Econometric-Process Models for Integrated Assessment of Agricultural Production Systems

John M. Antle
Professor

and

Susan M. Capalbo
Associate Professor

Agricultural Economics and Economics
Montana State University–Bozeman

Research Discussion Paper No. 40
March 2000

Abstract

This paper develops the conceptual and empirical basis for a class of empirical economic production models that can be linked to site-specific bio-physical models for use in integrated assessment research. Site-specific data are used to estimate econometric production models, and these data and models are then incorporated into a simulation model that represents the decision making process of the farmer as a sequence of discrete or continuous land use and input use decisions. This discrete/continuous structure of the econometric process model is able to simulate decision making both within and outside the range of observed data in a way that is consistent with economic theory and with site-specific bio-physical constraints and processes. An econometric-process model of the dryland grain production system of the Northern Plains demonstrates the capabilities of this type of model.

Econometric-Process Models for Integrated Assessment of Agricultural Production Systems

Quantitative integrated assessment – i.e., the use of linked disciplinary simulation models to evaluate complex natural and human systems – is becoming the standard methodology for analysis of many leading environmental issues. The purpose of integrated assessment is often to simulate behavior of a system outside the range of observed behavior, as in analysis of the impacts of global climate change. To simulate non-linearities or discontinuities of complex systems, biological and physical scientists use models based on their understanding of bio-physical processes. These process-based models are able to simulate behavior outside the range of observed data in ways that are consistent with established scientific understanding. Another feature of bio-physical processes in agriculture is that they are dependent on site-specific soil and climate conditions.

This paper develops the conceptual and empirical basis for a class of empirical economic production models — *econometric-process models* — that can be linked to site-specific bio-physical models for use in integrated assessment research. Site-specific data are used to estimate econometric production models, and these data and models are then incorporated into a simulation model that represents the decision making process of the farmer as a sequence of discrete or continuous land use and input use decisions. This discrete/continuous structure of the econometric-process model is able to simulate decision making both within and outside the range of observed data in a way that is consistent with economic theory and with site-specific bio-physical constraints and processes.

The next section reviews the production modeling approaches that have been used in integrated assessment research. The third section presents the conceptual foundations for the econometric-process simulation model and discusses the empirical procedures that can be used to implement the approach. The fourth section presents an application of the econometric-process approach to modeling the dryland grain production system in Montana. This application demonstrates the capabilities of the econometric-process model to reproduce within-sample distributions of land use and out-of-sample phenomena such as the conversion of cropland to a conserving use. The example also shows that the econometric-process model can simulate nonlinearilites and discontinuities in supply functions outside the range of observed data, such as the shutdown point on the supply function, that are consistent with economic theory but are not estimable with conventional, continuous econometric supply functions.

Economic Production Models for Integrated Assessment

The integrated assessment paradigm for agricultural production systems is presented as follows. Economic data are inputs into economic production models, and soils and climate data are inputs into crop or livestock process models that calculate site-specific productivity. The outputs of crop and livestock models may be inputs into economic models and environmental process models (e.g., models of chemical leaching and runoff, soil erosion, or changes in soil organic carbon). The outcomes of economic models also may be inputs into environmental process models. If the bio-physical and economic data are statistically representative of the population of land units and economic decision makers in a region, economic and environmental outcomes can be statistically aggregated to assess tradeoffs at a regional scale (Antle, Capalbo and Crissman, 1998).

A variety of production modeling approaches have been used in integrated assessment research. One approach utilizes representative farm programming models to estimate optimal resource allocations. Prato et al. use this approach to link economic models environmental process models; Kaiser et al. developed a representative farm model for Minnesota to examine the impacts of climate change; and Adams et al. (1995, 1998) utilize an aggregate model with representative farms for U.S. regions to study the regional impacts of climate change on U.S. agriculture. Kruseman et al. develop a bio-economic modeling approach that integrates bio-physical information with linear programming models. These models allow for discrete choices among technologies and land use. Another important feature of these models is that they represent production technology explicitly, so they can be linked to bio-physical process models of crop or livestock production. However, the reliance of these models on the representative farm construct limits their usefulness for explaining spatial variation in economic behavior and linking that behavior to spatially-explicit bio-physical process models. Another limitation of programming models is that their technology parameters and related data are not usually derived from statistically representative samples of the population.

A second class of production models is based on econometric models that explain observed outcomes, such as land use or net returns, as reduced-form functions of economic variables (output and input prices) and bio-physical characteristics of land units. Mendelsohn, Nordhaus and Shaw developed an econometric model that explains returns to land as a function of economic and climate variables, using U.S. county-level data. Recent studies of land allocation also use reduced-form models estimated with county-level data for regions of the U.S. (Wu and Segerson; Hardie and Parks.). Because reduced-form models do not explicitly represent the relationship between productivity and the physical environment, they cannot be linked to bio-physical process models of crop or livestock production. This feature limits the usefulness of these models for integrated assessment. For example, studies of global climate change show that increasing concentrations of atmospheric CO_2 are likely to have substantial impacts on crop yields, and thus should have significant impacts on economic decisions. This CO_2 fertilization effect cannot be incorporated into reduced-form economic models.

A third strand of the literature utilizes econometric methods to estimate neoclassical production, cost, or profit functions (e.g., Chambers; Segerson and Dixon). These models can be estimated and simulated with site-specific data, and thus can be used to represent spatial variation in both bio-physical conditions and economic behavior. They can also explicitly represent the impacts of bio-physical conditions on productivity. However, the parameters of econometric models can only represent the range of behavior observed within the spatial and temporal dimensions of the data used for their estimation. Moreover, econometric models that incorporate mixed discrete and continuous choices are difficult to estimate and simulate. Consequently, these models are not well suited to the simulation of discrete choices or behavior outside the range of observed data.

Most empirical economic production models do not incorporate biophysical data and information about growth processes. Leaving site-specific soil and climate variables out of a production function may lead to biased and inconsistent parameter estimates, analogous to the farm-specific management bias (Mundlak and Hoch). This bias can be avoided by estimating a dual representation of the technology in which environmental effects are relegated to error terms, but then cannot be linked to bio-physical data in simulation analysis. Some studies have included bio-physical data in a production function model (e.g., Kaufman and Snell). This type of procedure can introduce site-specificity into a production function model, and could be incorporated into an econometric-process simulation model.

Econometric-Process Simulation Models: Conceptual Foundations and Empirical Methods

The production process of activity j at site i in period t is defined by a technology set $\mathbf{v}_{ijt} \varepsilon \mathbf{T}(\mathbf{z}_{ijt}, \mathbf{e}_{it})$ where $\mathbf{v}$ is a vector of variable inputs, $\mathbf{z}$ is a vector of allocatable quasi-fixed factors of production and other fixed effects, and $\mathbf{e}$ is a vector of bio-physical characteristics of the site (soils, topography, climate, etc.). The production function for each activity is non-joint in inputs and can be written as $q_{ijt} = f(\mathbf{v}_{ijt}, \mathbf{z}_{ijt}, \mathbf{e}_{it})$, where $f(\bullet)$ and its parameters are independent of i and t (random terms are suppressed here for notational convenience). For expected output price p_{ijt}, the profit function is $\pi_{ijt} = \pi_j(p_{ijt}, \mathbf{w}_{ijt}, \mathbf{z}_{ijt}, \mathbf{e}_{it})$. If a crop is not grown, the crop is in a conserving use with a return of π_{ict}. Define $\delta_{ijt} = 1$ if the jth crop is grown at location i at time t and zero otherwise. The land use decision on site i at time t is

$$(1) \qquad \max_{(\delta_{i1t}, \ldots, \delta_{int})} \sum_{j=1}^{n} \delta_{ijt}\, \pi_j(p_{ijt}, \mathbf{w}_{ijt}, \mathbf{z}_{ijt}, \mathbf{e}_{it}) + (1 - \sum_{j=1}^{n} \delta_{ijt})\, \pi_{ict}.$$

The solution takes the form of a discrete step function

$$(2) \qquad \delta^*_{ijt} = \delta_j(\mathbf{p}_{it}, \mathbf{w}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \pi_{ict}),$$

where $\mathbf{p}_{it}$ is a vector of the p_{ijt} and likewise for the other vectors. Using Hotelling's lemma, the quantity of the jth output on the ith land unit is given by

$$(3) \qquad q^*_{ijt} = \delta^*_{ijt}\, \partial\pi_j(p_{ijt}, \mathbf{w}_{ijt}, \mathbf{z}_{ijt}, \mathbf{e}_{it})/\partial p_{ijt} = q_{ijt}(\mathbf{p}_{it}, \mathbf{w}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \pi_{ict}).$$

Variable input demands are likewise given by

$$(4) \qquad \mathbf{v}^*_{ijt} = -\,\delta^*_{ijt}\, \partial\pi_j(p_{ijt}, \mathbf{w}_{ijt}, \mathbf{z}_{ijt}, \mathbf{e}_{it})/\partial \mathbf{w}_{ijt} = v_{ijt}(\mathbf{p}_{it}, \mathbf{w}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \pi_{ict}).$$

Equations (2), (3) and (4) define the model of discrete, extensive margin (land use) decisions and continuous, intensive margin (supply and input use) decisions. This analysis assumes that net returns above variable cost are positive for at least one activity, otherwise the land is left idle as in the conventional analysis of the firm's shut-down decision.

The solution to (1) applies to a given land unit. Each land unit is managed separately under the assumptions that farmers are risk-neutral, sell their products into a well-functioning market (as opposed to subsistence farmers who produce for own consumption), and have access to well-functioning rental markets for land and capital. Risk aversion, or the lack of well functioning input or output markets (including rental markets for capital), may cause production decisions to be interrelated across land units (de Janvry, Fafchamps and Sadoulet; Just, Zilberman and Hochman). The above model can be generalized to address each of these cases. However, the simpler model presented above is likely to be a good first-order approximation under conditions of commercial agricultural production in countries like the United States where markets for products and inputs are well-developed.

Incorporating Crop Rotations

Crop rotations play a critical role in maintenance of soil quality and productivity. The effects of crop rotations can be accurately modeled only on a site-specific basis, because their representation requires site-specific data on the history of land use. Aggregation across fields, even at the farm level, prevents the dynamics of soil quality from being accurately represented in both economic and bio-physical process models. Accurate representation of some processes may even require sub-field analysis, as when soil erosion occurs at different rates within a field (Antle and Stoorvogel).

To illustrate, consider a simple crop/fallow rotation. The use of rotations and the length of rotations is an economic decision involving a tradeoff between the management and opportunity cost of the fallow and the productivity gains associated with the rotation. For each crop and location (deleting subscripts i and j for notational convenience) the production function takes the form $q_t = f(\mathbf{v}_t, \mathbf{z}_t, \mathbf{e}, \lambda_{t-1})$, where $\lambda_{t-1} = 1$ if the previous use was a crop and equals zero if the field was fallowed, hence $f(\mathbf{v}_t, \mathbf{z}_t, \mathbf{e}, 1) < f(\mathbf{v}_t, \mathbf{z}_t, \mathbf{e}, 0)$. If a unit of land was previously cropped, the decision to fallow this season with the intent to crop again next season is based on net returns above variable cost calculated as $(p_{t+1}q_{t+1} - vc_{t+1})(1/1+r) - fc_t$, where vc_{t+1} is variable cost of crop production, r is the interest rate, and fc_t is the variable cost associated with fallow. The profit function takes the form $\pi_{fal}(p_{t+1}, w_{t+1}, r, \mathbf{z}_t, \mathbf{e}, \lambda_{t-1})$. The returns to growing a crop in period t after a crop was grown in period t-1 is equal to $p_tq_t - vc_t$, giving the profit function $\pi_{crop}(p_t, w_t, \mathbf{z}_t, \mathbf{e}, \lambda_{t-1})$. The farmer will use fallow if $\pi_{fal} > \pi_{crop}$. The same logic can also be applied to the case wherein the field was fallowed in the previous period.

Incorporating Bio-Physical Data and Crop or Livestock Growth Processes

From an econometric point of view, we can interpret a process-based crop (or livestock) model as producing a yield or production estimate $q^p(\mathbf{v}_i^0, \mathbf{z}_i^0, \mathbf{e}_i)$ for the ith site, where $\mathbf{v}_i^0$ and $\mathbf{z}_i^0$ are the management practices specified for the crop simulation model. For example, yields in the widely-used DSSAT system of crop models are limited by water and nutrients but not pests, so an optimum level of pest management is implicitly assumed when these models are used (Tsuji, Uehara and Balas). Therefore, yield estimates from these models are not comparable to yields measured under field conditions where farmers' actual management practices are in use. However, $q^p(\bullet)$ does embody the process-based relationship between site-specific environmental conditions $\mathbf{e}_i$ and crop productivity, and thus can be interpreted as a site-specific index of productivity. The Appendix shows, using the concept of the frontier production function, that the site-specific production function can be expressed as

(5) $q_i = f(\mathbf{v}_i, \mathbf{z}_i, q^p(\mathbf{v}^0, \mathbf{z}^0, \mathbf{e}_i)),$

where the function $f(\bullet)$ has the properties of a production function and is increasing in q^p. This specification has two important implications for production analysis. First, it implies separability between **e** and the other variables in the production function, a testable hypothesis. Second, equation (5) shows that yield estimates may be appropriately interpreted as indicators of productivity at one site relative to another site or relative to the population mean. Therefore, as we shall discuss below, these yield estimates can be included in econometric production models as explanatory variables, and they can subsequently be used to incorporate the effects of unobserved bio-physical conditions on productivity and behavior outside the range of observed data used in the econometric analysis.

Econometric Procedures

Econometric specification and estimation of the production model described above must account for the discrete structure of land use decisions, the dynamics of crop rotations, the spatial variation in physical conditions, statistical properties of the spatial data, and features of the farmer's management behavior. Because of spatial variation in land use, a cross-section of data will have an unbalanced property, i.e., not all production activities will be undertaken by all producers. This means that a multi-product model that assumes positive values for all outputs on all land units, as is typically assumed in duality-based multi-output production models in the literature, is not appropriate. Huffman proposed an econometric methodology to address the zero output problem for aggregate multi-product models that are joint in inputs. This model is not appropriate under the assumption made here of non-joint production at the field level.

Another feature of production data, particularly at the farm-level and field-level, is that observed net returns may be negative even though farmers choose to enter a unit of land into production, thus implying that *expected* returns were positive. Consequently, most non-linear-in-parameters functional forms for profit functions, such as the Cobb-Douglas, constant elasticity of substitution, and translog, are not amenable to applied production analysis at the farm or field scale.

Another issue that arises is the specification of the model's error structure to account for the use of spatial data (Anselin). Climatic events may cause error terms to be correlated across land units, and correct specification of these spatial correlations is required to obtain accurate estimates of parameters' standard errors. However, a number of factors limit the practicality of using econometric models with non-diagonal error covariance matrixes to account for spatial correlation. First, with the single cross-section of data typically available from surveys of site-specific production practices, general covariance structures cannot be estimated, so some *ad hoc* structure must be imposed on the spatial covariance structure (e.g., spatial autoregression structures). These assumptions can be tested relative to the null hypothesis of no spatial correlation (Anselin et al.). However, a second complicating factor is that agricultural production does not occur uniformly distributed through space, so spatial covariances are not likely to follow a regular structure such as is typically assumed in the literature. Under these circumstances, one cannot argue that a model is necessarily better specified with an arbitrary non-zero spatial covariance than with the assumption of zero spatial correlation. The use of the standard diagonal covariance assumption is further reinforced by recent findings that show conventional F-tests and t-tests to be robust to covariance misspecification (Banerjee and Magnus).

Several econometric specifications are available for field-scale models. A first choice would be to specify a multi-output production model along the lines in the literature discussed above. These models generally assume that all outputs are produced by all farms, thus they fail to account for discrete choice among production activities and the unbalanced data that result from these discrete choices.

Another approach would be to utilize a multinomial discrete choice model (e.g., logit or probit) to estimate the vector of reduced-form land use functions analogous to equation (2), and a system of factor demand equations such as (4) to represent management decisions conditional on land use decisions. While this approach provides an efficient means to estimate the probability of each land use alternative in a consistent manner, it has several practical limitations.

First, the discrete choice model is in reduced form and thus faces the limitations of models in which productivity is not represented explicitly. This problem could be solved by replacing the vector **e** with a yield estimate from a crop model as discussed in the Appendix. A second, related problem is that yields are observed only for the crop that is produced at the site, whereas the decision between alternative crops involves a comparison of productivity between all possible crops. This difficulty is magnified in an integrated assessment that involves simulation of unobserved conditions, e.g., a perturbed climate scenario, in which existing crops may be uncompetitive and new crops may become profitable. The productivity of alternative crops could be estimated using either a statistical model (e.g., an econometric supply function) or a bio-physical simulation model, subject again to the limitations discussed earlier (simulated yields are not directly comparable to actual yields).

Several other practical problems complicate the use of an explicit discrete choice model. The estimation and simulation of multi-dimensional probability distributions required for the discrete land use decision problem poses difficult computational problems, and available software limits this type of analysis to the logistic and normal distributions without any clear justification of these functional forms. Adaptation of these models to account for spatial dependence would be even more challenging than with conventional linear statistical models. A final limitation is that this approach would provide only reduced-form estimates of land use choices, but would not provide estimates of production, cost and net returns that are of interest to policy makers.

Design of Econometric-Process Simulation Models

The approach to production modeling for integrated assessment proposed here is closely related to the models discussed above, with an important difference. The econometric-process approach combines econometric production models, represented by the system of supply and demand functions (3) and (4), with a process-based representation of the discrete land use decision represented by (1) and (2). This combination of econometric and process-based models overcomes the various limitations associated with other modeling approaches, as we now show. The general structure of an econometric-process simulation model is illustrated in Figure 1. The upper part of the figure shows the key steps in estimation of the econometric models, while the lower part shows the structure of the simulation model. The model simulates the farm manager's crop choice defined in equation (1), and the related output and cost of production for that crop choice, at the field scale, over space and time. This simulation model utilizes the stochastic properties of the econometric models and the sample data, so its output can be interpreted as providing a statistical representation of the *population* of land units in an agricultural region. By operating at the field scale with site-specific data, the simulation can represent spatial and temporal differences in land use and management, such as crop rotations, that give rise to different economic outcomes across space and time in the region.

The econometric production models are used to calculate expected net returns in the simulation of the land use decisions (more generally, objectives that incorporate risk and other decision criteria could be used). Since observed net returns may be negative, a system of a supply function, cost function and corresponding input demand functions is used rather than a profit function. The revenue component of expected returns is equal to the output price times the supply function derived from the restricted profit function, and cost computed with the cost function. This system of a supply function and a cost function is theoretically equivalent to a profit function, and joint estimation produces efficient parameter estimates.

For the simulation model, each field is described by total acres, location, and a set of location-specific prices paid and received by producers, and quantities of inputs. Using sample distributions estimated from the data, draws are made with respect to expected output prices, input prices, and any other site-specific management factors (e.g., previous land use). The econometric production models are simulated to estimate expected output, costs of production, and expected returns. The land use decision (1) for each site is made by comparing expected returns for each production activity. If the selected land use is a production activity, the corresponding system of factor demand equations is simulated to determine input use at the site. This system of factor demand equations may be specified in static form as in equation (4), or a dynamic system may be derived from a sequential decision model (Antle, Capalbo and Crissman, 1994). These spatially and temporally explicit land use and management decisions may be subsequently used as input into biophysical process models in an integrated assessment (e.g., erosion, chemical leaching, changes in soil carbon). Land use and other decisions in each period are used to initialize the computation of expected returns for the subsequent period as the decision making process is simulated through time.

By carrying out these simulations for a statistically representative set of fields or farms, the economic and environmental outcomes of the simulations can be interpreted as characterizing the empirical joint distribution of these outcomes in the population of land units or farms. This empirical joint distribution corresponds to the theoretical joint distribution of outcomes as described by Just and Antle, and by Antle and Just. The site-specific economic and environmental outcomes can be statistically aggregated to assess impacts at larger spatial scale. The impacts of alternative policy or technology scenarios can be analyzed by changing corresponding parameters in the economic and bio-physical process models.

Real-world behavior of farmers — in terms of land use and management — is highly spatially variable in most cases due to spatial variability in soil, climate and economic conditions. The econometric-process simulation model (Figure 1) differs fundamentally from a deterministic optimization model (e.g., a linear or non-linear programming model) in the way that land allocation decisions are represented, and as a result, it provides a more realistic representation of the spatial distribution of land use. In a deterministic optimization model, expected returns are compared for alternative activities as functions of prices, technology parameters and resource constraints. The same economically optimal activity is attributed to all land units that are represented by a given parameterization of the model, hence, the same decisions are attributed to these land units with probability one. In the econometric-process model, economic decisions are based on the *spatial and temporal distributions* of expected returns associated with each alternative land use or input choice. There is a positive probability that each feasible activity will be selected at each site. Thus, as repeated draws are made from the underlying statistical distributions, a realistic spatial and temporal distribution of competing activities is obtained. Unrealistic outcomes such as corner solutions can only be obtained as a limiting case in which one activity economically dominates other activities at all sites or all time periods.

An Application to Dryland Grain Production

This section describes the application of the modeling approach to the dryland grain production system of Montana, using farm- and field-level production data collected in a survey designed to be statistically representative of the grain producing areas of the state, stratified by the USDA's Major Land Resource Areas (MLRA). Detailed descriptions of the data and summary statistics are found in Johnson et al., and in Antle, Capalbo, Johnson and Miljkovic.

The econometric production models for Montana grain crops (winter wheat, spring wheat, and barley) were specified as log-linear supply functions and variable cost functions. A log-linear equation was also included to represent machinery operating costs. The supply and machinery cost equations are functions of prices for fertilizer and herbicide inputs normalized by output price, field size, and a land use indicator (fallow or crop). The MLRAs were stratified into high and low precipitation sub-zones according to historical climate data (Paustian et al.). Dummy variables for these zones were included to capture systematic differences in productivity across sites. A variety of herbicides are used by farmers, and a hedonic procedure was used to quality-adjust these input data (Antle, Capalbo and Crissman, 1994). Expected crop prices were defined as average prices received in a farmer's region net of transportation costs to the nearest grain elevator.

Sample Selection Bias and Production Risk

The problem of unbalanced data is solved in the econometric-process modeling approach by estimating a non-joint production model for each crop. The decision problem in (1) implies a sample selection process based on expected returns that may give rise to sample selection bias in the estimation of the individual crop production models. Application of the Heckman two-stage procedure to test for sample selection bias showed that there was sample selection bias only in the barley supply equation and in the machinery cost equation for spring wheat. However, the supply and cost function parameter estimates were virtually identical to the parameters estimates obtained without use of the Heckman procedure. It was concluded that sample selection bias did not have a discernible effect on the parameter estimates.

A plausible alternative to the risk-neutral (profit-maximization) model used here is an model with risk aversion. It is often suggested that farmers use a crop/fallow rotation to reduce production risk associated with low soil moisture. This hypothesis was tested applying the production risk models of Just and Pope (1978) and Antle (1983). This analysis did not support the hypothesis that crops grown on fallowed fields had lower yield risk than crops grown after a crop. This evidence implies that farmers use the crop/fallow rotation when and where the *expected* profitability is higher than continuous cropping, not because yield risk is lower for the crop/fallow rotation, so the risk-neutral model was maintained in the following analysis.

Econometric Production Model Estimates

The econometric models for winter wheat, spring wheat and barley crops were specified in log-linear form, and were estimated using non-linear three-stage least squares with linear homogeneity of the cost function and zero-degree homogeneity of the supply function imposed. The parameter estimates in Table 1 show that the quantity supplied and machinery costs are approximately proportional to field size. Supply (and thus yield) and machinery cost also vary significantly by sub-MLRA. Using sub-MLRA 52-low as the basline, the yields are significantly lower for less productive regions (MLRAs 53a, 54, and 58a) and higher for the more productive sub-MLRA 52 high. Costs differ by sub-MLRA although not in a systematic way. The fertilizer and pesticide price parameters in the supply functions have the theoretically predicted negative sign. Noting that the supply functions are estimated with linear homogeneity in prices imposed, these parameters imply short run supply elasticities with respect to the output price of about 0.36 for winter wheat, 0.14 for spring wheat, and 0.35 for barley.

Table 1 also shows that winter wheat, spring wheat and barley yields are about 31, 23 and 9 percent higher when the crop is grown after fallow (these percent changes are calculated as e^{d} - 1, where *d* is the parameter of the fallow dummy variable). In the cost function, the fallow dummy variable shows that variable costs of production are about 40 percent lower for all three crops after fallow. These results confirm the hypothesis that fallowed fields are more productive than those that are continuously cropped. When fallow costs are considered, the data show that the crop/fallow system and the continuous cropping system yield similar net returns on average, explaining the fact that land in the region is allocated to each type of system in roughly equal proportions.

Simulation Model Calibration and Validation

The simulation model was calibrated to predict the observed mean frequencies of crops produced in the sample data. The model was calibrated using three parameters: the expected yield variability, the discount rate, and the expected future crop price. The expected yield variability refers to the variance of yield expectations in the population. The estimated econometric supply function provides an estimate of the population mean supply and yield, but it seems doubtful that all farmers form the same yield or output expectations, even when they face the same economic and bio-physical conditions. Presumably, the variance of yield expectations in the population of farmers is less than the variance of observed yields. The base simulations used an expected yield variance that is 90 percent of the observed variance. Analysis showed that the simulation results were not highly sensitive to this parameter.

As discussed earlier, the expected present value of returns to a crop/fallow rotation depends on a nominal discount rate. In the base model simulations, this discount rate was set at 7 percent. The net returns for crops produced after fallow also involve expected future crop prices. Because these prices will not be realized until the end of the next crop year, they are highly uncertain. To represent this uncertainty, and to account for the fact that in 1995 prices were above the long-run trend in real crop prices, the expected future crop prices were assumed to be less than the average observed market price in 1995. The base simulations utilized the assumption that the future crop prices were random variables with a mean 10 percent below the 1995 average market price, and with a variance equal to the observed variance prices. Analysis showed that the choice between continuous cropping and a crop/fallow rotation is sensitive to both the discount rate and future expected prices. This finding reflects the fact that the two systems are competitive, so that small changes in expected future prices relative to current prices can induce a farmer to modify the choice between continuous cropping and the use of fallow.

To provide a validation of the model, the observed proportion of each land use in each sub-MLRA was computed and compared to the simulated proportions. Figure 2 shows that the plot of observed and simulated mean land use falls along a 45-degree line, an indication that the simulation model does reproduce the observed data without a systematic bias. It is useful to note that the site-specific land-use data follow a binomial distribution for each use (i.e., the data are coded 1 if use j occurs and zero otherwise). It follows that the sample proportions plotted in Figure 2 are sufficient statistics for the entire distribution (all of its moments are functions of this proportion). Thus, Figure 2 shows that the *distributions* of land use across the region are well represented by the econometric-process simulation model.

Additional validation of the model can be made by testing its ability to predict observed phenomena that were not represented in the data used to estimate and calibrate the model. For this purpose, the model was used to simulate the percentage of acreage allocated to conserving uses (as in the Conservation Reserve Program operated by the U.S. Department of Agriculture) as levels of payments for the conserving use are varied. This exercise showed that the model correctly predicts that larger amounts of land are allocated to conserving uses in the sub-MLRAs where crops are less profitable, and it also correctly predicts that approximately 20 percent of acreage in the region is allocated to the conserving use when simulated payments are in the range of actual payments for CRP contracts in Montana (Antle, Capalbo, Mooney, Elliott and Paustian).

Implications for Economic Models in Integrated Assessment

In this section we further explore the ability of the econometric-process model to represent (1) the spatial variability in economic behavior, and (2) discontinuities and non-linearities in behavior implied by the structure of the model but not observed in the data used to estimate the model. For this purpose the model described above was subjected to changes in relative output prices. A base simulation scenario (observed prices) is contrasted with output price distributions for which the mean price of spring wheat is 30 percent and 60 percent below and above the observed mean price. Each simulation consisted of a four-year production cycle replicated five times. A four-year cycle was found adequate to represent the dynamics of the crop rotation. Years 3 and 4 of the simulations were used to represent the equilibrium in land use in response to the changes in relative prices.

Spatial Variability in Net Returns

Figures 3, 4, and 5 show the means, coefficients of variation, and skewness of the distributions of net returns by sub-MLRA for the spring wheat price scenarios. Sub-MLRAs 52-high and 52-low exhibit higher productivity, and hence higher mean returns, than the other sub-MLRAs (Figure 3). The less productive areas exhibit a higher degree of variability for each price scenario, and that these differences are amplified by low prices (as prices decrease, means and standard deviations of net returns decrease, but means generally decrease faster than standard deviations) (Figure 4). In all cases net returns distributions are positively skewed. In the more productive regions, skewness tends to decrease as prices increase, because the means of the distributions increase and the distributions become more symmetric. Some of the less productive regions show an increase in skewness at higher prices (Figure 5).

Supply Functions

The econometric-process simulation model can be used to determine site-specific supply functions, and the data can be aggregated to the sub-MLRA level or to the level of the entire region. To illustrate the properties of the supply functions generated by the econometric-process model, points on the aggregate supply function for spring wheat were derived by varying the price from a low value of about $1.60 per bushel to a high value of about $6.70 per bushel and aggregating the results (a cubic function is fit to the points in the figure to approximate the shape of the continuous supply function) (Figure 6). At the low price of $1.60 per bushel, production of spring wheat approaches the shut-down point where most producers have substituted into other crops or taken the land out of production. At low prices, the econometric-process model generates a highly elastic supply curve for spring wheat, as land is substituted from other uses into spring wheat, with the elasticity in the range of 4.0. At higher prices, the elasticity declines and approaches the value near zero obtained in the spring wheat supply function of Table 1.

The inelastic supply response behavior is obtained with the econometric-process model at high prices because available land is in production and most of it is being allocated to spring wheat. Thus, at a high relative price of spring wheat the only source of supply response is through intensification of the spring wheat crop. Along this section of the supply function the supply response behavior is similar to the behavior estimated with econometric supply-response models that do not account for site-specific land use decision making. For purposes of comparison, a constant elasticity supply function with an elasticity of 0.5 (a value typical of the literature) is plotted in Figure 6 so that it intersects the econometric-process model's supply function at the base price level. However, as the spring wheat price declines relative to winter wheat and barley, the econometric-process model shows that land would be reallocated to the other crops and production would decline substantially before reaching the shut-down point at the price of $1.60. The constant-elasticity function underestimates price responsiveness relative to the econometric-process model, and predicts that as price declines production would decline little. This example demonstrates that the econometric-process model is capable of simulating a non-linear property of the supply function (the cubic shape shown in Figure 7) and discontinuities in behavior (the shut-down point on the supply function) in a way that a conventional econometric supply function cannot.

Conclusions

This paper develops a new approach to agricultural production analysis that combines conventional econometric production models with simulation models that embed both discrete and continuous choices of the farmers. These econometric-process models are well suited for use in integrated assessment of agricultural production systems: they can link land use and management decision making with bio-physical crop growth and environmental processes on a site-specific basis; they can realistically represent the spatial variability in economic behavior; and they can simulate discontinuities and non-linearities implied by the logic of the decision making process that are outside the range of observed behavior.

The application of this methodology to the dryland grain production system of Montana was used to demonstrate some of the properties and capabilities of this type of model. The model was validated in two ways. The model was able to reproduce the within-sample distributions of land use decisions, and also able to predict out-of-sample phenomena such as the conversion of crop land to a conserving use. The spatial variability in net returns was simulated, and it was found that the distributions of net returns vary according to the productivity levels between sub-MLRA zones. The econometric-process model produced a non-linear characterization of supply response that is different from a conventional constant-elasticity supply function. This difference is due to the explicit representation of the discrete land use decision in the econometric-process model. Finally, this example showed that the econometric-process model represented a discontinuity in behavior (the shut-down point on the supply curve) that is not observed in the data and not represented in a conventional, continuous econometric supply function model.

Several methodological issues raised in this paper deserve further attention by researchers. The validation of economic models for simulation analysis remains a largely unexplored but important topic. Also the question of how to extrapolate economic models beyond the range of observed data should be further explored. In this paper we showed that crop simulation models could be linked to economic models for both estimation and extrapolation. These procedures need to be investigated and their predictive capability needs to be compared to extrapolation based on purely statistical models. Finally, the statistical properties of spatial agricultural production data need to be further explored, using some of the recent results in the spatial econometrics literature.

Appendix: Frontier Production Interpretation of Crop Models

Following Aigner, Lovell and Schmidt, the frontier production function $h(\mathbf{v}_i, \mathbf{z}_i)$ is written as being related to observed output q_i by an efficiency factor $h(\varepsilon_i)$ such that $q_i = h(\mathbf{v}_i, \mathbf{z}_i)h(\varepsilon_i)$, where $\mathbf{v}_i$ and $\mathbf{z}_i$ are variable and fixed input vectors for sites i = 1,..., N and ε_i is an error term representing technical inefficiency at site i such that $0 \leq h(\varepsilon_i) \leq 1$. Under the hypothesis that site-specific productivity is determined by bio-physical characteristics $\mathbf{e}_i$, the conventional (mean) production function $f(\mathbf{v}_i, \mathbf{z}_i, \mathbf{e}_i)$ is related to the frontier function by

(A1) $f(v_i, z_i, e_i) = h(v_i, z_i)E[h(\varepsilon_i)|e_i]$,

where E[.] is the mathematical expectation operator over the probability distribution of ε given $\mathbf{e}_i$.

The genetic potential of the crop can be represented by the crop's maximum yield $q^p(\mathbf{v}^0,\mathbf{z}^0)$ under specified management $(\mathbf{v}^0, \mathbf{z}^0)$ and optimal environmental conditions. A crop growth model produces an estimate of yield under bio-physical conditions $\mathbf{e}_i$ and management $(\mathbf{v}^0,\mathbf{z}^0)$, and can be summarized by a model of the form $q^p(\mathbf{v}^0, \mathbf{z}^0, \mathbf{e}_i)$. It follows that $0 \leq q^p(\mathbf{v}^0, \mathbf{z}^0, \mathbf{e}_i)/q^p(\mathbf{v}^0, \mathbf{z}^0) \leq 1$. Thus, we can interpret the crop growth model for a specific site as providing an index of site-specific productive efficiency, and for some monotonic function g(.) such that g(0) = 0, g(1) = 1, and g' > 0, we hypothesize that

(A2) $E[h(\varepsilon_i)|\mathbf{e}_i] = g[q^p(\mathbf{v}^0, \mathbf{z}^0, \mathbf{e}_i)/q^p(\mathbf{v}^0, \mathbf{z}^0)]$.

Combining (A1) and (A2) it follows that the conventional mean production function $f(\mathbf{v}_i, \mathbf{z}_i, \mathbf{e}_i)$ can be expressed as a function of the frontier production function $h(\mathbf{v}_i, \mathbf{z}_i)$ and the crop growth model's yield estimate as

(A3) $f(\mathbf{v}_i, \mathbf{z}_i, \mathbf{e}_i) = h(\mathbf{v}_i, \mathbf{z}_i)g[q^p(\mathbf{v}^0, \mathbf{z}^0, \mathbf{e}_i)/q^p(\mathbf{v}^0, \mathbf{z}^0)] = f(\mathbf{v}_i, \mathbf{z}_i, q^p(\mathbf{v}^0, \mathbf{z}^0, \mathbf{e}_i))$.

Equation (A3) is the basis for equation (5) in the text.

References

Adams, R.M., R.A. Fleming, C.C. Chang, B.A. McCarl, and C. Rosenzweig. “A Reassessment of the Economic Effects of Global climate Change on U.S. Agriculture.” *Climatic Change* 30(1995): 14-167.

Adams, R.M., B.A. McCarl, K. Segerson, C. Rosenzweig, K.J. Bryant, B.L. Dixon, R. Conner, R.E. Evenson, and D. Ojima. “The Economic Effects of Climate Change on U.S. Agriculture.” Chap. 2 in *The Economics of Climate Change*, R. Mendelsohn and J. Neumann, eds. Cambridge UK: Cambridge University Press, 1998.

Aigner, D.J., A.K. Lovell, and P. Schmidt. “Fomulation and Estimation of Stochastic Frontier Production Functions.” *Journal of Econometrics* 6(1977):21-38.

Anselin, L. *Spatial Econometrics: Methods and Models.* Kluwer Acad. Pub., Boston, 1988.

Anselin, L., A.K. Bera, R. Florax, and M.J. Yoon. “Simple diagnostic tests for spatial dependence.” *Regional Science and Urban Economics* 26(Feb. 1996):77-104.

Antle, J.M. “Testing the Stochastic Structure of Production: A Flexible Moment-Based Approach.” *Journal of Business and Economic Statistics* 3(1983):192-201.

Antle, J.M., S.M. Capalbo, and C.C. Crissman. "Econometric Production Models with Endogenous Input Timing: An Application to Ecuadorian Potato Production." *Journal of Agricultural and Resource Economics* 19(July 1994):1-18.

Antle, J.M., S.M. Capalbo, and C.C. Crissman. “Tradeoffs in Policy Analysis: Conceptual Foundations and Disciplinary Integration.” In *Economic, Environmental, and Health Tradeoffs in Agriculture: Pesticides and the Sustainability of Andean Potato Production*. C.C. Crissman, J.M. Antle, and S.M. Capalbo, eds. Boston: Kluwer Acad. Pub., 1998.

Antle, J.M., S. Capalbo, J. Johnson, and D. Miljkovic. "The Kyoto Protocol: Economic Effects of Energy Prices on Northern Plains Dryland Grain Production." *Agricultural and Resource Economics Review* 28(April 1999), 96-105.

Antle, J.M., S.M. Capalbo, S. Mooney, E.E. Elliot, and K. Paustian. "Economics of Agricultural Soil Carbon Sequestration in the Northern Plains." Trade Research Center, Montana State University, Research Discussion Paper No. 38, October 1999.

Antle, J.M., and R.E. Just. "Effects of Commodity Programs on Resource Use and the Environment." In *Commodity and Resource Policies in Agricultural Systems*, eds R.E. Just and N. Bockstael, pp. 97-128. New York: Springer-Verlag Publishing Co., 1991.

Antle, J.M., and J. Stoorvogel. "Models for Integrated Assessment of Sustainable Land Use and Soil Quality." Paper presented at the conference on Land Degradation in Developing Countries, Wageningen Agricultural University, The Netherlands, July 1, 1999.

Banerjee, A.N., and J.R. Magnus. "On the Sensitivity of Usual t- and F-tests to Covariance Misspecification." *Journal of Econometrics* 95(March 2000):157-176.

Chambers, R.G. *Applied Production Analysis*, Cambridge England: Cambridge Univ. Press, 1988.

de Janvry, A., M. Fafchamps, and E. Sadoulet. "Peasant Household Behavior with Missing Markets: Some Paradoxes Explained." *Economic Journal* 101(Nov. 1991):1400-1417.

Hardie, I.W., and P.J. Parks. "Land Use with Heterogeneous Land Quality: An Application of an Area Base Model." *American Journal of Agricultural Economics.*79(May 1997):299-310.

Heckman, J. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1979): 153-161.

Huffman, W.E. “An Econometric Methodology for Multiple Output Agricultural Technology: An Application of Endogenous Switching Models.” In *Agricultural Productivity: Measurement and Explanation,* S.M. Capalbo and J.M. Antle, eds. Washington, D.C.: Resources for the Future Press, 1988.

Johnson, J.B., W.E. Zidack, S.M. Capalbo, J.M. Antle, and D.F. Webb. “Farm-Level Characteristics of Larger Central and Eastern Montana Farms with Annually-Planted Dryland Crops.” *Departmental Special Report #21*, Department of Agricultural Economics and Economics, Montana State University, 1997.

Just R. E., and J.M. Antle. “Interactions Between Agricultural and Environmental Policies: A Conceptual Framework.” *American Economic Review* 80(May, 1990): 197-202.

Just, R.E., and R.D. Pope. “Stochastic Specification of Production Functions and Economic Implications.” *Journal of Econometrics* 7(1978):67-86.

Just, R.E., D. Zilberman, and E. Hochman. “Estimation of Multicrop Production Functions.” *American Journal of Agricultural Economics* 65(Nov. 1983):770-80.

Kaiser, H.M., S.J. Riha, D.S. Wilks, D.G. Rossiter, and R. Sampath. “A Farm-Level Analysis of Economic and Agronomic Impacts of Gradual Climate Warming” *American Journal of Agricultural Economics* 75(May 1993):387-398.

Kaufmann, R.K., and S.E. Snell. “A Biophysical Model of Corn Yield: Integrating Climatic and Social Determinants.” *American Journal of Agricultural Economics* 79(February 1997):178-190.

Kruseman, G., R. Ruben, H. Hengsdijk, and M.K. van Ittersum. “Farm Household Modelling for Estimating the Effectiveness of Price Instruments in Land Use Policy.” *Netherlands Journal of Agricultural Science* 43(1995):111-123.

Mendelsohn, R., W.D. Nordhaus, and D. Shaw. "The Impact of Global Warming on Agriculture: A Ricardian Analysis." *American Economic Review* 84(1994) 753-771.

Mundlak, Y., and I Hoch. "Consequences of Alternative Specifications in Estimation of Cobb-Douglas Production Functions." *Econometrica* 33(1965):814-28.

Paustian, K., E. Elliott, and L. Hahn, "Agroecosystem Boundaries and C dynamics with Global Change in the Central United States." FY 1998/1999 Progress Report, National Institute for Global Environmental Change. (Available at www.nigec.ucdavis.edu).

Prato, T., C. Fulcher, S. Wu, and J. Ma. "Multiple-Objective Decision Making for Agroecosystem Management." *Agricultural and Resource Economics Review* 25(October 1996):200-212.

Segerson, K., and B.L. Dixon. "Climate Change and Agriculture: The Role of Farmer Adaptation." Chap. 4 in *The Economics of Climate Change,* R. Mendelsohn and J. Neumann, eds. Cambridge UK: Cambridge University Press, 1998.

Tsuji, G.Y., G. Uehara, G. and S. Balas. (eds). *DSSAT Version 3*. University of Hawaii, Honolulu, Hawaii, 1994.

Wu, J., and K. Segerson. "The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin." *American Journal of Agricultural Economics* 77(November 1995):1033-1047.

Table 1. NL3SLS Estimates Of Supply Function, Machinery Cost, and Cost Function Models

	Winter Wheat	Spring Wheat	Barley
Supply Function			
Intercept	2.448 (5.32)	3.185 (11.49)	3.615 (9.47)
Land	1.003 (20.73)	0.993 (38.89)	0.910 (16.76)
Fallow Dummy	0.268 (3.12)	0.211 (6.06)	0.089 (1.45)
Fertilizer Price	-0.350 (-2.70)	-0.124 (-1.56)	-0.320 (-2.61)
Pesticide Price	-0.014 (-0.43)	-0.015 (-0.61)	-0.033 (-1.10)
MLRA 52 - High	0.041 (0.53)	0.004 (0.06)	0.056 (0.69)
MLRA 53A - Low	-0.024 (-0.08)	-0.286 (-4.90)	-0.351 (-3.30)
MLRA 53A - High	-0.327 (-1.70)	-0.423 (-7.26)	-0.544 (-5.26)
MLRA 54 - Low	-0.348 (-2.63)	-0.679 (-9.82)	-0.506 (-3.42)
MLRA 54 - High	--	-0.486 (-5.31)	-0.318 (-1.57)
MLRA 58A - Low	-0.294 (-2.11)	-0.428 (-5.65)	-0.201 (-1.26)
MLRA 58A - High	-0.168 (-2.04)	-0.429 (-5.75)	-0.287 (-2.97)
Machinery Cost			
Intercept	-0.304 (-0.14)	-1.054 (-0.45)	2.447 (2.80)
Land	1.109 (22.93)	1.104 (33.44)	1.085 (20.35)
Fallow	0.022 (0.26)	-0.101 (-2.25)	-0.013 (-0.22)
Crop Price	1.716 (1.10)	2.170 (1.36)	0.186 (0.18)
Fertilizer Price	-0.063 (-0.47)	-0.147 (-1.35)	0.110 (0.84)
Pesticide Price	0.040 (1.12)	0.002 (0.05)	-0.001 (-0.04)
MLRA 52 - High	0.130 (1.59)	0.069 (0.87)	0.151 (1.79)
MLRA 53A - Low	0.613 (1.88)	0.082 (0.81)	-0.143 (-1.03)
MLRA 53A - High	0.272 (1.30)	-0.148 (-1.47)	-0.019 (-0.14)
MLRA 54 - Low	-0.117 (-0.76)	-0.294 (-2.58)	-0.111 (-0.64)
MLRA 54 - High	--	-0.031 (-0.20)	-0.145 (-0.61)
MLRA 58A - Low	0.053 (0.33)	-0.019 (-0.16)	-0.325 (-1.64)
MLRA 58A - High	0.106 (1.12)	0.133 (1.11)	-0.005 (-0.05)
Cost Function			
Intercept	0.784 (1.02)	-0.702 (-1.17)	-1.522 (-1.44)
Fertilizer Price	0.885 (55.06)	0.827 (46.80)	0.868 (52.49)
Output	1.014 (11.10)	1.129 (15.83)	1.200 (9.71)
MLRA 52 - High	-0.356 (-2.40)	-0.225 (-1.36)	-0.292 (-1.68)
MLRA 53A - Low	-0.053 (-0.09)	-0.079 (-0.49)	-0.200 (-0.84)
MLRA 53A - High	0.113 (0.31)	0.000 (0.00)	-0.181 (-0.74)
MLRA 54 - Low	-0.421 (-1.60)	0.304 (1.54)	0.111 (0.34)
MLRA 54 - High	--	0.570 (2.25)	-0.117 (-0.26)
MLRA 58A - Low	-0.060 (-0.22)	0.061 (0.28)	-0.139 (-0.41)
MLRA 58A - High	-0.053 (-0.33)	0.296 (1.40)	-0.048 (-0.23)
Fallow Dummy	-0.546 (-3.25)	-0.575 (-5.88)	-0.482 (-3.71)

Note: t-statistics in parentheses.

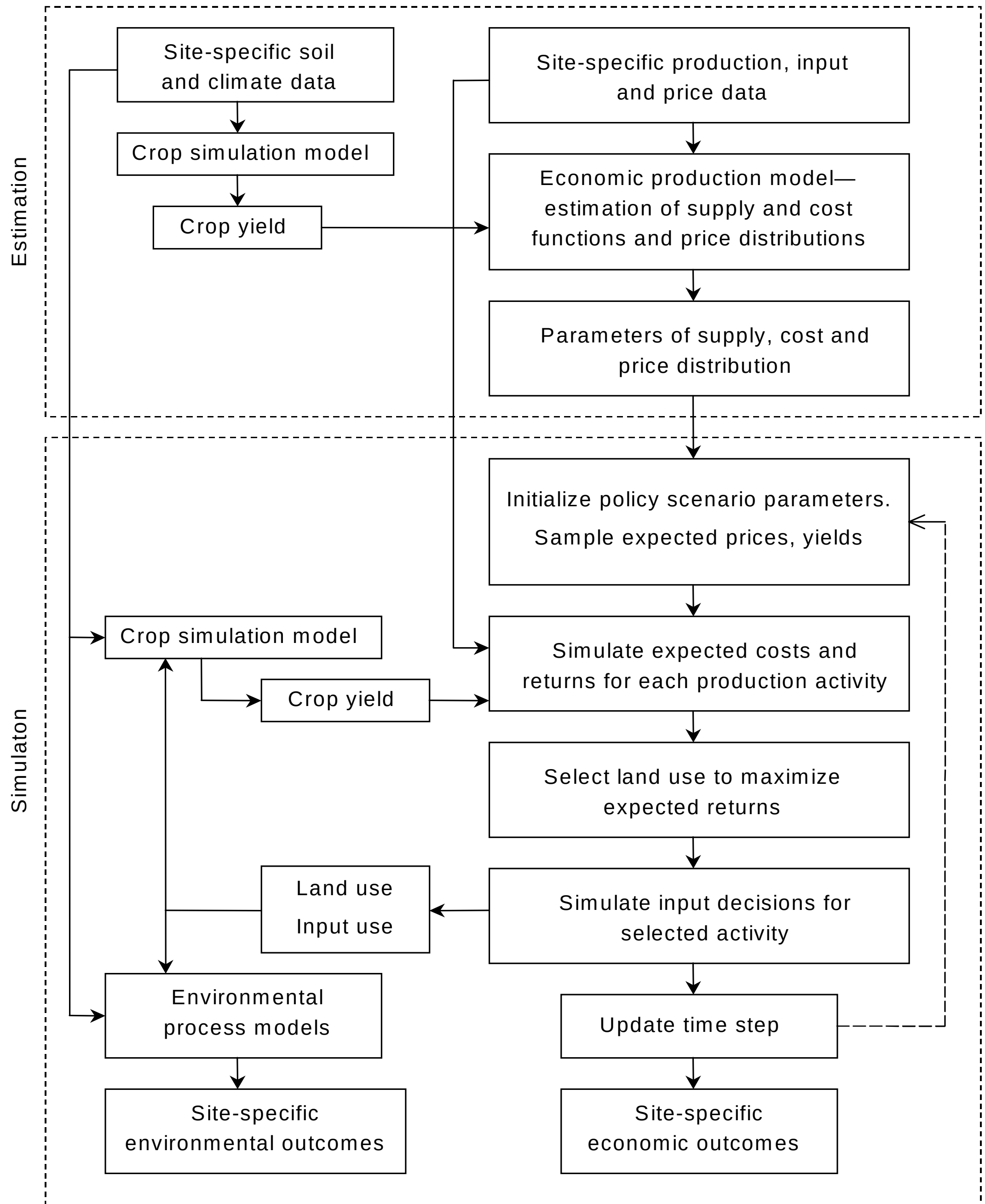

Figure 1. Structure of an Econometric Process Simulation Model

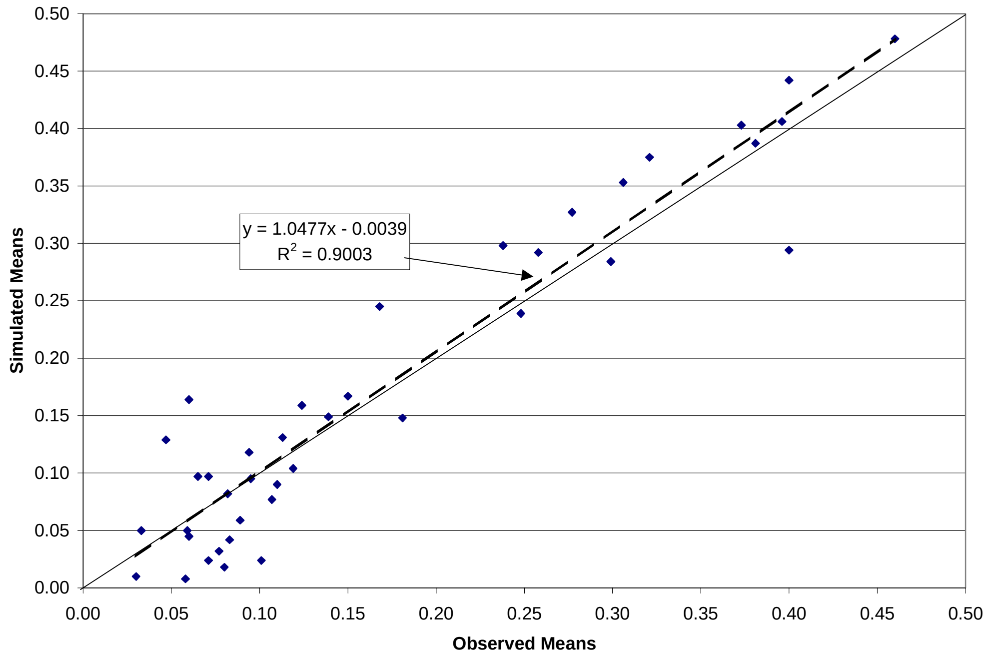

Figure 2. Observed vs. Simulated Land Use by Sub-MLRA

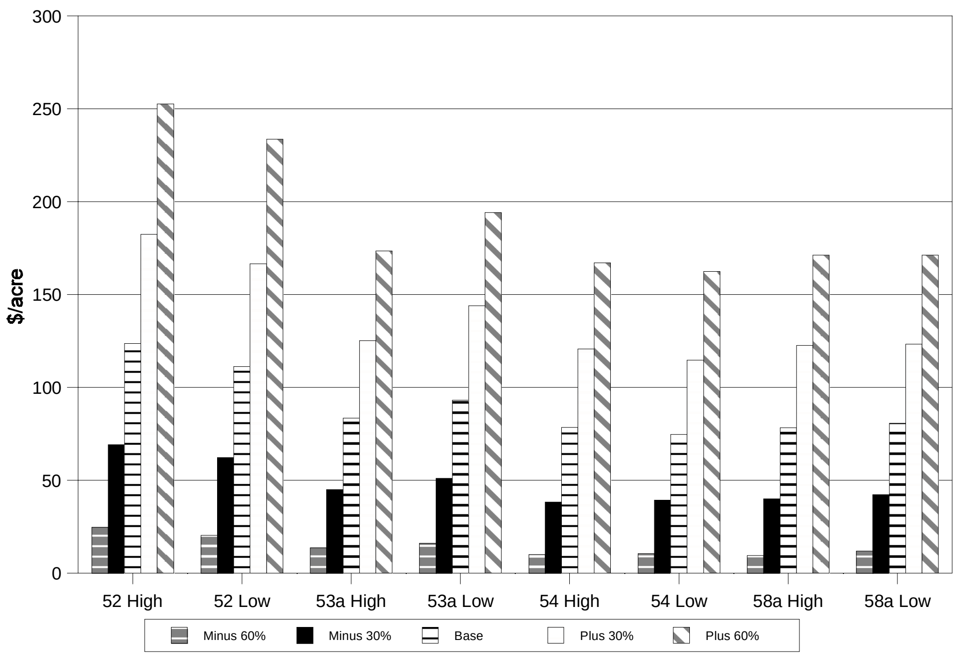

Figure 3. Mean Net Returns by Sub-MLRAs for Spring Wheat Price Scenarios

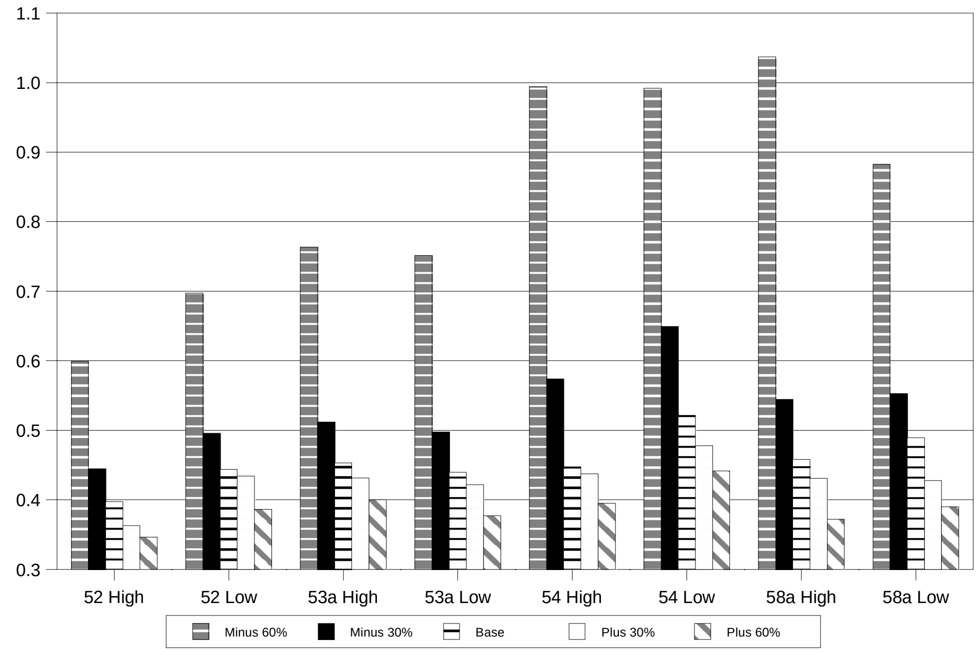

Figure 4. Coefficients of Variation of Net Returns for Spring Wheat Price Scenarios

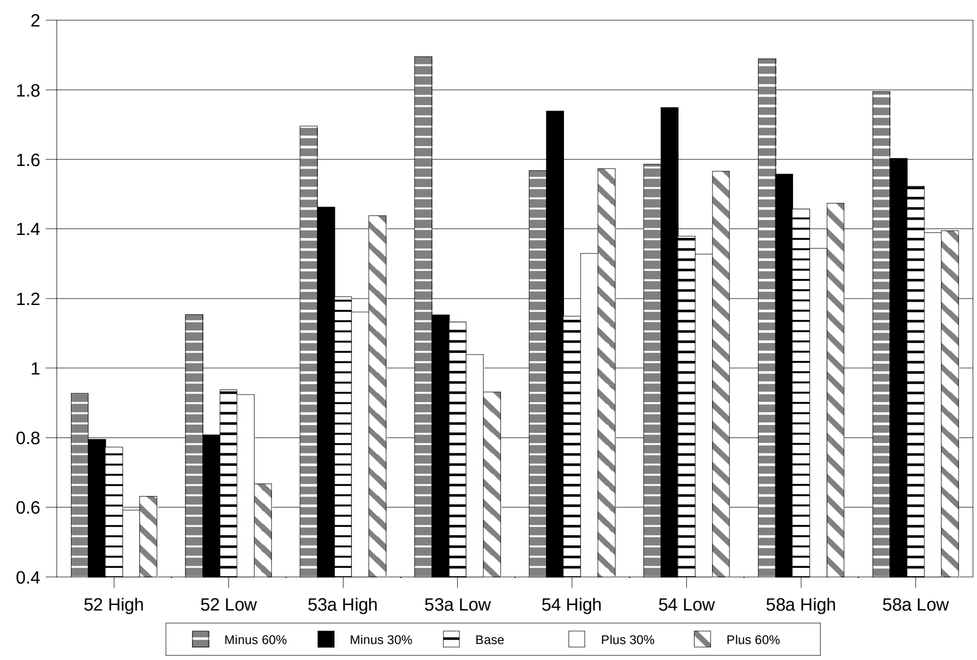

Figure 5. Skewness of Net Returns by Sub-MLRA for Spring Wheat Price Scenarios

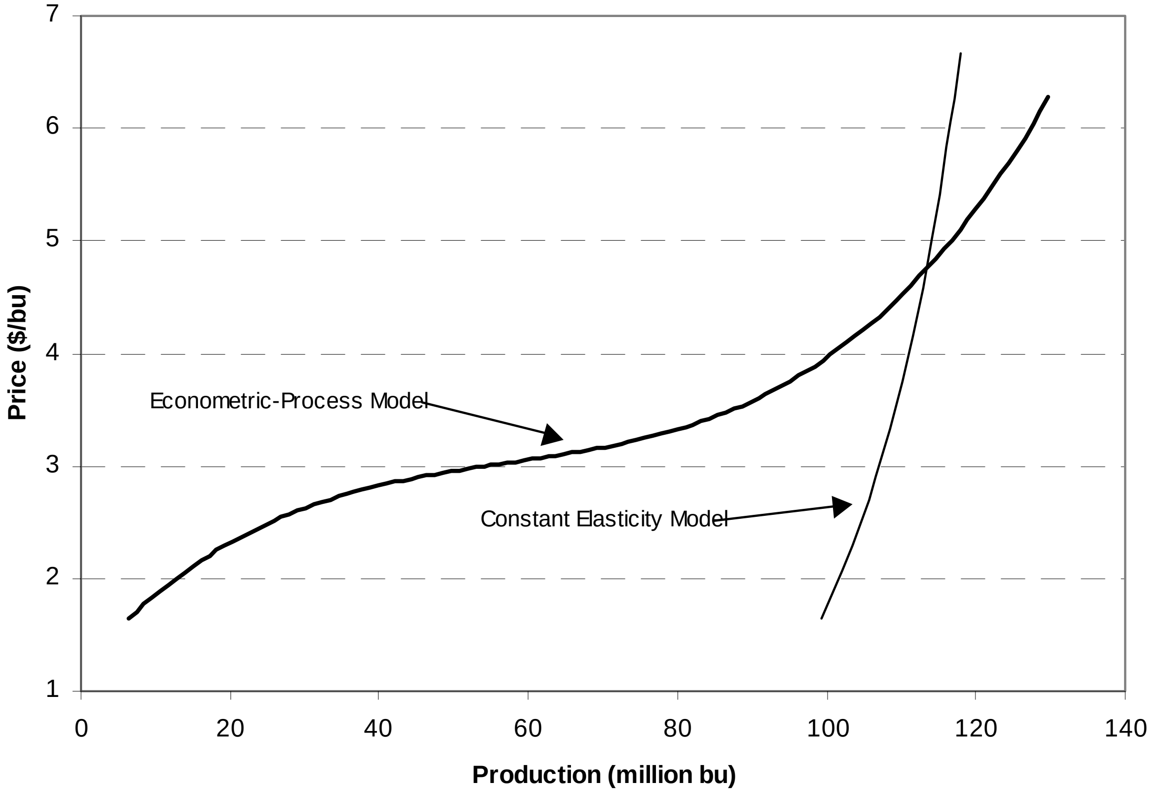

Figure 6. Simulated Spring Wheat Supply Functions from Econometric-Process Model and Constant Elasticity Model