ESTIMATING STATE-CONTINGENT PRODUCTION FUNCTIONS

by

Svend Rasmussen (sr@kvl.dk)
Kostas Karantininis (kok@kvl.dk)

Food & Resource Economics Institute
The Royal Veterinary & Agricultural University of Copenhagen (KVL)
Phone: (+45) 35282266, FAX: (+45) 35282295

Paper prepared for presentation at the XIth Congress of the EAAE
(European Association of Agricultural Economists)
Copenhagen 2005

ESTIMATING STATE-CONTINGENT PRODUCTION FUNCTIONS

By

Svend Rasmussen and Kostas Karantininis
Food and Resource Economics Institute
The Royal Veterinary and Agricultural University, KVL
Copenhagen, Denmark

Abstract
The paper reviews the empirical problem of estimating state-contingent production functions. The major problem is that states of nature may not be registered and/or that the number of observation per state is low. Monte Carlo simulation is used to generate an artificial, uncertain production environment based on Cobb Douglas production functions with state-contingent parameters. The parameters are subsequently estimated based on different sizes of samples using Generalized Least Squares and Generalized Maximum Entropy and the results are compared. It is concluded that Maximum Entropy may be useful, but that further analysis is needed to evaluate the efficiency of this estimation method compared to traditional methods.

Keywords*:* Maximum Entropy, State-Contingent, Uncertainty, Production, Monte Carlo Simulation
JEL classification: C13, C15, D80

1. Introduction

The classical approach to the problem of optimizing production under risk/uncertainty is the expected utility model (EU model). The EU-model is, in its basic form, a relatively general model. But as regards empirical application, the tradition has developed over time to the EU-model being the equivalent of a model, where utility is maximized as a function of the expected value and variance of profit (EV model) (Robison and Barry; Dillon and Anderson; Hardaker, Huirne, and Anderson).

This approach to decision making under uncertainty has been severely criticized by Chambers and Quiggin in their book on state-contingent production from 2000, as well as in subsequent papers (Chambers and Quiggin, 2002a; 2002b). The main problem being that the traditional approach typically does not consider the interaction between the uncontrolled (uncertain) variables and the decision variables controlled by the decision maker. Furthermore, although Dillon and Anderson realized the basic need for modelling this kind of interaction, they did not derive criteria for optimal production that went beyond maximizing utility, defined as a function of expected value and variance of profit

With the state-contingent approach developed by Chambers and Quiggin (2000), the foundation for alternative ways of describing and analyzing production under uncertainty, were made available. In a recent article, Rasmussen (2003) used the state-contingent approach to derive criteria for optimal production (input use) under uncertainty. While the article illustrates that the state-contingent approach has the merit of being based on well-known marginal principles and optimization tools, it also indicates that the state-contingent approach has its own weaknesses when it comes to empirical application. Thus, the basic problem of not knowing the decision makers' utility function still exists, and the problem of how to estimate state-contingent production functions, has not been solved. Thus, the question of how to apply the theory of state-contingent production to the real problems of actual decision making still has no clear answer.

This paper considers the problem of how to estimate state-contingent production functions. The problem is that empirical data have not traditionally been registered in a form which may be appropriate for estimating state-contingent production functions. Data from farm accounts do not normally include observations on the states of nature. Further, the number of different states of nature may be very large, and therefore the number of observations per state may be rather small. Estimation methods therefore call for methods that may handle small number of observations.

To compare alternative estimation methods, alternative samples of production data based on a pre-specified production function with state-contingent parameters are generated. The data are generated using Monte-Carlo simulation, and alternative sample sizes of both ordinary time series and panel data are generated. Parameters of the pre-specified production function are estimated using traditional estimation methods (OLS, and ITSUR), and Maximum Entropy (ME), and the estimation efficiency is compared measuring how well the alternative methods are able to estimate the correct parameter values.

The state-contingent approach to decision making under uncertainty may or may not involve an improvement compared to the traditional methods. It is not clear whether the state-contingent approach has the potential at all of providing a better framework for decision making than the traditional EV model. The further perspective of the research presented in this paper is to use the results on how to estimate state-contingent production functions, as a basis for answering the following question: With a given set of data, is it better to base decision making on estimated state-contingent production functions, or does the well-known EV model based on an estimated stochastic production function and variances provide just as good – or maybe even better decisions?

2. The problem of estimating state-contingent production functions

The state-contingent approach to describing production under uncertainty is based on the concept of state-contingent production functions. Depending on the state of nature, there is a specific (i.e. state-contingent) production function. Thus, the relation between controllable inputs and output (the production function) for production of wheat will depend on the climate (the state of nature): If it is a rainy season the production function will be different compared to a dry season.

More generally, consider an uncertain production environment consisting of a set of states $\Omega = \{1, 2, \ldots, s, \ldots, S\}$, from which 'nature' picks the state of nature independently of the decisions made by the decision maker. Nature picks the state of nature *after* the decision maker has made his production decision. The decision variables are the amounts of inputs and the choice of the technique[1)]. If the input vector is $\mathbf{x} = (x_1, \ldots, x_N)$, then the amount of output z produced if nature picks state s is: $z_s = f_s(x_1, \ldots, x_N)$, where $f_s(.)$ is the state-contingent production function in state s. With an output price of p_s in state s, and a vector of input prices $\mathbf{w} = (w_1, \ldots, w_N)$, the net-return from producing product z in state s is $y_s = p_s z_s - \mathbf{wx}$.

To determine the optimal production decision, the decision maker has to know *all S* production functions[2)]. In case he does, all the production uncertainty has been eliminated in the sense that there is no uncertainty concerning the production in state s. The only uncertainty left is the frequency by which state s occurs.

It is hard to imagine that this ideal case would take place in practice. First of all, the *number of possible real states (S) is often very large*. (This is indeed the case when the variables describing the states of nature, are continuous variables). Therefore, if state-contingent production functions are available, it will in practice typically be for only *some* of the possible real states. To illustrate, consider the simple decision problem of fertilizer application to a crop of barley. The yield of barley four months later depends both on the amount of fertilizer applied now, but also on the real state of nature during the growing season. Assume for simplicity that the real state of nature may be quantified by the amount of sunshine and rain during the growing season. Further assume that the relevant interval of possible amount of rainfall is between 10 and 50 centimetres, and that the relevant interval of the amount of sunshine is between 200 and 800 hours. With only these two state-variables describing the real states of nature, there would – if state-variables are measured in integer units - be 50x600 = 24,000 different states of nature. Imagine that state-contingent production functions are estimated based on experimental yields. Then, even in the unrealistic case that none of the states came out twice,

[1] The choice of different techniques may cause some problems of non-convexity. This is not consider here.
[2] The decision maker also has to know the utility function and the state-contingent output prices. The problem of determining these parameters is not considered here (see Rasmussen (2003)).

it would take at least 24,000 years/experiments[3] to collect enough observations to estimate the 24,000 state-contingent production functions!

Secondly, a state of nature is often characterized by a large number of state-variables. If only a few of these variables are in fact observed/registered when doing the experiments which create the data, then the state-description is *incomplete*. The variables registered could be e.g. monthly rainfall and hours of sunshine/month. However, other variables (like for instance wind velocity or CO_2 content of the atmosphere) influencing the output may not be observed (and registered). In that case the date set is incomplete, and it is only possible to estimate the production functions that refer to the registered states. Besides this, these functions will be stochastic production functions because the level of the non-registered state-variables may vary.

In empirical work the time series of data available for estimating production functions are often limited. With the objective of estimating state-contingent production functions, this problem grows drastically, because the number of observations for each state of nature typically will be very small if existing at all! For instance, with just 5 possible states with equal probabilities, one would need a time series of at least 50 experiments to have just 10 observations per state.

It is therefore critical to the empirical application of the state contingent approach that we consider estimation methods that will work even with very few empirical observations. One such method is the method of Maximum Entropy (Golan et al., 1996). Also the use of panel data will improve estimation efficiency considerably. Both approaches are considered in the following.

3. Generating the stochastic environment.

3.1. The real states of nature.

The uncertain production environment used as the basis for empirical analysis in the following was created using the following Cobb-Douglas state-contingent production function:

$$z_s = A_s x_1^{a_{1s}} x_2^{a_{2s}} x_3^{a_{3s}} \qquad (s = 1, \ldots, S) \qquad (1)$$

where z_s is the output in state s, x_1, x_2, and x_3 are three variable inputs, and A_s, a_{1s}, a_{2s}, and a_{3s} are the parameters in state s.

The various (S) states of nature were generated by combining the following values of the four parameters:

Table 1. Values and relative frequencies of parameters (state-variables)

A		a_1		a_2		a_3	
Value	Probability	Value	Probability	Value	Probability	Value	Probability
2	0.20	0.05	0.10	0.12	0.15	0.24	0.20
4	0.50	0.12	0.35	0.22	0.30	0.28	0.30
5	0.30	0.19	0.45	0.32	0.40	0.32	0.30
		0.26	0.10	0.42	0.15	0.36	0.20

Using all possible combinations of the parameter values in Table 1, a total of 3×4×4×4 = 192 possible states of nature (S = 192), and a corresponding number of state-contingent production functions is generated. In the following we shall for convenience refer to a specific state of nature as $s_{i,j,k,l}$, where i is an *index of state-variable* 1, j is an index of state-variable 2, k is an index of state-variable 3, and l is an index of state-variable 4[4]. Index i is in this example at the same time an index of the possible values of parameter A (i=1,…, 3), j is an index of the possible values of parameter a_1 (j=1,…, 4), k

[3] In the following it is appropriate to think of the data generating process to take place on an imaginative experimental station. Therefore, it is more appropriate to talk about "experiments" rather than "years", since one experiment may span for more than one calendar year, whereas many experiments can be carried out within one calendar year.

[4] The four state variables characterising a specific state of nature (weather) could be for instance be sun, rain, temperature, wind velocity.

is an index of the possible values of parameter a_2 (k=1,…, 4), and l is an index of the possible values of parameter a_3 (l=1,…, 4).

The four state-variables are assumed to be independent, and to occur with a relative frequency corresponding to the probabilities shown in Table 1. Thus, the probability of a specific state of nature is calculated simply by multiplying the probabilities of the individual state-variables. Thus, for instance, the probability of state $s_{2,3,1,4}$ is 0.50×0.45×0.15×0,20 = 0.00675.

In real life the state-variables will typically *not* be independent. Thus, if the state-variables are for instance sunshine, rain, and temperature, these three variables are typically not independent variables. However, in the context considered in this paper, this is not important, and the individual state variables are for convenience considered as being independent.

One should also notice that the number of states is subjectively determined by the scale of measurement of the individual state-variables and the number of state-variables. If the amount of rain is for instance measured in integer number of 100 mm intervals, then the state-variable rain may take only 4 or 5 discrete values. If instead rain is measured in integer number of 1 mm intervals, then the state-variable rain will include maybe 500 discrete values.

As with the dependency between state-variables, this subject will not be considered further. With the choice made here the number of possible states are 192, a number, which will be sufficient to serve as an example for the following illustrations.

3.2. The registered states of nature

In the example above, there are 192 *real states*, each state being generated by a certain combination of values of the four state-variables.

In practice, only *some* of the state-variables influencing production are *registered* (together with the production data (input and output)). Consider for instance an agricultural experimental station performing experiments with different levels of various inputs. Besides registering the amount of input (*controlled* input) and output, experimental stations typically also register the state of nature in the form of the level of some of the (none-controllable) state-variables influencing production. However, hardly *all* state-variables influencing production are registered. And even if they were, these data may not be available or may not be available in an appropriate form, so that the data can be used by the decision maker - i.e. the farmer. The typical case in empirical work is therefore that *only some of the state-variables are registered* and therefore are available for econometric analysis. In the extreme case, none of the state-variables are registered, and no information about the states of nature is available at all.

In the following, the situation with *only some of the state-variables (A and a_1) being registered* is illustrated by assuming that in the example above *only the first two state-variables are registered state-variables*. In this case the empirical production data therefore covers only 3×4 = *12 registered states* (as opposed to the total amount of 192 *real states*). And for each of these registered states, the other two state-variables (a_2 and a_3) could take on any of the other 4×4 = 16 values. Thus, with a given amount of input, the output in any of the 12 registered states will be a stochastic variable with 16 possible values distributed according to the probability distribution of each of these 16 states calculated by multiplying the probabilities of state-variables a_2 and a_3 according to the probabilities in Table 1[5].

4. Estimation of state-contingent, stochastic production functions.

4.1. Amount of input

The stage is now prepared for generating the empirical data to be used for estimating the state-contingent production functions. However, first one needs to consider the amount of input applied.

To generate the production data, the following "experimental plan" was used:

[5] The disturbance term generated in this way is heteroscedastic. The problem is dealt with accordingly in the estimation procedure in Section 4.

Table 2. Application of input (experimental plan)

x_1	x_2	x_3
10	40	15
30	80	45
50	120	75
70	160	105
90	200	135
110	240	165

The term “experimental plan” is in this case to be interpreted as follows: For each experiment and for each plot of land, the amount of each input applied is determined by drawing randomly an amount of input from the individual columns of the six possible amounts of input shown above. Thus, in any experiment and on any of the plots, the experimental station may have applied a combination of for instance 30 units of x_1, 160 units of x_2, and 105 units of x_3. This specific combination of these amounts of input occurs in the data set with a relative frequency of (1/6)×(1/6)×(1/6) = 1/216.

4.2. Generating data

The *data generating process* runs as follows:

For each experiment *t* the following steps are carried out:

1) The amount of input applied to a “plot” is determined by random choice of the possible input amounts in Table 2. These amounts of input applied to the plot in question are registered.

2) In the case of more plots per experiment, the procedure in (1) is repeated for every plot.

3) The state of nature in the experiment in question is determined by drawing individually the four state-variables in Table 2 randomly according to the probabilities in Table 2. *Only the value of two state-variables A and a_1 are registered, i.e. the value of state-variable a_2* and a_3 *are not registered).*

4) The amount of output y is calculated for each plot using (1) by inserting the relevant amounts of input determined in (1) and (2), and the parameter values determined in (3). The amount of output y is registered (for each plot).

5) The experiment number *t* is registered

A large number of experiments (in this case 40,000) were generated. The resulting data set with data registered as mentioned above is in the following referred to as the “population”. It is from this population that the samples of data used for estimating the state-contingent production functions in the following are drawn.

As only the first two state-variables, each with three and four possible values, respectively, have been registered, it is only possible to estimate - depending on the sample size - a maximum of 3×4 = 12 different state-contingent production functions. According to the definitions above, these production functions are thus in fact *stochastic*, state-contingent production functions.

In the following, two different methods will be used for estimation: The traditional Ordinary and Generalized Least Squares (OLS and ITSUR), and the Generalized Maximum Entropy (GME).The two methods will be compared on the basis of their ability to replicate the true parameter values shown in Table 1.

The production functions to be estimated is given in equation (1) above, where the number of (registered) states is now 3×4 = 12, where the number 3 refers to the number of possible values of state-variable one (here the parameter A), and the number 4 to the number of the possible values of state-variable two (here the parameter a_1). In the following we shall refer to the 12 registered states as $s_{i,j}$ (i=1,…,3; j=1,…, 4). Using this way of naming registered states of nature, the production functions to be estimated are:

$$z_{ij} = A_{ij} x_{1ij}^{a_{1ij}} x_{2ij}^{a_{2ij}} x_{3ij}^{a_{3ij}} \qquad (i = 1, \ldots, 3;\ j=1,\ldots, 4) \qquad (2)$$

Taking the logarithm and adding an error term ε to account for the fact that the value of state-variable 3 and 4 are not known and that therefore (2) are stochastic production functions; the econometric model has the following form:

$$\ln z_{ij\tau} = \ln A_{ij} + a_{1ij} \ln x_{1ij\tau} + a_{2ij} \ln x_{2ij\tau} + a_{3ij} \ln x_{3ij\tau} + \varepsilon_{\tau}$$
$$(i =1, \ldots,3;\ j=1,\ldots, 4;\ \tau=1,\ldots, T_j) \qquad (3)$$

where the index τ refers to the number of observations of each registered state, and where therefore $z_{ij\tau}$ is the τ'th observation of output in state s_{ij}. T_{ij} is the total number of observations of state s_{ij} in the sample.

To estimate these state-contingent production functions, a sample of observations (experiments) is drawn from the data set (the population) generated as described above. To analyse the consequence of having available different sample sizes, estimations were carried out for different number of experiments, in this case 100 experiments, 200 experiments and 400 experiments, respectively. At the same time, the number of plots was varied from one to three plots to measure the consequence of having available more observations for the same state of nature.

With more than one plot, the model (3) changes to:

$$\ln z_{ij\tau p} = \ln A_{ij} + a_{1ij} \ln x_{1ij\tau p} + a_{2ij} \ln x_{2ij\tau p} + a_{3ij} \ln x_{3ij\tau p} + \varepsilon_{\tau p}$$
$$(i =1, \ldots,3;\ j=1,\ldots, 4;\ \tau=1,\ldots, T_{ij};\ p=1, 2 ,3) \qquad (4)$$

where p is an index of plot

4.3. OLS and ITSUR -estimation.

As the econometric model in (4) is linear in the parameters, Ordinary Least Squares (OLS) can be applied directly to estimate each of the 12 state-contingent production functions.

However, the variance of the error term is not a constant. Using the information on how the data were generated, it is easy to show that the variance σ_{ε}^2 of the error term ε_{τ} is:

$$\sigma_{\varepsilon}^2 = (\ln(x_2))^2 \sigma_{a_2}^2 + (\ln(x_3))^2 \sigma_{a_3}^2 \qquad (5)$$

where $\sigma_{a_2}^2$ and $\sigma_{a_3}^2$ are the variances of the parameters a_2 and a_3, respectively. Thus, the error term is *heteroscedastic* and the estimation was performed accordingly

This information in (5) would not be available if the data were real empirical data. Therefore, it may be considered incorrect to use this information in this simulation case. On the other hand, an experienced researcher would probably test the real empirical data for heteroscedastic error terms, and in the case that heteroscedasticity was determined, a generalised least squares estimator would be used, for instance in the form of weighted least squares[6)].

With the assumption of only <u>one plot per experiment</u>, the estimation was carried out using PROC REG in SAS 8.02. The estimation was carried out as weighted regression with the inverse of the value in (6) as the weights.

With <u>more than one plot per experiment</u> it is assumed that the parameters are identical across plots for each experiment, i.e. that the soil quality, the management, the technology, and the state of

[6]) It was therefore decided to use weighted least squares, using the inverse of the square-root of variance estimated by using (6) as the weight (Judge et al., 1982, p 414).

nature is the same on every plot during an experiment. The only thing that varies *between* plots *within* experiments is therefore the amount of the three inputs x_1, x_2, x_3, and the corresponding yield z.

This situation illustrates the extreme case of having available perfectly correlated panel data. In the normal empirical cases, observations (from different farms) will typically be disturbance related, exhibiting some correlation. In the case described here the correlation between the observations (different plots) is perfect in the sense that the error term is exactly the same when the amount of input is the same. This information is valuable, and should of course be used when performing the estimation. This is done by considering the regression equations for each plot as a disturbance-related set of regression equations, and including the restriction that the parameters are equal across plots (Judge et al. (1982), Cha. 11). Estimation is carried out by using the iteratively seemingly unrelated regression (ITSUR), which is available in PROS SYSLIN in SAS 8.02.

Using the ITSUR facility means, the option of weighted regression is not directly available. Therefore, to perform weighted regression in this case, all the exogenous variables in the estimation model in (4) (including the intercept (1)) were transformed by dividing them by the square root of the variance in (5).

The OLS and ITSUR columns of Table 3 show the results of the estimation of the parameters of the stochastic, state-contingent production functions in (4) for different numbers of experiments (observations), for one plot and two plots. (The results for three plots are not shown as they differ only marginally from the two-plot results).

The first three columns of Table 3 describe the 12 states defined by the 12 possible combinations of the two first state-variables A and a_1 shown In Table 1. The values of the parameters a_2 and a_3 in the 'TRUE'-column are the expected values according to the values and probabilities in Table 1.

The numbers in Table 3 are the results of 25 simulation runs. The numbers in the "NUMBER" column are the average number of observations in the "sample". The numbers in the EST columns are the average values of the parameter estimates. And the numbers in the ER columns are the estimation errors calculated as $(\sum_{i=1}^{25}|\gamma_i - \hat{\gamma}_i|)/25$, where γ_i is the true parameter value (in the third column) and $\hat{\gamma}_i$ is the parameter estimate. GME refers to Generalized Maximum Entropy estimation (see following Section).

4.4. GME-estimation.

As mentioned above, to estimate the complete set of state-contingent production functions one needs to estimate 12×4=48 parameters (12 states and a 3-input Cobb-Douglas production function). The use of standard econometric techniques, such as maximum likelihood or generalized methods of moments, for the estimation of a production function for each state requires that there are at least 4-plus observations registered for each state. This is not always available in real world situations. Instead, the researcher has to resort to "recognize" a smaller number of states, imposing severe restrictions on their model, or alternatively, losing significant amount of information. Such restrictions are not necessary with the generalized maximum entropy (GME) formalism. The coefficients of the production functions are recovered for each state of nature, as long as there is at least one observation for the state in question.

In order to recover the parameters for the 12 state-contingent production functions using the GME formalism, equations (4) is re-written in the following form:

$$\begin{bmatrix} z_{111} \\ z_{112} \\ \vdots \\ z_{11p} \\ \vdots \\ z_{1\tau p} \\ \vdots \\ z_{s\tau p} \end{bmatrix} = \begin{bmatrix} \mathbf{x}'_{111} & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{x}'_{112} \cdots & & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & 0 \cdots \mathbf{x}'_{11p} & \cdots & & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 \cdots\cdots \mathbf{x}'_{1\tau p} & & \cdots & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ 0 & 0 \cdots 0 \cdots 0 \cdots \mathbf{x}'_{s\tau p} & & & \end{bmatrix} \begin{bmatrix} v_{1k} p_{11k} \\ v_{1k} p_{12k} \\ \vdots \\ v_{1k} p_{1pk} \\ v_{2k} p_{21k} \\ v_{2k} p_{22k} \\ \vdots \\ v_{sk} p_{spk} \end{bmatrix} + \begin{bmatrix} \beta_{s\lambda} w_{111\lambda} \\ \beta_{s\lambda} w_{112\lambda} \\ \vdots \\ \beta_{s\lambda} w_{11p\lambda} \\ \vdots \\ \beta_{s\lambda} w_{1\tau p\lambda} \\ \vdots \\ \beta_{s\lambda} w_{s\tau p\lambda} \end{bmatrix} \quad (6)$$

$$([s \times \tau \times p] \times 1) \quad ([s \times \tau \times p] \times 4[s \times \tau \times p]) \quad (4[s \times \tau \times p] \times 1) \quad ([s \times \tau \times p] \times 1)$$

Where $z_{s\tau p}$ is the logarithm of the output of the p^{th} plot at the τ^{th} experiment at the s^{th} state[7]. The matrices $\mathbf{x}'$ are the corresponding inputs in logarithms. Each matrix contains a column of [1] for the intercept. We follow the standard GME parameterization process where each parameter α_{sp} is expressed as a vector product of a vector of probabilities $[p_{sp1}\ p_{sp2} \ldots p_{spk}]$ times a support vector $[v_{s1}\ v_{s2} \ldots v_{sk}]$, where k=5. Similarly, the error term is parameterized as a vector product of probabilities $[w_{s\tau p1}\ w_{s\tau p2} \ldots w_{s\tau pk}]$ times a support vector $[\beta_{s1}\ \beta_{s2} \ldots \beta_{sk}]$, where λ=5. The extreme points for both support vectors were set to three standard deviations on each side, according to the 3σ rule (Golan et al., 1996). Therefore, the support vectors for each state s, were set to: $v_s=\beta_s=[-3\sigma_s\ \ -1.5\sigma_s\ \ 0\ \ +1.5\sigma_s\ \ +3\sigma_s]$, where σ_s is the standard deviation of the output $\ln(z_s)$ for each state.

To correct for heteroscedasticity all data were divided by σ_ε from equation (5). The constraint that all parameters were equal across plots for each state was also imposed:

$$\sum_k v_{sk} p_{s\mu k} = \sum_k v_{sk} p_{s\nu k}\text{, for } \forall\ (\mu, \nu) \in \text{p},\ \mu \neq \nu \quad (7)$$

Following the GME formalism, the entropy metric is minimized:

$$\text{Min } H(\mathbf{p},\mathbf{w}) = \mathbf{p}' \ln(\mathbf{p}) + \mathbf{w}' \ln(\mathbf{w}) \quad (8)$$

subject to constraints (7), (8), the non-negativity constraints:

$$\mathbf{p} \geq \mathbf{0}\text{ , } \mathbf{w} \geq \mathbf{0} \quad (9)$$

and the additivity constraints:

$$\mathbf{p}'\mathbf{1}=1\ ;\ \mathbf{w}'\mathbf{1}=1 \quad (10)$$

[7] Here, for ease of notation we use a single index s for each state instead of the combination of ij used previously.

Table 3. Results

100 EXPERIMENTS

STATE		TRUE	NUMBER	ONE PLOT				NUMBER	TWO PLOTS			
				OLS		ME			ITSUR		ME	
				EST	ER	EST	ER		EST	ER	EST	ER
1	A	2.00	5	0.35	1.65	2.01	1.79	3	0.35	1.65	1.72	2.00
	a1	0.05		-0.63	0.68	0.15	0.14		-0.63	0.68	0.15	0.14
	a2	0.28		-0.21	0.49	0.30	0.27		-0.21	0.49	0.30	0.26
	a3	0.30		1.71	1.41	0.22	0.33		1.71	1.41	0.23	0.33
2	A	2.00	9	22.60	22.39	2.05	0.13	7	5.08	3.35	2.06	0.12
	a1	0.12		0.10	0.29	0.15	0.03		0.12	0.02	0.16	0.04
	a2	0.28		0.31	0.45	0.33	0.06		0.24	0.05	0.33	0.05
	a3	0.30		0.38	0.26	0.25	0.06		0.26	0.05	0.24	0.06
3	A	2.00	9	4.12	3.51	2.18	0.21	8	2.00	0.25	2.28	0.28
	a1	0.19		0.34	0.27	0.18	0.02		0.19	0.02	0.19	0.01
	a2	0.28		0.19	0.29	0.33	0.06		0.26	0.05	0.33	0.06
	a3	0.30		0.36	0.22	0.28	0.03		0.30	0.02	0.27	0.03
4	A	2.00	6	0.44	1.56	2.40	1.81	3	1.56	0.44	2.24	1.91
	a1	0.26		1.80	1.54	0.23	0.21		0.26	0.00	0.24	0.20
	a2	0.28		-1.77	2.05	0.35	0.26		0.32	0.04	0.36	0.27
	a3	0.30		1.40	1.10	0.29	0.22		0.32	0.02	0.30	0.23
5	A	4.00	7	46.70	46.50	2.35	1.65	5	5.05	2.84	2.31	1.69
	a1	0.05		0.28	0.51	0.19	0.14		0.04	0.08	0.21	0.16
	a2	0.28		0.49	0.45	0.34	0.07		0.33	0.11	0.35	0.08
	a3	0.30		0.37	0.26	0.27	0.03		0.33	0.11	0.28	0.03
6	A	4.00	18	6.93	5.59	2.69	1.31	19	3.90	0.41	2.58	1.42
	a1	0.12		0.09	0.10	0.19	0.07		0.12	0.01	0.21	0.09
	a2	0.28		0.38	0.19	0.34	0.07		0.27	0.02	0.35	0.08
	a3	0.30		0.30	0.16	0.30	0.02		0.30	0.01	0.29	0.01
7	A	4.00	24	13.81	10.95	2.99	1.01	23	3.97	0.24	2.89	1.11
	a1	0.19		0.20	0.10	0.23	0.04		0.19	0.01	0.24	0.05
	a2	0.28		0.20	0.16	0.34	0.07		0.28	0.02	0.36	0.09
	a3	0.30		0.27	0.10	0.31	0.02		0.30	0.01	0.31	0.01
8	A	4.00	6	19.38	17.86	3.23	1.12	5	10.98	9.50	3.09	1.27
	a1	0.26		0.88	0.92	0.26	0.04		0.49	0.24	0.28	0.05
	a2	0.28		-0.10	0.64	0.38	0.14		0.28	0.20	0.38	0.14
	a3	0.30		0.13	0.44	0.34	0.07		0.24	0.24	0.34	0.07
9	A	5.00	6	88.28	86.61	2.42	3.53	4	4.15	1.53	2.16	3.79
	a1	0.05		0.68	0.63	0.20	0.16		0.59	0.62	0.22	0.18
	a2	0.28		0.11	0.41	0.37	0.15		-0.27	0.67	0.39	0.18
	a3	0.30		0.35	0.52	0.30	0.09		0.59	0.29	0.28	0.09
10	A	5.00	10	52.05	50.88	2.80	2.20	9	4.85	0.62	2.86	2.14
	a1	0.12		0.16	0.25	0.23	0.11		0.12	0.01	0.23	0.11
	a2	0.28		0.33	0.35	0.36	0.09		0.29	0.04	0.36	0.09
	a3	0.30		0.33	0.23	0.30	0.02		0.31	0.03	0.30	0.01
11	A	5.00	14	41.35	38.64	3.33	1.67	13	5.95	1.33	3.17	1.83
	a1	0.19		0.22	0.22	0.24	0.05		0.18	0.02	0.25	0.06
	a2	0.28		0.22	0.23	0.35	0.08		0.25	0.05	0.37	0.10
	a3	0.30		0.26	0.17	0.32	0.02		0.29	0.02	0.32	0.02
12	A	5.00	6	43.02	44.09	3.71	2.24	3	2.72	2.28	3.87	2.17
	a1	0.26		0.23	0.24	0.29	0.08		0.24	0.07	0.30	0.09
	a2	0.28		0.76	0.69	0.40	0.18		0.43	0.16	0.41	0.19
	a3	0.30		0.37	0.38	0.35	0.11		0.34	0.04	0.36	0.11

200 EXPERIMENTS

STATE		TRUE	NUMBER	ONE PLOT				NUMBER	TWO PLOTS			
				OLS		ME			ITSUR		ME	
				EST	ER	EST	ER		EST	ER	EST	ER
1	A	2.00	6	5.33	5.56	1.78	0.28	4	1.46	0.83	1.66	0.34
	a1	0.05		-0.02	0.43	0.14	0.09		0.22	0.46	0.14	0.09
	a2	0.28		0.55	0.57	0.31	0.05		0.56	0.30	0.30	0.04
	a3	0.30		0.23	0.76	0.23	0.07		0.09	0.43	0.24	0.07
2	A	2.00	14	6.96	6.37	1.99	0.18	14	1.93	0.25	2.04	0.07
	a1	0.12		0.09	0.16	0.15	0.03		0.12	0.01	0.15	0.04
	a2	0.28		0.30	0.19	0.30	0.04		0.28	0.03	0.30	0.03
	a3	0.30		0.33	0.14	0.26	0.05		0.30	0.01	0.25	0.05
3	A	2.00	19	4.43	3.45	2.06	0.19	18	1.97	0.14	2.16	0.24
	a1	0.19		0.26	0.14	0.18	0.02		0.19	0.01	0.18	0.02
	a2	0.28		0.19	0.17	0.33	0.06		0.28	0.02	0.33	0.06
	a3	0.30		0.29	0.15	0.27	0.03		0.30	0.01	0.27	0.03
4	A	2.00	6	169.14	167.76	2.27	0.76	5	42.27	40.68	2.23	0.75
	a1	0.26		0.10	0.30	0.20	0.09		0.19	0.08	0.22	0.07
	a2	0.28		0.30	0.36	0.37	0.14		0.26	0.13	0.36	0.12
	a3	0.30		-0.11	0.46	0.30	0.09		0.26	0.05	0.09	0.07
5	A	4.00	9	16.16	15.14	2.27	1.73	10	3.65	0.70	2.25	1.75
	a1	0.05		0.63	0.66	0.17	0.12		0.17	0.14	0.19	0.14
	a2	0.28		0.16	0.68	0.35	0.08		0.24	0.11	0.35	0.08
	a3	0.30		0.45	0.41	0.28	0.04		0.29	0.04	0.27	0.03
6	A	4.00	35	9.69	7.30	2.69	1.31	37	4.01	0.22	2.51	1.49
	a1	0.12		0.13	0.07	0.19	0.07		0.12	0.01	0.22	0.10
	a2	0.28		0.27	0.12	0.32	0.05		0.28	0.02	0.33	0.06
	a3	0.30		0.25	0.12	0.31	0.03		0.30	0.01	0.29	0.02
7	A	4.00	47	5.44	2.37	3.00	1.00	44	4.05	0.17	2.82	1.18
	a1	0.19		0.21	0.05	0.22	0.04		0.19	0.01	0.23	0.04
	a2	0.28		0.23	0.08	0.34	0.07		0.27	0.01	0.35	0.08
	a3	0.30		0.29	0.08	0.31	0.02		0.30	0.01	0.31	0.01
8	A	4.00	11	29.13	28.11	3.24	0.76	10	3.86	0.82	3.32	0.68
	a1	0.26		0.31	0.24	0.27	0.02		0.30	0.05	0.28	0.02
	a2	0.28		0.35	0.25	0.38	0.11		0.29	0.06	0.38	0.11
	a3	0.30		0.28	0.21	0.34	0.04		0.32	0.04	0.33	0.03
9	A	5.00	6	111.45	111.26	2.42	2.58	6	15.58	12.80	2.31	2.69
	a1	0.05		0.17	0.43	0.19	0.14		0.28	0.27	0.20	0.02
	a2	0.28		0.45	0.54	0.38	0.11		0.35	0.24	0.41	0.13
	a3	0.30		0.24	0.42	0.29	0.03		0.13	0.23	0.28	0.03
10	A	5.00	23	10.49	7.88	2.91	2.13	21	4.87	0.36	2.76	2.24
	a1	0.12		0.11	0.14	0.22	0.10		0.12	0.01	0.24	0.12
	a2	0.28		0.22	0.12	0.34	0.07		0.28	0.02	0.36	0.10
	a3	0.30		0.32	0.10	0.30	0.02		0.30	0.01	0.30	0.01
11	A	5.00	28	7.95	5.53	3.52	1.68	26	5.01	0.35	3.21	1.79
	a1	0.19		0.18	0.12	0.24	0.05		0.19	0.01	0.25	0.06
	a2	0.28		0.28	0.17	0.34	0.07		0.28	0.02	0.35	0.10
	a3	0.30		0.31	0.07	0.31	0.02		0.30	0.01	0.32	0.02
12	A	5.00	8	9.21	9.83	3.53	1.47	6	5.11	1.80	3.52	1.48
	a1	0.26		0.32	0.35	0.29	0.04		0.15	0.16	0.29	0.33
	a2	0.28		0.53	0.52	0.39	0.12		0.57	0.37	0.39	0.12
	a3	0.30		0.34	0.31	0.34	0.04		0.39	0.13	0.34	0.04

400 EXPERIMENTS

STATE		TRUE	NUMBER	ONE PLOT				NUMBER	TWO PLOTS			
				OLS		ME			ITSUR		ME	
				EST	ER	EST	ER		EST	ER	EST	ER
1	A	2.00	9	22.72	21.80	1.80	0.21	8	2.15	0.28	1.79	0.21
	a1	0.05		-0.02	0.19	0.11	0.07		0.03	0.03	0.12	0.07
	a2	0.28		0.33	0.38	0.30	0.05		0.31	0.07	0.32	0.05
	a3	0.30		0.22	0.31	0.25	0.06		0.27	0.04	0.23	0.07
2	A	2.00	28	3.50	2.15	1.96	0.16	27	2.03	0.11	1.99	0.09
	a1	0.12		0.15	0.09	0.16	0.04		0.12	0.01	0.16	0.04
	a2	0.28		0.23	0.09	0.29	0.03		0.28	0.02	0.30	0.03
	a3	0.30		0.27	0.10	0.25	0.06		0.30	0.01	0.24	0.06
3	A	2.00	39	3.13	1.90	2.03	0.23	35	1.95	0.12	2.14	0.25
	a1	0.19		0.17	0.07	0.19	0.02		0.19	0.01	0.18	0.02
	a2	0.28		0.25	0.10	0.32	0.05		0.28	0.02	0.31	0.04
	a3	0.30		0.32	0.10	0.28	0.03		0.30	0.01	0.27	0.03
4	A	2.00	8	10.62	10.08	2.30	0.49	8	2.23	0.63	2.41	0.61
	a1	0.26		0.16	0.32	0.22	0.07		0.39	0.17	0.22	0.06
	a2	0.28		0.71	0.59	0.35	0.10		0.33	0.11	0.36	0.10
	a3	0.30		0.17	0.33	0.29	0.05		0.27	0.07	0.27	0.05
5	A	4.00	20	7.08	5.34	2.30	1.70	19	3.98	0.28	2.24	1.76
	a1	0.05		0.09	0.14	0.22	0.12		0.05	0.01	0.19	0.14
	a2	0.28		0.27	0.17	0.35	0.05		0.28	0.03	0.32	0.05
	a3	0.30		0.31	0.11	0.29	0.04		0.30	0.02	0.27	0.03
6	A	4.00	71	5.18	2.08	3.04	1.16	72	4.01	0.14	2.88	1.45
	a1	0.12		0.13	0.06	0.17	0.06		0.12	0.00	0.19	0.07
	a2	0.28		0.26	0.07	0.31	0.04		0.28	0.01	0.32	0.07
	a3	0.30		0.27	0.06	0.30	0.02		0.30	0.01	0.29	0.03
7	A	4.00	92	4.89	2.08	3.10	91.00	90	4.02	0.14	2.98	1.07
	a1	0.19		0.20	0.05	0.22	0.03		0.19	0.00	22.67	0.04
	a2	0.28		0.26	0.08	0.31	0.04		0.27	0.01	0.32	0.05
	a3	0.30		0.29	0.05	0.32	0.03		0.30	0.01	0.32	0.02
8	A	4.00	19	5.89	5.00	3.76	1.27	20	4.11	0.26	3.47	0.71
	a1	0.26		0.29	0.12	0.27	0.02		0.26	0.01	0.26	0.01
	a2	0.28		0.29	0.17	0.35	0.10		0.27	0.03	0.36	0.11
	a3	0.30		0.39	0.16	0.33	0.03		0.30	0.01	0.32	0.03
9	A	5.00	12	6.99	6.36	2.53	2.47	13	4.95	0.68	2.44	2.56
	a1	0.05		0.09	0.19	0.18	0.13		0.06	0.02	0.19	0.14
	a2	0.28		0.53	0.43	0.25	0.08		0.28	0.03	0.36	0.09
	a3	0.30		0.38	0.20	0.30	0.03		0.30	0.02	0.30	0.02
10	A	5.00	41	6.28	3.09	2.93	2.07	43	4.90	0.23	2.81	2.19
	a1	0.12		0.11	0.07	0.21	0.09		0.12	0.01	0.23	0.11
	a2	0.28		0.27	0.09	0.32	0.06		0.27	0.01	0.32	0.07
	a3	0.30		0.30	0.07	0.31	0.03		0.30	0.01	0.31	0.01
11	A	5.00	51	5.35	2.78	3.35	1.65	52	4.95	0.18	3.21	1.79
	a1	0.19		0.20	0.06	0.23	0.04		0.19	0.00	0.24	0.04
	a2	0.28		0.28	0.08	0.34	0.07		0.27	0.01	0.34	0.08
	a3	0.30		0.32	0.07	0.33	0.03		0.30	0.01	0.34	0.04
12	A	5.00	12	8.53	6.62	3.46	1.54	12	5.01	0.55	3.52	1.48
	a1	0.26		0.28	0.13	0.29	0.03		0.27	0.01	0.28	0.03
	a2	0.28		0.33	0.24	0.37	0.10		0.27	0.03	0.39	0.12
	a3	0.30		0.33	0.20	0.35	0.05		0.31	0.01	0.35	0.05

5. Discussion

At this stage it has only been possible to make 25 simulation runs. Thus for each number of experiments, each number of plots, and each estimation method, 25 estimation of the four parameters for each state has been made.

When comparing the results of one and two plots estimation, it is interesting to notice the extreme improvement in estimation efficiency that is achieved when going from one to two plots per experiment. In many cases the error of the estimated parameters is reduced by a factor 10 or more. An example of the increased efficiency is shown by comparing the standard deviation of the OLS and the ITSUR parameter estimates for state 6 and 200 experiments in the following Table 4.

Table 4. Estimation Error (ER) for one and two plots in state 6

	35 experiments	
	One plot	Two plots
A	7.30	0.22
a_1	0.07	0.01
a_2	0.12	0.02
a_3	0.12	0.01

This result is not in itself surprising. But it underlines the fact that when estimating state-contingent production functions, it is of much higher importance to have a number of observations that are known to be from the same *real* state of nature (i.e. observations *within experiments*), instead of having observations from the same *registered states* (i.e. observations *over experiments*).

An example of the result of increasing the number of experiments (years), keeping the number of plots pr. experiment (year) unchanged, is shown in Table 5. As one can see, there is hardly any improvement in efficiency by increasing the number of experiments from 18 to 35 (years) with 2 plots.

Table 5. Estimation Error (ER) for 18 and 36 experiments in state 6 (2 plots)

	Two plots	
	18 expe-riments.	35 expe-riments.
A	0.41	0.22
a_1	0.01	0.01
a_2	0.02	0.02
a_3	0.01	0.01

When analysing the results in Table 3 there are three other significant observations:

First, even with a very high number of observations (experiments), some of the states only occur in the sample with very few observations. Thus, even with 400 experiments, four of the twelve registered states only have an average of 12 or less observations. These rare states are at the same time often the states with the most drastic consequences regarding production conditions, and therefore the states for which it may be important to have good information.

With few observations (100 experiments or less – which in the classical approach is not even few!), the number of observations per state is often so small that OLS/ITSUR cannot be used. And even if it can, the estimation error is very large. In this case GME may be applied (GME yield estimates even with just one observation).

However, when it is possible to use both approaches (GME and OLS/ITSUR), which of the two approaches are then the best – i.e. the most efficient? With very few observations GME is probably the most efficient.

The results in Table 3 show that with just *one plot* GME has a higher precision than OLS in almost all cases (i.e. independently of the number of observations). And especially with very few observations, GME is much more efficient than OLS.

With *two plots* the picture changes. Even with a relatively low number of observations, ITSUR seems to perform better than GME. The break even point is around 7-8 observations, which is a very low number of observations on which to base estimation of 4 parameters.

Further analyses based on more simulation runs are needed to answer this question of the relative efficiency of GME versus OLS/ITSUR. Golan *et al*., (1996) have shown, that as the number of observations increases, GME and maximum likelihood estimates converge asymptotically.

6. Conclusions

The major problem of empirical application of the state-contingent approach to production analysis is that the number of observations for each state is low – if available at all. At the same time one typically faces the problem that only few of the possible states of nature are in fact registered. Therefore, even the state contingent production functions that one may estimate are stochastic functions.

We have designed a computer simulated Monte Carlo experiment where several levels of three inputs were applied on a number of different plots under various states of nature. Output was "produced" with state-contingent production functions. Only some of the states of natures were "observable". The objective then was to recover the parameters of the production functions for each one of the observable states.

The Monte Carlo simulation experiment has shown that even with relatively few registered states and a long time series of data, the number of observations soon becomes critically low to use the traditional estimation methods. It is shown that using Generalised Maximum Entropy it is possible to estimate the parameters of state-contingent production functions. However, to compare the efficiency of alternative methods of estimation, further analysis is needed.

The main contribution of this paper is that it has shown that the state-contingent approach to decision making, is not only a theoretical model but it is an empirical possibility. We have laid out the methodology and have estimated state-contingent production functions. Certainly, more work is required. In a normative context, the ultimate test of the appropriateness of each method is the efficiency in relation to decision making.

REFERENCE LIST

Chambers,R.G. and J.Quiggin: "The state-contingent properties of stochastic production functions." *American Journal of Agricultural Economics* 84(May 2002a): 513-26.

Chambers,R.G. and J.Quiggin: "Optimal producer behavior in the presence of area-yield crop insurance." *American Journal of Agricultural Economics* 84(May 2002b): 320-34.

Chambers,R.G. and J.Quiggin: *Uncertainty, Production, Choice, and Agency.* Cambridge: Cambridge University Press, 2000.

Dillon,J.L. and J.R.Anderson *The Analysis of Response in Crop and Livestock Production,* Third ed. Oxford: Pergamon Press, 1990.

GAMS Development Corporation *GAMS. The Solver Manuals.* Washington DC: 2003.

Golan, A., G. Judge; and D. Miller. "Maximum Entropy Econometrics", Chichester: Wiley, 1996

Hardaker,J.B., R.B.M.Huirne, and J.R.Anderson *Coping with Risk in Agriculture,* 1st ed. Wallingford, UK: CAB International, 1997.

Oude Lansink, A. "Generalised Maximum Entropy Estimation and Heterogeneous Technologies" European Review of Agricultural Economics 26 (1999): 101-115

Rasmussen,S. "Criteria for optimal production under uncertainty. The state-contingent approach." *Australian Journal of Agricultural and Resource Economics* 47(December 2003): 447-76.

Robison,L.J. and P.J.Barry *The Competitive Firm's Response to Risk.* New York: Macmillan Publishing Company, 1987.