

# **Reconstructing Disaggregate Production Functions**

by  
Richard E Howitt\* and Siwa Msangi\*

Paper to be presented at the Xth Congress of the European Association of Agricultural Economists, Zaragoza (Spain) 28 - 31 August 2002.

\* Department of Agricultural and Resource Economics, University of California, Davis, California, 95616 email [howitt@primal.ucdavis.edu](mailto:howitt@primal.ucdavis.edu)

## Introduction

This paper develops a method for reconstructing flexible form production functions using minimal disaggregated data sets. We use the term reconstruction rather than estimation because many disaggregate data sets have insufficient observations to claim asymptotic properties for the resulting parameter estimates. Despite the very small sample size, we are able to calculate standard errors for the model parameters by bootstrapping the GME estimation program. We can therefore subject our reconstructed models to the three standard econometric tests, namely  $R^2$  on the production equation, tests of significance on the individual parameters, and the precision of out of sample forecasts of regional crop production.

Since we are interested in models that can address policy questions, the emphasis in this paper is on the ability of the model to reproduce the existing production system and predict the disaggregated outcomes of policy changes.

In many developed and developing agricultural economies there is considerable emphasis on the effect of agricultural policies and production on the environment, and conversely, the effect of environmental policies on the agricultural sector. This emphasis is likely to rekindle interest in the use of production function models for many policy problems. There are several reasons why production functions are suited to the analysis of agricultural- environmental policy. First, environmental values are measured in terms of the physical outcomes of agricultural activity. Second, primal data on crop yields, areas and input use is usually more readily available and often more accurate than cost data. Third, constraints and subsidies on some inputs require that shadow prices be added to nominal prices. Moreover some environmental policies are formulated as constraints on input use. Fourth, economic models of agricultural and environmental policy impacts often have to formally interact with process models of the physical systems. Such models require the economic output in terms of primal values.

Several authors have emphasized the need to spatially disaggregate models for environmental policy analysis (Antle & Capalbo, 2001; Just & Antle, 1990). However, such disaggregation is often made difficult either by the limited availability of disaggregate data or, if such data is present, the lack of enough degrees of freedom to identify disaggregate parameters within a classical estimation framework. Generalized Maximum Entropy (GME) estimation techniques (Golan et al., 1996(a)) have come into increasing use by researchers who seek to achieve higher levels of disaggregation in the face of these data problems (Lence & Miller, 1998(b); Lansink et al., 2001; Golan et al., 1994, 1996(b)). Given the inherent heterogeneity of soils and other agricultural resources, the researcher who wishes to disaggregate cross-sectional data must consider the trade-off between two possible sources of model error: namely error caused by aggregation bias versus error due to small sample bias. Aggregating across heterogeneous regions leads to aggregation bias, whereas ill-conditioned or ill-posed GME estimates may be biased due to the small samples on which they are based. An additional advantage that favors maximum entropy based alternatives is the ability to formally incorporate additional data or informative priors into the estimation process, in a Bayesian fashion.

Substitution at the intensive and extensive margins is a key focus of agricultural-environmental policy analysis. A basic policy approach is to provide incentives or penalties that lead to input substitution under a given agricultural technology. Such substitutions at the intensive margin can reduce the environmental cost of producing

traditional agricultural products or that of jointly producing agricultural and environmental benefits. These policies cannot be evaluated without explicit representation of the agricultural production process. It follows, therefore, that the potential for substitution should be explicitly modeled within a multi-input multi-output production framework.

The reconstructed production function (RPF) modeling approach proposed in this paper can be implemented with much smaller data sets than those required by conventional econometric approaches. However, RPF models still result in production functions that have all the properties enjoyed by flexible functional forms, such as multi-input, multi-output quadratic, square root, generalized Leontieff and trans-log specifications. This is achieved by using a combination of GME estimation and positive mathematical programming (PMP) calibration techniques (Howitt, 1995), so as to reconstruct a well-behaved production function that is consistent with micro theory and is calibrated to the base conditions in the data. Hansen and Heckman (1996) address the empirical foundations of calibration in the context of micro based calibration of macro models. They emphasize that simulated behavior is an important criterion in the evaluation any policy model.

The combination of calibration and GME distinguishes our approach from other GME production analyses in the literature (Zhang & Fan, 2001; Lence & Miller, 1998(a)). The GME estimates given in this paper do, however, converge to consistent estimates when the sample size is increased and have been shown to have the same asymptotic properties as conventional likelihood estimators (Mittlehammer et al., 2000). In our approach we ensure consistency by the including moment constraints in the estimation procedure. These constraints are not in the formulations presented in other applications of GME, such as the aforementioned studies. In addition, we generate the finite sample distribution properties of the resulting GME estimates by bootstrapping (Efron & Tibsharani, 1993). To our knowledge, only one other paper (Marsh & Mittelhammer, 2001) has used the bootstrap method to obtain GME parameter distributions from sample data. Previous work has tested GME results for sensitivity to their support spaces (Lansink et al, 2001; Leon et al, 1999), or has derived the approximate asymptotic parameter distributions from either the dual (Fraser, 2000) or by Monte Carlo replications within an experimental setting (Lence & Miller, 1998(a,b)). However, since our aim is to use small data samples, bootstrapping seems a natural method to generate the finite sample properties of the parameter estimates, and can be simply implemented.

The ability to simulate policy alternatives reliably with constrained profit maximization requires a model that satisfies the marginal and total product conditions and has stability in the second order profit maximizing conditions and decreasing returns to scale. It is our belief that those who use policy models are more interested in reproducing observed behavior and simulating beyond the base scenario, than in testing for the curvature properties of the underlying production function. Therefore, the RPF models presented here are reconstructed subject to parameter restrictions that result in locally concave profit functions and decreasing returns to scale. Within our programming-based reconstruction and simulation framework, we can also impose policy restrictions in the form of constraints on the reconstructed model.

Section II of the paper briefly reviews modeling methods used to estimate the effect on land use of agricultural and environmental policies. Section III develops the production model reconstruction and bootstrap procedure within the GME framework.

Section IV contains an empirical application that measures the economic effect of environmental policy changes to irrigation water supplies on California's irrigated crop sector, and is followed, in conclusion, by Section V.

## II. **Methods for Modeling Disaggregated Agricultural Production and Land Use.**

The approach that we use in this paper addresses some of the shortcomings of representative farmer models enumerated by Antle & Capalbo (2001), when they cite the limited range of response in the typical representative farm model. In our model, the embedded PMP parameters capture the individual heterogeneity of the local production environment, be it in terms of land quality or other site-specific effects, and allow the estimated production function to replicate the input usage and outputs produced in the base year.

Love (1999) made the point that the level of disaggregation matters in terms of the degree of firm-level heterogeneity and other localized idiosyncrasies that get averaged out of aggregated data samples. This affects the likelihood of observing positive results for tests of neo-classical behavior, such as cost minimization or profit maximization. In our approach, we impose curvature conditions on the reconstructed production function, since we are aiming for models that reproduce behavior rather than test for it. The relative stability we observe within cropping systems, despite the presence of substantial yield and price fluctuations is strong empirical evidence that farmers act as if their profit functions are convex in crop allocation. The gradual adjustment of agricultural systems to changes in relative crop profitability suggests that farmers adjust by progressive changes over time, along all the margins of substitution, rather than going from one corner solution to the next.

Zhang & Fan (2001) conclude that the behavioral assumptions of profit maximization are too strong for the example to which they applied a GME production function estimation. While their level of aggregation was severe, they made the case for using GME on the basis of its ability to incorporate non-sample information and to deal with imperfectly observed activity-specific inputs. Within our framework, we are able to implement more flexible functional forms for production than that used by Zhang & Fan and a greater degree of disaggregation.

Just et al (1983), stated in their classic production paper that their:

“Methodology is based on the following assumptions that seem to characterize most agricultural production:

- (a) *Allocated inputs.* Most agricultural inputs are allocated by farmers to specific production activities..
- (b) *Physical constraints.* Physical constraints limit the total quantity of some inputs that a farmer can use in a given period of time ...
- (c) *Output determination.* Output combinations are determined uniquely by the allocation of inputs to various production activities aside from random, uncontrollable forces.”

Just et al's specification admits jointness in multioutput production only through the common restrictions on allocatable inputs. The specification in this paper has constraints on the land available, but also allows for jointness between crops in a

region as reflected by the deviations of crop value marginal products from the opportunity cost of restricted land inputs.

The current range of approaches to agricultural production modeling and the associated analysis of environmental impacts, seems to fall into three groups, namely, disaggregated calibrated or constrained programming models (McCarl, 2000 ; Alig et al., 1998; CVPM<sup>1</sup>, 1997; CAPRI<sup>2</sup>, 2000) disaggregated logistic land use models (Wu & Babcock, 1999), and aggregate econometric land use models (Mendelsohn et al., 1994).

In this paper we hope to straddle the current divide between programming and econometric approaches to production analysis, not only by using constraints and output determination in our formulation, but also by using flexible functional forms and data in which the principle explanatory variables of yields, prices, and crop land allocation are based on small stochastic samples.

### **III Using Generalized Maximum Entropy to Reconstruct Production Functions**

The term reconstruction was developed in the field of image processing where the problem of reconstructing images from incomplete data is, by definition, ill-posed. Image reconstruction uses structural information about the image to generate a complete image from sparse or incomplete data observations. A simple example is using curvature criteria to reconstruct an image of the rings around Saturn based on relatively few pixels sent from outer space. The analogy in production models is to reconstruct the production surface for a specific crop at a specific location by using a small set of observations of the average product and marginal input allocations.

Reconstruction methods are now used in several scientific fields, among them are tomography, astronomy, and the earth sciences. We view the problem facing agricultural policy modelers as one of trying to reconstruct the behavioral and technical relationships that drive agricultural production decisions, at a scale that has policy relevance for those environmental resources affected by agriculture. Often the scale of analysis that is required for meaningful environmental policy, differs widely from that of the economic data set available. It follows that a flexible form production function model on the same level of disaggregation as that desired for environmental policy may suffer from low degrees of freedom or be ill-posed. Hence the need for a reconstruction approach. Fortunately, maximum entropy estimation is a commonly used basis for reconstruction algorithms in the physical sciences (Desmedt 1991), and is coming into increasing use in agricultural economics.

The nature of the data set defines the precise reconstruction method to be used. For disaggregated policy models, the available data usually takes the form of short time-series at the desired level of disaggregation, or a cross-sectional survey sample taken over each disaggregated region. The GME reconstruction approach advanced in this paper is completely in accord with classical econometric estimators for large sample problems and uses a standard bootstrap approach to estimate the distributions

---

<sup>1</sup> Central Valley Production Model , used in the 1997 Programmatic Environmental Impact Statement of the Central Valley Project Improvement Act (see references).

<sup>2</sup> Common Agricultural Policy Regional Impact (<http://www.agp.uni-bonn.de/agpo/rsrch/capri/>)

of key parameters. The novelty of the paper lies in the idea that the modeler does not have to accept the stricture of non-negative degrees of freedom, but may specify a complex model at the level of disaggregation that is thought to minimize the net effect of estimation and aggregation bias on the model outcome. The modeler can specify flexible multi-input production functions for any number of observations and calibrate closely to the base conditions. Essentially we show that a minimal level of data that would, in the past, have restricted the modeler to a simple linear programming model, can now be calibrated and reconstructed as a set of multi-input quadratic production functions. Other functional forms that have continuity and local concavity can also be used, namely the trans log and generalized Leontieff form. In the discussion below we use a quadratic production function, and model the multi-outputs from a disaggregated unit using the specification in Just et al (1983).

The first order conditions for optimal allocation have to incorporate the shadow value of any constraints on inputs. Since the allocatable inputs are restricted in quantity, and rotational interdependencies can exist between crops, we use a modified PMP model (Howitt 1995) on each data sample to obtain a numerical value for a shadow price that may exist above or below the allocatable input cash cost. Specifically, we impose upper and lower perturbed calibration constraints on the crop allocation in each sample to generate the additional values that may be observed due to rotational interdependencies or land heterogeneity<sup>3</sup>. For cases where the data is collected from micro samples that include reliable information on rental markets, the reconstruction proceeds directly, without the intervening PMP stage, whose only role is to generate shadow values for rotations and allocatable inputs that are consistent with the data.

Assume that we have "n" observations over time on a farm unit that produces "j" crops, each of which has "i" inputs. There is a subset of restricted, but allocatable inputs such as land or irrigation water. The data set consists of n observations on crop price, input price, crop yield, and input use by crop. This data set and other agronomic data can be used to define the implicit Leontieff matrix A and specify the following calibrated linear programming problem.

$$\begin{aligned}
 (1) \quad & \text{Max } \sum_j p_j y_{L,j} x_{L,j} - \sum_{ij} \omega_j a_{ij} x_{L,j} \\
 \text{s.t.} \quad & Ax \leq b \\
 & Ix_L \leq \tilde{x} + \varepsilon \\
 & Ix_L \geq \tilde{x} - \varepsilon \\
 & x_{L,j} \geq 0
 \end{aligned}$$

where b is the vector of available input quantities, p and  $\omega$  are the output and input price vectors (respectively) and  $y_{L,j}$  is the yield per acre for crop j, and  $x_{L,j}$  is the land area allocated to crop j.

The first set of allocatable resource constraints generates the shadow values for those constraints that influence the observed crop and input allocations. The perturbed upper and lower bound calibration constraints ensure that the crop allocation is within  $\varepsilon$  of the observed data<sup>4</sup>, and in addition, provide measures of the rotational cost

<sup>3</sup> We are indebted to Wolfgang Britz for the original idea, and other helpful comments.

<sup>4</sup> The  $\varepsilon$  also prevents a degenerate dual solution (see Howitt, 1995)

interdependence between the crops based on the equi-marginal principle for land allocation.

A generic GME reconstruction problem of a production system  $\bar{y} = f(X; \bar{\beta})$  can be written as

$$\begin{aligned} \max_{\bar{p}_\beta, \bar{p}_e} \quad & -\bar{p}_\beta^T \ln \bar{p}_\beta - \bar{p}_e^T \ln \bar{p}_e \\ \text{s.t.} \quad & \bar{g}(\bar{p}_\beta, \bar{p}_e, \bar{z}_\beta, \bar{z}_e; \bar{y}, X) = \bar{0} \\ & \iota \cdot \bar{p}_\beta^T = 1 \\ & \iota \cdot \bar{p}_e^T = 1 \end{aligned}$$

where the resulting probability vectors  $\bar{p}_\beta, \bar{p}_e$  define the production function parameters  $\bar{\beta}$  and stochastic sampling errors  $\bar{e}$  in terms of expectation (i.e.  $E(\bar{\beta}) = \bar{p}_\beta^T \bar{z}_\beta$ ,  $E(\bar{e}) = \bar{p}_e^T \bar{z}_e$ ). The estimating equations  $\bar{g}(\bar{p}_\beta, \bar{p}_e, \bar{z}_\beta, \bar{z}_e; \bar{y}, X) = \bar{0}$  are appropriate functions of the data  $(\bar{y}, X)$  and user-defined ‘support’ vectors  $(\bar{z}_\beta, \bar{z}_e)$  that define the space that can contain non-zero probability mass.  $\iota$  is a unit vector that defines the usual adding up property of the probability measures.

Before the GME reconstruction program can be solved, however, these support values have to be defined for each parameter and error term. To ensure that the set of support values spans the feasible solution set, we define the support values for the production function parameters  $(\bar{z}_\beta)$  as the product of a set of five weights and functions of the average Leontieff yield over the data set, for a particular crop/input combination. The support values for the error terms  $(\bar{z}_e)$  are defined by positive and negative weights that multiply the right-hand side values of the equation defined above for the expected vector of sampling errors<sup>5</sup>.

Curvature is added by solving for the parameters of the Cholesky decomposition of the quadratic matrix L where  $Z = LL'$ , and constraining the diagonal Cholesky parameters to be nonnegative, for details see (Paris & Howitt, 1998). If the quadratic production function is defined as:

$$(2) \quad y_j = \alpha_j' x_j - x_j' Z_j x_j$$

Where  $x_j$  is a  $i \times 1$  vector of inputs to crop  $j$ , and  $y_j$  is the total product of crop  $j$ , the GME reconstruction problem becomes:

<sup>5</sup> While these are usually spaced symmetrically about zero, by practitioners of GME, this is not enough to ensure that the model error is zero in expectation. Further comment on this point is made in the paper.

$$(3) \quad \text{Max} \quad \sum_{i,p} pa_{i,p} - \ln pa_{i,p} + \sum_{i',p} pl_{i',p} - \ln pl_{i',p} + \sum_{n,i,p} pel_{n,i,p} - \ln pel_{n,i,p} \\ + \sum_{n,p} pe2_{n,p} - \ln pe2_{n,p}$$

Subject to:

$$(3a) \quad cpr_{n,i} = \sum_p pa_{i,p} * za_{i,p} - 2 * [ \sum_{i'} \sum_p (pl_{i',p} * zl_{i',p}) * \sum_p (pl_{i',p} * zl_{i',p}) * x_{n,i'} ] \\ + \sum_p pel_{n,i,p} * zel_{n,i,p}$$

$$(3b) \quad tprod_n = (\sum_p pa_{i,p} * za_{i,p}) * x_{n,i} - [ \sum_i \{ \sum_{i'} \sum_p (pl_{i',p} * zl_{i',p}) * \sum_p (pl_{i',p} * zl_{i',p}) * x_{n,i'} \} * x_{n,i} ] \\ + \sum_p pe2_{n,p} * ze2_{n,p}$$

$$(3c) \quad \frac{ \left[ \sum_i \{ \sum_p pa_{i,p} * za_{i,p} - 2 * [ \sum_{i'} \sum_p (pl_{i',p} * zl_{i',p}) * \sum_p (pl_{i',p} * zl_{i',p}) * x_{n,i'} ] \} * x_{n,i} \right] }{ tprod_n } \leq 0.98$$

$$(3d) \quad \sum_n \{ \sum_p pe2_{n,p} * ze2_{n,p} \} = 0$$

$$(3e) \quad \sum_p pa_{i,p} = 1, \sum_p pl_{i,ii,p} = 1, \sum_p pel_{n,i,p} = 1, \sum_p pe2_{n,p} = 1.$$

The objective function (3) is the usual sum of the entropy measures for the parameter probabilities for the Cholesky decomposition of the quadratic matrix and the vector of linear terms. Following the normal GME procedure, the entropy of the error term probabilities is also maximized. The first equations (3a) are the first order conditions that set the cost-price ratio<sup>6</sup> (cpr) equal to the marginal physical product. If some inputs are restricted and the PMP calibration stage is used, the input cost in the first order equation will include the resulting shadow values as well as the nominal input price.

The second set of equations (3b) fit the production function to the observations on total production. While it is not normal in econometric models to include both the marginal and total products as estimating equations, we think that the information in the total product constraint is particularly important for two reasons. First, information on crop yields and areas is likely to be the most precisely known by farmers. While farmers are often doubtful and reluctant about stating their costs of production to surveyors, they always know their yields and are usually proud to tell you. Second, while the marginal conditions are essential for behavioral analysis, policy models also have to accurately fit the total product to be convincing to policy makers and correctly estimate the total impact on the environment and the regional economy of policy changes. Fitting the model to the integral as well as the marginal conditions improves the policy precision of the model.

The third constraint equation (3c) ensures that the resulting production functions have decreasing returns to scale. Decreasing returns to scale are important if the resulting model is to simulate multioutput production by optimization. While the crop first order conditions ensure calibration at the intensive margin, the returns to scale conditions ensure that the extensive margin between crops is also calibrated within the

<sup>6</sup> Defined as the ratio of the nominal input cost plus shadow value divided by the output price

variation of the observed crop allocations. If a given crop has constant or increasing returns to scale, a multi crop optimization will result in a corner solution for that crop. The scale function coefficient can be shown to be equal to the sum over the inputs, of the product of the marginal physical product and the input level, divided by the total product.<sup>7</sup>

The next equation (3d) is not found in the standard GME specification. Since the unbiased constraint, as we term it, requires that the errors sum to zero over the observations, it can be thought of as a moment restriction that forces the resulting small sample total product estimate to be unbiased. Even though the support values for the error terms are centered around zero, this does not ensure that the resulting maximum entropy solution is centered around zero, as the relative weight of the error term probabilities in the solution depends on the number of parameters and observations in the reconstruction<sup>8</sup>.

The remaining equations (3e) in the reconstruction program are the standard adding up constraints on the parameter and error probabilities. Due to the separability assumption on the production functions, if the shadow value of the constraining allocatable resources is included in the input cost, the reconstruction problem can be solved rapidly by looping through individual production functions.

### Evaluating Finite Sample Properties of GME Results Using the Bootstrap

One of the principle problems in the adoption of GME and entropy methods is the frequent question from users of conventional estimates. "I accept that maximizing entropy calculates an efficient distribution of the parameter, but how do I know that the expected value of the parameter is a reliable point estimate". In short, the potential user is understandably asking for the standard error of the coefficient. To date, the response from ME advocates is to reassure the potential user that the asymptotic properties are consistent (Golan et al., 1996). This asymptotic response is not very reassuring for an estimator whose use and comparative advantage is with small finite samples. It follows that there is a need to generate GME parameter error bounds from the small data sets in which GME excels.

Bootstrap methods have been used for the past twenty years to approximate the finite sample distribution of a statistic by systematically resampling the original sample data in a Monte Carlo fashion (Efron & Tibsharani, 1993). The GME bootstrap uses a uniform random distribution to select observations from the original sample of "n" observations with replacement. Having generated the bootstrap observations, the GME program described above calculates the GME estimates of the production function coefficients  $\alpha_{g,j,i,B}$  and  $z_{g,j,i',B}$ , where the are "g" regions and "j" crops, each of which has "i" inputs. Using the comparative static results from the

<sup>7</sup> Given a production function  $y = f(x_1 \dots x_n)$  and a scale proportion  $\lambda$ , the function coefficient  $\phi$  is

$$\text{defined as } \phi = \frac{\frac{dq}{q}}{\frac{d\lambda}{\lambda}} \quad \text{since } dq = f_1 dx_1 \dots f_n dx_n = f_1 x_1 \frac{dx_1}{x_1} \dots f_n x_n \frac{dx_n}{x_n}$$

$$\text{since } \lambda \text{ varies all inputs } \frac{dx_i}{x_i} = \frac{d\lambda}{\lambda} \quad \text{and} \quad dq = \left( \sum_i f_i x_i \right) \frac{d\lambda}{\lambda} \quad \text{and} \quad \phi = \frac{\sum_i f_i x_i}{y}$$

<sup>8</sup> See additional development of this point in Howitt & Msangi (2002)

next section, and inverse of the Hessian ( $Z^{-1}$ ), we calculate the bootstrapped elasticities of demand and supply  $\eta_{S,j,B}$  and  $\eta_{D,i,B}$ . We run the bootstrap loop for B iterations. The estimated asymptotic variance for a given GME parameter estimate, for instance the supply elasticity for the jth crop  $\hat{\eta}_{S,j}$ , can be estimated from the B bootstrapped estimates  $\hat{\eta}_{S,j,B}$  as:

$$Var \hat{\eta}_{S,j} = \frac{1}{B} \sum_{b=1}^B [\hat{\eta}_{S,j,b} - \hat{\eta}_{S,j}] [\hat{\eta}_{S,j,b} - \hat{\eta}_{S,j}]'$$

Following this approach, we are able to generate standard errors for the supply and demand elasticities and their corresponding pseudo t values. In addition, we can then apply statistical tests for significant differences between the policy-relevant parameters and thus implicitly, the value added by regional disaggregation.

We recognize the sensitivity of our GME parameter estimates to the choice of supports that we specify for the parameter space. Some examples of GME estimation overlook the importance of this issue, and assume a general insensitivity of parameter estimates to the researcher's choice of support specification. Paris & Caputo (2001) have demonstrated the importance of support specification, both analytically, as well as through Monte Carlo studies. While we have not implemented the techniques proposed by others to get around this shortcoming of GME estimation (Paris, 2001; Marsh & Mittlehammer, 2001), we have placed a priority on exploring them in future work. We do, however, pay close attention to an even more serious source of bias in GME estimates, by imposing moment constraints, similar to those used in GMM, Empirical Likelihood or Maximum Entropy Empirical Likelihood (Mittelhammer et al.) estimation. We find that this has not been explicitly addressed in the empirical literature employing GME methods, and should be of concern to practitioners.

### **Calculating Comparative Static Parameters for the Model**

The quadratic production function model has convenient properties for calculating policy parameters. Note that the Hessian resulting from the constrained profit maximization problem (1), is simply:  $\frac{\delta^2 \Pi}{\delta x_{ij}^2} = -Z$ , which will be useful in the following derivations.

#### Calculating the Derived Demands for Inputs

For simplicity, we will use the unconstrained profit term for a single crop.

$$(4) \quad \Pi = p (\alpha'x - 0.5 x' Z x) - \omega'x$$

$$\frac{\partial \Pi}{\partial x_i} = p(\alpha - Z x) - \omega = 0$$

$$Z x_i = \alpha - \frac{\omega_i}{p}$$

$$x_i^* = Z^{-1}\alpha - Z^{-1}\frac{\omega_i}{p}$$

Define the demand slope  $g_i = -Z^{-1}\frac{1}{p}$

$$x_i^* = a_i + g_i \omega_i$$

From the above equations, it is clear that if we can invert the Hessian of the maximized profit expression (  $-Z^{-1}$  ) we can calculate the derived demand for each input for each crop as a linear function of input and output price.

The elasticity of the input demand follows directly from :

$$(5) \quad \text{The elasticity is } \eta_i = g_i \frac{\omega_i}{x_i^*}$$

This elasticity is based on a single crop. For the usual multi-output case we weight the individual crop contribution by their relative resource use to arrive at a weighted elasticity for the resource.

### Calculating Supply Functions and Elasticities

Since production is a function of optimal input allocation and we now have the input demands as a function of input and output price, we can derive the output supply function by substituting the optimized input derived demands into the production function and simplify in terms of the output price. Going back to the derived demand and production function formulae:

$$x_i^* = Z^{-1}\alpha - Z^{-1}\frac{\omega_i}{p}$$

(6) *and*

$$y^* = \alpha'x^* - 0.5 x^{*'} Z x^*$$

Defining  $r_i = \frac{\omega_i}{p}$  and Using the vectors  $\alpha$  and  $r$

$$(7) \quad y^* = \alpha' Z^{-1}(\alpha - r) - 0.5 (\alpha - r)' Z^{-1} Z^{-1}(\alpha - r)$$

multiplying out and collecting  $\alpha$  and  $r$  terms yields

$$y^* = 0.5 \alpha' Z^{-1} \alpha - 0.5 r' Z^{-1} r$$

$$y^* = \hat{\alpha} - 0.5 \left( \frac{\omega_1}{p} \dots \frac{\omega_I}{p} \right)' Z^{-1} \left( \frac{\omega_1}{p} \dots \frac{\omega_I}{p} \right)$$

$$(8) \quad y^* = \hat{\alpha} - \frac{1}{2p^2} \omega' Z^{-1} \omega$$

$$\frac{\delta y^*}{\delta p} = \frac{1}{p^3} \omega' Z^{-1} \omega$$

$$\eta_s = \frac{\delta y}{\delta p} \frac{p}{y} = \frac{1}{p^3} \omega' Z^{-1} \omega \frac{p}{y}$$

$$\eta_s = \frac{1}{p^2 y} \omega' Z^{-1} \omega$$

### Calculating Elasticities of Input Substitution

There are many elasticities of substitution with different advantages and disadvantages. To demonstrate that we can obtain crop and input specific elasticities of substitution we use the classic Hicks elasticity of substitution defined by Chambers (1988, p. 31) as:

$$(9) \quad y^* = f(x_1 \dots x_I)$$

$$\sigma_{1,2} = \frac{-f_1 f_2 (x_1 f_1 + x_2 f_2)}{x_1 x_2 (f_{11} f_2^2 - 2f_{12} f_1 f_2 + f_{22} f_1^2)}$$

where  $f_i$  and  $f_{ij}$  are derivatives

The above derivations show that if the production function is quadratic and the Hessian is invertible, then it is possible to calculate all the standard econometric comparative static qualitative production measures that apply to the calibrating data set and reflect the average shadow values for restricted or incompletely priced inputs. We note that the supply functions, derived input demands, their associated elasticities, and the elasticities of substitution are obtainable from a data set of any size from one observation upwards. Clearly the reliance on the support space values and the micro theory structural assumptions is much greater for minimal data sets. However the approach does enable a formal approach to disaggregation of production estimates, since the specification of the problem is identical for all sizes of data sets.

This analysis uses the quadratic production function. Two other functional forms that are widely used are the Generalized Leontieff and Translog production functions. As would be expected, the ME empirical reconstruction methods outlined in this

chapter also can be used with these production functions, provided the input data for the trans-log function is bounded for a local optima.

#### **IV. The Empirical Reconstruction of Regional Crop Production in California.**

The empirical setting in which we will present our reconstruction approach is that of the California Statewide Water and Agricultural Production (SWAP) model (Howitt et al., 2001). In this paper, however, we will go beyond the deterministic Maximum Entropy estimation based on a single year of statewide data currently used in SWAP, and employ GME over several years of data from a subset of the original SWAP regions.

SWAP is a multi-input, multi-output economic optimization model that is disaggregated into 24 production regions that span the main agricultural regions of California. This level of disaggregation is based on the way that agricultural data is collected in California, and how water allocation institutions and agencies vary by regions. This specification allows for a robust representation of alternative water management policies that can be interpreted on a statewide basis. Regional crop prices, yields and areas grown were based on annual county agricultural commissioner reports. The data used in our reconstructed GME-based model includes the years 1994 to 2000, with prices normalized to 1992 levels.

##### Data Restrictions

Ideally, production models are reconstructed from a consistent time series of regional data, which includes all the crop inputs and outputs and their associated prices. Unfortunately, such rich, consistent data sets are rarely available. In some cases, comprehensive cross-section survey data is available, but it is rarely collected for more than one year. Given these restrictions, we have to use data collected annually by regional public agencies.

The data available is similar to many disaggregated production data sets collected consistently by public agencies. Typically, such data sets contain data that is available on a regional basis for crop yields, prices, and acres harvested, but rarely is available for other input use on a crop and regional basis. In this example that focuses on irrigation water use and its derived demand we have to generate estimates of the crop water allocation and associated capital inputs. These are derived by combining the crop acreage data and total surface water allocated to regions with crop water requirements based on the annual regional climate, and the efficiency of water application in the region. Other authors (Lence & Miller, 1998b) have found the maximum entropy principle to be useful in recovering activity-specific input usage from aggregate data.

The production functions in this paper are reconstructed from four years of data (1996 – 1999). However, we are faced with data inconsistencies that are commonly encountered when using empirical farm data for estimation. Specifically, the problem is that the observed crop allocations, based on the raw data, sometimes violate the assumption of efficiency and optimal decision-making. Since SWAP, like the majority of economic models, is predicated on the assumption of efficient decisions and bounded rationality, the empirical data has to be reconciled with

economic efficiency before we can proceed with reconstruction of the parameters. The empirical task is to calibrate our model to account for all the crops that are observed to be grown, even those that, based on the raw data, appear to lose money. In short, we have to adjust the raw data to make it consistent with informed and rational allocations.

The most logical explanation for growing crops that appear to lose money is that there is some added benefit from the crop to the producers that is not reflected in the observed data. This increased profitability can take the form of either reduced costs for that crop or increased returns in other crops. The marginal opportunity cost of production for apparently "inefficient" crops that require labor and machinery at a different time than the main "cash" crops will be below the input opportunity costs for the main "cash" crops. Crops with this characteristic are colloquially termed a "filler" crops by some farmers. A more common reason for growing apparently "inefficient" crops is the rotational benefits on crop yields that many low value "rotational" crops confer on cash crops in subsequent years. Technically this effect is an agronomic complementarity that adds to long-term revenue, but is not reflected in the annual crop costs and returns. In both econometric and programming models, these inter-temporal effects are usually addressed by arbitrary data adjustments to the cost or returns of the relevant crops. An alternative, and equally unsatisfactory method, is to impose fixed rotational proportions as constraints among the crops. These ad hoc adjustments do not use a consistent definition of inefficiency, or use a specified measure of adjustment across all crops and inputs. Our approach is to augment the output price of the crops that have binding lower calibration constraints. The augmentation is determined by the shadow values on the constraints and represents the revealed rotational value of these crops.

The capital costs used to calibrate SWAP are restricted to those used in irrigation since this is the particular focus of the model. The annual variable cost of capital in production therefore represents the annual irrigation system cost per acre, and is a combination of labor, management, capital costs and an associated irrigation technology that yields a given irrigation efficiency. The variable capital cost defines a functional relationship between the cost of irrigation technology and improvements in water use efficiency. Thus, investments in "better" irrigation technology result in an increase in irrigation efficiency that uses less applied water to achieve the same yield. In the SWAP model, prior estimates of CES isoquant functions ( USBR & Hatchett, 1997) are used to calculate the variable capital cost that is implied by a given physical efficiency of water use for a specific region and crop. The base year irrigation efficiency is calculated from the ratio of the regional ET divided by the observed applied water. This efficiency value is used in the CES function to solve for the appropriate capital cost using parameter estimates presented in CVPM.

The cost per acre-foot of water for each region uses a weighted average cost of the aggregate supply of water for all districts within the region. Solving the production optimization problem generates the annual amount of applied water that is allocated by crop and region. The annual quantity is apportioned by month, based on monthly crop water requirements, which have been identified for each crop and region using the Department of Water Resources Consumptive Use Model (1997).

### Production Function Specification

In the original SWAP specification, each region has a different production function for each of the crops produced. Within a region the production of different crops is connected by the restrictions on the total land and water inputs available and the land cost function. The wide range of agricultural production inputs has been aggregated and simplified to land, water, and capital.

The production function is written, in general, as:

$$( 10 ) \quad y = f( x_1, x_2, x_3 )$$

The specific quadratic form used in the SWAP model has the form:

$$( 11 ) \quad y = [\alpha_1 \alpha_2 \alpha_3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \\ z_{31} & z_{32} & z_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

where  $y$  is the total regional output of a given crop and  $x_i$  is the quantity of land, water or capital allocated to regional crop production. The quadratic matrix zeta ( $Z$ ) captures decreasing marginal productivity of inputs, as well as interaction effects between inputs. Second order conditions for the production problem require that the zeta matrix is positive definite, and are implemented by imposing necessary conditions on the Cholesky decomposition of the zeta matrix.

The land allocated to different crops is subject to substitution and complementary relationships between crops. These effects are due to the interdependence of crop production rotations, the heterogeneity of land and its restricted quantity for many farm businesses.

The full problem defined over  $G$  regions and  $i$  crops in each region for a single year is:

$$( 12 ) \quad \text{Max } \sum_G \sum_i p_i f_{Gi}( x_1, x_2, x_3 ) - \omega_1 x_1 - \omega_2 x_2 - \omega_3 x_3$$

$$\text{subject to } \sum_{Gi} x_{1Gi} \leq X_1 \quad (\text{Land})$$

$$\sum_{Gi} x_{2Gi} \leq X_2 \quad (\text{Water})$$

where the total annual quantities of irrigated land and water ( $X_1$  and  $X_2$ ) are limited in each region and must be optimally allocated across crops grown in that region. By changing the RHS quantity of water available on the constraint, we generated a derived demand function for each region<sup>9</sup>, which were then used to define agricultural demand nodes for water within the economic –engineering model (CALVIN<sup>10</sup>). Since CALVIN requires that water is valued on a monthly basis, the model specification was modified to give monthly valuations of water, by specifying monthly crop water requirements, as a proportion of the total annual consumptive use applied in each

<sup>9</sup> note that these can be obtained directly from the inverted Hessian, as explained previously

<sup>10</sup> California Value Integrated Network ( <http://cee.engr.ucdavis.edu/faculty/lund/CALVIN/> )

month. For the purposes of this paper, we have kept the optimization on the basis annual resource requirements and allocations.

### **Reconstruction of the Economic Model**

Calibration of the full set of parameters for the production function with three inputs requires that each regional crop be parameterized in terms of nine parameters, three for the linear terms, and six for the symmetric quadratic matrix. After imposing the required symmetry restrictions for the off-diagonal terms, in a single set of base year data the number of equations are limited to three first order conditions which represent the underlying behavioral assumption of optimizing behavior with respect to the three production inputs, and a total output equation. Given the small sample of four years, we have twelve observations on input allocations, and four observations on output, for a total of 16 observations. The resulting disaggregated production function has 7 degrees of freedom.

In addition, given the methods used to generate the water application and irrigation capital input levels, there is likely to be considerable collinearity between the inputs for a given crop and region. Fortunately, the generalized maximum entropy (GME) approach that we use for the reconstruction is robust under collinearity and low degrees of freedom.

The SWAP model is reconstructed in three stages. In the first stage, a linear programming ( LP ) model is constructed for each region that incorporates all the available data on cropping acreages, annual water use, yields, output prices and input costs and quantities. The LP model is maximized subject to land constraints and also a set of constraints that calibrate the model to the observed land use and production quantities in each region. This initial stage is the same as that used in the positive programming approach ( Howitt 1995 ). Stage two consists of reconstructing the production function by Maximum Entropy, which is followed by the solution of the non-linear constrained maximization problem (12) in the third stage.

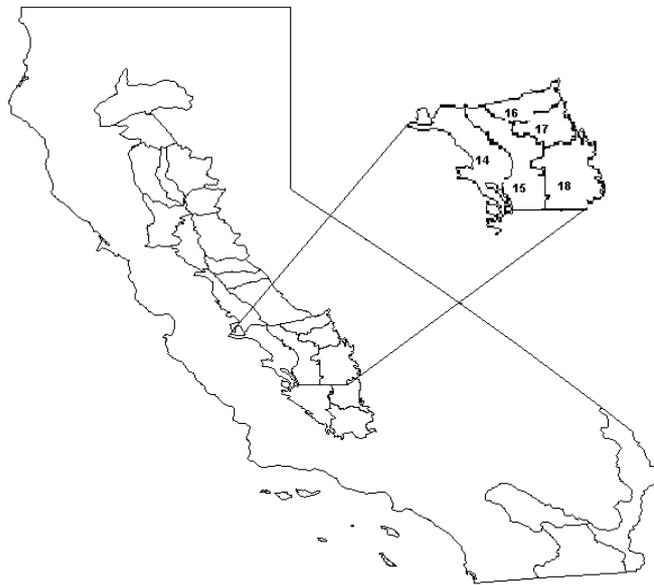
Where land is limiting and an adequate measure of the rental rate for land is not available, the opportunity cost of land must be inferred from the shadow value of the crops grown. However, for multicrop systems where rotational interdependencies among crops change their marginal contribution to rental returns, the crops need to be divided into those that are net users of attributes ( weed and disease control, or fertility) from other crops and those that are net contributors to the farm productivity. The third group is assumed to neither contribute to, nor reduce farm productivity. We can term these three groups of crops as cash crops, rotational crops, and filler crops.

#### Reconstructing a Four Region Production Model

Data from four regions in California's central San Joaquin valley are used to illustrate disaggregated reconstruction process. The reconstruction sequence proceeds by first solving the constrained calibration in equation (1) and using the resulting shadow values to define the LHS of the first-order conditions for input allocation in equation (3), and solving the generalized maximum entropy problem. The resulting probabilities from the GME solution are used to generate the expected values of the production function parameters  $\alpha_i$  and  $z_{ii'}$ . The regional production problem defined in equation (12) is solved using the production function specified in equation (11).

We use data from four of the five contiguous regions shown in figure( 1). Regional irrigated land ranges from 1,022 to 51 thousand irrigated acres and the regions have different institutions, water prices, and drainage conditions. The reconstruction problem was defined as a series of linked nonlinear optimization problems and solved by GAMS (Brooke et al. 1988), taking 90 seconds on a 0.75 megahertz PC. The problem requires the reconstruction of twenty-two three-input quadratic production functions, one for each regional crop grown.

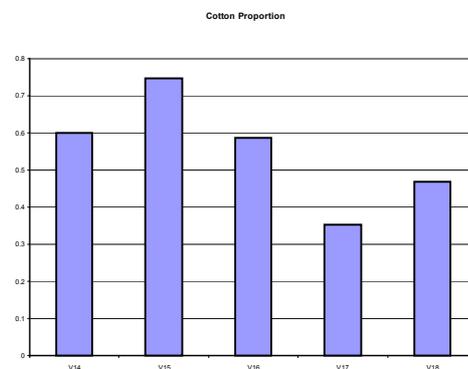
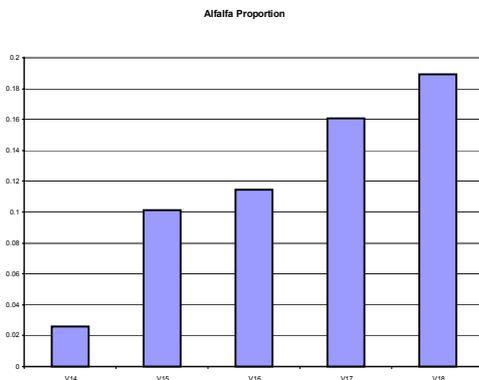
Figure 1



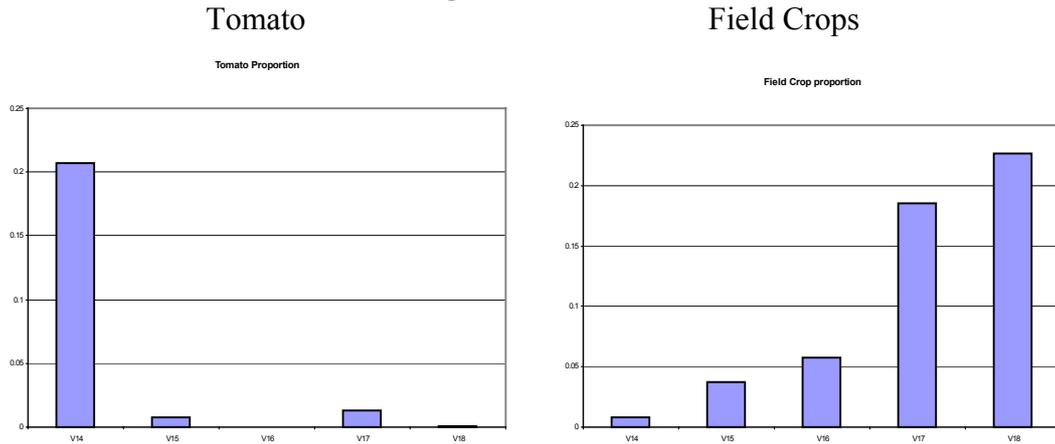
Figures 2a and 2b

Alfalfa

Cotton



Figures 2c and 2d



Despite being contiguous regions in the same valley, there are considerable differences in cropping patterns. Figures 2 (a) – (d) show the regional differences for four of the principle crops.

The four regions despite their proximity, show a wide variation in crop selection due to differences in soil-type, microclimate, water quality, and water constraints.

The goodness-of-fit of the estimated GME regional crop production models can be measured by an equivalent to the familiar  $R^2$  parameter<sup>11</sup> applied to the GME estimation defined in equations 3 –3d. The results are shown in Table 1.

When the  $R^2$  parameters are weighted by the regional crop proportion to the total cropped area they show the overall fit of the GME estimation. The weighted  $R^2$  parameter for the disaggregated crop production is 0.597.

In absolute terms the disaggregated simulation model captures the regional input allocation less well, with an average percentage error of 21.4% for land, 23.1% for water, and 21.3 % for capital. The average absolute percentage error of simulated sample crop yield prediction ranged from 9.5% to 24.9% .

<sup>11</sup> Defining  $\hat{y}_i$ ,  $y_i$ , and  $\bar{y}$  respectively as the estimated production, actual production, and mean

production, the simulation  $R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (y_i - \bar{y})^2}$

Table 1.  
Disaggregate Model  $R^2$  Values for Regional Production

	Cotton	Alfalfa	Field	Grain	Tomato	S-Beets
Region						
V14	0.644	-0.317	0.953	0.704	0.274	0.669
V15	0.671	0.464	0.697	0.658	0.699	0.638
V16	0.659	0.282	0.265	0.678	--	--
V17	0.703	0.612	0.732	0.585	0.782	0.638

The production  $R^2$  measures show that for most regions and crops the GME production function has good explanatory power. What is surprising, and currently inexplicable, is how the regional disaggregated model can fit a crop well for three regions and be non-informative for the fourth region. See, for example, the case of Alfalfa in table 1. This problem needs further investigation, and the explanation may be that the common simple price and yield expectation scheme are inappropriate for that region. The weighted percent deviation for input allocation and crop production generated by the optimization model is shown below in table's 2a and 2b. The results in table 2a show similar large errors in the input allocation results, and slightly better results in the production results in table 2b.

Table 2.a  
Average % Absolute Error in Regional Input Allocation

Land	Water	Capital
21.4%	23.1%	21.3%

Table 2.b  
Average % Absolute Error in Crop Production

Cotton	Alfalfa	Field Crops	Grain Crops	Tomatoes	Sugar Beets
24.7%	18.8%	20.7%	13.9%	9.5%	24.9%

Out of sample simulations for regional crop production in 2000 were calculated in two ways. The first method uses analytical supply functions derived earlier couples with the expected output and input prices for 2000. The opportunity cost component of the land input cost is based on the average values over the four-year data set. The second method of obtaining out of sample production forecasts, is to run the regional optimization models using the GME production functions and forecasted output and input prices. In an initial trial, neither method produced results with a satisfactory precision. The weighted error in production forecasts using the analytical functions was 37%, while the weighted out of sample production error using the optimization model was 28.7%. Neither of these results is acceptable when compared with the error from a naïve extrapolation of 1999 production levels of 9.4%. Clearly, the out of sample prediction should be better given the reasonable  $R^2$  values for the in-sample production fit. We will investigate alternative auto regressive approaches in the future.

Being that the primary use of this regional model is to estimate the response by farmers to changes in water price or availability, it follows that the elasticity of

demand for water is the most important policy parameter for this study. If the elasticity differs over regions there will be gains in policy precision for the disaggregated regional model.

Table 3.  
Regional Input Demand Elasticities.

	LAND	WATER	CAPITAL
V14	2.44	1.24	0.91
V15	1.80	2.35	0.88
V16	4.09	0.52	0.85
V17	3.29	0.78	0.80

Table 3 shows that the water derived demands are in the expected elasticity ranges, with substantial regional variation in elasticity. An average regional elasticity would distort the response to water policy given the substantial differences in the size of the regions and the elasticities. Land shows a smaller range of elasticities, and capital inputs are inelastic and fairly uniform over regions, reflecting the uniform underlying data manipulations and the absence of a constraint on capital.

Table 4.  
Regional Output Supply Elasticities

	Cotton	Fodder	Field Crops	Grain Crops	Tomatoes	Sugar Beets
V14	2.12	2.10	1.84	2.07	0.56	0.81
V15	2.09	1.86	2.14	1.37	0.71	0.60
V16	2.48	1.78	2.24	2.57	--	--
V17	2.38	0.80	2.22	2.47	0.64	0.63

The supply elasticities shown in table 4 are within the expected range, with the key cotton crop being quite elastic. The inelastic supply response of tomatoes and sugar beets may be due to the contracts under which tomatoes are grown, and the current restriction on sugar beet processing capacity in the region. Some crops, such as cotton have a fairly uniform elasticity across regions indicating there's little advantage in crop response to having a disaggregated model. However, other crops such as fodder and grain crops show upwards of a three-fold difference in elasticity values, indicating that a significant benefit could be gained from a disaggregated analysis.

From equations 4, 5, and 8, we see that the supply and demand elasticities heavily depend on the inverse Hessian matrix. As noted earlier, the values of this matrix depend on the support values specified for the individual parameter probabilities. While we have reasonably strong priors on functions of the production function parameters, we do not have priors on the parameters themselves, especially when they are specified in terms of a matrix decomposition. In this empirical example we used the resulting elasticities to balance the support values between the linear and quadratic terms. This explicit use of prior information is necessary and useful when trying to reconstruct reasonable policy models from limited data.

Table 5 shows the Hicks elasticities of substitution between land, and water and irrigation capital for a specific region (region V15). All the substitutions are surprisingly inelastic. Since the common requirement to calculate elasticities of substitution is the ability to reconstruct the Hessian, other more sophisticated substitution measures can also be calculated.

Table 5.  
Regional Hicks Elasticities of Substitution

	Water	Capital
Cotton.Land	0.0982	0.0300
Cotton.Water	0.0330	
Fodder.Land	0.0420	0.0272
Fodder.Water	0.0434	
Field.Land	0.0442	0.0262
Field.Water	0.0393	
Grain.Land	0.0634	0.0292
Grain.Water	0.0363	
Tomato.Land	0.0874	0.0532
Tomato.Water	0.0572	
SBeets.Land	0.0650	0.0347
SBeets.Water	0.0745	

### **Testing for Regional Policy Parameter Differences by Bootstrapping**

Since this paper is motivated by the hypothesis that disaggregated models of agricultural production will yield a more precise policy response, we use the bootstrap method to create an empirical distribution that will allow us to test for significant regional differences in critical policy parameters. The critical policy parameters are the supply and demand elasticities. The bootstrap results are used to calculate the standard errors of the elasticity estimates, which then enable us to do hypothesis testing with these elasticities. By re-sampling from our data for 130 bootstrap replications we generate the standard errors for our parameters and conduct pairwise t-tests for their equality across regions. Thus, we obtain tables 6 and 7, below.

Table 6.  
Bootstrap t-values on Pairwise Tests across Regional Supply Elasticities

	Cotton	Fodder	Field	Grain	Tomatoes	S Beets
V14.V15	0.423	0.948	-0.019	2.785	-1.238	1.088
V14.V16	-2.572	1.115	-0.020	-1.526	3.287	2.876
V14.V17	-1.627	3.926	-0.020	-1.665	-1.076	1.012
V15.V16	-2.829	0.329	-0.386	-3.402	5.939	2.953
V15.V17	-1.763	3.253	-0.333	-4.377	0.590	-0.258
V16.V17	0.579	3.869	0.062	0.319	-4.043	-2.693

In table 6, the hypothesis of equal supply elasticities is rejected at the 5% level in 14 out of 36 cases. The equality of field crop supply elasticity fails to be rejected in all pairwise comparisons, while in all other crops the hypothesis fails to be rejected in approximately half of the cases. Thus the advantage of disaggregated models for estimating supply elasticities is supported in 39% of the cases.

Table 7.  
Bootstrap t-values on Pairwise Tests across Regional Input Demand Elasticities

	LAND	WATER	CAPITAL
V14.V15	4.236	-9.156	0.535
V14.V16	-8.098	10.696	1.468
V14.V17	-3.898	6.620	2.559
V15.V16	-8.866	11.748	0.958
V15.V17	-7.502	10.551	2.989
V16.V17	2.845	-7.115	1.204

Table 7 of regional demand elasticity tests shows much greater evidence of regional heterogeneity. For land and water the equality of demand elasticities is rejected in all cases, showing that there is a clear advantage from model disaggregation in the estimation of these policy parameters. For the capital input, the equality hypothesis is rejected in only two out of six cases. These results suggest that the policy bias induced by aggregating over the regions will likely exceed the small sample bias in the GME estimates.

## V. Conclusions

This paper shows that by using a combined PMP, GME and Bootstrap approaches, it is possible to reconstruct flexible form production function models from a data set of modest size. In addition, the bootstrap method provides a measure of the precision of the results, and the ability to test critical policy parameters. Using these methods, a researcher can reconstruct a similar theoretically consistent flexible form production model using a data sets that range from “LP budget data” to full econometric data sets with standard degrees of freedom. The convergence of GME estimates to conventional estimates as the sample size increases means that as the data set is expanded there is a continuum between optimization and econometric models.

The reconstructed production models yield all the comparative static properties and parameters of large sample models, thus enabling the input demands, output supplies and elasticities of substitution to be calculated directly instead of the parametric methods required by programming models. The effect of any constraints on production is directly incorporated in the estimates through the addition of the constraint shadow values to the nominal prices of the allocatable resources.

We want to emphasize a cautionary note on the interpretation of the results from the reconstruction method. Users should be aware of where the information in the model results comes from. In general, reconstruction methods rely strongly on prior

structural information that is combined with sparse sample information. This does not invalidate the results, but should emphasize that they must be interpreted as reconstruction results rather than econometric estimates that get something for nothing. For example, the performance of the production function model changes dramatically if the “unbiasedness constraint” is removed from the GME stage. The model results change from those with a good fit to sample data, but a poor out of sample prediction precision, to ones with a very poor  $R^2$ , but an out of sample precision that is greatly improved for both the analytical supply function and optimization model predictions. Clearly, there is additional work to be done on the stability properties of reconstruction models.

Due to the sparse data set used in this paper ( four years), the resulting production model significantly depends on the definition and range of prior support values for the parameters, and the expectation structure assumed for the farmer’s expected yields and prices. However, we feel that this use of prior information on the farming system and process of production is a valuable source of modeling information that should be formally included. An advantage of modeling production functions is that they are readily understood by other disciplines, which are thus able to add information for the prior support values or constraints.

In this example the demand and supply elasticities from the disaggregated models were shown to be significantly different. These results suggest that the gain from disaggregation outweighs the small sample bias. However, the gain from disaggregation of production models is an empirical result that needs substantially more testing before one can conclude that it is a common phenomenon.

### References

- Alig, R.J., D.M. Adams and B.A. McCarl. " Impacts of Incorporating Land Exchanges Between Forestry and Agriculture in Sector Models." *Journal of Agricultural and Applied Economics*, 30:2 (1998): 389-401.
- Antle, J.M. & S. Capalbo. “Econometric-Process Models of Production.” *American Journal of Agricultural Economics*, 83 (May 2001): 389-401.
- Brooke, A., D. Kendrick, & A. Meeraus. "GAMS: A Users's Guide". The Scientific Press, 1988.
- Chambers, R. “Applied Production Analysis: The Dual Approach”, Cambridge University Press, Cambridge, 1988.
- Desmedt, P. " Maximum Entropy and Bayesian Image reconstruction: Overview of Algorithms" LEM Technical report DG91-14, University of Ghent, 1991.
- Efron, B. and J. Tibsharani “ An Introduction to the Bootstrap” London, Chapman & Hall ( 1993 ).
- Fraser, I. “An Application of Maximum Entropy Estimation: The Demand for Meat in the United Kingdom”, *Applied Economics*, 32(2000): 45-59.

Golan, A., G. Judge, D. Miller. "Maximum Entropy Econometrics: Robust Estimation with Limited Data", John Wiley & Sons, Chichester, England, 1996(a).

Golan, A., G. Judge & J. Perloff. "Estimating the Size Distribution of Firms Using Government Summary Statistics." *J. Indust. Econ.*, 44 (Mar 1996(b)):69-80.

Golan, A., G. Judge, S. Robinson. (1994) "Recovering Information from Incomplete or Partial Multisectoral Economic Data." *Rev. Econ. Statistics.*, 76 (Aug): 541-49.

Hansen.L.P. and J.J.Heckman "The Empirical Foundations of Calibration." *Journal of Economic Perspectives*, 10:1, (Winter 1996), pp 87-104.

Heckelei, T. & W. Britz. "Concept and Explorative Application of an EU-wide, Regional Agricultural Sector Model (CAPRI-Project)". Presented at the 65th EAAE-Seminar, Bonn. March, 2000.

Howitt, R. E. "Positive Mathematical Programming." *American Journal of Agricultural Economics*, 77(Dec 1995):329-342.

Howitt, R.E., K.B. Ward, S. Msangi. "Statewide Water and Agricultural Production Model". Written as Appendix N to the report for the California Value Integrated Network (CALVIN) model., 2001. (see: <http://cee.engr.ucdavis.edu/faculty/lund/CALVIN/>).

Howitt, R.E. & S. Msangi. "Consistency of GME Estimates Through Moment Constraints." *forthcoming Working Paper*, Department of Agricultural and Resource Economics, University of California at Davis, 2002.

Just,R.E., D Zilberman, and E. Hochman "Estimation of Multicrop Production Functions" *American Journal of Agricultural Economics*, 65 (1983): 770-780.

Just, R.E., J.M. Antle.(1990) "Interactions Between Agricultural and Environmental Policies - A Conceptual Framework". *American Economic Review*, 80:2(May):197-202.

Lansink, A.O., E. Silva & S. Stefanou. "Inter-Firm and Intra-Firm Efficiency Measures." *Journal of Productivity Analysis*, 15:185-199.

Lence, S. H. & D.J. Miller. "Estimation of Multi-Output Production Functions with Incomplete Data: A Generalized Maximum Entropy Approach." *Eur. Rev. Agr. Econ.*, 25(Dec 1998a):188-209.

Lence, S. H. & D.J. Miller. "Recovering Output-Specific Inputs from Aggregate Input Data: A Generalized Cross-Entropy Approach." *American Journal of Agricultural Economics*, 80(Nov 1998b):852-867.

Leon, Y., L. Peeters, M. Quinqu & Y. Surry. "The Use of Maximum Entropy to Estimate Input-Output Coefficients From Regional Farm Accounting Data." *Journal of Agricultural Economics*, 50(Sep, no.3):425-439.

Love, H.A. "Conflicts Between Theory and Practice in Production Economics." *American Journal of Agricultural Economics*, 81 (Aug 1999): 696–702.

Marsh, M.L. & R. Mittlehammer. "Adaptive Truncated Estimation Applied to Maximum Entropy." *Presented at the Western Agr. Econ. Assoc. Annual Meetings*, Logan, Utah, (July 2001) 30 p.

McCarl, B.A., D.M. Adams, R.J. Alig & J.T. Chmelik, "Analysis of Biomass Fuelled Electrical Power plants: Implications in the Agricultural and Forestry Sectors." *Annals of Operations Research*, 94(2000): 37-55.

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

Mittelhammer, R.C., G.G. Judge, D. J. Miller. "Econometric Foundations" New York : Cambridge University Press, 2000.

Paris, Q. & R.E. Howitt. "Analysis of Ill-Posed Production Problems Using Maximum Entropy." *American Journal of Agricultural Economics*, 80 (Aug 1988):124-138.

Paris, Q. "Multicollinearity and Maximum Entropy Estimators." *Economics Bulletins*, 3 (no. 11, 2001):1-9.

Paris, Q. & M. Caputo. "Sensitivity of the GME Estimates to Support Bounds." *Working Paper*, Department of Agricultural and Resource Economics, University of California at Davis, (Aug 2001) 18 p.

US Department of the Interior, Bureau of Reclamation.(1997) " Central Valley Project Improvement Act: Draft Programmatic Environmental Impact Statement", Sacramento, California.. (S. Hatchett main author of CVPM model description in *Technical Appendix*)

Wu, J.J. & B. Babcock. "Meta-modeling Potential Nitrate Water Pollution in the Central United States." *Journal of Environmental Quality*, 28( Nov-Dec, 1999):1916-28.

Zhang, X. & S. Fan. "Crop-Specific Production Technologies in Chinese Agriculture." *American Journal of Agricultural Economics*, 83 (May 2001): 378-388.