25.24	5.66	18.63
Soils information rent, group 1 ($/acre)	3.68-4.24	18.91	2.05	15.03
Expected net payment, group 2 ($/acre)	9.13-13.69	25.24	7.30	18.63
Soils information rent, group 2 ($/acre)	0.00	13.62	0.00	11.34

Table 3. Mean Optimal Payments and Information Premiums, Self-selecting Policies

Item	Risk Aversion (SSD)		Risk Neutrality	
	Output Monitoring	Input Monitoring	Output Monitoring	Input Monitoring
Assigned Policies by Soil				
Output payment, group 1 ($/ton)	-9.86	0.00	-8.49	0.00
Acreage payment, group 1($/acre)	230.81	6.33	199.24	3.60
Expected net payment, group 1 ($/acre)	3.61-6.53	6.33	3.61	3.60
Output payment, group 2 ($/ton)	-9.00	0.00	-8.08	0.00
Acreage payment, group 2 ($/acre)	237.01	5.15	213.31	5.15
Expected net payment, group 2($/acre)	20.82-22.01	5.15	19.25	5.15
Self-Selected Policies				
Output payment, group 1 ($/ton)	-9.86	0.00	-8.49	0.00
Acreage payment, group 1($/acre)	256.85	11.48	222.77	8.75
Expected net payment, group 1 ($/acre)	29.65-32.58	11.48	27.14	8.75
Soils information rent, group 1 ($/acre)	26.04	5.15	23.53	0.00
Output payment, group 2 ($/ton)	-9.00	0.00	-8.08	0.00
Acreage payment, group 2 ($/acre)	237.01	5.15	213.31	5.15
Expected net payment, group 2 ($/acre)	20.82-22.01	5.15	19.25	5.15
Soils information rent, group 2 ($/acre)	0.00	0.00	0.00	0.00