AAEA Long Beach Meetings
Principal Paper
Please do not cite.

Production Efficiency In The von Liebig Model

Garth Holloway[a] and Quirino Paris[b]

Abstract

The von Liebig hypothesis of non-substitution between inputs and yield plateau has received renewed attention in the literature. The von Liebig model was introduced around 1840 but was downplayed in many circles because it was an idea too sophisticated for the analytical knowledge of that time. The von Liebig technology, revisited in the 1970's, has refocused attention on constraining factors that limit nutrient response as the defining feature of crop growth. But von Liebig estimation has brought with it a host of new challenges for empirical production investigations. This paper compares two new approaches to von Liebig estimation. The first approach, based on a finite-mixture-of-normals formulation, has not been applied before to von Liebig estimation. We show how routine application of Gibbs sampling and data augmentation provide robust estimates of the production surface given typical input and output data. The second approach, based on a hierarchical, normal-linear model formulation, provides robust estimates of the production frontier. The techniques are illustrated using experimental and real and demonstrate routine application of Markov Chain Monte Carlo (MCMC) methods. Extensions are discussed.

Keywords: von Liebig model, Bayes estimation, Markov Chain Monte Carlo methods. (10 words)

Journal of Economic Literature Classifications: Q12, D24, C11.

[a]Department of Agricultural and Food Economics, School of Agriculture, Policy and Development, University of Reading, Earley Gate, Whiteknights House, Reading RG6 6AR, United Kingdom; phone: (+44)-(0)-118-931-6775; fax: (+44)-(0)-118-975-6467; email: garth.holloway@reading.ac.uk. [b]Department of Agricultural and Resource Economics, University of California-Davis, phone: (530)-752-1528; fax: (530)-752-5614; email: paris@primal.ucdavis.edu. We are extremely grateful to Richard Tiffin for several beneficial suggestions.

Production Efficiency in the von Liebig Model

Introduction

The von Liebig hypothesis, formulated circa 150 years ago, continues to intrigue researchers in the fields of production economics and econometrics due to its sophisticated simplicity that requires sophisticated tools for its analysis. The conjecture deals with the way crop yields---according to von Liebig---respond to differential fertilizer applications. In von Liebig's time, lime and manure were the sole or prelevant soil additives. The use of these natural and complex fertilizer compounds did not prevent von Liebig's intuition to formulate the conjecture that chemical elements combine in fixed proportions not only in an alembic but also in the soil. Toward the end of the 19th century and on the strength of von Liebig contributions, elemental nitrogen, phosphorus and potassium began to be recognized as profitable additives for achieving higher levels of productivity. In 1906, a poster was circulating among German farmers propounding the benefits of potassium (kalium) as a crop enhancer and illustrating the working of von Liebig's "law of the minimum" by means of a barrel with staves of different height and where the shortest stave corresponds to the actual level of production. This poster is the best available illustration of the non-substitution hypothesis that characterizes von Liebig's theory of crop response in an essential way.

From the beginning of its enunciation, von Liebig's law of the minimum received enormous attention, splitting the camp of agronomists and soil scientists. When agricultural economists came onto the scene, toward the first quarter of the past century, they too took side in the debate with J. D. Black (p. 275)—the distinguished economist from Harvard University—delivering one of the most scathing criticism of the law of the minimum. Black was grossly mistaken in its criticism of von Liebig, but this fact did not prevent the aura of his professional authority to cast a serious doubt among the fledging agricultural economics profession on the validity of von Liebig's hypothesis.

Von Liebig's law of the minimum can be expressed by the following analytical relation:

$$y = \min\{f_N(N,u_N),\, f_P(P,u_P),\, f_K(K,u_K),\ldots,\, f_C(C,u_C)\} \tag{1}$$

where y is the realized level of crop production, N,P,K and C are the levels of fertilizer nutrients (say, nitrogen, phosphorus, potassium and calcium) and u_N,u_P,u_K,u_C are random disturbances associated with the corresponding potential yield functions $f_N(N,u_N),\, f_P(P,u_P),\, f_K(K,u_K)$ and $f_C(C,u_C)$ representing the individual response to the various fertilizer nutrients.

Until 1977, the von Liebig hypothesis did not received any serious analytical scrutiny and empirical testing. The most likely reason is that equation (1) poses non-trivial mathematical and statistical issues for its estimation. Furthermore, the non-nested hypothesis testing procedure

necessary for discriminating among different models was not available to empirical researchers until that time.

The objectives of this paper is to revisit the von Liebig hypothesis by re-examining five sample of experimental data and by applying to it recent advances in Bayesian techniques. The samples were published by Hexem and Heady as described in a further section. Prior to outlining the estimation strategy, we discuss the intuition underlying our approach and, briefly, the literature on which it is based. We present an algorithm for the basic von Liebig formulation and demonstrate its application using simulated data (table 1). We then discuss modifications needed to the basic model that facilitate estimation of a von Liebig frontier and we demonstrate the extended algorithm using simulated data (table 2). We then explore, empirically, relationships between limiting water and nitrogen in the Hexem and Heady corn samples and compare the results between the two formulations (table 3). Finally, some conclusions and suggestions for further research are offered.

Previous Work

The recent revival of interest in the von Liebig's hypothesis is due to Lanzer who, in his PhD dissertation of 1977, laid out the algorithm for estimating a linear version of equation (1) using a mathematical programming approach. The crucial idea was to convert the "min" operator in equation (1) into a set of analytical constraints suitable for solution with available nonlinear programming packages. In the seventies, nonlinear programming packages were in their beta-testing phase, and the estimation of a von Liebig model as specified by Lanzer was a significant achievement. The maximum likelihood estimates so derived were presented by Lanzer and Paris.

The first comparison between a von Liebig model with linear regimes and more traditional production function specifications such as polynomial models was conducted by Ackello, Paris and Williams. By using a non-nested hypothesis testing procedure, their work showed that the von Liebig model interpreted the sample data better than several other polynomial specifications. In all these studies, the potential yield functions in equation (1) have a common additive error term which can, therefore, be taken outside the "min" operator. Grimm, Paris and Williams continued in the same line of research and applied the same procedure to different samples of experimental data. Paris and Knapp extended the specification of the von Liebig model to include the random disturbances under the "min" operator, as in equation (1). The von Liebig model thus assumed the structure of a switching regression model with endogenous sample separation.

Frank, Beattie and Embleton used the famous sample of experimental data collected by Heady and Pesek to make a non-nested hypothesis pairwise comparison among various models including a

quadratic polynomial, a linear response and plateau (LRP), and a Mitscherlich-Baule model. They concluded that the Mitscherlich-Baule formulation rejects all the other alternatives, including the LRP specification of the von Liebig model. Paris challenged that conclusion by showing that, when the von Liebig hypothesis is interpreted in its more general framework using nonlinear regimes in place of the linear responses, the von Liebig model rejects all the other alternatives, including the Mitscherlich-Baule specification.

More recently, Chambers and Lichtenberg, Berck, Geoghegan and Stohs, applied non-parametric methods to test the von Liebig hypothesis using the Heady-Pesek sample of data. While Chambers and Lichtenberg concluded in favor of the hypothesis, Berck, Geoghegan and Stohs rejected it. It is difficult to compare the results obtained with parametric and non-parametric methods. We leave this endeavour for future research.

Our Approach

The starting point for our approach is to combine various concepts (potential estimation problems) from previous work. First, there is the problem of correctly accounting for the statistical implications of the information conveyed by the "min" operator in equation (1). Second, because any one of several possible regimes could be limiting at one time, we need account for the possibility of "swiching" between regimes at any point in the data. Third—and noteworthy because it appears neither to have been recognized nor fully exploited in previous work and, crucially, because it supplies a link to the Markov Chain Monte Carlo (MCMC) literature that solves estimation problems for us—we note that the two preceding estimation difficulties lend the estimation work toward one of correctly classifying and discriminating within the entire sample. To be precise, we view von Liebig estimation as a straight-forward problem of classification within the sample and, hence, the problem of discriminating between mixtures in a finite set of normal-mixture regimes. The classification problem arises because—by virtue of the "min" operator, and ignoring "ties,"—only one of the several candidate regimes can dominate the alternatives at any particular sample point. Thus the problem of regime estimation, or discriminating within the sample, is a problem of sample sub-classification, and our estimation technique evolves with this simple fact in mind. But this problem is the only impeding practical problem in our way. To see why, consider the following logic. Suppose we know the correct classification of the data within the sample. That is, suppose we know with probability equal to one that a particular sample observation, say y_t, belongs to a particular category or mixture component in the sample, where each component of the mixture is specified in probability terms in relation to one von Liebig regime (say, water limiting, or nitrogen limiting). Once the

classification is known, and an error specification is selected, then estimation is standard within each regime. In other words, estimation is routine.

The significant problem arises due to the fact that the specification of each regime cannot be made with certainty, but must be probabilistically ascertained. When this is the case, we need to augment the observed data with a latent, classification variable and make estimates based on a particular classification of the observed data. In addition, we need to specify a way of relating the observed and augmented data, the production-surface parameters in each of the limiting regimes, and a relationship between them and the error structure proposed for the model. When such a relationship exists and leads to fully conditional distributions of well-known forms, a procedure known as Gibbs sampling can be used, alternating between the latent data and the estimated values of the model parameters, to simulate a draw from the joint posterior distribution corresponding to the von Liebig model. Fortunately, such a relationship does exist and we exploit it in each of the algorithms that we propose in this paper. This relationship between these unknown quantities is the modus operandi for our estimation procedure and has been formalized neatly in an earlier contribution that studies mixtures of simple means (Lavine and West). Except for the focus on frontier production, our extensions of this paper are relatively modest because they involve only the inclusion of a set of appropriate covariates into the mean-mixtures formulation. Subsequently, we provide a less trivial extension to the basic algorithm in order that a frontier can be estimated. But here, too, some earlier work (Koop, Osiewalski and Steel) provides crucial input, and much of the algebra that remains is a fairly simple extension of details contained in the appendix to that work (see appendix, pp. 100-103). In short, by combining two seminal pieces on mixture modelling (Lavine and West) and Bayesian estimation of production frontiers (Koop, Osiewalski and Steel), we are able to produce a robust algorithm for estimating von Liebig response regimes.

We now turn to a more detailed discussion of the algorithm. Readers interested in additional details concerning the Gibbs sampling procedures and hierarchical models used in estimation should read papers by Gelfand and Smith, Casella and George and Chib and Greenberg and a book by Gelman et al.

The Basic Algorithm

Because it provides the basis for all subsequent work, consider the normal-linear regression model

$$\mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \mathbf{u}, \tag{2}$$

where $\mathbf{y} \equiv (y_1, y_2, .., y_T)'$ denotes an n-vector of observed; $\mathbf{x} \equiv (\mathbf{x}_1, \mathbf{x}_2, .., \mathbf{x}_T)'$, $\mathbf{x}_1 \equiv (x_{11}, x_{12}, .., x_{1K})$, $\mathbf{x}_2 \equiv (x_{21}, x_{22}, .., x_{2k})$, .., $\mathbf{x}_T \equiv (x_{T1}, x_{T2}, .., x_{Tk})$ are observations on the covariates; and $\mathbf{u} \sim N(\mathbf{0}_T, \sigma^2\mathbf{I}_T)$ denotes

random error. We observe $\mathbf{x}$ and $\mathbf{y}$ but neither σ nor $\boldsymbol{\beta}$. Inferences about the locations and scales of the elements of σ and $\boldsymbol{\beta}$ are, however, easily obtained be deriving their marginal posterior distributions. Following standard work (Zellner), these distributions are available in closed form, respectively as inverse-gamma and (multivariate) normal distributions. In view of this fact, there is no need to make any approximation to the full posterior, which we denote in a standard notation $\pi(\boldsymbol{\Theta}|\mathbf{y})$, $\boldsymbol{\Theta} \equiv (\sigma,\boldsymbol{\beta})'$. But suppose that these marginal distributions are not available. Where Gibbs sampling becomes important is the case where the fully conditional posterior distributions $\pi(\boldsymbol{\sigma}|\mathbf{y})$ and $\pi(\boldsymbol{\beta}|\mathbf{y})$ have well-known forms. In this particular case, these forms are the inverse-gamma and normal distributions, which we denote, again, in a generic notation, $f^{ig}(\sigma|v, (\mathbf{y}\text{-}\mathbf{x}\boldsymbol{\beta})'(\mathbf{y}\text{-}\mathbf{x}\boldsymbol{\beta}))$ and $f^{N}(\boldsymbol{\beta}|(\mathbf{x}'\mathbf{x})^{-1}, \sigma^2(\mathbf{x}'\mathbf{x})^{-1})$, where the notation $f^{a}(b|c, d, .., e)$ denotes a probability density function of type 'a' for random variable 'b' that is dependent (conditional) on parameters 'c', 'd', .., 'e.' Given this dependence, we can select an arbitrary starting value, say $\sigma^{(0)}$, simulate draws for σ and β, alternating between the inverse-gamma and normal distributions, namely $\beta^{(1)}|\sigma^{(0)}$, $\sigma^{(1)}|\beta^{(1)}$, $\beta^{(2)}|\sigma^{(1)}$, .., $\beta^{(S)}|\sigma^{(S-1)}$, $\sigma^{(S)}|\beta^{(S)}$. The resulting samples $\{\beta^{(s)}, s = 1, 2, .., S\}$ and $\{\sigma^{(s)}, s = 1, 2, .., S\}$ can then be used to compute means and variances, or any posterior moment of interest, or can be used to plot densities in order to characterize locations and scales of any σ and $\boldsymbol{\beta}$, or, importantly, <u>any function thereof</u>. The Gibbs sample, therefore, provides a powerful tool for posterior explorations of data.

Now, in terms of the extended problem of classification, suppose that it is known with some probability that a particular observation, say y_t, comes from one of several components in a mixture of these normal, linear models. In other words, suppose that data, t = 1, 2, ..T is distributed $y_t \sim f^{N}(y_t|\mathbf{x}_t^{(j)}\boldsymbol{\beta}^{(j)}, \sigma^{2(j)})$ with probability $\theta^{(j)}$, where j = 1, 2, .., M denote the M components of the mixture. Unfortunately, the complete posterior for this model $\pi(\boldsymbol{\Theta}|\mathbf{y})$, $\boldsymbol{\Theta} \equiv (\sigma^{(j)}, \beta^{(j)}, \theta^{(j)}, j = 1, 2, .., M)'$ has an intractable form and so some form of approximation is necessary in order to characterize locations and scales of its parameters. However, by augmenting the known data $\mathbf{y} \equiv (y_1, y_2, .., y_T)'$ with latent classification data $\mathbf{z} \equiv (\mathbf{z}_1, \mathbf{z}_2, .., \mathbf{z}_T)'$, $\mathbf{z}_1 \equiv (z_{11}, z_{12}, .., z_{1M})$, $\mathbf{z}_2 \equiv (z_{21}, z_{22}, .., z_{2M})$, .., $\mathbf{z}_T \equiv (z_{T1}, z_{T2}, .., z_{TM})$, where $z_{ij} = 1$ if observation i comes from component j of the mixture and $z_{ij} = 0$, otherwise, the data-augmented posterior distribution $\pi(\boldsymbol{\Theta}, \mathbf{z}|\mathbf{y})$, has a tractable form in the sense that each of the component conditional distributions are easy to sample from (Lavine and West, p. 452). In particular, each of the M-vectors in $\mathbf{z} \equiv (\mathbf{z}_1, \mathbf{z}_2, .., \mathbf{z}_T)'$ has a multinomial distribution; and conditional on this classification implied by the multinomial distribution, the equation parameters $\sigma^{(j)}$ and $\beta^{(j)}$, j = 1, 2, .., M, have their (now familiar) inverse-gamma and normal forms. Further, conditional on classification, the "mixing weights," $\theta^{(j)}$, j = 1, 2, .., M have Dirichlet forms. Because it is easy to sample from each of these distributions, a Gibbs-sampling algorithm can be implemented that samples for a "burn-in"

phases and then collects samples $\{\mathbf{z}^{(s)}, s = 1, 2, .., S\}$, $\{\beta^{(j)(s)}, j = 1, 2, .., M; s = 1, 2, .., S\}$, $\{\sigma^{(j)(s)}, j = 1, 2, .., M; s = 1, 2, .., S\}$ and $\{\theta^{(j)(s)}, j = 1, 2, .., M; s = 1, 2, .., S\}$, from which posterior conclusions can be drawn. It is essentially this algorithm that we implement as our basic von Liebig estimation procedure.

Basic von Liebig Algorithm

Because, in the Hexem and Heady data to follow, we consider two inputs, water and nitrogen and, further, because in this initial foray we restrict attention to linear specifications, consider a special case of equation (1), namely

(3) $\quad y_t = \min\{ \alpha_o + \alpha_1 w_t, \beta_o + \beta_1 n_t, \pi \} + u_t, \qquad t = 1, 2, ..., T;$

where y_t is crop yield; w_t and n_t are, respectively, water and nitrogen application; α_o and α_1 are parameters in the water-limiting regime; β_o and β_1 are parameters in the nitrogen-limiting regime; and π denotes a yield plateau. In terms of the notation just developed, there are j = 1, 2, 3 regimes, with $\alpha_o + \alpha_1 w_t \equiv \mathbf{x}_t^{(1)}\boldsymbol{\beta}^{(1)}$; $\beta_o + \beta_1 n_t \equiv \mathbf{x}_t^{(2)}\boldsymbol{\beta}^{(2)}$; and $\pi \equiv \mathbf{x}_t^{(3)}\boldsymbol{\beta}^{(3)}$. We consider that each data point y_t, t = 1, 2, .., T, has probability $\theta^{(1)}$ that water is limiting, probability $\theta^{(2)}$ that nitrogen is limiting, and probability $\theta^{(3)}$ that the plateau is limiting and, hence, that one of the three regimes governs data generation at that point. Consequently, a five-step algorithm for implementing the basic von Liebig model follows from sampling in turn from multinomial, inverse-gamma, normal and Dirichlet distributions and we now consider its accuracy.

We illustrate the procedure with an informative prior on the mixing weights but a flat prior over the regression coefficients and the other parameters in the model. There are two reasons for this. First, experience with the algorithm suggests that non-informative analysis sometimes works "too well." By this we mean that some of the cell probabilities comprising the multinomial distribution are assigned zero mass at successive iterations. This leads to divisions by zero. The second reason we prefer to implement an informative prior is to illustrate to those readers troubled by its use that the algorithm is robust to prior parameter choice—always a weakly informative prior can be selected so that the data dominate the outcome. Hence, in our two-regime problem we select the prior mixing weights in favor of each regime to be exactly one half, with numerical values selected so that the prior and data information are combined in proportions 10:100. (This prior corresponds to fixing $\mathbf{a}_o$ = (10, 10, 10)) in Lavine and West, section 2.) This choice enables quick convergence and accurate estimation of parameters at relatively little cost in terms of real execution time.

In the reports of the results that follow, we choose highly conservative 'burn-in' and sample sizes of t = 10,000 and s = 10,000 observations, respectively. Experiments with alternative starting values

suggest that the Gibbs sequence converged after about 50 iterations. However, even with this highly conservative burn-in phase the entire estimation procedure took only about 14 minutes of real time on a DELL™ Precision 340 machine running a Pentium™ IV processor at 2.0 gigahertz with commands executed in MATLAB™ version 5.1.0.421. All computer code is available upon request.

We generate data from the simple von Liebig specification in equation (3), with parameters set at $(\alpha_o, \alpha_1, \beta_o, \beta_1, \pi) = (0, 1, 0, 1, 100)$ and data w_t and n_t generated from uniform distributions in the intervals [0, 200]. Because the plateau and nutrient-response regimes intersect at the covariate value 100, the covariate values depict "ideal" circumstances for the experiments. Across the experiments we hold constant the number of observations at T = 100 but vary the error variance parameter to be $\sigma = 1$, $\sigma = 10$ and $\sigma = 100$, respectively. The first case reflects an error variance one one-hundreth the size of the plateau yield; the second case reflects an error variance equal to the plateau yield; and the third situation reflects an error variance one-hundred times the plateau yield. These choices are sufficient to reflect a wide array of sampling conditions from which to evaluate the performance of the Gibbs sampler.

The row entries in table 1 are estimates, respectively, of the constant term in the water-regime, the response term in the water regime, the constant term in the nitrogen regime, the response rate in the nitrogen regime, the plateau yield, the error variance parameter, the mean number of misclassified observations in the sample and the execution time in seconds. Not surprisingly, at $\sigma = 1$, the fit of the model is near perfect. Both the nutrient-response and the plateau-yield regimes are predicted to a high degree of accuracy. However, such conditions are unlikely to prevail in field situations or in experimental data. Turning to the second experiment, estimation performance does not appear to be markedly reduced when the error variance is increased to a more realistic level. And in the third experiment, although the actual precision of estimation drops quite dramatically, the pattern of significance among the estimated coefficients remains robust. Hence—at least in terms of the prototype von Liebig model—the Gibbs sampler performs well under varied sampling circumstances and relatively limited information (N = 100) and appears to be relatively inexpensive in terms of execution time.

The results motivate three conclusions. First, the Gibbs sequence appears to converge quite quickly, even under relatively imprecise sampling conditions. Second, considering the complex form of the response relationship, the relatively few observations in the sample and the relatively high levels of error variance in the second and third experiments, the parameter estimates are surprisingly accurate. Third, Gibbs sampling appears to provide a promising, alternative avenue for deriving estimates of von Liebig production.

Frontier von Liebig Estimation

With the results for the basic algorithm now at hand, it is relatively straight-forward to introduce frontier effects into the estimation. This work follows closely the path-breaking contribution by Koop, Osiewalski and Steel, and especially their "marginally independent efficiency distribution" (MIED) formulation (pp. 83-84). One important issue in this development that must not be overlooked is the need to avoid the problem of 'over-parameterization,' which arises when more than T quantities are being estimated from the T available observations (Fernandez, Osiewalski and Steel). For this reason, we partition the data into N = 5 sub-samples, and restrict attention to efficiency recommendations between these samples. The choice of five divisions is motivated by the fact that an obvious 5-sample division is available in the real-data analysis to follow.

Thus, with i = 1, 2, .., N units in question and t = 1, 2, .., T_i observations on each of the sub-units, the data are of the form of a panel, and the von Liebig panel data production function is

(4) $\quad y_{it} = \min\{ \alpha_o + \alpha_1 w_{it}, \beta_o + \beta_1 n_{it}, \pi \} + \varepsilon_i + u_{it},$

i = 1, 2, .., N, t = 1, 2, ..., T. Except for the introduction of the "inefficiency terms," $\varepsilon_1, \varepsilon_2, .., \varepsilon_5$, the specification shares an obvious proximity with the specification in equation (3). This proximity is more than semantic because it follows, naturally, that the extensions required over the basic algorithm are few.

The inefficiencies $\varepsilon_1, \varepsilon_2, .., \varepsilon_5$, although they are sub-sample specific, provide an additional source of variation and are important because their one-sidedness restriction forces the frontier of the basic von Liebig specification. Many specifications of these inefficiencies have been postulated in the literature, including half-normal, gamma, and beta forms. However, arguably the most simple, and the one to which we restrict attention henceforth, is the exponential form. Some algebra (available upon request) reveals that the resulting posterior has component conditional distribution that are easy to work with. In particular, whereas the basic model retains the properties outlined above, the inclusion of exponential (one-sided) errors implies that the resulting full conditional distribution for the "inefficiencies" has a truncated-normal form and the at the conditional distribution for the parameter λ in the exponential distribution for the errors has a gamma distribution. Because sampling from these distributions is also straight-forward, the inclusion of "frontier-effects" in the basic von Liebig algorithm does not appear to pose any considerable difficulty and we turn to examine its accuracy using simulated data.

One note prior to estimation is that, because of collinearity problems (cf. Koop et al., equation (2), p. 80), the constant components within the "min" operator and the "inefficiency terms" are combined into single elements, ε_i^* for estimation purposes. Although there is a potential loss of information

here, it does not appear to be too troublesome. With this modification, the data are generated according to the model, $(\alpha_1, \beta_1, \pi, \lambda) = (1, 1, 100, 0.01)$ and the terms ε_1^*, ε_2^*, ε_3^*, ε_4^* and ε_5^* are random draws from the exponential distribution with mean $1/\lambda$.

Table 2 contains the estimation of the frontier von Liebig with the same experimental data used in table 1. The results reflect the degree of uncertainty built into the three sets of data according to the increasing value of σ. As this parameter's value increases, the precision of the frontier decreases substantially. The values of the parameters ε_1^*, ε_2^*, ε_3^*, ε_4^* and ε_5^* represent the inefficiencies of the five sub-sets of data in which the sample of was divided. The row labeled "missclassified" indicates the average number of observations that were allocated to the "wrong" regime by the Gibbs procedure, to be compared to the sample size of 220.

Empirical Evidence Using Hexem and Heady Data

The empirical data used in this paper was taken from Hexem and Heady and deal with corn production at five different locations in the United States using water and nitrogen as fertilizers. The locations are Fort Collins, CO, Mesa 1971, AR, Mesa 1970, AR, Yuma Mesa 1970, AR, and Yuma Valley 1970, AR. The agronomic experiments that generated these data utilized an incomplete factorial design. Water and nitrogen were applied in various combinations of five different levels. The five data sets were combined in a single sample to create a panel data of 220 observations.

The estimation of the von Liebig response function and the von Liebig frontier production function was executed using the Gibbs algorithm explained above without any significant difficulty. Table 3 presents the results of the two versions of the von Liebig model. The traditional von Liebig specification corresponds to equation (3). The three different regimes (water, nitrogen, and plateau) are limiting for the reported average number of observations in the Gibbs sample. The results in the frontier von Liebig correspond to equation (4). The intercept for the water and nitrogen regimes was confounded with the one-sided error of the frontier and appears implicitly in the reported values of ε_1^*, ε_2^*, ε_3^*, ε_4^* and ε_5^* that measure the inefficiencies of the five sub-samples from the frontier. The significant variation between the two models is somewhat surprising at this stage and requires further inquiry into the specification of the von Liebig model in relation to the Gibbs sampling procedure. In particular, it seems odd that the value of the plateau in the frontier model is lower that the corresponding value in the production function model.

Conclusion

This paper has presented a Bayesian procedure to estimate traditional von Liebig production function models and frontier specifications. The algorithm is based on an application of the Gibbs sampler as

applied in the Markov Chain Monte Carlo methodology. This preliminary work indicates that the procedure constitutes a promising way to tackle a difficult estimation problem as that represented by a von Liebig model.

Further lines of research have emanated from this work. First of all, the algorithm could be modified to handle a nonlinear specification of the various regimes in the von Liebig model. Second, the algorithm could be extended to deal with an endogenous sample separation of the von Liebig model as estimated in Paris. This specification corresponds to the most general structure of the von Liebig hypothesis as stated in equation (1). Third, the marginal-likelihood algorithm suggested in Chib can be applied in this case for comparing alternative models when the Gibbs sampler is used to obtain the parameter estimates of the various alternatives.

References

Ackello-Ogutu, C., Q. Paris and W. A. Williams. "Testing A Von Liebig Crop Response Function Against Polynomial Specifications." Am. J. Agr. Econ. 67(4)(August 1985):873-80.

Black, J. D. Review of "The Law of Diminishing Returns" by W. J. Spillman and E. Lang. Journal of Farm Economics, (1925):274-280.

Berck, P. J. Geoghegan and S. Stohs. "A Strong Test Of The Von Liebig Hypothesis." Am. J. Agr. Econ. 82(4)(November 2000):948-55.

van den Broeck, J., G. Koop, J. Osiewalski and M. F. J. Steel. "Stochastic Frontier Models: A Bayesian Perspective." Journal of Econometrics 61(1994):273-303.

Casella, G. and E. George. "Explaining the Gibbs Sampler." *American Statistician* 46(1992):167-74.

Chambers, R. G. and E. Lichtenberg. "A Non-Parametric Approach To The Von Liebig-Paris Technology." Am. J. Agr. Econ. 78(2)(May 1996):373-86.

Chib, S. "Marginal Likelihood From the Gibbs Output." *Journal of the American Statistical Association* 90(1995):1313-21.

Chib, S. and E. Greenberg. "Understanding the Metropolis-Hastings Algorithm. *American Statistician* 49(1995):327-335.

Fernandez, C., G. Koop and M. F. J. Steel. "A Bayesian Analysis of Multiple-Output Production

Frontiers." Mimeograph, Department of Economics, University of Edinburgh, October 1999.

Fernandez, C., J. Osiewalski and M. F. J. Steel. "On the Use of Panel Data in Stochastic Frontier Models with Improper Priors." Journal of Econometrics 79(1997):169-93.

Gelfand, A. and A. F. M. Smith. "Sampling Based Approaches To Calculating Marginal Densities." J. Am. Stat. Assoc. 85(1990):398-409.

Gelman, A., J. B. Carlin, H. S. Stern and D. B. Rubin. Bayesian Data Analysis. London: Chapman and Hall, 1995.

Grimm, S. S., Q. Paris and W. A. Williams. "A Von Liebig Model For Water And Nitrogen Crop Response." West. J. Agr. Econ. 12(2)(December 1987):182-92.

Hexem, R. W. and E. O. Heady. "Water Production Functions for Irrigated Agriculture." Iowa State University Press. Ames, Iowa, 1978.

Koop, G., J. Osiewalski and M. F. J. Steel. "Bayesian Efficiency Analysis Through Individual Effects: Hospital Cost Frontiers." Journal of Econometrics 76(1977):77-105.

Lanzer, E. A. and Q. Paris. "A New Analytical Framework For The Fertilization Problem." Am. J. Agr. Econ. 63(1)(February 1981):93-103.

Lavine, M. and M. West. "A Bayesian Procedure for Classification and Discrimination." Can. J. Stat. 20(December 1992):451-6 1.

Liebig, J. von. Organic Chemistry In Its Application To Agriculture And Physiology. London: Playfair, 1840.

Paris, Q. "The Von Liebig Hypothesis." Am. J. Agr. Econ. 74(4)(November 1992):1019-28.

_____ and K. Knapp. "Estimation Of Von Liebig Response Functions." Am. J. Agr. Econ. 71(1)(February 1989):178-186.

Zellner, A. An Introduction To Bayesian Inference In Econometrics. New York: Wiley and Sons, 1971.

Table 1. Experimental evidence – von Liebig model.

Parameter	Experiment 1 ($\sigma = 1$)	Experiment 2 ($\sigma = 10$)	Experiment 3 ($\sigma = 100$)
α_0	0.65	11.08	50.58
	(0.23)	(2.64)	(42.32)
α_1	0.99	0.63	0.75
	(0.00)	(0.06)	(0.31)
β_0	0.21	9.63	-68.85
	(0.23)	(4.57)	(33.55)
β_1	1.00	0.78	0.81
	(0.00)	(0.12)	(0.37)
π	99.93	94.75	62.07
	(0.16)	(3.16)	(40.37)
σ	1.05	10.72	75.87
	(0.05)	(0.96)	(9.02)
missclassified	2.37	65.37	144.07
	(1.20)	(10.89)	(6.79)
Burn-in sample size	10,000	10,000	10,000
Gibbs sample size	10,000	10,000	10,000
Data sample size	220	220	220
Execution time (sec.)	1511	1603	1601

Numbers in parentheses are standard errors in the Gibbs sample.

Table 2. Experimental evidence – frontier von Liebig model.

Parameter	Experiment 1 ($\sigma = 1$)	Experiment 2 ($\sigma = 10$)	Experiment 3 ($\sigma = 100$)
α_1	1.00	0.78	0.47
	(0.03)	(0.05)	(0.24)
β_1	1.00	0.80	1.10
	(0.02)	(0.06)	(0.21)
π	100.13	91.36	92.74

	(4.34)	(2.96)	(29.58)
σ	1.49	10.69	83.75
	(4.78)	(0.71)	(5.96)
ε_1^*	249.85	156.24	18.06
	(6.12)	(2.78)	(15.11)
ε_2^*	41.62	208.54	172.63
	(2.83)	(2.83)	(21.53)
ε_3^*	41.20	11.42	60.08
	(2.28)	(2.69)	(22.28)
ε_4^*	144.37	107.92	67.56
	(7.80)	(2.79)	(22.70)
ε_5^*	141.78	393.72	103.75
	(6.11)	(2.30)	(22.95)
λ	0.01	0.01	0.01
	(0.00)	(0.00)	(0.01)
missclassified	5.21	65.35	143.94
	(13.26)	(0.78)	(7.01)
Burn-in sample size	10,000	10,000	10,000
Gibbs sample size	10,000	10,000	10,000
Data sample size	220	220	220

Numbers in parentheses are standard errors in the Gibbs sample

Table 3. Empirical evidence on Hexem and Heady corn data.

Parameter	Estimation	
	von Liebig	Frontier von Liebig
α_0	691.55	
	(346.33)	
α_1	375.51	828.16
	(84.33)	(74.94)
β_0	1420.33	
	(247.36)	
β_1	259.92	698.71
	(73.67)	(68.35)
π	4462.93	2373.25
	(208.90)	(251.22)
σ	886.80	778.09
	(75.22)	(60.09)
water limiting	85.58	81.55
	(5.06)	(8.94)
nitrogen limiting	70.16	74.40
	(6.18)	(7.75)
plateau limiting	64.26	64.04
	(7.19)	(8.00)
ε_1^*		4721.92
		(208.43)
ε_2^*		1992.61
		(203.08)
ε_3^*		87.68
		(78.47)
ε_4^*		1846.45
		(239.64)
ε_5^*		4946.66
		(225.07)
λ		4.4×10^{-4}
		(1.8×10^{-4})
Burn-in sample size	10,000	10,000
Gibbs sample size	10,000	10,000
Data sample size	220	220
Execution time (seconds)	1511	1603

Numbers in parentheses are standard errors in the Gibbs sample.