Australian Journal of Agricultural Economics, Vol. 30, Nos 2 and 3 (Aug/Dec 1986), pp. 128-152

A BAYESIAN FRAMEWORK FOR OPTIMAL INPUT ALLOCATION WITH AN UNCERTAIN STOCHASTIC PRODUCTION FUNCTION

WILLIAM E. GRIFFITHS*
University of New England, Armidale, NSW 2351

A Bayesian approach is used to provide a framework for optimal input allocation for a stochastic production function with uncertain parameters. The chosen production function is a single-input version of a function which can exhibit positive or negative marginal risk. The sensitivity of optimal input allocation to the degree of risk aversion, the marginal risk parameter, and the assumed level of uncertainty about the parameters is examined. Using an example problem it is shown that parameter uncertainty will not always lead to a lower input level when a producer is risk averse. Some problems with the production function specification are uncovered, as is the incompatibility of some expected utility and stochastic assumptions.

In most empirical studies of production under uncertainty (see, for example, Anderson, Dillon and Hardaker 1977 and references therein) there has been a tendency to concentrate on the effects of stochastic prices and stochastic output where the functional form of the production function is assumed to be known, and the estimated parameters of the production function are treated as though they are known to be equal to the 'true underlying values'. Under this set of conditions it is straightforward, in principle, to use the classical or sampling theory approach to inference to derive the probability distribution (or the moments may be sufficient) of profit. The probability distribution of profit and a specific utility function provide a framework for finding input levels which maximise expected utility. The ease with which optimal input levels can be found depends on the functional forms of the production and utility functions and on the stochastic assumptions about price and output; the task can be computationally difficult (see, for example, Anderson and Griffiths 1982).

A natural extension of previous studies on input allocation under uncertainty is to allow for uncertainty in the values of the production function parameters. In a typical empirical study these parameters are estimated using econometric techniques and, consequently, the estimates are subject to sampling error. Under the assumptions of most models, any future realisations of output will be determined by the true underlying parameters, not by the estimates based on past realisations of output. Thus, there is a need to allow for the uncertainty associated with the parameter estimates when finding future optimal input levels. The classical approach to inference does not provide a convenient

* I would like to acknowledge Shahidur Rahman's enthusiastic and efficient research assistance and useful conversations with Peter Pope and Howard Doran on the convergence properties of some integrals.

framework for incorporating parameter uncertainty. It does not permit the inclusion of sampling error without making the moments of output, and hence the moments of profit, a function of some unknown parameters. The Bayesian approach to inference, however, provides a natural, convenient and logical framework for pursuing the problem. In this approach, uncertainty about the unknown parameters is expressed in terms of a Bayesian subjective probability distribution which depends only on known prior parameters and/or past observations on output and inputs. From the joint probability distribution on all the parameters, and the probability distributions on other stochastic elements such as the production function disturbance and output price, it is conceptually possible to derive the probability distribution (or the moments) of profit, and to use this information in conjunction with a utility function to find input levels which maximise expected utility. Furthermore, if additional information on the parameters becomes available, Bayesian methods for including this additional information are convenient.

The purpose of this paper is to illustrate how a Bayesian framework can be used to find a utility maximising input level when a producer is faced with a stochastic production function with uncertain parameter values. Allowance for parameter uncertainty is made by deriving the joint posterior distribution of the parameters and considering the consequent effects on the mean and variance of output. Examples of similar treatments with related models can be found in Zellner (1971, Ch. 11).

The vehicle used to illustrate the Bayesian framework is a single-input version of a Just and Pope (1978) type of production function which can exhibit positive or negative marginal risk. Estimation procedures for such functions have been outlined by Just and Pope (1978) and Griffiths and Anderson (1982); some specific estimates, and some results from constrained utility maximisation, have been presented by Anderson and Griffiths (1981, 1982). However, little attention has been given to the characteristics of unconstrained utility maximising input levels for production functions of this type. A second objective in this paper, is therefore, to examine the nature of optimal input levels for Just–Pope type production functions and to explore their sensitivity to marginal risk, the degree of risk aversion, and the assumed level of uncertainty about the parameters.

In the first section the production function is specified and the utility maximisation framework is outlined. The parameter values for an illustrative example which is employed in later sections are given. The next section examines the problem of maximising expected utility when the parameters are known with certainty. It is shown that the behaviour of the expected utility function is quite different for different parameter values. Some properties of the production function which might be considered undesirable are outlined. The Bayesian framework is introduced in the third section for the special case of zero marginal risk. An expression for expected utility is derived in terms of the posterior moments of the parameters and optimal solutions for an example problem are computed. The next section contains a digression where some results on the non-existence of expected utility are discussed. It is pointed out that many results are not robust to assumptions about the utility function and the probability distribution of profits. The Bayesian framework is extended to the non-zero marginal risk case in the second last section. The differences which occur between the Bayesian solution and the solution at a set of point estimates are highlighted. The paper is concluded with a section on the implications of the findings and possible extensions.

The Framework

Consider a producer who wishes to maximise expected utility, $E(U)$, where utility is assumed to be some function of profit, Π, which in turn is given by:

(1) $\Pi = py - wx$

where p is the price of output y, and w is the price of the single variable input x. Fixed costs are not explicitly included but can be allowed for by viewing Π as variable profit. It is assumed that the firm faces a competitive market and hence that its levels of input and output have no bearing on p and w. Output is assumed to be related to the input through the production function:

(2) $y = \alpha x^{\beta} + u x^{\gamma/2}$

where α, β and γ are unknown parameters; u is a random error term which has zero mean, constant variance σ^2, and is uncorrelated over time. Output and hence profit are stochastic because u is stochastic; the producer's decision concerning the optimal level of x is made before u has been realised.

The function in (2) is a single-input 'Cobb–Douglas type' with an additive heteroscedastic disturbance; it can exhibit positive marginal expected product and either positive or negative marginal risk. Marginal expected product is given by:

(3) $$\frac{\partial E(y)}{\partial x} = \beta \alpha\, x^{\beta - 1}$$

and will be positive under the usual assumptions that both α and β are positive. Marginal risk is given by:

(4) $$\frac{\partial V(y)}{\partial x} = \gamma \sigma^2 x^{\gamma - 1}$$

and will be positive or negative depending on the sign of γ. Just and Pope (1978) argue that functions such as (2) which possess this property are preferable to the traditional multiplicative error specification $y = \alpha\, x^{\beta} e^{u}$ where marginal risk is constrained to be positive (the same sign as β).

As a simplifying assumption both input and output price will be treated as non-stochastic and known. Also, uncontrollable but *ex post* measurable variables which are likely to enter the production function will be ignored. It would, of course, be more realistic to allow for stochastic prices, particularly output price, and to include other variables in the production function. It was decided, however, to defer these complications to future work.

The utility function, or approximation to expected utility, which will be used is:

$$(5)\quad E(U)=E(\Pi)-\frac{\phi}{2}V(\Pi)$$

This function is convenient and gives a simple trade-off between mean and variance. When Π is normally distributed it is straightforward to use properties of the log-normal distribution to show that maximising equation (5) yields the same input level as maximising the expectation of the constant absolute risk aversion utility function:

$$(6)\quad U=-\exp(-\phi\Pi)$$

Most of the analysis will use equation (5), but some reference will also be made to equation (6).

To illustrate the framework some specific parameter values and prices will be used, namely:

$$(7)\quad \alpha=80 \quad \beta=0.2 \quad \sigma^2=400 \quad p=1000 \quad w=120$$

These values are based on the function estimated by Anderson and Griffiths (1981) for wool production in the pastoral zone of eastern Australia, and on the prices used by Anderson and Griffiths (1982). All inputs other than labour have been taken as fixed at their geometric means. The coefficient β has been rounded to 0.2 and α has been increased from 20 to 80. The purpose of increasing α was to increase the variation in the mean function to compensate for variation lost by fixing all inputs other than labour, and to increase the risk free optimum which was about 80 work weeks and corresponded to $\gamma=-0.3$. The sample range for labour was 60 to 1400 and, since one objective was to examine the sensitivity of results to changes in γ, it seemed desirable to make the risk free optimum more toward the middle of the sample range.

Given the values in (7), the risk free optimum obtained by maximising (1), or by maximising (5) with $\phi=0$, is given by:

$$(8)\quad x^*=\left(\frac{\alpha\beta p}{w}\right)^{1/(1-\beta)}=453$$

Maximisation with Known Parameters

Before turning to the Bayesian framework for finding the optimum input level when the parameters are uncertain, it is informative to consider some results when the parameters are known. It will be assumed that the disturbance u is normally distributed and hence that, given known parameters, output y and profit Π are also normally distributed. Under these circumstances, both utility specifications (5) and (6) lead to identical results and so only (5) will be considered. Thus, the function to be maximised is:

$$\begin{aligned}(9)\quad E(U) &= E(\Pi) - (\phi/2)V(\Pi)\\ &= pE(y) - wx - (\phi/2)p^2V(y)\\ &= p\alpha x^{\beta} - wx - (\phi/2)p^2\sigma^2x^{\gamma}\end{aligned}$$

Differentiating equation (9) with respect to x and setting the derivative to zero yields:

$$(10)\quad \alpha\beta x^{\beta-1} - (\phi/2)p\sigma^2\gamma x^{\gamma-1} = (w/p)$$

This equation is not amenable to analytic solution and so equation (9) must be maximised numerically. However, it is possible to consider two special cases. First, if marginal risk is zero ($\gamma=0$), or if the producer is risk neutral ($\phi=0$), the solution to (10) is identical to that given in (8). Secondly, if $\gamma=\beta>0$, then the solution to (10) is:

$$(11)\quad x^* = \left(\frac{p\beta[\alpha-(\phi/2)p\sigma^2]}{w}\right)^{1/(1-\beta)}$$

Comparing (8) with (11) shows how the presence of risk aversion and a positive marginal risk associated with x lead to a reduction in the optimal level of x. When the quantity in square brackets in (11) is negative, a real solution to (11) does not exist and the boundary point $x^*=0$ leads to the maximum.

The second-order condition for maximising (9) can be derived as:

$$(12)\quad x^{\beta-\gamma} > \frac{\phi p\sigma^2\gamma(\gamma-1)}{2\alpha\beta(\beta-1)}$$

If $0<\beta<1$ and marginal risk is negative ($\gamma<0$), then the right hand side of this inequality will be negative and hence (12) will be satisfied for all non-negative input levels. On the other hand, for $\gamma>0$, the condition in (12) will not always be satisfied. For example, when $\beta=\gamma>0$, it reduces to $\phi p\sigma^2<2\alpha$, the condition already noted above in connection with the first-order condition.

Some examples of the behaviour of the $E(U)$ function in (9) are illustrated in Figure 1 for the parameter values in (7) and selected values of ϕ and γ. When γ is negative ($\gamma=-0.23$ for the graph), $V(y)\rightarrow\infty$ and $E(U)\rightarrow-\infty$ as $x\rightarrow 0$. Thus, for $\gamma<0$, the optimal x will always be positive. Whether or not the maximum achievable $E(U)$ is positive or negative when $\gamma<0$ has no significance. In the graph it is negative; for lower values of ϕ it would be positive.

For positive γ the behaviour can be quite different. In this case when $x=0$, $E(U)=0$, and so if $E(U)$ is negative for all positive x the optimal solution is the boundary point $x=0$. The graph for $\phi=0.0005$ and $\gamma=0.15$ is an example of such a case; the function reaches a maximum at $x=80$, but $x=0$ is still optimal because $E(U)$ is negative at $x=80$. When γ is reduced to 0.12, $E(U)$ reaches a positive maximum at $x=195$ and so this level becomes optimal. In both cases ($\gamma=0.12$ and $\gamma=0.15$) the function exhibits a minimum at a small value of x, as well as the maximum. The final illustration is for $\gamma=0.2$ and $\phi=0.00005$. In this

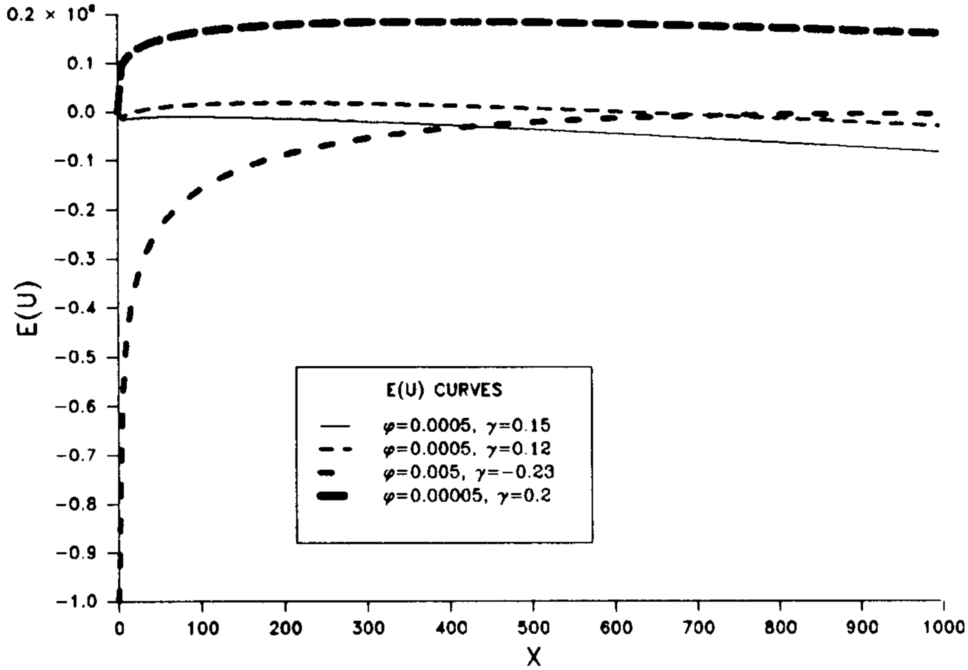

FIGURE 1—E(U) Function for Selected ϕ and γ.

case $E(U)$ is always positive and reaches a maximum at $x=385$.

The behaviour of the $E(U)$ function is a direct consequence of the properties of the production function, some of which may be considered undesirable. In particular, it might be considered desirable for output to be zero when the input level is zero. When $\gamma>0$, this property is satisfied. However, for $\gamma=0$ and $x=0$, $y=u$; and for $\gamma<0$ and $x\to 0$, $y\to\pm\infty$. Thus, when $\gamma\leqq 0$ the production function is unlikely to be a good approximation to reality in the region of $x=0$. The situation may be improved if, instead of normality, an alternative distributional assumption about the disturbance u is made. Some preliminary efforts with the gamma and log-normal distributions suggest that work in this direction may be difficult. It should also be noted that some of the estimation results in Just and Pope (1978) and Griffiths and Anderson (1982) depend on the normality assumption.

The sensitivity of the optimal input level to γ and ϕ is illustrated in Figure 2 where it is graphed against γ over the interval $-1\leqq\gamma\leqq 0.5$, for the three values $\phi=\{0.005, 0.0005, 0.00005\}$. When $\gamma=0$ the optimal solution is the risk free one of 454 work weeks (correct to the nearest even number of weeks), irrespective of the value of ϕ. For $\gamma>0$ the following comments can be made:

$\phi=0.005$: This relatively high weight placed on the variance term makes a zero input level the best option for all $\gamma>0$.

$\phi=0.0005$: A positive input level is optimal for small values of γ (up to 0.12); otherwise $x=0$ is best.

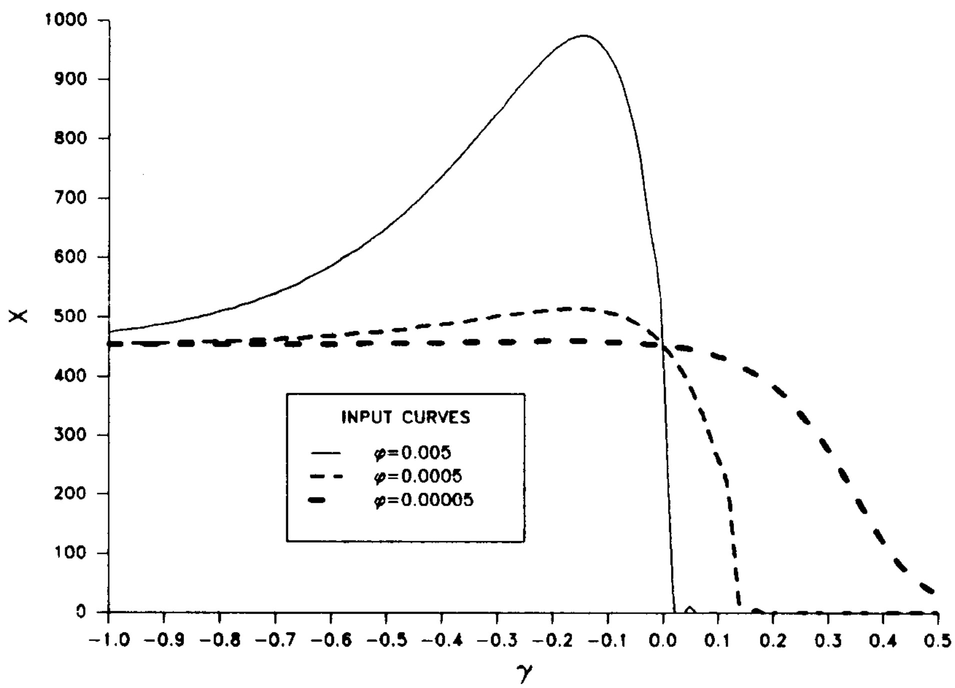

FIGURE 2—Optimal Input Levels with Known Parameters.

$\phi=0.00005$: This relatively small weight on the variance term means a positive input level is optimal for all values of γ considered, and the optimal level declines as the marginal risk parameter γ increases.

As expected, in all cases the optimal input level approaches zero as $\gamma \rightarrow \infty$, because the term $(\phi/2)p^2\sigma^2x^\gamma$ becomes dominant in the computation of $E(U)$.

For $\gamma<0$ and a decreasing γ (becoming more negative), the optimal input level increases at first, reaches a maximum at $\gamma=-0.14$, and then decreases, becoming asymptotic to the risk free ($\gamma=0$) solution of 454. When $\phi=0.005$ the maximum optimal level is relatively large (more than twice the risk free solution), but for the small value of $\phi(=0.00005)$ it barely deviates from the risk free solution. The value of γ for which the optimal input level reaches a maximum does not depend on ϕ.

The behaviour of the optimal solution as $\gamma \rightarrow -\infty$ could perhaps be counter-intuitive. It might be expected that the greater the risk reducing effect of an input (the more negative the value of γ), the greater would be the optimal level of that input. As has been noted, however, the optimal input level increases only over a limited range of γ (as far as -0.14 in the example), and then declines. The reason for this behaviour is apparent after examination of the $E(U)$ function in (9). As $\gamma \rightarrow -\infty$, the last term $(\phi/2)p^2\sigma^2x^\gamma$ becomes negligible relative to the remainder of the function and thus maximisation of $E(U)$ becomes equivalent to maximisation of $E(\Pi)$. Thus, if there are two risk averse producers who are identical in every respect, except that the first producer has a value of the parameter γ which is more negative than that of the other producer, then it is possible for the utility maximising input level of the first producer to be less than that for the other producer. It is questionable whether this property is desirable. This raises further questions about the appropriateness of the production function specification.

Maximisation with Uncertain Parameters and Zero Marginal Risk

To introduce the Bayesian framework for handling uncertain parameters it will first be assumed that $\gamma=0$, and then, in the next section, the more complicated case of non-zero marginal risk will be considered. In contrast to the parameter certainty case, setting $\gamma=0$ does not imply that the variance of output is irrelevant for maximisation of $E(U)$; as will be demonstrated, with uncertain parameters the variance of output depends on the input level even when $\gamma=0$.

The production function considered in this section is, therefore,

$$(13) \quad y=\alpha x^{\beta}+u$$

where $u \sim N(0, \sigma^2)$. It is assumed that $(\alpha, \beta, \sigma^2)$ are unknown, but that the producer's uncertain knowledge about these parameters can be expressed in terms of a Bayesian subjective probability density function. Furthermore, this probability density function is assumed to be a posterior density function derived from an initial non-informative prior density and T subsequent observations on y and x.

The past observations will be subscripted as $\{(y_t, x_t), t=1,2,\ldots,T\}$ and the future values of y and x for which expected utility will be maximised will bear no subscripts. For shorthand notation the past data will be denoted by D. That is,

$$(14) \quad D=\{y_1,y_2,\ldots,y_T,x_1,x_2,\ldots,x_T\}$$

The initial non-informative prior that is used is obtained by assuming α, β and σ^2 are a priori independent with marginal density functions

$$(15) \quad g(\alpha) \propto \text{constant} \qquad -\infty<\alpha<\infty$$

$$(16) \quad g(\beta)=1 \qquad 0<\beta<1$$

$$(17) \quad g(\sigma^2)\propto 1/\sigma^2 \qquad \sigma^2>0$$

The density functions in (15) and (17) for α and σ^2 are standard non-informative priors for parameters with ranges $(-\infty,\infty)$ and $(0,\infty)$, respectively (Zellner 1971). For parameters with a range from 0 to 1, such as β in this case, a number of non-informative priors have been suggested in the literature. A simple uniform one has been chosen; an alternative could be used with few additional complications.

The assumed range for α in (15) could be regarded as unrealistic. It is likely that most producers would specify $\alpha>0$ for even an uninformative prior. However, the specification in (15) does greatly simplify the analysis and, as an approximation, extending the range of α

to negative values is unlikely to be of much consequence for relatively large T. In future research, specifying an alternative to (15) should be considered in conjunction with relaxation of the assumption of a normally distributed disturbance.

Before proceeding to maximisation of $E(U)$, expressions for the posterior density function for (α,β,σ^2) and a number of related quantities need to be derived. Working in this direction, the likelihood function can be written as:

$$(18)\quad l(\alpha,\beta,\sigma^2|D) \propto (\sigma^2)^{-T/2} \exp\{-\Sigma\,(y_t-\alpha x_t^\beta)^2/2\sigma^2\}$$

In (18) and all future equations, the summation Σ is assumed to be indexed from $t=1$ to $t=T$. Using Bayes' theorem to combine the likelihood function in (18) with the joint prior density function $g(\alpha,\beta,\sigma^2)$ obtained from (15) to (17) yields the joint posterior density function:

$$(19)\quad g(\alpha,\beta,\sigma^2|D) \propto (\sigma^2)^{-T/2-1} \exp\{-\Sigma(y_t-\alpha x_t^\beta)^2/2\sigma^2\}$$

From this posterior density function a number of quantities useful for the computation of expected utility can be derived. Those of particular interest are listed in equations (20) to (23). Details are given in Appendix 1.

The posterior mean and variance for α conditional on β are given, respectively, by:

$$(20)\quad \hat{\alpha}=E(\alpha\,|\,\beta,D)=\Sigma y_t x_t^\beta/\Sigma x_t^{2\beta}$$

and

$$(21)\quad \hat{\sigma}_\alpha^2=V(\alpha\,|\,\beta,D)=\hat{\sigma}^2/\Sigma x_t^{2\beta}$$

where $\hat{\sigma}^2$ is the posterior mean for σ^2 conditional on β and is given by:

$$(22)\quad \hat{\sigma}^2=E(\sigma^2|\beta,D)=(\Sigma y_t^2-\hat{\alpha}\Sigma y_t x_t^\beta)/(T-3)$$

The marginal posterior density function for β can be derived as:

$$(23)\quad g(\beta|D) \propto (\Sigma x_t^{2\beta})^{-0.5}\,[(T-3)\hat{\sigma}^2]^{-(T-1)/2}$$

It is also useful to consider the predictive density function for the future disturbance u. In the sampling theory approach to inference a traditional predictive density is conditional on the unknown parameters. In the Bayesian approach, where the unknown parameters are treated as random variables, the joint density function for the parameters and the future random variable is found as $g(\alpha,\beta,\sigma^2,u|D)$ in this case; then, the predictive density function is the marginal density obtained by integrating out the parameters. Details are given in Appendix 1. The results useful for this analysis are:

$$(24)\quad E(u\,|\,\beta,D)=E(u)=0$$

$$(25) \quad V(u|\beta,D)=\hat{\sigma}^2$$

and

$$(26) \quad \mathrm{Cov}(u,\alpha|\beta,D)=\mathrm{Cov}(u,\alpha)=0$$

Enough results have now been established to consider an expression for expected utility. In the derivation of this expression unconditional expectations are separated into expectations conditional on β and expectations with respect to β. Such an approach is necessary because any expectations with respect to β require numerical integration.

Rewriting (9) to recognise that the new problem is to maximise expected utility given past observations yields:

$$(27) \quad E(U|D)=pE(y|D)-wx-(\phi/2)p^2V(y|D)$$

Using the results in (20) to (27), the mean and variance of y can be derived as:

$$\begin{aligned}(28) \quad E(y|D) &= E(\alpha x^{\beta}|D)+E(u|D)\\ &= E_{\beta}[x^{\beta}E(\alpha|\beta,D)]+0\\ &= \int_0^1 x^{\beta}\hat{\alpha}g(\beta|D)d\beta\end{aligned}$$

and

$$\begin{aligned}(29) \quad V(y|D) &= V(\alpha x^{\beta}|D)+V(u|D)+2\mathrm{Cov}[(\alpha x^{\beta},u)|D]\\ &= E(\alpha^2x^{2\beta}|D)-[E(\alpha x^{\beta}|D)]^2+E_{\beta}[V(u|\beta,D)+\\ &\quad 2x^{\beta}E(\alpha u|D)]\\ &= E_{\beta}[x^{2\beta}\{V(\alpha|\beta,D)+[E(\alpha|\beta,\mathrm{D})]^2\}]-[E(y|D)]^2+\\ &\quad E_{\beta}\hat{\sigma}^2+0\\ &= \int_0^1 x^{2\beta}(\hat{\sigma}_{\alpha}^2+\hat{\alpha}^2)g(\beta|D)d\beta-[E(y|D)]^2+\\ &\quad \int_0^1 \hat{\sigma}^2g(\beta|D)d\beta\end{aligned}$$

Thus, the optimal input level can be found by finding that value of x which maximises $E(U|D)$, where $E(y|D)$ and $V(y|D)$ are defined in (28) and (29), respectively. Note that two integrals (one in $E(y|D)$ and one in $V(y|D)$) need to be evaluated numerically for *every* value of x for which $E(U|D)$ is computed. Note also that the last term in the last line of (29) does not depend on x and so can be ignored when finding the optimal x.

To illustrate the above approach and to examine the sensitivity of the optimal input level to parameter uncertainty, some data were artificially generated and the corresponding optimal input levels found. Sample sizes of $T = 5$, 10, 20, 30 and 50 were considered. The values for x_t were drawn from a uniform distribution with range (60, 1500) and corresponding values for y_t were computed using the production function relationship and the parameter values used earlier (see equation (7)). For a given sample size T and a given value of ϕ the optimal x is obtained by first computing $E(U|D)$ for all x between 5 and 1500 at 5-unit intervals and then selecting that value for which $E(U|D)$ is a maximum.

In Table 1 the optimal input levels for the different sample sizes and different levels of risk aversion are reported. It should be remembered that these results depend on the particular sample which was generated, and, therefore, any general conclusions should be made cautiously. Nevertheless the values in the table do give an idea of the possible variation due to parameter uncertainty and it is possible to make some observations. First, when sample size is small and hence the degree of uncertainty about the parameters is greatest, a zero input level is optimal for high levels of risk aversion. Second, it is possible for an optimal value to be greater than the solution $x = 453$ which occurs under parameter certainty. A more relevant comparison, however, is between each optimal solution and the solution which would be obtained if α and β were set at their posterior means (a reasonable pair of point estimates) without any recognition of sampling variation. For $T \geqq 20$ such a comparison shows that most optimal solutions are greater than the corresponding solution at the posterior means and the difference is greater the larger the sample size and the greater the degree of risk aversion. A possible explanation for this behaviour is the size of the sample mean, $\bar{x}$, relative to the solution at the posterior means. In the traditional linear regression model the variance of a prediction error is smallest when the explanatory variables are set equal to their means over the sample period. The model in this case is by no means directly analogous, and this explanation is not satisfactory for $T \leqq 10$, but for

TABLE 1

Optimal Input Levels with Uncertain Parameters and Zero Marginal Risk[a]

Risk aversion (ϕ)	Sample size (T) 5	10	20	30	50
0.01	0	0	565	595	570
0.005	0	0	550	580	540
0.001	0	470	510	515	460
0.0005	270	470	490	495	440
0.0001	275	465	465	465	420
0.00005	280	465	465	465	415
0.00001	290	465	460	460	415
$\phi = 0$, $\beta = 0.2$	453	453	453	453	453
$\phi = 0$, α, β set at posterior means	338	486	466	464	415
$\bar{x}$	342	588	686	713	728

[a] The values in the upper part of the table are within an accuracy of 5 work weeks.

$T \geqq 20$ there does seem to be a tendency to reduce variance by increasing x above the risk free solution toward the sample mean of the x_t. For small sample sizes ($T=5, 10$) the optimal solutions are below the corresponding solution at the posterior means, suggesting that reducing x will decrease the variance when parameter uncertainty is greatest.

The way in which some of the posterior moments and density functions change as T changes are illustrated in Table 2 and Figure 3. The posterior means for α and β used to calculate the solutions in the second last row of Table 1 and the standard deviations of β are given in Table 2. The way the posterior density function $g(\beta|D)$ changes as a larger sample yields more information is illustrated in Figure 3.

Before turning to the case of uncertain parameters and non-zero γ, the consequences of changing the utility function will be considered. Also, some results for the traditional multiplicative error production function specification criticised by Just and Pope (1978) will be described.

TABLE 2

Some Posterior Moments for α and β

	Sample size (T)				
	5	10	20	30	50
E(α|D)	109.87	74.10	76.61	78.23	89.95
E(β|D)	0.1521	0.2120	0.2059	0.2038	0.1833
S.D. (β|D)	0.0917	0.0406	0.0224	0.0207	0.0147

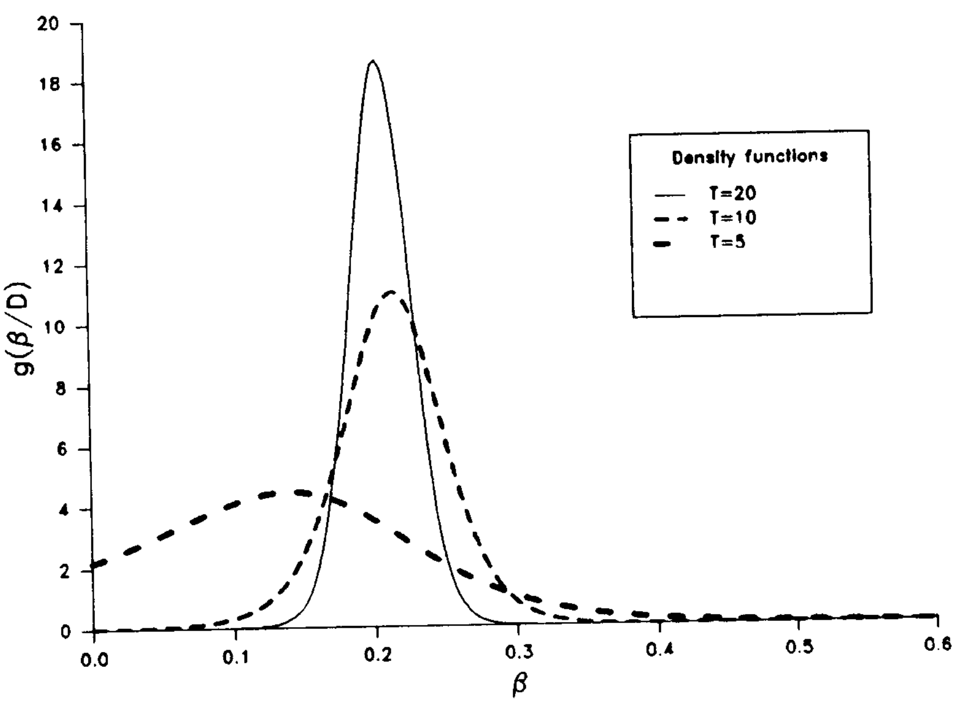

FIGURE 3—Posterior Density Functions for β when $\gamma=0$.

Two Examples Where Expected Utility Does Not Exist

If the mean–variance formulation of expected utility is replaced by the constant absolute risk aversion utility function and an attempt is made to maximise:

$$(30)\quad E(U|D) = -E[\exp(-\phi\Pi)|D]$$

a problem is encountered. To describe this problem it is convenient to write the expectation on the right hand side of (30) as:

$$(31)\quad E[\exp(-\phi\Pi)|D] = \exp(\phi wx)E[\exp(-\phi py)|D]$$

In Appendix 1 it is shown that the density function $g(\alpha,u|\beta,D)$ is a bivariate t-density. Using properties of the t-distribution it follows that the predictive density for $y = \alpha x^{\beta} + u$, conditional on β, is a univariate t-density. This result follows because, given β, y is a linear combination of α and u. Since ϕ and p are constants it also follows that $g(z|\beta,D)$, where $z = \phi py$, is a t-distribution. The last expectation in equation (31) can now be written as:

$$(32)\quad E[\exp(-\phi py)|D] = E_{\beta}E[\exp(-z)|\beta,D]$$

A problem arises because the integral defined by $E[\exp(-z)|\beta,D]$ does not converge. Thus, expected utility does not exist (or is infinite). The difficulty arises as z(or y)$\rightarrow -\infty$. As $z\rightarrow -\infty$, $\exp(-z)$ grows at too rapid a rate to be compensated for by the decline in $g(z|\beta,D)$.

What are the implications of this result? Since the difficulty occurs as $y\rightarrow -\infty$, the obvious problem lies with the production function specification which permits negative values of output. Certainly, it can be concluded that the expected utility specification in (30), and the production function with an additive normally distributed error, are not compatible within the Bayesian framework considered here. Also, the lack of robustness to the expected utility formulation and the distributional assumptions could be of some concern. When Π is normally distributed, both definitions of expected utility lead to the same result; when Π follows a t-distribution, one definition of expected utility leads to a result but the other is infinite.

The other example where expected utility does not exist is somewhat of a digression since it involves the traditional production function specification with multiplicative error, namely:

$$(33)\quad y = \alpha x^{\beta}\exp(u)$$

where $u \sim N(0,\sigma^2)$. In this case negative values of output are precluded but a problem with $E(U|D)$ still arises.

Taking logs of equation (33) and adding a t subscript to indicate observations ($t = 1,2,\ldots,T$) over a sample period yields the normal linear regression model:

$$(34)\quad q_t = \log y_t = \log\alpha + \beta\log x_t + u_t$$

Beginning with the usual non-informative prior on $(\log \alpha, \beta, \sigma^2)$ it is well known that, after T observations, the predictive density function for a future observation on $q=\log y$ will be a t-distribution. Alternatively, if a special prior is placed on β to restrict it to the range (0,1), the predictive density for q conditional on β will be a t-distribution. In both cases $E(y|D)$ and $V(y|D)$ do not exist, and, therefore, expected utility does not exist when it is defined as $E(U)=E(\Pi)-(\phi/2)V(\Pi)$. The non-existence of the moments is clear when $E(y|D)$ is written as:

$$(35) \quad E(y|D)=E[\exp(q)|D]$$

where q has a t-distribution. It is a similar problem to that encountered earlier, but, in this case, the difficulty arises at the upper limit of the integral as $q \to \infty$. The quantity $\exp(q)$ increases at too rapid a rate to be compensated for by the decline in the t-density $g(q|D)$.

If the constant absolute risk aversion utility function is employed in conjunction with the production function in equation (33), the problem boils down to the existence or otherwise of:

$$(36) \quad E[\exp(-y)|D]=E[\exp\{-\exp(q)\}|D]$$

In this case $\exp\{-\exp(q)\}$ does not approach infinity and the integral does converge. Specifically,

$$\text{as } q \to \infty, \ \exp\{-\exp(q)\} \to 0$$

and

$$\text{as } q \to -\infty, \ \exp\{-\exp(q)\} \to 1$$

Thus, for the multiplicative error production function considered within the Bayesian framework of this paper, maximisation of expected utility will yield a solution for the constant absolute risk aversion utility function, but does not yield a solution for the mean–variance expected utility specification because, in this latter case, expected utility does not exist. More generally, if the log of profit follows a t-distribution, then the utility function in (6) will give a solution but that in (5) will not.

Maximisation with Uncertain Parameters and Non-zero Marginal Risk

Returning to the original model with non-zero marginal risk, the production function can be written as:

$$(37) \quad y=\alpha x^{\beta}+ux^{\gamma/2}$$

where $u \sim N(0,\sigma^2)$. As in the previous case, it will be assumed that T past observations $(y_t, x_t, t=1,2,\ldots,T)$ are available, and that the problem is to find that value of x which maximises the expected utility (defined in terms of the mean–variance formulation) from a future realisation on y.

In addition to the prior density functions in (15) to (17), it will be assumed that γ is a priori independent of the other parameters and that its prior density function is the non-informative one:

$$(38) \quad g(\gamma) \propto \text{constant} \qquad -\infty < \gamma < \infty$$

The introduction of stronger prior information would not make the analysis any more difficult. The likelihood corresponding to the model in equation (37) is given by:

$$(39) \quad l(\alpha,\beta,\sigma^2,\gamma|D) \propto (\sigma^2)^{-T/2}\,\Pi x_t^{-\gamma/2} \exp\{-\Sigma[(y_t - \alpha x_t^\beta)/x_t^{\gamma/2}]^2/2\sigma^2\}$$

Details of the posterior analysis for this model are given in Appendix 2. Those results which differ from the previous model ($\gamma = 0$) and which are useful for the computation of expected utility are:

$$(40) \quad \hat{\alpha} = E(\alpha|\beta,\gamma,D) = \Sigma y_t x_t^{\beta-\gamma}/\Sigma x_t^{2\beta-\gamma}$$

$$(41) \quad \hat{\sigma}^2 = E(\sigma^2|\beta,\gamma,D) = [\Sigma(y_t^2/x_t^\gamma) - \hat{\alpha}\Sigma y_t x_t^{\beta-\gamma}]/(T-3)$$

$$(42) \quad \hat{\sigma}_{\hat{\alpha}}^2 = V(\alpha|\beta,\gamma,D) = \hat{\sigma}^2/\Sigma x_t^{2\beta-\gamma}$$

and

$$(43) \quad g(\beta,\gamma|D) \propto (\Sigma x_t^{\beta-\gamma})^{-0.5}\,\Pi x_t^{-\gamma/2}[(T-3)\hat{\sigma}^2]^{-(T-1)/2}$$

Expected utility can now be defined in terms of these quantities as follows:

$$(44) \quad E(U|D) = pE(y|D) - wx - (\phi/2)p^2 V(y|D)$$

where

$$(45) \quad E(y|D) = \int_{-\infty}^{\infty}\int_0^1 x^\beta \hat{\alpha} g(\beta,\gamma|D)d\beta d\gamma$$

and

$$(46) \quad V(y|D) = \int_{-\infty}^{\infty}\int_0^1 x^{2\beta}(\hat{\sigma}_{\hat{\alpha}}^2 + \hat{\alpha}^2)g(\beta,\gamma|D)d\beta d\gamma - [E(y|D)]^2 + \int_{-\infty}^{\infty}\int_0^1 x^\gamma \hat{\sigma}^2 g(\beta,\gamma|D)d\beta d\gamma$$

It is worth pointing out the difference between the terms in (45) and (46) and those which would be obtained if the traditional approach of treating estimated parameters as true parameters was adopted. For the traditional approach, $E(y|D)$ would be taken as $x^\beta\hat{\alpha}$, with $\hat{\alpha}$ being evaluated at a pair of estimates (say maximum likelihood) of γ and β. In the Bayesian framework, $E(y|D)$ can be viewed as a weighted average of all possible values of $x^{\beta}\hat{\alpha}$, with weights given by the posterior density $g(\beta,\gamma|D)$. The first two terms in $V(y|D)$ appear because of the Bayesian framework's recognition of uncertainty in the parameters; these terms would not appear in the traditional approach. The last term is analogous to the term in $E(y|D)$. In the traditional approach, it would appear as $x^{\gamma}\hat{\sigma}^2$ with $\hat{\sigma}^2$ evaluated at a pair of estimates for β and γ. In the Bayesian framework, $x^{\gamma}\hat{\sigma}^2$ is averaged over all possible values of β and γ.

Computationally, $E(U|D)$ appears daunting because of the need to evaluate numerically three bivariate integrals for all values of x. Fortunately, some short cuts can be found. Since x^{β} and $x^{2\beta}$ do not depend on γ, the first two integrals can be integrated with respect to γ before considering any values of x, leaving only univariate numerical integration with respect to β for every value of x. Similarly, in the last integral, x^{γ} does not depend on β so that integration with respect to β can occur before x^{γ} is considered, leaving only univariate numerical integration with respect to γ for every value of x. Note that, unlike the previous model, the last term in the expression for the variance now depends on x.

To illustrate the above approach, optimal solutions for different values of T, ϕ and γ were computed using the example and parameter values considered earlier. Before examining these optimal solutions, it is instructive to consider the marginal posterior density functions for γ(for $T=5, 10, 20$) presented in Figures 4a, 4b and 4c. These densities correspond to a true value of γ of -0.09. Three points are worth noting.

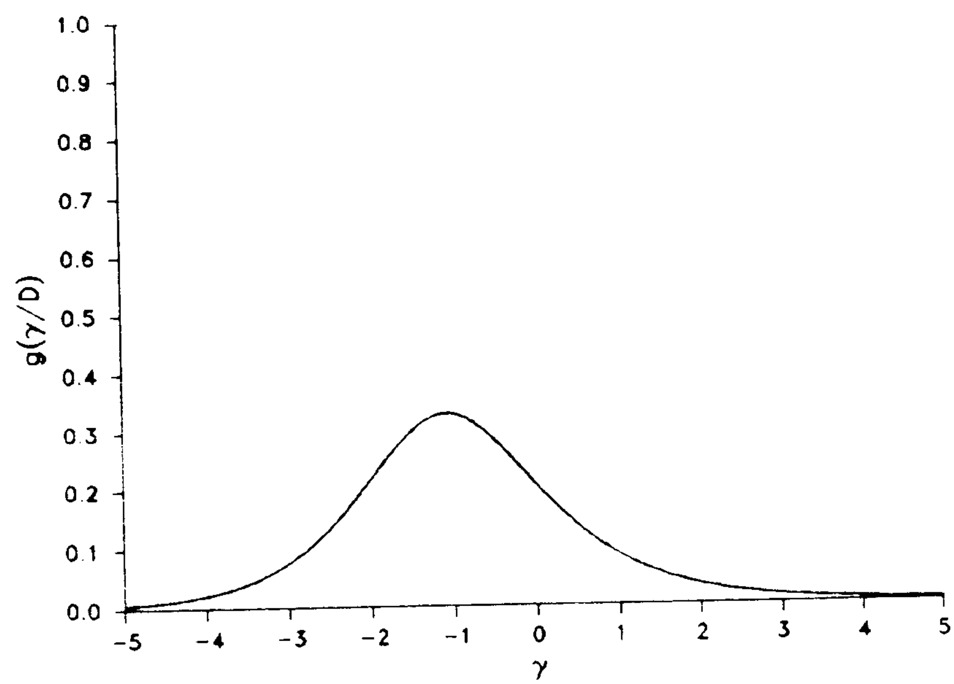

FIGURE 4a—Posterior Density for γ: $T=5$, $\gamma=-0.09$.

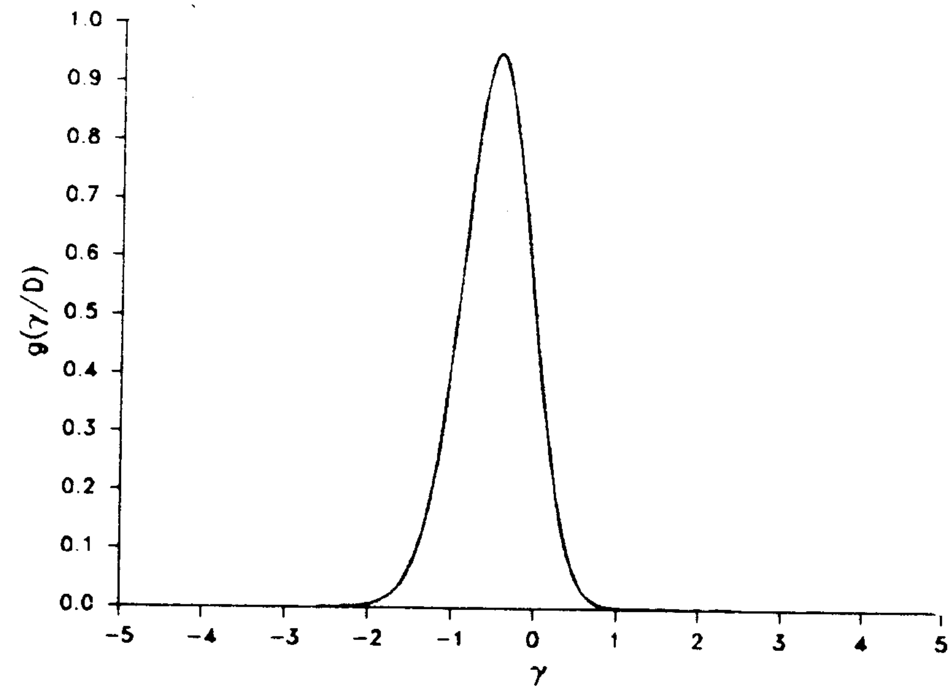

FIGURE 4b—Posterior Density for γ: $T=10$, $\gamma=-0.09$.

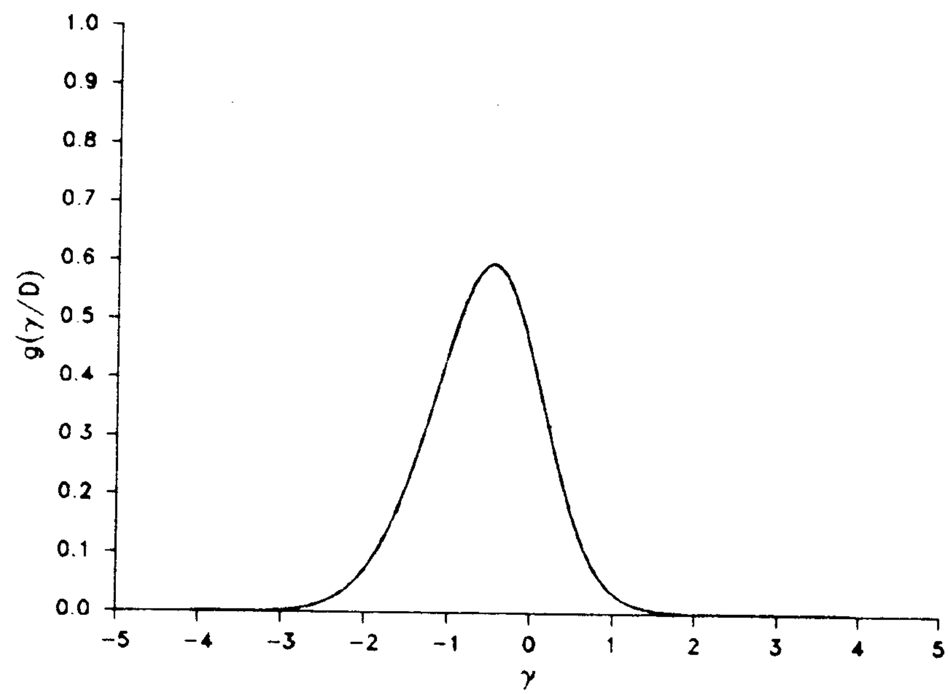

FIGURE 4c—Posterior Density for γ: $T=20$, $\gamma=-0.09$.

First, with respect to the spread of the distributions, it is very clear how the information about γ becomes more precise at T gets larger. However, despite the impression given by the horizontal scale, even for $T=20$ the distribution is still not extremely informative. It indicates that γ could reasonably be anywhere between -1.45 and 0.4, values over which the optimal solution for x can change drastically (with parameter certainty). The second point to note is the tendency for γ to be underestimated. All three distributions are centred to the left of the true value of -0.09; the means for $T=20$, $T=10$ and $T=5$ were -0.50, -0.57 and -0.85, respectively. When the posterior densities for other values of γ were considered, the posterior means for $T=10$ and $T=20$ did not become positive until $\gamma \geqq 0.45$. The final point to note is with respect to the posterior density when $T=5$. This density function is extremely non-informative. Also, although the one presented in Figure 4a appears well behaved, functions computed for other values of γ were sometimes bimodal and trimodal, with spikes quite distant from the main centre of the distribution. It is clear that it is very difficult to estimate the model when only a few observations are available.

The optimal solutions for $\phi=0.0005$, 0.00005 and $T=10,20$ are graphed in Figures 5a, 5b, 5c and 5d for $-1<\gamma<0.5$. Two solutions are given for each case: one is a traditional solution obtained at a set of point estimates and the other is the solution from the Bayesian framework. For the traditional solution, the point estimates used were the posterior means for α, β and γ and the conditional posterior mean for σ^2, evaluated at the posterior means of β and γ. It was impossible to use the unconditional posterior mean for σ^2 because this mean does not exist.

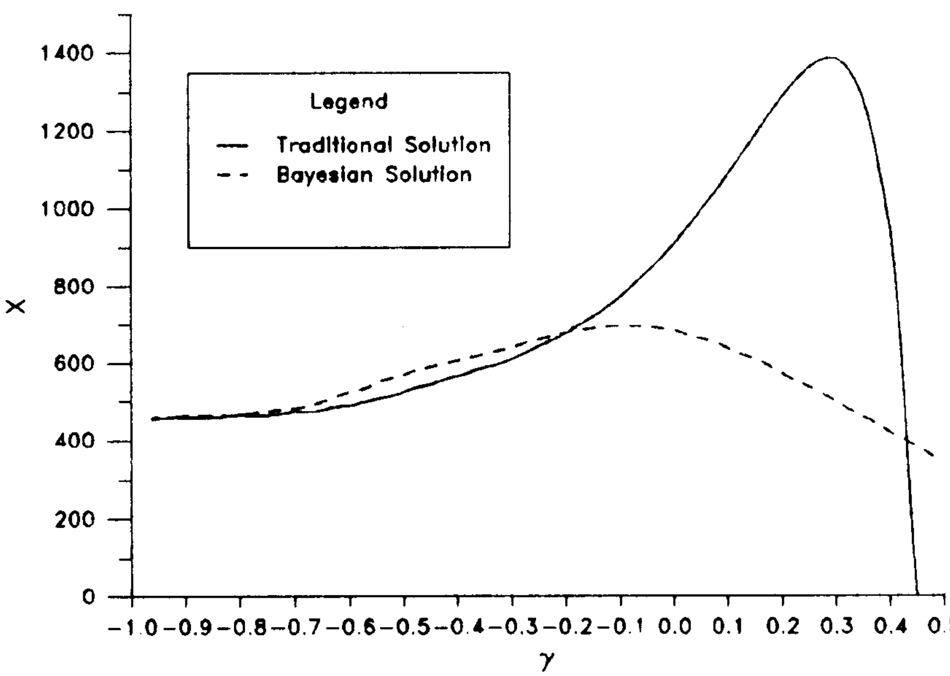

FIGURE 5a—Optimal Input Levels: $T=10$, $\phi=0.0005$.

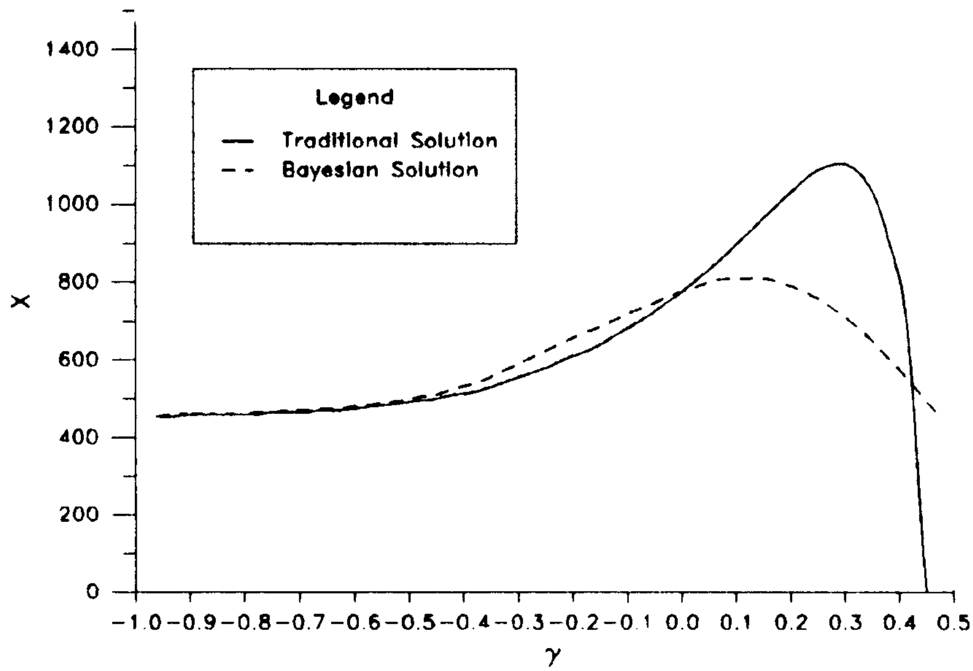

FIGURE 5b—Optimal Input Levels: $T=20$, $\phi=0.0005$.

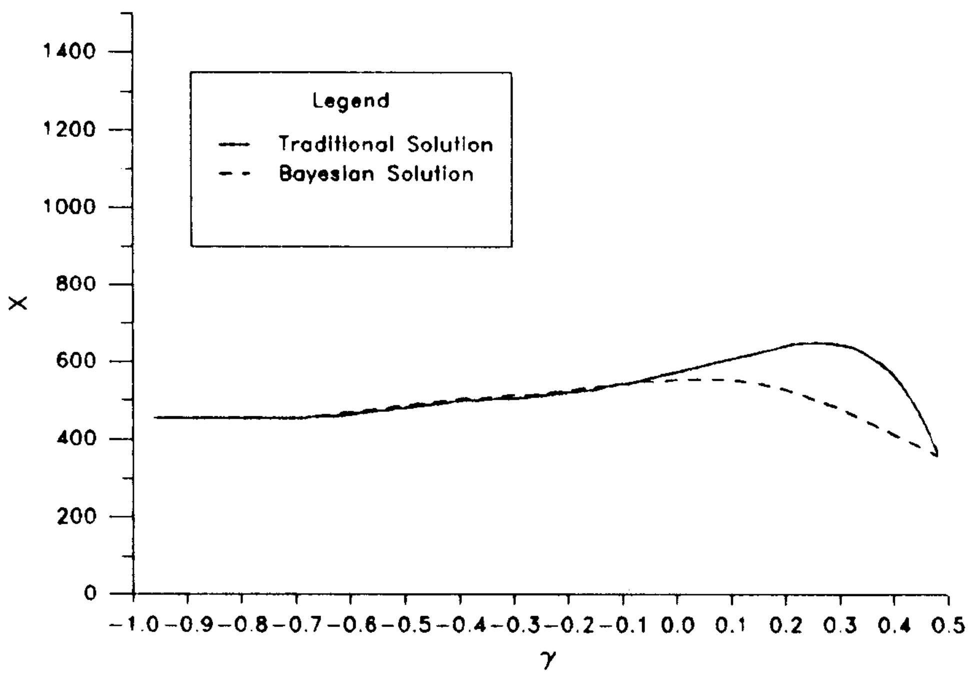

FIGURE 5c—Optimal Input Levels: $T=10$, $\phi=0.00005$.

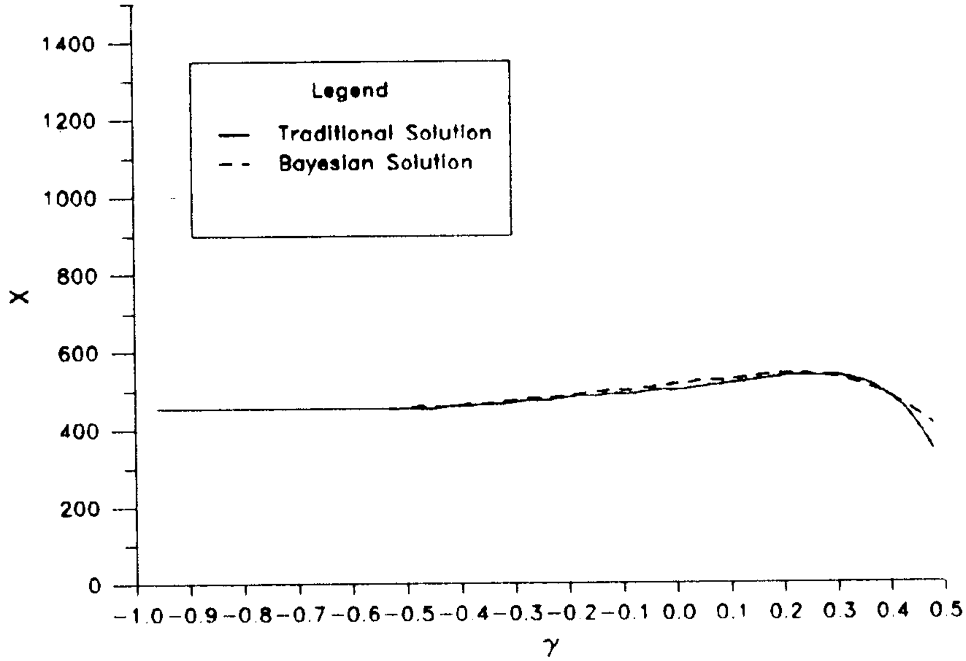

FIGURE 5d—Optimal Input Levels: $T=20$, $\phi=0.00005$.

The integral defining this mean diverges as $\gamma \to -\infty$. However, all integrals required for the Bayesian solution converge. Wherever $\hat{\sigma}^2$ appears in (45) and (46) there is a compensating term which depends on γ and which leads to convergence.

In Figures 5a and 5b where $\phi=0.0005$ (relatively large), the optimal solution from the traditional approach becomes very high as γ increases, reaches a maximum at $\gamma=0.3$, and then begins to decline. The difference between this graph and that in Figure 2 for the parameter certainty case is dramatic. The cause of this behaviour can be traced to the underestimation of γ. Up to $\gamma=0.42$, the posterior mean for γ is always negative. Thus, this estimate for γ suggests that x should be increased to gain a risk reducing effect, even when $\gamma>0$. Furthermore, a negative estimate of γ tends to lead to a large estimate of σ^2 (tied in with the fact that the conditional posterior mean for σ^2 approaches infinity as $\gamma \to -\infty$); this large estimate of σ^2 reinforces the risk reducing effect from increasing x. The optimal solution's dramatic drop to zero at $\gamma=0.45$ occurs because this is the point at which the posterior mean for γ becomes positive.

When $\phi=0.00005$, the traditional solution changes in a similar way to that for $\phi=0.0005$, but the effect is not as pronounced because the variance term carries less weight. Also, when the posterior mean for γ becomes positive, the optimal solution falls, but not to zero.

Turning to the Bayesian results, it can be seen that when $\phi=0.00005$ there is little difference between the two solutions, the greatest difference being when $T=10$ and $\gamma>0$. It appears that if T is sufficiently large to obtain reasonable estimates of the parameters, and ϕ is sufficiently small for the parameter uncertainty to carry little weight, the traditional approach is not far from optimal.

When $\phi = 0.0005$ the difference between the traditional and Bayesian solutions is quite large. The Bayesian solution is much less when $\gamma > 0$, reflecting the fact that positive values of γ carry some weight in the computation of $V(y|D)$, even when the posterior mean for γ is negative. Finally, note that $x = 0$ is never an optimal solution in the Bayesian framework. As $x \rightarrow 0$, $V(y|D) \rightarrow \infty$, because negative values of γ carry some weight in the computation of the last term in equation (46). Thus, it is always optimal to set x at some positive value to increase $E(U|D)$ away from $-\infty$.

Conclusions

A number of important conclusions have emerged from this study, some of which did not arise from the original objectives, but were discovered incidentally during the research process. First, the Just–Pope type production function has some undesirable properties which appear to have been overlooked. In particular, when the marginal risk of an input is negative and the level of that input approaches zero, output becomes infinite (either positive or negative). Also, a counter-intuitive result which occurs is that, for large negative values of the risk parameter γ, the optimal input level for a risk averse producer is identical to the risk free level. A second incidental conclusion concerns the lack of robustness of optimal input levels to the specifications of expected utility and the probability distribution of output. Simply changing an assumption of normally distributed output to one where output follows a t-distribution has drastic implications for the maximisation of expected utility and for the compatibility of utility function specifications with distributional assumptions. Difficulties occur because expected utility sometimes becomes infinite (or is not defined).

The convenience of the Bayesian framework for incorporating parameter uncertainty into input allocation has been demonstrated. It might be expected that allowing for parameter uncertainty would decrease the optimum input level for a risk averse producer. However, such is not always the case. Whether or not parameter uncertainty leads to a lower input level will depend on the model specification and on the settings of past input levels (and the resulting outputs) which have supplied the uncertain information about the parameters. If the degree of risk aversion is very small and the number of past observations large, then, as expected, there is little difference between the Bayesian solution and the traditional solution which is based on point estimates of the parameters. However, a moderate degree of risk aversion, or a moderate sample size, can cause a marked divergence. In particular, point estimation of the risk parameter can be extremely unreliable, in which case the traditional solution can be markedly different from that obtained under parameter certainty. Because the Bayesian solution uses the complete posterior density of the risk parameter, it includes information on the unreliability of any estimate and does not yield a completely unrealistic result.

A number of questions or problems which suggest possible directions for future research have been uncovered. One direction is to examine alternative production functions and alternative distributional assumptions in conjunction with different classes of utility functions. It would be desirable to develop an overall set of assumptions which are mutally compatible and which lead to robust results. Other directions are created by relaxing some of the restrictive assumptions of the analysis, such as a non-stochastic price, and the presence of only a single-variable input. Also, maximisation of expected utility has been considered only within a time horizon of one period. If the analysis is extended to more than one period then other considerations, such as setting x in one period to achieve low parameter variance in subsequent periods, become important. Finally, another question which could be examined is the loss in expected utility from ignoring the parameter uncertainty and using the traditional solution. In this paper the differences in optimal input levels have been considered without any regard for how the maximum values of the objective function differ.

APPENDIX 1

Posterior and Predictive Moments with Zero Marginal Risk

The joint posterior density function for $(\alpha, \beta, \sigma^2)$ for the special case where $\gamma = 0$ was given in (19) as:

(A1) $g(\alpha, \beta, \sigma^2|D) \propto (\sigma^2)^{-T/2-1} \exp\{-\Sigma(y_t - \alpha x_t^\beta)^2/2\sigma^2\}$

From (A1) it is possible to complete the square on α and use properties of the normal distribution to integrate it out, yielding:

(A2) $g(\beta, \sigma^2|D) \propto (\sigma^2)^{-(T+1)/2} (\Sigma x_t^{2\beta})^{-0.5} \exp\{-(T-3)\hat{\sigma}^2/2\sigma^2\}$

where

(A3) $(T-3)\hat{\sigma}^2 = \Sigma y_t^2 - (\Sigma y_t x_t^\beta)^2/\Sigma x_t^{2\beta}$

The marginal posterior density function for β can be obtained by using properties of the gamma function to integrate σ^2 out of (A2). This procedure yields:

(A4) $g(\beta|D) \propto (\Sigma x_t^{2\beta})^{-0.5} [(T-3)\hat{\sigma}^2]^{-(T-1)/2}$

Also of interest is the joint density for (α,β) obtained by integrating σ^2 out of (A1). It is given by:

(A5) $g(\alpha, \beta|D) \propto [(T-3)\hat{\sigma}^2 + \Sigma x_t^{2\beta}(\alpha - \hat{\alpha})^2]^{-T/2}$

where

(A6) $\hat{\alpha} = \Sigma y_t x_t^\beta / \Sigma x_t^{2\beta}$

The conditional posterior means for σ^2 and α given β are useful for the computation of expected utility. By viewing (A2) as conditional on β and using properties of the inverted gamma distribution (Zellner 1971) it can be shown that:

$$\text{(A7)} \quad E(\sigma^2 | \beta, D) = \hat{\sigma}^2$$

From (A5) the posterior density $g(\alpha | \beta, D)$ is a t-distribution with mean

$$\text{(A8)} \quad E(\alpha | \beta, D) = \hat{\alpha}$$

and variance

$$\text{(A9)} \quad V(\alpha | \beta, D) = \hat{\sigma}_\alpha^2 = \hat{\sigma}^2 / \Sigma x_t^{2\beta}$$

Next, consider the joint density function $g(u, \alpha | \beta, D)$ where u is the disturbance for the next time period. This density function is useful for finding the moments of the predictive density $g(u | \beta, D)$. First, note that the joint posterior density function in (A1) can be rewritten as:

$$\text{(A10)} \quad g(\alpha, \beta, \sigma^2 | D) \propto (\sigma^2)^{-T/2-1} \exp\{-(\Sigma x_t^{2\beta}/2\sigma^2)(\alpha - \hat{\alpha})^2\} \exp\{-(T-3)\hat{\sigma}^2/2\sigma^2\}$$

and the density function for u conditional on the parameters is:

$$\text{(A11)} \quad g(u | \sigma^2) \propto (\sigma^2)^{-0.5} \exp\{-u^2/2\sigma^2\}$$

Then,

$$\text{(A12)} \quad g(\alpha, u, \beta | D) = \int_0^\infty g(\alpha, \beta, \sigma^2 | D) \cdot g(u | \sigma^2) d\sigma^2$$

Carrying out the integration in (A12) and conditioning on β yields:

$$\text{(A13)} \quad g(\alpha, u | \beta, D) \propto \left[(T-3)\hat{\sigma}^2 + \begin{pmatrix} \alpha - \hat{\alpha} \\ u \end{pmatrix}' \begin{pmatrix} \Sigma x_t^{2\beta} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha - \hat{\alpha} \\ u \end{pmatrix}\right]^{-(T+1)/2}$$

This is a bivariate t-distribution with $(T-1)$ degrees of freedom and covariance matrix:

$$\begin{pmatrix} \hat{\sigma}^2/\Sigma x_t^{2\beta} & 0 \\ 0 & \hat{\sigma}^2 \end{pmatrix}$$

It can be seen that $(u | \beta, D)$ and $(\alpha | \beta, D)$ are uncorrelated (but not independent) with the following moments for u:

$$\text{(A14)} \quad E(u | \beta, D) = E(u) = 0$$

(A15) $V(u|\beta,D)=\hat{\sigma}^2$

(A16) $\mathrm{Cov}(u,\alpha|\beta,D)=\mathrm{Cov}(u,\alpha)=0$

The moments for α agree (A13) agree (of course) with the results in (A8) and (A9).

APPENDIX 2

Posterior and Predictive Moments with Non-zero Marginal Risk

When $\gamma\neq 0$ and a uniform prior density for $g(\gamma)$ is employed, the joint posterior density for all parameters is given by:

(A17) $g(\alpha,\beta,\gamma,\sigma^2|D)\propto$
$$(\sigma^2)^{-T/2-1}\,\Pi x_t^{-\gamma/2}\exp\{-\Sigma[(y_t-\alpha x_t^{\beta})/x_t^{\gamma/2}]^2/2\sigma^2\}\propto$$

$$(\sigma^2)^{-T/2-1}\,\Pi x_t^{-\gamma/2}\exp\{-(\Sigma x_t^{2\beta-\gamma}/2\sigma^2)(\alpha-\hat{\alpha})^2\}\exp\{-(T-3)\hat{\sigma}^2/2\sigma^2\}$$

where

(A18) $\hat{\alpha}=\Sigma y_t x_t^{\beta-\gamma}/\Sigma x_t^{2\beta-\gamma}$

and

(A19) $(T-3)\hat{\sigma}^2=\Sigma(y_t^2/x_t^{\gamma})-\hat{\alpha}\Sigma y_t x_t^{\beta-\gamma}$

It should be kept in mind that the definitions of $\hat{\alpha}$ and $\hat{\sigma}^2$ are different from those in Appendix 1, now that γ has been included. Also, $\hat{\alpha}$ and $\hat{\sigma}^2$ will now depend on both β and γ.

It is possible to integrate analytically α and σ^2 out of (A17). Carrying out this process yields:

(A20) $g(\beta,\gamma,\sigma^2|D)\propto$
$$(\sigma^2)^{-(T+1)/2}(\Sigma x_t^{2\beta-\gamma})^{-0.5}\,\Pi x_t^{-\gamma/2}\exp\{-(T-3)\hat{\sigma}^2/2\sigma^2\}$$

(A21) $g(\alpha,\beta,\gamma|D)\propto \Pi x_t^{-\gamma/2}[(T-3)\hat{\sigma}^2+\Sigma x_t^{2\beta-\gamma}(\alpha-\hat{\alpha})^2]^{-T/2}$

(A22) $g(\beta,\gamma|D)\propto(\Sigma x_t^{2\beta-\gamma})-0.5\Pi x_t^{-\gamma/2}[(T-3)\hat{\sigma}^2]^{-(T-1)/2}$

By viewing (A20) and (A21) as conditional on (β,γ), it is possible to show that:

(A23) $E(\alpha|\beta,\gamma,D)=\hat{\alpha}$

(A24) $V(\alpha|\beta,\gamma,D)=\hat{\sigma}_\alpha^2=\hat{\sigma}^2/\Sigma x_t^{2\beta-\gamma}$

and

(A25) $E(\sigma^2|\beta,\gamma,D)=\hat{\sigma}^2$

These quantities will be useful for computing expected utility. Further analysis of $g(\beta,\gamma|D)$ in (A22) must be done using bivariate numerical integration techniques.

It is also useful to derive the joint density function $g(\alpha,u|\beta,\gamma,D)$. Following a similar line of reasoning to that in Appendix 1 it can be shown that:

$$\text{(A26)} \quad g(\alpha,u|\beta,\gamma,D) \propto \left[(T-3)\hat{\sigma}^2 + \begin{pmatrix}\alpha-\hat{\alpha}\\ u\end{pmatrix}' \begin{pmatrix}\Sigma x_t^{2\beta-\gamma} & 0\\ 0 & 1\end{pmatrix} \begin{pmatrix}\alpha-\hat{\alpha}\\ u\end{pmatrix}\right]^{-(T+1)/2}$$

The useful information from this density function is:

$$\text{(A27)} \quad E(u|\beta,\gamma,D) = E(u) = 0$$

$$\text{(A28)} \quad E(u^2|\beta,\gamma,D) = \hat{\sigma}^2$$

$$\text{(A29)} \quad \mathrm{Cov}(u,\alpha|\beta,\gamma,D) = \mathrm{Cov}(u,\alpha) = 0$$

References

Anderson, J. R., Dillon, J. L. and Hardaker, J. B. (1977), *Agricultural Decision Analysis*, Iowa State University Press, Ames.

Anderson, J. R. and Griffiths, W. E. (1981), 'Production risk and input use: Pastoral Zone of eastern Australia', *Australian Journal of Agricultural Economics* 25(2), 149–59.

—— (1982), 'Production risk and efficient allocation of resources', *Australian Journal of Agricultural Economics* 26(3), 226–32.

Griffiths, W. E. and Anderson, J. R. (1982), 'Using time-series and cross-section data to estimate a production function with positive and negative marginal risk', *Journal of the American Statistical Association* 77(379), 529–36.

Just, R. E. and Pope, R. D. (1978), 'Stochastic specification of production functions and economic implications', *Journal of Econometrics* 7(1), 67–86.

Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons, New York.