Empirical Calibration of a Least-Cost Conservation Reserve Program

Glenn Sheriff*
Department of International and Public Affairs
Columbia University

May 18, 2006

Abstract

Mechanism design models typically conclude by characterizing an optimal allocation schedule based on the principal's beliefs regarding agent value functions and the distribution of agent types. This article addresses the question of how a principal can develop these beliefs given a standard cross-sectional data set in which agents' input-output choices are observable, but their underlying heterogeneity is not. I employ the methodology to evaluate strategies for reducing the cost of a voluntary program that reduces cultivation on environmentally-sensitive farmland.

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Long Beach, California, July 23-26, 2006

*420 W. 118th St., Room 1405, New York, NY 10027, Ph: (212) 854-0027, Fax: (212),864-4847, email: gs2096@columbia.edu

Robert G. Chambers, Jean-Paul Chavas, Bruce Gardner, W. Bentley MacLeod, Tigran Melkonyan, Valerie Mueller, Isabelle Perrigne, and Michael Roberts provided valuable comments and suggestions. I also benefited greatly from data-related discussions with V. Eldon Ball and William McBride and other staff at the USDA Economic Research Service (ERS). This article was screened by ERS to ensure that no confidential information has been revealed.

I. Introduction

This article deals with an adverse selection problem of the general class considered by Baron and Myerson (1982). Agents have private information regarding a "type" parameter. The uninformed principal acts as a Stackelberg leader, designing the best contract possible, given the information asymmetry. There is a rich theoretical literature extending this basic model. Typically the precise terms of an optimal contract schedule depend on the principal's prior beliefs regarding the probability distribution of types and the manner in which type affects agents' reservation utility. What is lacking in the literature is an easily implemented methodology by which the principal can develop robust beliefs regarding these two items with data that do not include agents' private information. This article develops such a methodology and applies it to the problem of reducing the cost of one of the United States' largest environmental policies, the Conservation Reserve Program (CRP). Administered by the U.S. Department of Agriculture (USDA), this program pays agricultural producers to refrain from cultivating environmentally-sensitive cropland. The CRP is one of the largest environmental programs in the U.S., both in terms of budget ($1.7 billion per year) and scope (34.7 million acres) (Farm Service Agency, 2005b).

The key econometric problem is estimating the parameters of the type-dependent production technology and the probability distribution of agent types when type is inherently unobservable both to the principal and the econometrician. One branch of the empirical contract theory literature, such as Wolak (1994), Thomas (1995), and Lavergne and Thomas (2005) conducts estimation under the hypothesis that existing contracts are optimally designed by a sophisticated principal. The authors first derive an optimal mechanism. This analysis provides equations that can be econometrically estimated with observable data to infer the principal's beliefs regarding the distribution of types and the impact of type on the agent. By its nature, this line of research is descriptive rather than prescriptive. It allows the econometrician to infer the principal's beliefs, but provides no guidance to the principal regarding how to develop them.

A second branch deals with how to estimate parameters of a contracting model without assuming a sophisticated principal. Work in this area has used the contracting framework of Laffont and Tirole (1986), focusing on heavily regulated industries such as urban transport where existing regulations create moral hazard problems in addition to adverse selection. In an initial attempt along this line, Dalen and Gómez-Lobo (1997) interpreted regression residuals as an indication of the type of an individual agent. That model was rather restrictive since it did not allow for any random error. In their study of French urban transport firms, Gagnepain and Ivaldi (2002) employ a similar approach, assuming that labor efficiency (type in their model) is an unobserved time-invariant random variable with a beta distribution. Importantly, they also append a normally-distributed random error. With panel data, they are able to use maximum likelihood techniques to estimate the parameters of firms' cost functions as well as the parameters of the distribution of agent types.

Like Gagnepain and Ivaldi, the estimation strategy employed here allows for a random distribution of agent types as well as a stochastic noise. Rather than analyzing a regulated monopoly, it analyzes a sector in which all production decisions are chosen by the agents themselves.[1] As a result, differences in regulatory framework (e.g., cost-plus versus fixed price) do not affect agent actions. In addition, output cannot be considered exogenous.

The key difference, however, lies in econometric methodology. In the spirit of the stochastic frontier models pioneered by Aigner et al. (1977) and Meeusen and Broeck (1977), the specification employed here allows use of a two-part additive composite error structure.

One part is stochastic noise. The second part is a strictly positive type variable. This approach estimates the parameters of a cost frontier. Intuitively, observed deviations from the frontier can occur for two reasons. Random estimation errors can place an observation either above or below the frontier. Agent type can only place an observation above the frontier, since a less effective firm can only perform worse than the best, i.e., have higher than minimum cost. Generalized method of moments (GMM) techniques are used to identify the parameters of each of the distributions of the composite error as well as the technological parameters.[2] This methodology has several advantages over earlier approaches: it easily accommodates robust covariance matrix estimation, allows estimation of a full system including expenditure share equations, is consistent with profit-maximizing behavior, is implementable with cross-sectional data, and is computationally undemanding.

As an illustration, I evaluate alternatives for reducing the cost of the CRP. From its inception in 1985, there have been concerns that asymmetric information regarding producers' reservation values for idled land could inflate the cost of the program. In an effort to overcome information problems, the program uses an auction format. Producers submit bids on the rental value they are willing to accept in order to idle an eligible parcel of land. Initially, the USDA accepted almost all bids that were below a maximum per acre rental rate, or bid cap, set at a regional level. Producers soon learned the cap for their region and bids converged to that level (Smith, 1995).

In practice, the CRP thus functioned more like a per-acre Pigouvian subsidy than an auction. As a result, farmers with relatively unproductive land received surplus payments above their reservation values. One strategy for reducing these surplus payments is to redesign the allocation mechanism itself so it does a better job of minimizing the cost of attaining its objective, given the information asymmetry. Another is to gather data to reduce the degree of hidden information. Over the years, program administrators have chosen the latter course. The USDA uses soil maps and local land rents to establish county-level bid caps rather than the previous regional ones.

I calculate potential savings to taxpayers from these two policy alternatives. First, I evaluate the cost reduction from changing the bidding process so that it functions as a second-best mechanism rather than a Pigouvian subsidy. Second, I obtain an indication of the usefulness of efforts undertaken by the government to overcome information asymmetries by comparing the cost of the first-best (full information) mechanism to the second-best mechanism.

In Section II, I describe and implement the econometric strategy for developing consistent beliefs regarding the production technology and distribution of agent types when agent heterogeneity is unobservable. In Section III, I characterize the theoretical least-cost land set aside program. In Section IV, I use this information to simulate policy alternatives for reducing the cost of the CRP.[3] Section V contains concluding comments.

II. Empirical Model

A. Specification

Producers are characterized by two exogenously fixed factors, one observable and one not. The observed fixed factor is acres of land, denoted $a \in \Re_{++}$. The unobserved factor is a type productivity index $\theta \in \Theta \equiv (0,1]$. Let $\mathbf{x} \in \Re_{+}^{N}$ denote the variable-input vector and $q \in \Re_{+}$ denote aggregate output. The variable-input requirement set is $V(q,a,\theta) \equiv \{\mathbf{x} : \mathbf{x} \text{ can produce } q \text{ given } a,\theta\}$. In addition to $V(q,a,\theta)$ being a closed, convex set, it is assumed to satisfy the following properties:

V1. $\mathbf{x} \in V(q,a,\theta) \Rightarrow \lambda\mathbf{x} \in V(q,a,\theta), \lambda \geq 1$ (weak disposability of inputs);

V2. $V(q,a,\lambda\theta) = \lambda^{-1}V(q,a,\theta), \lambda > 0$ (type is neutrally input augmenting).

Property V1 indicates that inputs can expand along a ray from the origin without reducing feasible output. Property V2 specifies the effect of θ on production. An increase in type implies a proportional radial expansion of $V(q,a,\theta)$. For example, referring to Figure 1, if θ_2 is twice θ_1, it can produce the same output with half of each variable input, given a. Together with V1, V2 implies that a producer can do no worse than a lower type since for any given q its set of feasible input bundles completely includes the set of feasible input bundles of all lower types.

For a vector of variable input prices $\mathbf{w} \in \Re^N_{++}$ the minimum variable cost function is $C(\mathbf{w},q,a,\theta) \equiv \inf_{\mathbf{x}} \{\mathbf{w}'\mathbf{x} : \mathbf{x} \in V(q,a,\theta)\}$. It follows from V2 that a proportional change in type by a factor $\lambda > 0$ implies that the minimum cost of producing q with a changes by a factor of λ^{-1}. Consequently,

$$\ln C(\mathbf{w},q,a,\theta) = \ln C(\mathbf{w},q,a,1) - \ln\theta, \tag{1}$$

where $C(\mathbf{w},q,a,1)$ is the cost frontier; i.e., the minimal cost function across all types. For an interior solution, Shephard's Lemma yields the expenditure share equations for a cost-minimizing producer:

$$\frac{w_n x_n{}^*}{C(\mathbf{w},q^*,a,\theta)} = \frac{\partial \ln C(\mathbf{w},q,a,1)}{\partial \ln w_n}, \ n = 1,...,N, \tag{2}$$

where $x_n{}^*$ is the cost minimizing level of input n. Note that a further consequence of V2 is that these expenditure shares do not vary across types.

For a given output price $p \in \Re_{++}$, the variable profit function is $\pi(p,\mathbf{w},a,\theta) \equiv$

$\sup_q \{pq - C(\mathbf{w}, q, a, \theta)\}$. Let q^* be the profit-maximizing output quantity. Algebraic manipulation of the first order condition for an interior solution to the profit maximization problem yields the following expression of ratio of revenue to cost:

$$\frac{pq^*}{C(\mathbf{w}, q, a, \theta)} = \frac{\partial \ln C(\mathbf{w}, q, a, 1)}{\partial \ln q}. \tag{3}$$

Like expenditure shares (2), the ratio (3) is independent of θ.

Following Diewert (1982), Eqs. (1), (2), and (3) provide the basis for estimating a parametric technology for profit-maximizing producers. For the cost function, I use a modified Cobb-Douglas specification that allows the marginal return to land to vary with farm size:

$$C(\mathbf{w}, q, a, \theta) = \theta^{-1} \exp(\beta_0)\, q^{\beta_q} \textstyle\prod_n w_n^{\beta_n} a^{\beta_a + \beta_{aa} \ln a}. \tag{4}$$

For this specification, with cross-sectional data the system of estimating equations for a typical observation is:

$$\begin{aligned}
\ln C(\widetilde{\mathbf{w}}, q^*, a, \theta) &= \beta_0 + \sum_{n=1}^{N-1} \beta_n \ln \tilde{w}_n + \beta_q \ln q^* + \beta_a \ln a + \beta_{aa} (\ln a)^2 + v_0 - \ln \theta & (5)\\
\frac{w_i x_i{}^*}{C(\mathbf{w}, q^*, a, \theta)} &= \beta_n + v_n\,, \qquad n = 1, ..., N-1 & (6)\\
\frac{pq^*}{C(\mathbf{w}, q^*, a, \theta)} &= \beta_q + v_N\,, & (7)
\end{aligned}$$

where $\widetilde{\mathbf{w}} \equiv (\tilde{w}_1, ..., \tilde{w}_N)' \equiv (w_1/w_N, ..., w_N/w_N)'$ is the vector of input prices normalized by w_N and $C(\widetilde{\mathbf{w}}, q^*, a, \theta) = C(\mathbf{w}, q^*, a, \theta)/w_N$ by the linear homogeneity of the cost function in input prices. The vector of stochastic noise for producer s is $\mathbf{v}_s \equiv (v_0, v_1, ..., v_N)'$.

Since output is endogenous under the assumption of profit maximization, in estimation output price p acts as an instrument for q^*. Let $\mathbf{z}_s \equiv (\ln a, \ln a^2, \ln \widetilde{\mathbf{w}}', \ln p)'$ denote the vector of exogenous variables for producer s.

Due to the information asymmetry, θ is inherently unobservable to the government and the econometrician. Following the stochastic frontier approach, I treat $\ln\theta$ as a random variable. Assume that $\mathbf{v}$ and θ satisfy the following moment conditions:

M1. $E[\mathbf{v}_s|\mathbf{z}_s] = \mathbf{0}$;

M2. $E[v_0^3] = 0$;

M3. $E[\ln\theta|\mathbf{z}_s] = -\sigma\sqrt{2/\pi}$;

M4. $E\left[(\ln\theta - E[\ln\theta])^3\right] = \sigma^3(1-4/\pi)\sqrt{2/\pi}$.

Under M1, the vector $\mathbf{v}_s$ has a mean-zero disturbance vector uncorrelated with the instruments. In addition, M2 states that the disturbance for Eq. (5) is symmetrically distributed. Assumptions M3 and M4 require that $\ln\theta$ be uncorrelated with the instruments, and that its mean and skewness correspond to those of $-|y|$, where y is a random variable distributed $N(0,\sigma^2)$.

As noted by Aigner et al. (1977), these distributional assumptions have two practical implications. First, consider a least squares estimator that ignores the type-dependent component of the error structure, mistakenly using the moment conditions $E[v_0 - \ln\theta|\mathbf{z}_s] = 0$. By treating the expected compound error as mean zero, rather than mean $\sigma\sqrt{2/\pi}$, this regression effectively replaces Eq. (5) with

$$\begin{aligned}\ln C(\tilde{\mathbf{w}}, q^*, a, \theta) \quad = \quad & \left(\beta_0 + \sigma\sqrt{2/\pi}\right) + \sum_{n=1}^{N-1}\beta_i \ln\frac{w_n}{w_N} + \beta_q \ln q^* + \beta_a \ln a \\ & + \beta_{aa}(\ln a)^2 + \left(v_0 - \ln\theta - \sigma\sqrt{2/\pi}\right). \end{aligned} \tag{8}$$

Such an estimator generates consistent estimates of all parameters in the cost function except β_0, which is upwardly biased by $\sigma\sqrt{2/\pi}$. The second practical implication is that the compound error term $(v_0 - \ln \theta)$ is positively skewed with third central moment equal to $-\sigma^3 (1 - 4/\pi) \sqrt{2/\pi}$.

Let $\mathbf{e} \equiv (e_1, ..., e_S)'$ denote the residuals for Eq. (8) of a regression of the system comprising Eqs. (6)-(8). The third moment of the residuals is a consistent estimator for the third moment of the combined error term. This suggests an additional equation that can be estimated sequentially after Eqs. (6)-(8):

$$e_s^3 = -\sigma^3 (1 - 4/\pi) \sqrt{2/\pi} + v_{N+1}, \tag{9}$$

where v_{N+1} is random noise. The estimate $\hat{\sigma}^3$ can then be used to correct the initial bias in β_0. Adapting Newey (1984), one can employ the residuals from Eqs. (6)-(8) to compute the asymptotic covariance matrix for the entire system.

Using these results, estimation of the system proceeds in three steps. In the first step, I ignore $\ln \theta$, and estimate Eqs. (6)-(8) by system two-stage least squares (2SLS). This procedure generates consistent estimates of all parameters, except β_0. Although otherwise consistent, the 2SLS estimator is likely to be inefficient and generate inconsistent estimates of the covariance matrix. In addition to correlation of errors for the same observation across equations, the noise component may be heteroskedastic or influenced by unobserved shocks commonly affecting all producers in the same geographic area. Such shocks may be short-lived or persist across time.

I account for these potential problems in the next step. Following Pepper (2002), Wooldridge (2002), and Wooldridge (2003), the 2SLS residuals are used to construct a more robust GMM estimator. This estimator is asymptotically efficient in the presence of arbitrary heteroskedasticity and arbitrary county-level correlations both within and between time periods. Finally, the third moment of the GMM residuals from Eq. (8) provide the necessary information to obtain consistent estimates of σ, β_0, and the covariance matrix.

B. Data and Estimation Results

Producer cost and returns data come from 1997-2000 Agricultural Resource Management Study (ARMS) surveys conducted by the USDA's National Agricultural Statistics Service (NASS). The surveys are independent annual cross-sections in which it is not possible to identify individual producers across time. They collect producer-level data on input expenditures, output quantities, and land. Disaggregated input and output price data come from the Bureau of Labor Statistics for capital and labor, the Federal Reserve for interest rates and NASS for other inputs and crop and livestock outputs.

I aggregate outputs into a single category and variable inputs into capital services, energy, materials, labor using a multilateral Tornqvist index (see Caves et al., 1982).[4] Since ARMS surveys record capital assets as estimated market value at year end, I calculate capital services adapting the methodology of Hall and Jorgenson (1969). Table 1 contains the summary statistics for the data set.

Since the estimation procedure implicitly assumes all producers have the same general production technology (up to the type parameter), I focus attention on one relatively homogenous area, the "Heartland" Farm Resource Region.[5] This region comprises the entire states of Illinois, Indiana, and Iowa, as well as portions of Kentucky, Minnesota, Missouri, Nebraska, Ohio, and South Dakota. It is the region with most farms, most cropland, and greatest value of production (Economic Research Service, 2000).

To control for systemic production shocks such as weather and technological change that affect the whole region in a given year, I replace the constant in Eq. (5) with an annual dummy vector $\mathbf{d} \equiv (d_{1997}, ..., d_{2000})'$, and β_0 with the parameter vector $\boldsymbol{\delta} \equiv (\delta_{1997}, ..., \delta_{2000})'$.

Table 2 reports the estimation results.

With these parameter estimates, profit-maximizing output can be calculated as:

$$q^* = \left[(\theta p)^{-1} \beta_q \exp\left(\sum_{i=1997}^{2000} \delta_i d_i \right) \prod_{n=1}^{N} w_n^{\beta_n} a^{\beta_a + \beta_{aa} \ln a} \right]^{\frac{1}{1-\beta_q}}, \tag{10}$$

with the corresponding profit function

$$\pi(\cdot) = \left[\beta_q - 1\right] \cdot \left[\theta^{-1} \left(\frac{\beta_q}{p} \right)^{\beta_q} \exp\left(\sum_{i=1997}^{2000} \delta_i d_j \right) \prod_{n=1}^{N} w_n^{\beta_n} a^{\beta_a + \beta_{aa} \ln a} \right]^{\frac{1}{1-\beta_q}}. \tag{11}$$

For the parameter estimates in Table 2, marginal returns to land are increasing until about 400 acres, after which the profit function is strictly concave in land cultivated. The marginal profit from land is positive for all farms less than 6,350 acres. The marginal effect of type on profit is positive, and the single-crossing condition $\pi_{a\theta} > 0$ is satisfied for all farms for which the marginal return to land is positive. In addition, this profit function satisfies theoretical monotonicity and curvature conditions with respect to prices.

Although model specifications and data sets differ, input own-price elasticities are comparable to results from earlier studies of U.S. agriculture such as Ray (1982). The expected farm size for any randomly chosen acre of land in the sample is approximately 2,000 acres. The estimated average annual return to land for the typical 2,000 acre farm over the four years is approximately \$33 per acre. When combined with lump-sum government payments of about \$54 per enrolled acre, this figure is reasonably close to average farmland rental rates for the region of \$91.[6]

As we shall see in the next section, empirical characterization of the optimal contract schedule requires two components: a profit function for each type of producer and a probability distribution for type. The procedures described in this section provide precisely this information. The estimated technological parameters of the cost function can be used to generate a parametric profit function for each type. In addition, the estimate $\hat{\sigma}_\theta$ can be used to obtain a probability density function of producer types.

III. Alternative Environmental Contracts

The environmental policies examined here are three stylized versions of the CRP. In all three versions, the program has two salient features. First, the government must ensure that the sector idles a targeted quantity of land. Second, the program must be voluntary.[7] The task of the government is to design a policy mechanism for allocating land and transfer payments such that these conditions are satisfied at least cost to taxpayers.

The three versions of the program differ in additional constraints faced by the government. In the least constrained "first-best" case, the government has full information regarding producer types, and can directly use this information to design contracts. In the "second-best" case, the government cannot observe type, but can offer producers different per-acre payments depending on the quantity of land enrolled in the program. In the final "Pigouvian subsidy" case, the government cannot observe type nor can it price discriminate among producers. Instead, it can only offer a linear payment per unit of land retired. The difference in cost between the first and second best programs provides an upper bound on the value of actions designed to overcome the information asymmetry. The difference between the second best and Pigouvian programs indicates the maximum benefits obtainable by redesigning the program itself, without obtaining new information.

The variables $a(\theta)$ and $t(\theta)$ denote the terms of a contract for type θ, where $a(\theta)$ is the amount of land cultivated and $t(\theta)$ is the transfer. Suppressing price arguments, the market profit function $\pi(a(\theta), \theta)$, completely characterizes the production technology. All producers are assumed to be equally endowed with an initial acreage allocation $\overline{a}$ equal to 2,000 acres. This maximum farm size is assumed to be fixed in the short run. I also assume that for reasons of overall efficiency the government does not wish farmers to produce in the range of increasing returns to land. Therefore the minimum amount of land cultivated, $\underline{a}$, is 500 acres.[8] These restrictions ensure that the profit function satisfies $\pi_a > 0$, $\pi_{aa} < 0$ and $\pi_{a\theta} > 0$ over the relevant range of a. Consistent with the previous sections, assume the probability density function of types, $f(\theta) \equiv dF(\theta)/d\theta$, is such that $\ln\theta$ has a normal distribution truncated at zero from above.

Assuming all farmland is eligible for participation, the environmental constraint is a requirement that the average quantity of land idled across all producers be at least A acres:

$$\int_{\Theta} [\overline{a} - a(\theta)]\, dF(\theta) \geq A \tag{12}$$

To ensure that program participation is voluntary, producers must be compensated for the opportunity cost of idled land:

$$\pi(a(\theta), \theta) + t(\theta) \geq \pi(\overline{a}, \theta), \quad \text{for all } \theta. \tag{13}$$

Define surplus payments received by a firm in excess of the minimum necessary to satisfy (13) by:

$$s(\theta) \equiv \pi(a(\theta), \theta) + t(\theta) - \pi(\overline{a}, \theta). \tag{14}$$

The participation constraint (13) can then be expressed more succinctly by:

$$s(\theta) \geq 0, \quad \text{for all } \theta. \tag{15}$$

The expected expenditure of the program is:

$$E \equiv \int_{\Theta} \{s(\theta) - \pi(a(\theta), \theta) + \pi(\overline{a}, \theta)\} \, dF(\theta). \tag{16}$$

Finally, let $\lambda \geq 0$ and $\mu(\theta) \geq 0$ respectively denote the Lagrange multipliers for (12) and the restriction $a(\theta) \leq \overline{a}$.

The first-best program chooses $s(\theta)$ and $a(\theta)$ to minimize Eq. (16), subject to (12), (15), and the boundaries of $a(\theta)$. The first-best Lagrangian is:

$$L^{FB} = \int_{\Theta} \left\{ s(\theta) - \pi(a(\theta), \theta) + \pi(\overline{a}, \theta) - \lambda[\overline{a} - a(\theta) - A] - \frac{\mu(\theta)}{f(\theta)}[\overline{a} - a(\theta)] \right\} dF(\theta) \tag{17}$$

At the optimum $s(\theta)$ is clearly set equal to zero for all types. In addition to the constraints already discussed, the optimal land allocation satisfies the following conditions:

$$\lambda + \frac{\mu(\theta)}{f(\theta)} - \pi_a(a(\theta), \theta) \geq 0; \tag{18}$$

$$[a(\theta) - \underline{a}] \left[\lambda + \frac{\mu(\theta)}{f(\theta)} - \pi_a(a(\theta), \theta) \right] = 0; \tag{19}$$

$$a(\theta)[\overline{a} - a(\theta)] = 0; \qquad \lambda \int_{\Theta} \{[\overline{a} - a(\theta)] - A\} \, dF(\theta) = 0. \tag{20}$$

Consequently, for an interior solution, the optimal first-best program satisfies the equimarginal principle. It equates the marginal profit from cultivating an additional acre of land for each producer to the shadow cost of tightening the environmental constraint for the entire sector. With full information, the first-best program could be implemented by offering a uniform price for idled land equal to λ, combined with a type-dependent lump-sum tax that recovers each producer's surplus.

The optimal second-best policy is slightly more complicated. Without loss of generality, I appeal to the Revelation Principle and restrict attention to direct revelation mechanisms satisfying incentive compatability (see Myerson, 1979):

$$\theta \in \arg\max_{\tilde{\theta}} \left\{ \pi \left(a \left(\tilde{\theta} \right), \theta \right) + t \left(\tilde{\theta} \right) \right\} \quad \text{for all} \quad \left(\theta, \tilde{\theta} \right) \in \Theta^2. \tag{21}$$

This requirement, combined with the participation constraint (15), imposes two restrictions on the set of feasible contract allocations (both follow directly from results in Baron and Myerson (1982)). First, for an interior solution, a truthful mechanism requires that land use be monotonically non-decreasing in type:

$$a' (\theta) \geq 0. \tag{22}$$

Second, a truthful mechanism requires the change in expected surplus over type be decreasing at the rate

$$s' (\theta) = \pi_\theta \left(a (\theta), \theta \right) - \pi_\theta \left(\overline{a}, \theta \right). \tag{23}$$

Since surplus is decreasing, the best the principal can do while satisfying (15) and (23) is to set $s \left(\bar{\theta} \right) = 0$. Using Eq. (23), surplus is then a function of land allocations:

$$s (\theta) = \int_\theta^{\bar{\theta}} \pi_\theta \left(a (\omega), \omega \right) - \pi_\theta \left(\overline{a}, \omega \right) d\omega. \tag{24}$$

Temporarily ignoring (22), substitution of Eq. (24) into Eq. (16) and integrating by parts yields the following Lagrangian for the government's second-best problem:

$$
\begin{aligned}
L^{SB} &= \int_{\Theta} \left\{ \frac{F(\theta)}{f(\theta)} \left[\pi_\theta(\overline{a}, \theta) - \pi_\theta(a(\theta), \theta) \right] - \pi(a(\theta), \theta) + \pi(\overline{a}, \theta) \right\} dF(\theta) + \\
&\quad -\lambda \int_{\Theta} \left[\overline{a} - a(\theta) - A \right] dF(\theta) - \int_{\Theta} \mu(\theta) \left[\overline{a} - a(\theta) \right] d\theta.
\end{aligned} \tag{25}
$$

Consequently, an optimal second-best land allocation satisfies the following conditions:

$$
\lambda - \frac{F(\theta)\pi_{a\theta}(a(\theta), \theta) + \mu(\theta)}{f(\theta)} - \pi_a(a(\theta), \theta) \geq 0; \tag{26}
$$

$$
\left[a(\theta) - \underline{a} \right] \left[\lambda - \frac{F(\theta)\pi_{a\theta}(a(\theta), \theta) + \mu(\theta)}{f(\theta)} - \pi_a(a(\theta), \theta) \right] = 0; \tag{27}
$$

$$
\mu(\theta) \left[\overline{a} - a(\theta) \right] = 0; \qquad \lambda \int_{\Theta} \left\{ \left[\overline{a} - a(\theta) \right] - A \right\} dF(\theta) = 0; \tag{28}
$$

The properties of $F(\theta)$ and $\pi(a(\theta), \theta)$ ensure that this solution satisfies (22) (see Guesnerie and Laffont, 1984). The impact of asymmetric information can be easily seen for interior solutions. Unlike the first-best case, rather than having the marginal profit of land be equated for all farms, there is a distortion created by the term $F(\theta)\pi_{a\theta}/f(\theta)$. As a result, the equimarginal principal is never satisfied. This program could be implemented by the government requesting that producers choose a land allocation and total transfer payment from a menu of possible choices. For an interior solution, such a scheme would not result in a linear price per acre of land idled.

The Pigouvian subsidy program can be thought of as a hybrid of the first and second best programs. With this program, the transfer is $ta(\theta)$ rather than $t(\theta)$. The first order condition of incentive compatibility condition (21) requires that an interior solution satisfy:

$$
\pi_a(a(\theta), \theta) = t. \tag{29}
$$

Therefore, like the first-best program, for an interior solution the Pigouvian land allocation scheme equates marginal profit of land across producers. Moreover, in order to satisfy the environmental constraint (12), this land allocation must be exactly the same as for the first-best program. It is like the second-best program, however, in the sense that Eq. (23) must still be satisfied. With asymmetric information, the government cannot recover surplus payments from producers. The best it can do is ensure that the highest participating type receives zero surplus.

IV. Policy Simulations

Using the structural parameters estimated in Section II and the theoretical results from Section III, one can empirically characterize the three programs in terms of amount of land idled and transfer received by each type. In this section, I conduct simulations to evaluate two policy decisions. The first is the value of removing the information asymmetry. Suppose type were completely embodied in a measurable soil quality index. By comparing the cost of the first and second best mechanisms, we obtain the maximum amount the government should be willing to pay to collect the soil quality information. The second policy decision involves the value of policy reform. Revising policy inherently involves some degree of transition cost. By comparing the second-best program with the optimal linear-price program mechanism, we obtain the maximum amount the government should be willing to incur to develop an optimal policy without collecting additional information.

For purposes of the simulation, the profit function is defined as the average of $\pi(a, \theta)$ given prices of the four year period 1997-2000. Since about 5 percent of cropland in the Heartland region participates in the CRP, the expected land set aside target A is set to 100 acres for the 2,000-acre farms.[9] The lower bound of $\ln \theta$ with a half-normal distribution function is $-\infty$. To make the model tractable, I truncate the distribution from below at -1, thereby excluding roughly one percent of farms. The remaining farm types are assumed to be distributed with a mean-zero log-normal distribution truncated from above by zero and below by -1. The variable $\ln \theta$ is measured with a precision of 0.001. For each program, the numerical solution of the relevant necessary conditions characterizes the contract terms.

Figure 2 depicts the contract terms for the three programs. For all programs the majority of producer types fall into a corner solution. For the first-best and Pigouvian programs, the highest 88.6 percent of types do not participate, and the lowest 3.4 percent of types enroll the maximum amount of land. For the second-best program more types fall into both corner solutions with the corresponding figures of 89.8 and 4.2 percent respectively. The distortion in the land allocation caused by the information asymmetry is clearly visible in panel (a). By Eq. (24) total surplus payments decrease as the lowest non-participating type decreases. Consequently, relative to the first best the second-best program shifts enrolled acres away from higher types towards lower types.

Panel (b) shows that the cost difference caused by the information asymmetry is large compared to that of using the sub-optimal Pigouvian program. As shown in panels (b) and (c) the choice between the second best and Pigouvian pricing mechanisms also has implications for income distribution within the agricultural sector. The second-best mechanism reduces total surplus payments made to the sector, relative to a linear subsidy. This reduction in surplus comes exclusively at the expense of the more profitable farms. Relatively less profitable producers (those with a type lower than $e^{-0.8}$) would actually benefit from a policy switch from a Pigouvian to the second-best mechanism.

Assuming 58,000 farms (in 1997 there were approximately 116 million acres of cropland in the Heartland region), the total costs of the first-best, second-best, and Pigouvian programs are \$24.098, \$38.489, and \$38.547 million. Gains from policy reform are small, averaging about one penny per acre enrolled. Efforts undertaken to eliminate information asymmetry could be much more valuable, achieving at most a 37.5 percent reduction in total program costs relative to the second best.

V. Conclusion

Several recent articles have shown the usefulness of stochastic frontier techniques for analysis of regulation under asymmetric information. Here, I extend earlier results in two directions. First, I develop a GMM-based methodology for estimating a stochastic cost frontier for a profit-maximizing producer. This approach differs from earlier GMM and ML frontier techniques in that it accommodates multiple equations (in this case a cost equation, expenditure share equations, and the ratio of revenue to cost) and is robust to arbitrary cross-equation correlation, heteroskedasticity, and geographic clustering. Further, it is easy to estimate as it does not require non-linear optimization.

Second, I extend the empirical contract theory literature by using stochastic frontier analysis to estimate the technology of a heterogeneous sample of producers in an unregulated sector. Although the econometrician cannot directly observe producer type, this approach permits estimation of the probability distribution of types in the population. This information provides the necessary material for empirically specifying an optimal contract mechanism in the Baron and Myerson framework.

Application of this methodology to a simulated model of a voluntary environmental program for land retirement in the agricultural sector yields some interesting results. The simulation permits comparison of an optimal second best mechanism to the optimal full information mechanism and a sub-optimal Pigouvian subsidy mechanism. In reality, actual policy is a mix of the three. Producer types are unobservable, but the government uses soil maps and county average land rents to proxy for type. The policy is nominally structured as an auction in order to induce producers to reveal the true opportunity cost of their land. In practice, however, the program is similar to a Pigouvian subsidy since most producers bid at or near the maximum permissible rental rate.

By examining three "pure" hypothetical programs, the analysis conducted here provides guidance to policy makers interested in reducing the cost of the environmental policy. For example, if the second-best program closely approximates the full information program, efforts may be better spent redesigning the program to make it function more like a true auction rather than gathering detailed soil data and developing agronomic/economic models linking soil type to overall profit. The results of the exercise indicate the contrary, however. The second-best mechanism performs only slightly better than a linear Pigouvian subsidy. Even absent any transition costs, the savings from perfecting the program (given that type is unobservable) are in the order of one penny per acre. More significant cost reductions can only be obtained by gathering data to overcome the information asymmetry.

Notes

[1]Although agriculture may not leap to mind as an unregulated profit-maximizing sector, since 1996 government payments have largely been independent of farm production decisions.

[2]Kopp and Mullahy (1990) was the first to use GMM techniques to estimate a cost frontier. Kleit and Terrell (2001) and Knittel (2002) have used stochastic frontier analysis to infer the impact of regulatory framework on costs in the electricity sector.

[3]All computations in Sections II and IV were programmed in Gauss 5.0.

[4]Disaggregated outputs include barley, canola, cotton, fruit, hay, oats, potatoes, rice,

sorghum, soybeans, sugar beets, tobacco, vegetables, wheat, cattle, dairy, eggs, hogs, and poultry. Capital includes farm machinery and vehicles. Labor includes both farm household and hired labor. Energy inputs are diesel, gasoline, liquid propane gas, and electricity. The materials category consists of seed, fertilizer, chemicals, supplies, feed, livestock inputs, poultry inputs, custom work, and repairs.

[5]The USDA Economic Research Service divides the country into "Farm Resource Regions" with similar physiographic, soil, and climatic characteristics.

[6]Data on government payments, acreage enrollment, and rental rates come from Environmental Working Group (2005), Farm Service Agency (2005a), and National Agricultural Statistics Service (2001), respectively.

[7]For more complex model incorporating additional aspects of the program, see Smith (1995).

[8]The actual CRP has a cap limiting enrollment to 25 percent of farmland in any county.

[9]Information on regional cropland and CRP participation comes from Vesterby (2002) and Economic Research Service (2003).

References

Aigner, Dennis, C. A. Knox Lovell, and Peter Schmidt, "Formulation and Estimation of Stochastic Frontier Production Function Models," Journal of Econometrics, 6(1) (1977), 21–37.

Baron, David P. and Roger B. Myerson, "Regulating a Monopolist with Unknown Costs," Econometrica, 50(4) (1982), 911–930.

Caves, Douglas W., Lauritis R. Christensen, and W. Erwin Diewert, "Multilateral Comparisons of Output, Input and Productivity Using Superlative Index Numbers," The Economic Journal, 92(365) (1982), 73–86.

Dalen, Dag M. and Andrés Gómez-Lobo, "Estimating Cost Functions in Regulated Industries Characterized by Asymmetric Information," European Economic Review, 41(3-5) (1997), 935–942.

Diewert, W. Erwin, "Duality Approaches to Microeconomic Theory," in Kenneth J. Arrow and Michael D. Intriligator, (Eds.) Handbook of Mathematical Economics (New York: North-Holland, 1982).

Economic Research Service, Farm Resource Regions, Agricultural Information Bulletin 760 (Washington, DC: United States Department of Agriculture, 2000).

Economic Research Service, Agricultural Resource and Environmental Indicators, Agriculture Handbook No. (AH722) (Washington, DC: U.S. Department of Agriculture, 2003).

Environmental Working Group, "Farm Subsidy Database," (2005), http://www.ewg.org/farm.

Farm Service Agency, Comparison of 2002 PFC and DCP Bases and Yields by State (Washington, DC: United States Department of Agriculture, 2005a), http://www.fsa.usda.gov/pas/farmbill/2002_2003_enroll.htm.

Farm Service Agency, Conservation Reserve Program: Enrollment Statistics and Program Information, 2004 Fiscal Year (Washington D.C.: U.S. Department of Agriculture, 2005b), http://www.fsa.usda.gov/dafp/cepd/stats/FY2004.pdf.

Gagnepain, Philippe and Marc Ivaldi, "Incentive Regulatory Policies: The Case of Public Transit Systems in France," RAND Journal of Economics, 33(4) (2002), 605–629.

Guesnerie, Roger and Jean-Jacques Laffont, "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm," Journal of Public Economics, 25 (1984), 329–369.

Hall, Robert E. and Dale W. Jorgenson, "Tax Policy and Investment Behavior," American Economic Review, 57(3) (1969), 391–414.

Kleit, Andrew N. and Dek Terrell, "Measuring Potential Efficiency Gains from Deregulation of Electricity Generation: A Bayesian Approach," Review of Economics and Statistics, 83(3) (2001), 523–530.

Knittel, Christopher R., "Alternative Regulatory Methods and Firm Efficiency: Stochastic Frontier Evidence from the U.S. Electricity Industry," Review of Economics and Statistics, 84 (2002), 530–540.

Kopp, Raymond J. and John Mullahy, "Moment-Based Estimation and Testing of Stochastic Frontier Models," Journal of Econometrics, 46(1-2) (1990), 165–183.

Laffont, Jean-Jacques and Jean Tirole, "Using Cost Observation to Regulate Firms," Journal of Political Economy, 94(3) (1986), 614–641.

Lavergne, Pascal and Alban Thomas, "Semiparametric Estimation and Testing in a Model of Environmental Regulation with Adverse Selection," Empirical Economics, 30(1) (2005), 171–192.

Meeusen, Wim and Julien Van Den Broeck, "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error," International Economic Review, 18(2) (1977), 435–444.

Myerson, Roger B., "Incentive Compatibility and the Bargaining Problem," Econometrica, 47(1) (1979), 61–73.

National Agricultural Statistics Service, Agricultural Cash Rents 2001 Summary (Washington, DC: United States Department of Agriculture, 2001).

Newey, Whitney K., "A Method of Moments Interpretation of Sequential Estimators," Economics Letters, 14(2-3) (1984), 201–206.

Pepper, John V., "Robust Inferences from Random Clustered Samples: An Application Using Data from the Panel Study of Income Dynamics," Economics Letters, 75(3) (2002), 341–345.

Ray, Subhash C., "A Translog Cost Function Analysis of U.S. Agriculture, 1939-77," American Journal of Agricultural Economics, 64(3) (1982), 490–498.

Smith, Rodney B., "The Conservation Reserve Program as a Least-Cost Land Retirement Mechanism," American Journal of Agricultural Economics, 77(1) (1995), 93–105.

Thomas, Alban, "Regulating Pollution under Asymmetric Information: The Case of Industrial Wastewater Treatment," Journal of Environmental Economics and Management, 28(3) (1995), 357–373.

Vesterby, Marlow, "Land Use, Value, and Management: Major Land Uses," in Economic Research Service Briefing Rooms (Washington, DC: U.S. Department of Agriculture, 2002), http://www.ers.usda.gov/Briefing/LandUse/Gallery/croplandfrreg.htm.

Wolak, Frank, "An Econometric Analysis of the Asymmetric Information, Regulator-Utility Interaction," Annales d'économie et de statistique, (34) (1994), 13–69.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data* (Cambridge: MIT Press, 2002).

Wooldridge, Jeffrey M., "Cluster-Sample Methods in Applied Econometrics," *American Economic Review*, 93(2) (2003), 133–138.

Figure 1: Type and Productivity

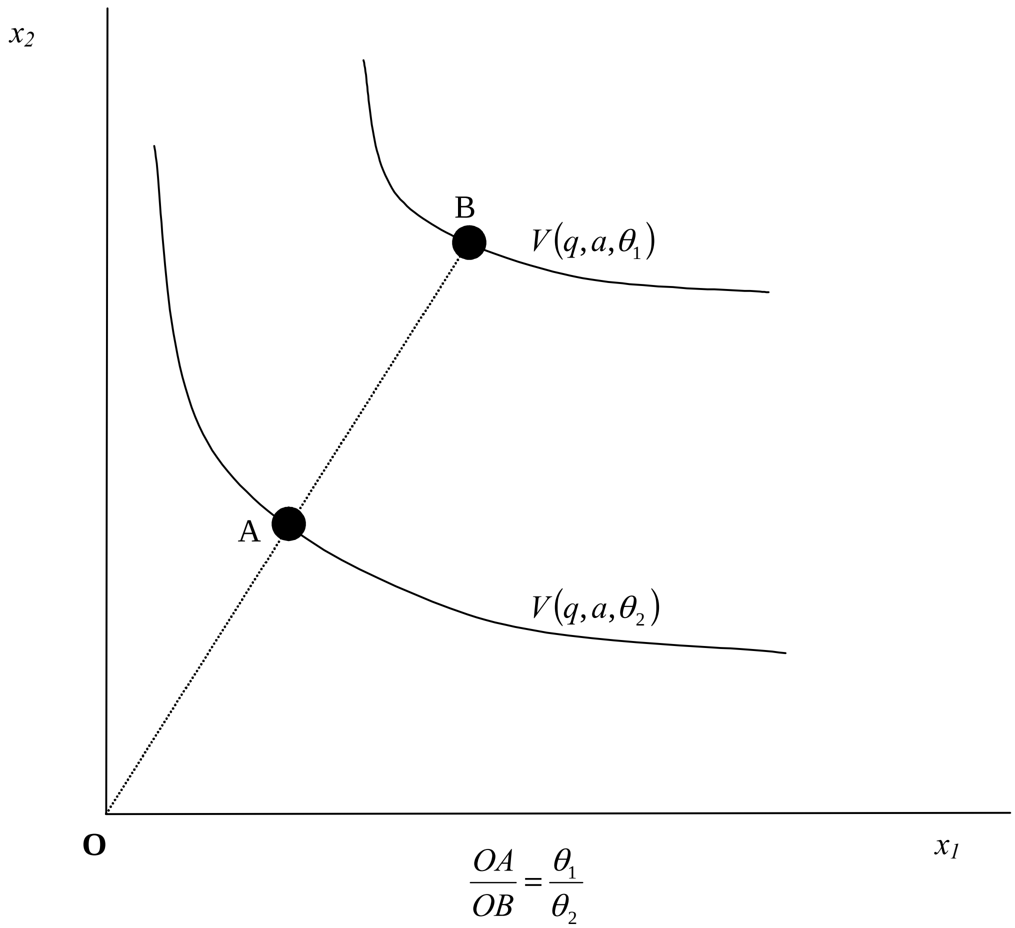

Table 1: Summary Statistics

	Mean	Standard deviation
Cost in current dollars (x1,000)	433	2,813
Input expenditure shares		
Capital	0.379	0.151
Labor	0.173	0.115
Energy	0.042	0.029
Materials	0.405	0.161
Revenue/cost	1.20	0.831
Output index (x1,000)	525	2,743
Acres	988	984
Price indices		
Output	1.005	0.098
Capital	0.999	0.033
Labor	1.013	0.129
Energy	1.001	0.029
Materials	1.001	0.099
Number of observations	5,528	

Table 2: Parameter Estimates

Variable	Value	Standard error
1997	0.6508	0.0457***
1998	0.7022	0.0455***
1999	0.6385	0.0452***
2000	0.3020	0.0453***
ln Capital Price	0.3906	0.0024***
ln Labor Price	0.1652	0.0043***
ln Energy Price	0.0443	0.0029***
ln Materials Price	0.3999	0.0026***
ln Output	1.2021	0.0024***
ln Acres	-0.8756	0.1374***
ln Acres squared	0.0500	0.0105***
σ	0.4274	0.3228*

Notes: The first four variables are year dummies. Standard errors adjusted for county clustering and and heteroskedasticity.
***Significantly greater than zero at 1-percent level.
*Significantly greater than zero at 10-percent level.

Figure 2: First Best, Second Best, and Pigouvian Contract Characteristics