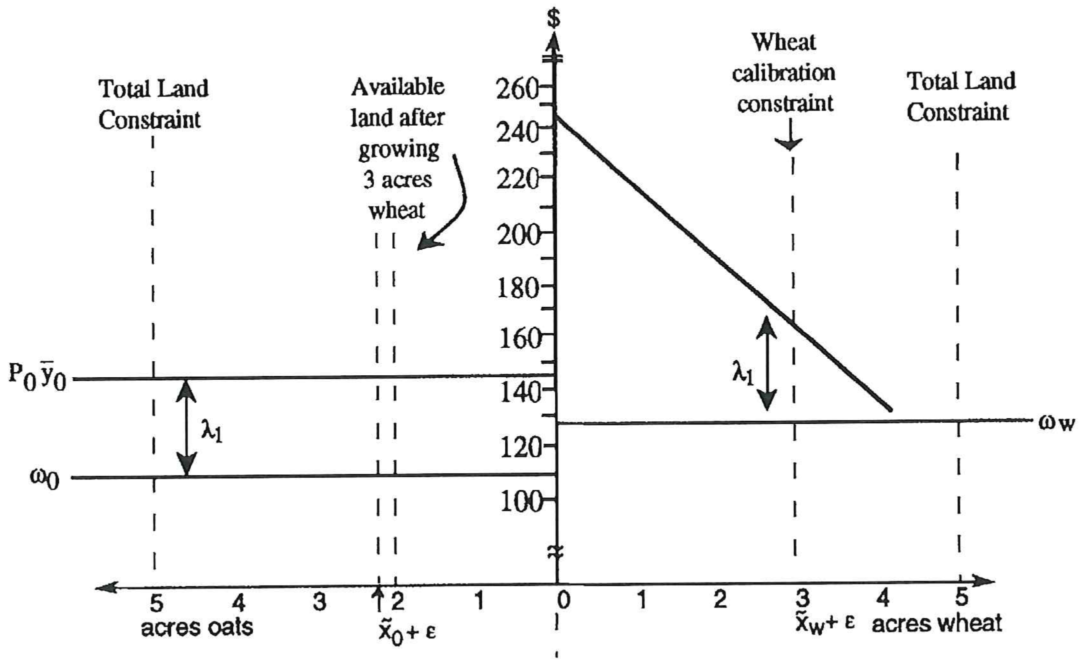

Figure 2. PMP yield function on wheat

ence, PMP can be thought of as revealed efficiency based on observed land allocations.

Equation (12) substantiates the dual values shown in figure 1, where the duals for the calibration constraint set (λ_2) in the stage I problem are equal to the divergence between the LP average net value product per acre and opportunity cost per acre. Since the value λ_2 represents the difference between VAP and VMP for the more profitable crops, and given the linear yield function in (8), a single element of λ_2 can be expressed as

$$(13) \quad \lambda_{2i} = P_i(\beta_i - \delta_i x_i) - P_i(\beta_i - 2\delta_i x_i) = P_i \delta_i x.$$

Using (13) the yield slope coefficient can be solved as

$$(14) \quad \delta_i = \frac{\lambda_{2i}}{P_i x_i}.$$

Using equation (10) the intercept coefficient (β_i) for crop i can be solved in terms of δ_i and $\bar{y}_i$.

Despite all the notation, the basic concept of PMP is numerically simple and easy to solve automatically, even on desktop computers. A numerical example applied to the problem in equation (1) and figures 1 and 2 demonstrates this simplicity.

The problems shown in figures 1 and 2 have a single land constraint (500 acres) and two crops, wheat and oats. The following parameters are used:

	Wheat (w)	(Oats) (o)
Crop prices	P_w = \$2.98/bu	P_0 = \$2.20/bu
Variable cost/acre	ω_w = \$129.62	ω_0 = \$109.98
Average yield/acre	$\bar{y}_w$ = 69 bu	$\bar{y}_0$ = 65.9 bu

The observed acreage allocation in the base year is 300 acres of wheat and 200 acres of oats. The problem in figure 1 is

$$(15) \quad \max\ (2.98 * 69 - 130)x_w + (2.20 * 65.9 - 110)x_0$$

subject to

$$\begin{array}{llll} (i) & x_w + x_0 & \leq & 500 \\ (ii) & x_w & \leq & 300.01 \\ (iii) & x_0 & \leq & 200.01 \end{array}$$

Note the addition of the ε perturbation term (0.01) on the right-hand side of the calibration constraints. The average gross margin from

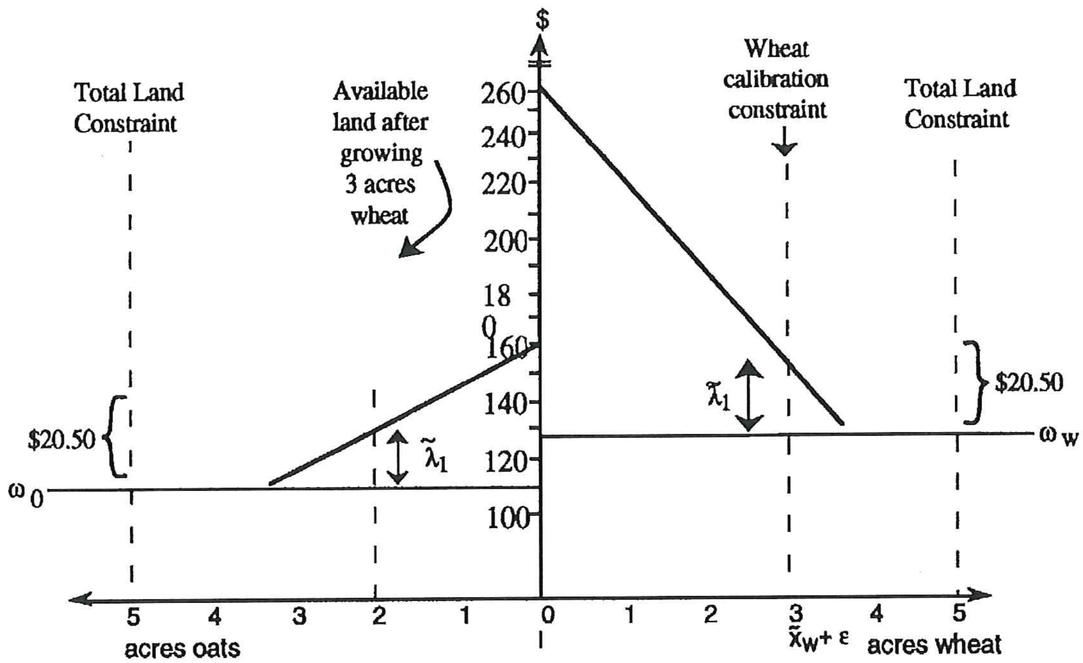

Figure 3. PMP model—quadratic yields on all crops

wheat is \$76/acre and for oats is \$35/acre. The optimal solution to the stage 1 problem (1) is when the wheat calibration constraint is binding at a value of 300.01 and constraint (i) is binding when the oat acreage equals 199.99. The oat calibration constraint is slack.

The dual value on land (λ_1) is \$35 and on the two calibration constraints (λ_2) = [41 and 0]. Using equation (14), the λ_2 value for wheat and the base-year data, the yield function slope for wheat is calculated as

$$(16)\quad \delta_w = 41/(2.98*300.01) = 0.04586.$$

Quantity δ_w is now substituted into equation (10) to calculate the yield slope intercept β_w

$$(17)\quad \beta_w = 69 + (0.04586*300.01) = 82.76.$$

Using the yield function parameters, the Stage II primal PMP problem becomes (see figure 2)

$$(18)\quad \max\ [2.98(82.76 - 0.04586*x_w) - 130]x_w + (2.20*65.9 - 110)x_0$$

subject to

$$x_w + x_0 \leq 500.$$

A quick empirical check of the calibration of problem (18) to the base values can be performed by calculating the VMP of wheat at 300 acres. If it is close to the VMP (VAP) of oats and convergent, the model will calibrate without additional calibration constraints.

The marginal yield per acre of wheat is

$$y|_{300} = 82.76 - 2*0.04586*300 = 55.25$$
$$\mathrm{VMP}_{w,300} = 2.98*55.25 - 130 = 34.65$$

The VMP for wheat at 300 acres of \$34.65 is marginally below the VMP for oats (\$35). Thus, the unconstrained PMP model will calibrate within the rounding error of this example.

This numerical example shows that PMP models can be calibrated using simple methods. The three-stage process and calculation of the parameters is easily programmable as a single process using GAMS/MINOS.[4] Thus, given the initial data and specifications, the PMP model is automatically calibrated in the time it takes to solve an LP and QP solution for the model.

The PMP model specified in (18) calibrates in all aspects. That is, the optimal solution, binding constraints, objective function value

[4] A PMP program written for the GAMS/MINOS optimization package is available from the author by e-mail (rehowitt@ucdavis.edu). The program can be used to automatically calibrate and run a range of agricultural production problems by PMP.

and dual values will all be within rounding error of the original LP in (15) that is constrained by the calibration constraints.

A valid objection to the simple PMP specification in (15) is that we assume a decreasing yield/acre function for the more profitable unconstrained crops $\mathbf{x_N}$, but the crop set $\mathbf{x_B}$ that is constrained by resources is assumed to have constant yields.

Calibrating the marginal crops ($\mathbf{x_B}$) with decreasing yield functions requires additional empirical information. The independent variables, as $\mathbf{x_N}$ are termed, use both the constrained resource opportunity cost ($\boldsymbol{\lambda}_l$) and their own calibration dual ($\boldsymbol{\lambda}_2$) (figure 1) to solve for the yield function parameters implied by the observed crop allocations. However, the marginal crops ($\mathbf{x_B}$) have no binding calibration constraint, and thus cannot empirically differentiate marginal and average yield of the observed calibration acreage using the minimal LP data set specified.

Clearly some additional data are needed. The simplest source of additional data are measurements on the expected yield variation of the marginal crops ($\mathbf{x_B}$) within a given region and year. Regional acreage response elasticities would supply the equivalent information, but it would seem that yield variation is an easier empirical value to obtain from farmers, particularly if it is simplified into percentage deviations above and below the mean yields in the region.

Returning to the simple pedagogical example in equation (15) and figure 2, the stage 1 calibrated problem is run exactly as before. One of the important pieces of information from the optimal solution of the stage 1 problem is the activities which are in the $\mathbf{x_N}$ and $\mathbf{x_B}$ groups. This information is unlikely to be known beforehand.

In the example, assume that the a priori information on oats is that expected yield variation is plus or minus 10% of the mean. The reduced marginal yield information now causes a recalculation of the opportunity cost of land. Given an average yield ($\bar{y}_0$) for oats of 65.9 bu/acre and a price of \$2.20, the marginal return given 10% yield reduction will now be based on a yield of \$59.31 bu/acre; therefore, the dual value on land (11) is reduced by \$14.50 to \$20.50. The PMP dual (λ_2) must also be increased by this same amount to ensure the first-order conditions (12) hold. The new value for $\lambda_2 = \$55.50$.

The calculations for the yield coefficients in (16) and (17) are now applied to all activities, both marginal (x_B) and independent (x_N). Note that the adjusted λ_2 values are used for the independent activities, and the MVP based on the prior data is used for the marginal crops.

The PMP problem, given the information on marginal yields for the oat crop, is

$$
(19)\quad \max\ [2.98(87.63 - 0.0621 * x_w) - 130]x_w + [2.20(72.49 - 0.0329 * x_0) - 110]x_0
$$

subject to

$$
x_w + x_0 \le 500.
$$

The problem is shown in figure 3. The calibration acreage can be checked by calculating the VMP for each crop at the calibration acreages of $\tilde{x}_w = 300$ and $\tilde{x}_0 = 200$.

$$
(20)\quad (i)\ VMP_w\big|_{\tilde{x}_{w=300}} = 2.98 * 50.37 - 130 = 20.10
$$

$$
(ii)\ VMP_0\big|_{\tilde{x}_{0=200}} = 2.20 * 59.33 - 110 = 20.53
$$

With the VMP's equal, aside from rounding error, the PMP with endogenous yield functions will calibrate arbitrarily close to the base-year acreages.

The resulting model will calibrate acreage allocation and input use, and the objective function value precisely. However, the dual value on resources will be lower reflecting the additional, and presumably more accurate, data on the yield variation among the marginal crops.

Policy Modeling with PMP

The purpose of most programming models is to analyze the impact of quantitative policy scenarios which take the form of changes in prices, technology, or constraints on the system. The policy response of the model can be characterized by its response to sensitivity analysis and changes in constraints.

Advantages of the PMP specification are not only the automatic calibration feature, but also its ability to respond smoothly to policy scenarios. Paris shows that input demand functions and output supply functions obtained by parameterizing a PMP problem satisfy the Hicksian conditions for the competitive firm. In addition, the input demand and supply functions are continuous and differentiable with respect to prices, costs, and right-hand side quantities. At the point of a change in basis, the supply and demand functions are not differentiable. This is in contrast to LP or stepwise problems, where the dual values, and sometimes the optimal solution, are unchanged by parameterization until there is a discrete change in basis, when they jump discontinuously to a new level.

The ability to represent policies by constraint structures is important. The PMP formulation has the property that the nonlinear calibration can take place at any level of aggregation. That is, one can nest an LP subcomponent within thc quadratic objective function and obtain the optimum solution to the full problem. An example of this is used in technology selection where a specification that causes discrete choices may be appropriate. Suppose a given regional commodity can be produced by a combination of five alternative linear technologies, whose aggregate output has a common supply function. The PMP can calibrate the supply function while a nested LP problem selects the optimal set of linear technology levels that make up the aggregate supply (Hatchett, Horner, and Howitt).

Since the intersection of the convex sets of constraints for the main problem and the convex nested subproblem is itself convex, then the optimal solution to the nested LP subproblem will be unchanged when the main problem is calibrated by replacing the calibration constraints with quadratic PMP cost functions. The calibrating functions can thus be introduced at any level of the linear model. In some cases, the available data on base-year values will dictate the calibration level. Ideally, the level of calibration would be determined by the properties of the production functions, as in the example of linear irrigation technology selection. The PMP approach does not replace all linear cost functions with equivalent quadratic specifications, but only replaces those that data or theory suggest are best modeled as nonlinear.

If one has prior information on the nature of yield externalities and rotational effects between crops, they can be explicitly incorporated by specifying cross-crop yield interaction coefficients in equations (13) and (14). The PMP yield slope coefficient matrix is positive definite, $k \times k$, and has rank k. Without the cross-crop effects the matrix is diagonal.

Resource-using activities such as fodder crops consumed on the farm may be specified with zero valued objective function coefficients. Where an activity is not resource-using, but merely acts as a transfer between other activities, there is no empirical basis or need to modify the objective function coefficients.

Conclusions

Programming models have a strong role to play in agricultural policy analysis, particularly where reliable time-series data are absent, or shifts in market institutions or constraints have changed substantially over time. The problem addressed in this paper is one of calibrating programming models without adding constraints that cannot be justified by economic theory or agricultural technology. The solution proposed by the PMP approach is based on the derivation of nonlinear yield functions from the base-year data and prior crop yield data. The derivation is achieved by a simple three-step procedure.

Calibration of a model to the base-year data set and constraints is a necessary, but not sufficient, condition for a meaningful policy model. The ultimate test of a policy model is its ability to predict behavioral responses out of the sample base-year. If the yield response functions calibrated in the PMP method have a basis in regional soil variation and farmer behavior, then they should be relatively stable over time and can provide additional structural information for policy response. Empirical tests of the stability of the PMP values are required to evaluate the stability of the calibrated models. Initial tests in Kasnakoglu and Bauer are encouraging.

The PMP approach is shown to satisfy the main criteria for calibrating sectoral and regional models. Using PMP, the model calibrates precisely to output and input quantities, the objective function value, dual constraint values, and output prices. In addition, the PMP approach can incorporate priors on yield variability or supply elasticities.

[Received July 1991;
final revision received November 1994.]

References

Bauer, S., and H. Kasnakoglu. "Non Linear Programming Models for Sector Policy Analysis." *Econ. Model.* (1990):275–90.

Day, R.H. "Recursive Programming and the Production of Supply." *Agricultural Supply Functions.* Heady et al., eds. Iowa State University Press, 1961.

Hatchett, S.A., G.L. Horner, and R.E. Howitt. "A

Regional Mathematical Programming Model to Assess Drainage Control Policies." *The Economics and Management of Water and Drainage in Agriculture*. A. Dinar and D. Zilberman, eds. pp. 465–89. Boston: Kluwer, 1991.

Hazell P.B.R., and R.D. Norton. *Mathematical Programming for Economic Analysis in Agriculture*. New York: Macmillan, 1986.

Horner, G.L., J. Corman, R.E. Howitt, C.A. Carter, and R.J. MacGregor. *The Canadian Regional Agriculture Model: Structure, Operation and Development*. Agriculture Canada, Technical Report 1/92, Ottawa, October 1992.

House, R.M. *USMP Regional Agricultural Model*. Washington DC: U.S. Department of Agriculture. National Economics Division Report, ERS, 30 pp., July 1987.

Howitt, R.E. "Calibration Methods for Agricultural Economic Production Models." *J. Agr. Econ.*, in press.

Just, R.E. "Discovering Production and Supply Relationships: Present Status and Future Opportunities." *Rev. Market. and Agr. Econ.* 61(1993):11–40.

Just, R.E., D. Zilberman, and E. Hochman. "Estimation of Multicrop Production Functions." *Amer. J. Agr. Econ.* 65(November 1983):770–80.

Kasnakoglu, H., and S. Bauer. "Concept and Application of an Agricultural Sector Model for Policy Analysis in Turkey." *Agricultural Sector Modelling*. S. Bauer and W. Henrichsmeyer, eds. Wissenschaftsuerlag: Vauk-Kiel, 1988

Luenberger, D.G. *Linear and Nonlinear Programming*. Reading MA: Addison-Wesley, 1984.

McCarl, B.A. "Cropping Activities in Agricultural Sector Models: A Methodological Proposal." *Amer. J. Agr. Econ.* 64(November 1982):768–71.

Meister, A.D., C.C. Chen, and E.O. Heady. *Quadratic Programming Models Applied to Agricultural Policies*. Ames IA: Iowa State University Press, 1978.

Oamek, G., and S.R. Johnson. "Economic and Environmental Impacts of a Large Scale Water Transfer in the Colorado River Basin." Paper presented at the WAEA annual meeting, Honolulu HI, 10-12 July 1988.

Paris. Q. "PQP, PMP, Parametric Programming, and Comparative Statics." Chap. 11 in "Notes for AE253." Dept. Agr. Econ., University of California, Davis, November 1993.

Paris, Q., and K. Knapp. "Estimation of von Liebig Response Functions." *Amer. J. Agr. Econ.* 71(February 1989):178–86.

Peach, T. *Interpreting Ricardo*. Cambridge UK: Cambridge University Press, 1993.

Quinby, B., and D.J. Leuck. "Analysis of Selected E. C. Agricultural Policies and Dutch Feed Composition Using Positive Mathematical Programming." Paper presented at AAEA annual meeting, Knoxville, TN, 31 July 1988.

Ribaudo, M.O., C.T. Osborn, and K. Konyar. "Land Retirement as a Tool for Reducing Agricultural Nonpoint Source Pollution." *Land Econ.* 70(February 1994):77–87.

Rosen, M.D., and R.J. Sexton. "Irrigation Districts and Water Markets: An Application of Cooperative Decision-Making Theory." *Land Econ.* 69(February 1993):39–53.

Appendix A

Proposition 1. *Given an agent maximizing multi-output profit subject to linear constraints on some inputs or outputs, if the number of nonzero nondegenerate production activity levels observed (k) exceeds the number of binding constraints (m), then a necessary and sufficient condition for profit maximization at the observed levels is that the profit function be nonlinear (in output) in some of the (k) production activities.*

Proof. Define the profit function in general as a function of input allocation $\mathbf{x}$, $\mathbf{f}(\mathbf{x})$.

(a1) problem is max $\mathbf{f}(\bar{\mathbf{x}})$

subject to

$$\bar{\mathbf{A}}\bar{\mathbf{x}} \leq \mathbf{b} \quad \bar{\mathbf{x}} = n \times 1$$
$$\bar{\mathbf{A}} = m \times n \quad m < n$$

At the observed optimal solution (nondegenerate in primal and dual specifications) there are k non-zero values of $\bar{\mathbf{x}}$. Drop the zero values of $\bar{\mathbf{x}}$ and define the $m \times m$ basic partition of $\mathbf{A}$ as the $(m \times m)$ optimal solution basis matrix $\mathbf{B}$ and the remaining partition of $\mathbf{A}$ as $\mathbf{N}$ $(m \times k - m)$. Partitioning the $k \times 1$ vector x into the $m \times 1$ vector $\mathbf{x_B}$ and $(k - m)$ x1 vector $\mathbf{x_N}$, the problem (a1) is written as

$$\text{(a2)} \quad \max \mathbf{f}(\mathbf{x}) \qquad \text{subject to } [\mathbf{B} : \mathbf{N}]\begin{bmatrix}\mathbf{x_B}\\ \mathbf{x_N}\end{bmatrix} = \mathbf{b}$$

or

$$\text{(a3)} \quad \max \mathbf{f}(\mathbf{x_B}, \mathbf{x_N}) \qquad \text{subject to } \mathbf{Bx_B} + \mathbf{Nx_N} = \mathbf{b}$$

Given the constraint set in (a3), $\mathbf{x_B}$ can be written

$$\text{(a4)} \quad \mathbf{x_B} = \mathbf{B^{-1}b} - \mathbf{B^{-1}Nx_N}.$$

Since the binding constraints are implicit in (a4), substituting (a4) into the (a3) objective function gives (a5) $\max \mathbf{f}(\mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N, \mathbf{x}_N)$.

Taking the gradient of (a5) with respect to $\mathbf{x}_N$ yields the reduced gradient ($\mathbf{r}_{x_N}$)

(a6) $\mathbf{r}_{x_N} = \nabla \mathbf{f}_{x_N} - \nabla \mathbf{f}_{x_B}\mathbf{B}^{-1}\mathbf{N}$.

A zero reduced gradient is a necessary condition for optimality (Luenberger). Without loss of generality we define the basic part of the objective function as linear with coefficients $\mathbf{c}_B$, which yields the optimality condition

(a7) $\mathbf{r}_{x_N} = \nabla \mathbf{f}_{x_N} - \mathbf{c}'_B\mathbf{B}^{-1}\mathbf{N} = \mathbf{0}$.

The objective function associated with the independent ($\mathbf{x}_N$) variables has either zero coefficients, linear coefficients, or a nonlinear specification. If $f(\mathbf{x}_N)$ had zero coefficients, $\mathbf{x}_N$ would have to be zero at the optimum given the positive opportunity cost of resources. If $f(\mathbf{x}_N)$ was linear, say $\mathbf{c}_N$, then (a7) would be the reduced cost of the activity. A zero reduced cost of a nonbasic activity implies degeneracy when coupled with a zero activity level $\mathbf{x}_N$. Since $\mathbf{x}_N > 0$ at the optimum, $f(\mathbf{x}_N)$ cannot be linear and hence must be nonlinear for (a7) to hold.

PROPOSITION 2. *A necessary condition for the exact calibration of a $k \times 1$ vector $\mathbf{x}$ is that the objective function associated with the $(k - m) \times 1$ vector of independent variables $\mathbf{x}_N$ contain at least $(k - m)$ linearly independent instruments that change the first derivatives of* $\mathbf{f}(\mathbf{x}_N)$.

Proof. By proposition 1 $f(\mathbf{x}_N)$ is nonlinear in $\mathbf{x}_N$. Each element of the gradient $\nabla \mathbf{f}(\mathbf{x}_N)$ has a component that is a function of $\mathbf{x}_N$, and probably also a constant term. The optimality conditions in equation (a7) are modified by subtracting the constant components in the gradient ($\bar{\mathbf{k}}$) from both sides to give

(a8) $\nabla \bar{\mathbf{f}}_{x_N} = \mathbf{c}^*$

where

$$\nabla \bar{\mathbf{f}}_{x_N} = \nabla \mathbf{f}_{x_N} - \bar{\mathbf{k}}'$$

and

$$\mathbf{c}^* = \mathbf{c}'_B\mathbf{B}^{-1}\mathbf{N} - \bar{\mathbf{k}}'$$

The $1 \times (k - m)$ vector $\nabla \bar{\mathbf{f}}_{x_N}$ can be written as the product of $\mathbf{x}_N$ and a $(k - m) \times (k - m)$ matrix $\mathbf{F}$, where the ith column of $\mathbf{F}$ has elements

$$\frac{\partial f(x_N)}{\partial x_i}\frac{1}{x_i}$$

as in equation (4).

Using this decomposition

(a9) $\nabla \bar{\mathbf{f}}_{x_N} \equiv \mathbf{x}'_N\mathbf{F}$

the necessary reduced gradient condition (a8) can now be rewritten as

(a10) $\mathbf{x}'_N\mathbf{F} = \mathbf{c}^*$

Calibration of an optimization model requires that the observed solution vector $\tilde{\mathbf{x}}$ results from the optimal solution of the calibrated model. From equation (a4) the independent values $\tilde{\mathbf{x}}_N$ imply the dependent values $\tilde{\mathbf{x}}_B$. Since from (a8), $\mathbf{c}^*$ is a vector of fixed parameters, the necessary condition (a10) can only hold at $\tilde{\mathbf{x}}_i$ if the values of $\mathbf{F}^{-1}$ can be calibrated to map $\mathbf{c}^*$ into $\tilde{\mathbf{x}}_N$. Thus the matrix of calibrating gradients $\mathbf{F}^{-1}$ must span $\tilde{\mathbf{x}}$ such that

(a11) $\tilde{\mathbf{x}}'_N = \mathbf{c}^*\mathbf{F}^{-1}$

It follows that the rank of $\mathbf{F}$ must be $(k - m)$ and there have to be $(k - m)$ linearly independent instruments which change the values of $\mathbf{F}$ to exactly calibrate $\tilde{\mathbf{x}}$.

Example. Let x_n be a 2×1 vector

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and

(a12) $\mathbf{f}(\mathbf{x}_N) = \boldsymbol{\alpha}'\mathbf{x}_N - \mathbf{x}'_N\mathbf{Q}\mathbf{x}_N$

where

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

and symmetric. Writing (a7) as

(a13) $$[\alpha_1 - 2x_1q_{11} - 2x_2q_{12}, \alpha_2 - 2x_2q_{22} - 2x_1q_{21}] - c'_BB^{-1}N = 0$$

defining the $1 \times (k - m)$ row vector $\mathbf{c}^*$ as in equation (a8) results in

(a14) $[2x_1q_{11} + 2x_2q_{12}, 2x_1q_{21} + 2x_2q_{22}] = c^*$

By definition, the left-hand side of equation (a14) can be written as the product of $\mathbf{x}'_N$ and a matrix $\mathbf{F}$ where

(a15) $$\mathbf{F} = \begin{bmatrix} 2q_{11} & 2q_{21} \\ 2q_{12} & 2q_{22} \end{bmatrix}.$$

Therefore the optimality condition that the reduced gradient equals 0 requires that $\mathbf{x}_N\mathbf{F} = \mathbf{c}^*$. If particular values of $\mathbf{x}_N$, say $\tilde{\mathbf{x}}_N$, are required by changing the coefficients of $\mathbf{F}$, then $\tilde{\mathbf{x}}_N = \mathbf{c}^* \mathbf{F}^{-1}$.

Note from equation (a8) that $-\mathbf{c}^*$ is the difference between the constant linear term in the objective function $\bar{\mathbf{k}}$ and the opportunity cost of the resources. Thus $-\mathbf{c}^*$ is equal to the vector of PMP dual values $\boldsymbol{\lambda}_2$. Solving for the parameters of $\mathbf{F}$, given $\mathbf{c}^*$ and $\tilde{\mathbf{x}}_N$ is computationally identical to solving for the vector of δ_i parameters which requires the necessary condition that $\mathbf{F}$ is linearly independent and of rank $(k - m)$.

COROLLARY. *The number of calibration terms in the objective function must be equal to or greater than the number of independent variables to be calibrated.*

Appendix B

Perturbation of the calibration constraints is shown to preserve the primal and dual values.

Constraint Decoupling

Constraint decoupling is shown given the degenerate problem where the binding and slack resource constraints under values $\tilde{\mathbf{x}}$ are separated into groups I and II.

Problem P1.

$$\text{(b1)}\quad \begin{array}{lll} \text{maximize} & \mathbf{f(x)} & \\ \text{subject to} & \mathbf{Ax} = \mathbf{b} & \text{(I)} \\ & \hat{\mathbf{A}}\mathbf{x} < \hat{\mathbf{b}} & \text{(II)} \\ & \mathbf{Ix} = \tilde{\mathbf{x}} & \text{(III)} \end{array}$$

$$\mathbf{x} = k \times 1,\ \mathbf{A} = m \times k\ \ \hat{\mathbf{A}} = (l - m) \times k$$

$$\tilde{\mathbf{x}} = k \times 1\ \ k > m\ \ \mathbf{b} = m \times 1\ \ \hat{\mathbf{b}} = (l - m) \times 1.$$

$\tilde{\mathbf{x}}$ is a $k \times 1$ vector of activities that are observed to be nonzero in the base-year data; $k > m$ implies that there are more nonzero activities to calibrate than the number of binding resource constraints (I).

We assume that $\mathbf{f(x)}$ is monotonically increasing in $\mathbf{x}$ with first and second derivatives at all points, and that problem P1 is not primal or dual degenerate.

PROPOSITION 3. *There exists a $k \times 1$ vector of perturbations $\boldsymbol{\varepsilon}$ ($\boldsymbol{\varepsilon} > 0$) of the values $\tilde{\mathbf{x}}$ such that*

(a) The constraint set (I) in equation (b1) is decoupled from the constraint set (III), in the sense that the dual values associated with constraint set I do not depend on constraint set III;

(b) The number of binding constraints in constraint set III is reduced so that the problem is no longer degenerate; and

(c) The binding constraint set I remains unchanged.

Proof. Define the perturbed problem with the calibration constraints defined as upper bounds without loss of generality.

Problem P2.

$$\text{(b2)}\quad \begin{array}{lll} \text{maximize} & \mathbf{f(x)} & \\ \text{subject to} & \mathbf{Ax} = \mathbf{b} & \text{(I)} \\ & \hat{\mathbf{A}}\mathbf{x} < \hat{\mathbf{b}} & \text{(II)} \\ & \mathbf{Ix} \le \tilde{\mathbf{x}} + \mathbf{e} & \text{(III)} \end{array}$$

Any row of the nonbinding resource constraints (II) $\hat{\mathbf{A}}\mathbf{x} < \hat{\mathbf{b}}$ in problem P1 can be written

$$\text{(b3)}\quad \sum_{j=1}^{k} \left|\hat{a}_{ij} x_j\right| < \hat{b}_i \qquad i = 1, \ldots, (1 - m)$$

Select the constraint $i = 1, \ldots, (l - m)$ such that

$$b_i - \sum_{j=1}^{k} \hat{a}_{ij}\tilde{x}_j$$

is minimized. If $\varepsilon_j > 0$, $j = 1, \ldots, k$ are selected such that

$$\text{(b4)}\quad \sum_{j=1}^{k} \left|\hat{a}_{ij}\varepsilon_j\right| < \left[b_i - \sum_{j=1}^{k} \hat{a}_{ij}\tilde{x}_j\right].$$

By rearranging (b4), an inequality holds for the constraint when $\mathbf{x} = \tilde{\mathbf{x}} + \boldsymbol{\varepsilon}$, but x cannot exceed $\tilde{\mathbf{x}} + \boldsymbol{\varepsilon}$ from constraint set (III); therefore, those constraints in $\mathbf{Ax} \le \mathbf{b}$ that are inactive under the values $\tilde{\mathbf{x}}$ will remain inactive after the perturbation to $\tilde{\mathbf{x}} + \boldsymbol{\varepsilon}$.

The invariance of the binding resource constraints for (I) under the perturbation $\boldsymbol{\varepsilon}$ can be shown using the reduced gradient approach (Luenberger). Using (b4) we can write problem P2 using only constraint sets I and III.

$$\text{(b5)}\quad \begin{array}{ll} \text{maximize} & \mathbf{f(x)} \\ \text{subject to} & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{Ix} \le \tilde{\mathbf{x}} + \boldsymbol{\varepsilon} \end{array}$$

where $\mathbf{A}(m \times k)$, and $\mathbf{I} = k \times k$. Invoking the nondegeneracy assumption for $\mathbf{A}$ and starting with the solution for problem P1 $\tilde{\mathbf{x}}$, the constraints can be partitioned

$$\text{(b6)}\quad \begin{bmatrix} \mathbf{B} & \mathbf{N} \\ \mathbf{I}_1 & \\ & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} \begin{array}{c} = \\ \le \\ \le \end{array} \begin{bmatrix} \mathbf{b} \\ \tilde{\mathbf{x}}_B + \boldsymbol{\varepsilon}_B \\ \tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon}_N \end{bmatrix}$$

For brevity, the partition of **A** has been made so that the $(k - m)$ activities associated with **N** have the highest value of marginal products for the constraining resources. The reduced gradient for changes in $\tilde{\mathbf{x}}_N$ is therefore

$$\text{(b7)}\quad \mathbf{r}_{x_N} = \nabla \mathbf{f}_{\tilde{x}_N} - \nabla \mathbf{f}_{\tilde{x}_B} \mathbf{B}^{-1}\mathbf{N}.$$

Since $\mathbf{f}(\bullet)$ is monotonically increasing in $\mathbf{x}_N$ and $\mathbf{x}_B$, the resource constraints will continue to be binding since the optimization criterion will maximize those activities with a nonnegative reduced gradient until the reduced gradient is zero or the upper-bound calibration constraint $\tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon}$ is encountered. Since $m < n$, the model overspecializes in the more profitable crops when subject only to constraint sets I and II. Under the specification in problem P2 the most profitable activities will not have a zero-reduced gradient before being constrained by the calibration set II at values of $\tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon}$. Thus, the binding constraint set I remains binding under the $\boldsymbol{\varepsilon}$ perturbation.

The resource vector for the resource constrained crop activities ($\mathbf{x}_B$) now is

$$\text{(b8)}\quad \mathbf{b} - \mathbf{N}(\tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon})$$

and from (b6)

$$\mathbf{x}_B = \mathbf{B}^{-1}[\mathbf{b} - \mathbf{N}(\tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon})].$$

Since **B** is of full rank m, exactly m values of $\mathbf{x}_B$ are determined by the binding resource constraints, which depend on the input requirements for the subset of calibrated crop acre values $\tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon}$.

The slackness in the m calibration constraints associated with the m resource constrained output levels $\mathbf{x}_B$, follows from the monoticity of the production function in the rational stage of production. Since the production function is monotonic, the input requirement functions are also monotonic, and expansion of the output level of the subset of crop acreage to $\tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon}$ will have a nonpositive effect on the resource vector remaining for the vector of crop acreages constrained by the right-hand side, $\mathbf{x}_B$. That is

$$\text{(b9)}\quad \mathbf{b} - \mathbf{N}(\tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon}_N) \le \mathbf{b} - \mathbf{N}\tilde{\mathbf{x}}_N \text{ for } \boldsymbol{\varepsilon}_N > 0.$$

But since the input requirement functions for the $\mathbf{x}_B$ subset are also monotonic, (b9) and (b6) imply that

$$\text{(b10)}\quad \mathbf{x}_B \le \tilde{\mathbf{x}}_B \text{ or } \mathbf{x}_B < \tilde{\mathbf{x}}_B + \boldsymbol{\varepsilon}_B \text{ for } \boldsymbol{\varepsilon}_B > 0.$$

From (b10) it follows that the m perturbed upper bound calibration constraints associated with $\mathbf{x}_B$ will be slack at the optimum solution. Given (b4) and (b10), the constraints at the optimal solution to the perturbed problem P2 are

$$\text{(b11)}\quad \begin{bmatrix} \mathbf{B} & \mathbf{N} \\ \hat{\mathbf{A}}_1 & \hat{\mathbf{A}}_2 \\ \mathbf{I}_1 & \\ & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_B \\ \tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon}_N \end{bmatrix} \begin{matrix} = \mathbf{b} \\ < \hat{\mathbf{b}} \\ < \tilde{\mathbf{x}}_B + \boldsymbol{\varepsilon}_B \\ = \tilde{\mathbf{x}}_N + \boldsymbol{\varepsilon}_N \end{matrix}.$$

Thus, there are k binding constraints, $\mathbf{b}(m \times 1)$ and $\mathbf{x}_n + \boldsymbol{\varepsilon}_N$ $[(k - m) \times 1]$.

The dual constraints to this solution are

$$\text{(b12)}\quad \begin{bmatrix} \mathbf{B}' & \mathbf{0} \\ \mathbf{N}' & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_1^* \\ \hline \boldsymbol{\lambda}_2^* \end{bmatrix} = \begin{bmatrix} \nabla_{x_B} \mathbf{f}(\mathbf{x}^*) \\ \nabla_{x_N} \mathbf{f}(\mathbf{x}^*) \end{bmatrix}$$

using the partitioned inverse,

$$\text{(b13)}\quad \begin{bmatrix} \boldsymbol{\lambda}_1^* \\ \boldsymbol{\lambda}_2^* \end{bmatrix} = \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{Q} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \nabla_{x_B} \mathbf{f}(\mathbf{x}^*) \\ \nabla_{x_N} \mathbf{f}(\mathbf{x}^*) \end{bmatrix}$$

where $\mathbf{P} = \mathbf{B}'^{-1}$ and $\mathbf{Q} = -\mathbf{N}'\mathbf{B}'^{-1}$.

Thus, the $\boldsymbol{\varepsilon}$ perturbation on the upper-bound constraint set II decouples the dual values of constraint set I from constraint set II. This ensures that k constraints are binding and the partitioning of **A** into **B** and **N** is the unique outcome of the optimal solution to problem P2 in the first stage of PMP.

A CALIBRATION METHOD FOR AGRICULTURAL ECONOMIC PRODUCTION MODELS

Richard E. Howitt*

A method for calibrating agricultural production models is presented. The data requirements are those for a linear programming model with the addition of elasticities of substitution. Using these data, production models with a CES production function can be simply and automatically calibrated using small computers. The resulting models are shown to satisfy the standard microeconomic conditions. When used for analysis of policy changes, the CES models are able to respond smoothly to changes in prices or constraints. Prior estimates of elasticities of substitution, supply or demand can be incorporated in the models.

1. Introduction

Agricultural models that are used for policy analysis are often required to be disaggregated by region, commodity and input use. The level of disaggregation depends on the policy, but for analysis of the interaction between agricultural price supports and environmental outcomes, the model requirements frequently exceed the capacity of the data base for direct estimation. In this case, the modeller has to use formal or informal calibration methods to match the model outcome to the available data base. In microeconomic modelling the process of calibrating models is widely practised, but rarely formally discussed. In contrast, calibration methods for macroeconomic models have stimulated an emerging literature. Hoover (1995) provides a survey and analysis of the contending viewpoints. Gregory and Smith (1993) conclude that "Studies which use calibration methods in macroeconomics are now too numerous to list, and it is safe to say that the approach is beginning to predominate in the quantitative application of macroeconomic models". In an earlier paper these same authors (Gregory and Smith, 1990) define calibration as involving the choice of free parameters in a model by matching certain moments of simulated models to those of the data.

In this paper, a new method for calibrating partial-equilibrium agricultural production models on a national, regional or individual scale is presented. The

* Richard Howitt is a professor in the Department of Agricultural Economics at the University of California, Davis, California, USA. The work described in this paper was supported by grants 3AEM-0-80037 and 3AEL-8-000-85 from the Economic Research Service of the US Department of Agriculture.

Journal of Agricultural Economics
46 (2) (1995) 147-159

ability formally to model input substitution makes the model particularly suitable for the analysis of agricultural input policies where substitution is an important avenue of adjustment for farmers.

Regional modellers often face the added difficulty of a severely restricted data set which requires a compromise between the specification complexity of the model and the degree of disaggregation. The trade-off required to model the preferred specification with less than optimal data usually determines the economic modelling methodology used. The calibration method in this paper is able to calibrate nonlinear CES production functions in agricultural models using a minimum data set that usually restricts the modeller to a linear programme.

In the following section the calibration approach to model specification is outlined. This calibration approach has some characteristics of both econometric and programming models in that it has a more flexible production specification than linear or quadratic programming (LP, QP) models, but the free parameters in the model are based on observed farmer behaviour subject to resource and policy constraints.

The paper concludes with an overview of the properties of the models which can be termed calibrated production equilibrium (CPE) models, owing to their conceptual similarities to computable general equilibrium (CGE) models. A simple empirical example of the model calibration and response to input price changes is shown.

2. Modelling Production Microeconomics in Agriculture

Linear programming models have a long and well-established tradition in the regional analysis of agricultural production systems. They have significant advantages in that they can be generated using minimal data sets and can explicitly show how resources are used and the effect of policy constraints. However, the specification of programming models raises a number of problems. The root cause of the problem is that the production technology in all programming problems is locally linear in all inputs, including land. Quadratic (QP) specifications which include endogenous prices and risk terms add some nonlinearities but do not change the linear stepwise specification of regional production (Howitt, 1995).

The linearity in programming models results in the following empirical problems. First, the methods used to calibrate linear programmes against the base-year data have to strike a balance between poor base-year calibration and fully constrained models that may bias policy results. The second problem with using linear production specifications for agricultural policy analysis is that changes in input costs or commodity support prices in the model do not cause changes in the dual values or types of output unless they precipitate a change of basis. This leads to the well-known stepwise response of LP models to parameterisation. For models based on aggregate data, the range between steps may be larger than many levels of policy change, thus making the models inflexible for some types of policy analysis. A third shortcoming of LP models for analysing the interaction of agricultural policy and environmental consequences is that the Leontief technology, inherent in the linear response, cannot reflect the gradual substitution of inputs as their costs or quantities are changed.

Primal econometric models of production systems raise a different set of empirical problems for the regional policy modeller. Unlike programming models, the problems with primal econometric models arise not from restrictions on the specification which is usually theoretically consistent, but from the empirical compromises that have to be made to accommodate the limited data sets available. Aggregation over regions or time periods to allow degrees of freedom may mask important regional resource differences, with resulting distortions in predicted policy response.

CPE models use the basic calibration concepts from CGE models to calculate the equilibrium production function coefficients for variable inputs. The allocable resource inputs, such as land, are calibrated in a different manner using the basic price data, the dual values on crop allocations and the implicit costs of production generated using the positive mathematical programming (PMP) approach (Howitt, 1995).

Using relationships based on the first-order conditions, CPE models can calibrate regional crop-specific CES or Cobb-Douglas production functions without imposing arbitrary calibration constraints. The resulting models have the capacity to simulate detailed regional changes in agricultural policy or environmental constraints. In addition, because they have the same technology as more aggregated models, CPE models can be aggregated to sector-level production functions in CGE or econometric models. CPE models are designed to nest into one sector of a more general CGE or econometric model.

The ability to disaggregate from a national level has two advantages. First, it enables the effect of broad agricultural policy changes to be expressed on a regional agricultural basis. Similarly, national agricultural policy effects on regional environmental variables can also be calculated. Often regional differences are notable, and the political impact of regional diversity is important. While agriculture is not a large component of many industrial economies, it does have a disproportionate effect on environmental impacts, and often has a strong political role. In less-developed economies, the agricultural sector is usually dominant in terms of resources used and labour employed.

The second advantage of regional disaggregation of the agricultural sector is that it enables the agricultural economy to be directly linked to its regional resource base. Thus economic policies at any level can be linked to specific environmental impacts. For example, a change in the exchange rate can be linked to changes in the export demand for a given agricultural crop in a national model, and the shift in crop demand due to the exports could be translated by the CPE model into changes in the levels of regional herbicide use.

Over the years there have been several different approaches to defining calibrating constraints in linear models. CPE models use the observed regional crop-land allocations to deduce the first-order conditions. The empirical values are then combined with a cost (or yield) function that is nonlinear in the regional crop-land allocation. The changing cost of production is based on the Ricardian concept of heterogeneous inputs (Peach, 1993) in a given region or farm. Examples of this heterogeneity are differing soil qualities, or the fixed amount of seasonal operation time and management available in most farm businesses. Both these factors lead to increasing marginal costs for regional crop production.

Production economists have often noted that crop yields are stochastic (Anderson, 1974; Antle, 1983) but, owing to the aggregation of land in most economic models, the linkage between expected yield and land quality is not usually formally defined. Agronomists and soil scientists have compiled tables that group soils by yield classification for most established agricultural areas. While information on the variation in yield potential is hard to quantify on a farm level, farmers are acutely aware of which fields have the most profit potential for a given crop and weather situation. The 'positive' modelling approach assumes that the farmer uses this knowledge of the effect that expansion or contraction of acreage will have on profit per acre. The marginal conditions that reflect this knowledge are revealed in the crop-land allocation made by the farmer.

For the reasons given above, the gross margin per acre is assumed to fall as the acreage in a particular crop is increased. By using the data on crop-acreage selection under given expected prices and costs, the modeller can deduce the first-order conditions for land allocation.

The following section develops an empirical calibration method. The method uses the crop-land allocations, the basic LP data set and an estimate of the elasticity of substitution to calibrate a regional CES model.

3. Calibrated Production Equilibrium Models

The empirical calibration procedure uses a three-stage approach. A constrained linear programme is specified for the first stage. In the second stage, the regional production and cost parameters that calibrate the nonlinear CES model to the base-year data are derived from the numerical results of the linear programme. The resource and policy constraints that reflect the empirical data are also included in the calibration process. The third-stage model is specified with a nonlinear objective function that incorporates the nonlinear production functions and land costs. The CES model also has resource and policy constraints. However, the calibration constraints used in the first stage are absent.

The initial development of positive mathematical programming (PMP) used nonlinear cost functions and Leontief technology to calibrate a range of models. Over the past ten years the PMP method of calibrating has been applied to national models of the US, Canadian and Turkish agricultural economies and several regional models (Bauer and Kasnacoglu, 1990; Horner *et al.*, 1992; House, 1987).

Analysis of a wider response to agricultural policy requires the introduction of more flexible production functions. The PMP and CGE calibration approaches can be combined to calibrate agricultural production models consistently and simply. In this example we will use the simplest crop-production data set possible, although this approach can be easily applied to mixed or pure livestock production. For an example of calibration methods applied to mixed livestock and crop production see Bauer and Kasnacoglu (1990).

The data set, which can be termed the minimum LP data set, is a single cross-section observation of regional production over i crops. Observations include product prices P_i, acreage allocation $\tilde{x}_{il}$, crop input use x_{ij}, cost per unit input ω_j, and average yields $\bar{y}_i$. Allocable resource limits or policy constraints are defined as b_j, the right-hand side values of inequality constraints on the production activities. Regional subscripts have been omitted for simplicity. The first stage LP model is defined in equations (1a) to (1c). Because the linear technology specification is suboptimal for some policy changes, does not mean that the numerical dual values for the base-data LP model are incorrect. The generation of the dual values for the two types of constraint in model (1) is an essential step in the derivation of adjusted factor costs that will allow the more complex CES specification to be calibrated from the simple data base.

$$\text{Max}\ \Sigma_i\, p_i \bar{y}_i x_i - \Sigma_j\, \omega_j\, a_{ij} x_i \tag{1a}$$

$$\text{s. t. } Ax \leq b, \tag{1b}$$

$$Ix \leq \bar{x} + \epsilon. \tag{1c}$$

The model differs from the usual LP format by the set of calibration constraints shown as (1c). The ϵ perturbation on the calibration constraints decouples the true resource constraints (1b) from the calibration constraints, and ensures that the dual values on the allocable resources represent the marginal values of the resource constraints. The two constraint sets will yield two sets of dual values. $\lambda 1$ are the resource shadow value duals associated with constraint set (1b). The vector of elements $\lambda 2$ are the PMP duals from the calibration constraint set (1c). The dual values on these calibration constraints are the additional marginal 'implicit' costs that are needed for the equimarginal conditions for land allocation among crops to hold. In other words, the imperfect market for land and its heterogeneity do not, in general, allow the marginal allocation conditions to hold for each crop grown. A marginal cost in addition to the average land cost is required if the first-order conditions for optimal land allocation are to hold for the observed cropping pattern.

These two sets of dual values are used to calculate the equilibrium opportunity cost of land and other fixed but allocable inputs. These values are then used in the derivation of the production function coefficients.

CGE models are by definition and convention based on Walras' law for factor allocation, which defines the set of prices that equate excess supply and demand (Dervis, *et al.*, 1982). For partial-equilibrium models, the fixed resource endowment and local adjustment costs result in resource factors having scarcity costs that may not be fully reflected in the nominal resource or rental prices. While CGE calibration methods can use market prices and quantities to define the share equations and production function parameters, partial-equilibrium agricultural models have to augment the nominal prices by the resource and crop-specific shadow values generated in the first LP stage of the calibration.

Equation (2) shows a three-input CES production function for a single crop, i.

$$y_i = \alpha_i(\beta_1 x_{i1}^{\gamma} + \beta_2 x_{i2}^{\gamma} + \beta_3 x_{i3}^{\gamma})^{\frac{1}{\gamma}} \tag{2}$$

where $\gamma = \frac{\sigma - 1}{\sigma}$, $\beta_3 = 1 - \beta_1 - \beta_2$ and σ = a prior on the elasticity of substitution.

The production function is specified as having constant returns to scale for a given quality of land, since use of the two sets of dual values and the nominal factor prices exactly allocates the total value of production among the different inputs. If the modeller needs to specify groups of inputs with differing elasticities of substitution, perhaps zero for some inputs, the nested approach suggested by Sato (1967) can be incorporated. The Cobb-Douglas production function or restricted quadratic specifications can be used instead of the CES.

The definition of model calibration in the introduction, and over a decade of empirical practice with calibrating CGE models, has established the precedent of using robustly estimated parameters from other studies for calibration. Elasticity parameters are often used as they represent underlying preferences or technologies and, as such, are less likely to vary over specific model applications. The use of exogenously estimated demand elasticities to calibrate demand functions in quadratic programming models is well established. This more general calibration approach extends this concept to incude elasticities of substitution, and in some other applications, elasticities of supply (House, 1987).

Given the data, equation (2) with J inputs has J unknown parameters to calibrate. Namely, (J − 1) share parameters β_j and one scale parameter, α. Following the usual practice in econometric specifications and CGE calibrations the (J − 1) unknown share parameters are expressed in terms of the factor cost and input shares. The first-order conditions for input allocation equate the value marginal product to the nominal input cost plus any shadow costs for constrained resources. Algebraic manipulation of the first-order conditions yields the recursive set of equations in (3a)-(3c) below that are solved for the crop and regional-specific share coefficients. The algebraic derivation of equations (3a)-(3c) is shown in the Appendix.

$$\frac{1}{\beta_1} = 1 + \frac{\overline{\omega_2}}{\overline{\omega_1}}\left(\frac{x_1}{x_2}\right)^{-\frac{1}{\sigma}} + \frac{\overline{\omega_3}}{\overline{\omega_1}}\left(\frac{x_1}{x_3}\right)^{-\frac{1}{\sigma}} \tag{3a}$$

$$\beta_2 = \beta_1\frac{\overline{\omega_2}}{\overline{\omega_1}}\left(\frac{x_1}{x_2}\right)^{-\frac{1}{\sigma}} \tag{3b}$$

$$\beta_3 = 1 - \beta_1 - \beta_2 \tag{3c}$$

where $\overline{\omega_j}$ = factor plus opportunity cost and σ = elasticity of substitution.

Share equations for variable factor inputs whose supply functions are assumed elastic are calibrated similarly to those in CGE model production functions. An important difference between CPE and CGE models is in the specification of the resource share equations. In regional partial-equilibrium models the physical limits on the availability of these resources has to be reflected in the allocations. In most partial-equilibrium models these fixed resources will have a market price, but it is likely that the physical limits will also result in a dual value for the resource. Accordingly, the share equations for allocable resource inputs other than land have the resource shadow cost, measured by the dual for constraint group (b) in model 1, $\lambda 1_j$, added to the market price of the input to yield $\overline{\omega}_j$. Owing to changes in quality, the cost of land inputs is derived by adding the market price, shadow value ($\lambda 1_l$) and the marginal crop-specific PMP cost, $\lambda 2_i$ to yield the land factor cost $\overline{\omega}_{li}$. This crop-specific cost of land reflects both the scarcity value of land and the quality differences in land allocated to different crops.

The differences in land-quality value reflected in the PMP costs enable multiple crop outputs with different average returns to land to be calibrated against a single supply of land. This approach requires the solution of the LP calibration problem in equations (1a)-(1c), and is one way in which this partial-equilibrium calibration method differs from CGE methods. In CGE models the same calibration of multiple crops is usually achieved by defining different land-supply functions for individual crops. This specification is not convincing for the disaggregated models addressed in this paper.

The adjusted factor costs $\bar{\omega}_j$ exactly exhaust the total revenues for each cropping activity and are used in equations (3a)-(3c) to calibrate the share coefficients.

The crop and regional scale coefficient α in equation (2) is calibrated by substituting the values of β, σ, y, and x back into equation (2), as shown in equation (10) in the Appendix.

Since the marginal implicit cost of changing crop acreage is included in the share equations via the parameter $\bar{\omega}_{li}$, the cost function must also be explicitly represented in the objective function. Following Occam's razor, we specify the implicit cost function for each crop in equation (4a) as quadratic in the acreage allocated to the crop.

$$\text{Implicit cost} = \Psi_i\, x_{il}^2 \tag{4a}$$

$$\lambda_{i2} = 2\Psi_i x_{il} \tag{4b}$$

$$\text{therefore } \Psi_i = \frac{\lambda_{i2}}{2x_{il}}. \tag{4c}$$

Defining the quadratic cost function in equation (4a) as the implicit cost of increasing regional crop acreage, the marginal implicit cost is calibrated using the crop-specific PMP dual value. Equation (4b) shows how $\lambda 2_i$ from problem (1) is used to calibrate the implicit cost function coefficient Ψ_i in equation (4c).

Using the coefficients calibrated above, a general CES representation of the agricultural resource production problem is shown in equation (5).

$$\text{Max } \Sigma_i\, p_i y_i - \Sigma_j\, \omega_{ij} x_{ij} - \Sigma_i\, \Psi_i x_{il}^2 \tag{5a}$$

$$\text{s.t. } y_i = \alpha_i (\Sigma_j\, \beta_{ij}\, x_{ij}^{\gamma})^{\frac{1}{\gamma}} \tag{5b}$$

$$Ax \leqslant b. \tag{5c}$$

The model in equation (5) differs from that in the first stage, equation (1), in three significant ways. First, the production technology is more general and has the empirical elasticity of substitution incorporated in it. This means that the model in (5) solves for the optimal input proportions in conjunction with the land allocation, but not in fixed proportions to it as in the Leontief specification in model (1).

Second, the objective function has the additional implicit cost function specified for each land allocation. The basis of this cost is in the heterogeneity of land, other inputs, and the fixed nature of some farm inputs such as family labour and major machinery units.

Third, the set of calibration constraints (1c) are omitted from the CPE model in (5). The CPE model still calibrates with the base-year inputs and outputs since the dual values from model (1) are incorporated in the first-order condition used to calibrate the production and cost coefficients. Thus the CPE model calibrates exactly to the base-year data without any arbitrary or empirically insupportable constraints.

To summarise, this section has shown how a minimal data set for a constrained LP model can be used to generate a more general self-calibrating CES model. The calibration process may sound complex, but with modern algorithms such as GAMS/MINOS (Brooke *et al.*, 1992) the whole process can be written in code that performs swiftly and automatically on desktop machines. The GAMS/MINOS code to perform these operations in one sequence for this general class of problems is available from the author by e-mail (rehowitt@ucdavis.edu).

4. Microeconomic Properties of Calibrated Production Models

In generalising the production specification to the CES class of functions, CPE models show properties consistent with microeconomic theory that are not exhibited in LP or input/output models. The ability for unconstrained calibration has been addressed in the previous section.

With the specification of a nonlinear profit function in land in PMP models, the standard Hicksian microeconomic properties can be derived. By specifying the primal-dual model formulation, and making the usual assumption that the matrix of implicit cost coefficients Ψ is positive definite, it can be shown (Paris, 1993 Ch. 11) that the slopes of the supply and demand functions derived from the CPE model are respectively positive and negative, as in equations (6a) and (6b). The Hicks symmetry conditions shown in equation (6c) also hold for the CPE model.

$$\frac{\delta y}{\delta p} = \text{PSD} \tag{6a}$$

$$\frac{\delta x}{\delta \omega} = \text{NSD} \tag{6b}$$

$$\frac{\delta y}{\delta \omega} = -\frac{\delta x}{\delta p}. \tag{6c}$$

The problem of stepwise response to policy changes in linear programming models is solved by the nonlinear specification in CPE models. The response of the model output to changes in price, or input use to changes in cost, is a continuous function, even though the basis may not change. When the basis of linear constraints changes, the parametric response function changes slope but is still continuous with the next basis. The importance of this property is that politically acceptable agricultural policies are usually constrained to relatively small changes in costs or policy constraints. The continuous functions in CPE models can reflect these small policy changes and simulate their economic and physical impact on a regional scale.

A simple empirical example illustrates the above points. The data for a greatly simplified and aggregated model of US irrigated crop production is shown in Table 1. The model is specified as having two regions (California, rest

Table 1 Data for the Illustrative Model

Crop Production	*Price ($/bu)*	*Average Yield (bu/acre)*
Cotton (CA)	2.924	220.0
Cotton (RUS)	2.924	151.0
Wheat (CA)	2.98	85.0
Wheat (RUS)	2.98	69.0
Rice (CA)	7.09	70.1
Rice (RUS)	7.09	48.1

Regional Resource Constraints

LAND (CA)	(Million Acres)	2.65
LAND (RUS)	(Million Acres)	14.99
WATER (CA)	(Million Acre Ft)	8.69
WATER (RUS)	(Million Acre Ft)	28.33

Resource Costs Per Unit ($)

	Land	*Water*	*Capital*	*Chemical*
Cotton (CA)	66.0	25.6	10.0	10.0
Cotton (RUS)	28.0	28.4	10.0	10.0
Wheat (CA)	33.0	25.6	10.0	10.0
Wheat (RUS)	11.0	28.4	10.0	10.0
Rice (CA)	49.0	25.6	10.0	10.0
Rice (RUS)	39.0	28.4	10.0	10.0

Base Year Resource Allocation

	Land	*Water*	*Capital*	*Chemical*
Cotton (CA)	1.49	4.47	3.960	2.640
Cotton (RUS)	5.75	5.23	1.680	1.120
Wheat (CA)	0.62	1.14	1.980	1.320
Wheat (RUS)	6.50	6.89	0.660	0.440
Rice (CA)	0.54	3.08	2.940	1.960
Rice (RUS)	2.74	7.95	2.340	1.560

Notes: CA: California; RUS: Rest of USA; elasticity of substitution: 0.7.

of USA), three irrigated crops (cotton, wheat, rice) and four inputs per crop (land, water, capital, chemicals). The data required for the CES model is the minimum set required for a linear programme plus an estimate of the elasticity of substitution obtained from prior econometric studies. Table 1 shows the data, expected output price, average regional yields, expected input costs, constraints on the allocable resources and the input allocations to regional crop production observed in the base year of the model.

Table 2 contains the parameters calibrated for the CES production function and the regional quadratic land-cost function. The scale parameters are coincidentally very similar for cotton production in the two regions. The wheat coefficients differ slightly, and rice production shows marked differences between regions. The input share parameters in Table 2 differ widely among crops in a given region, and also for the same crop between regions. These differences do not have empirical meaning given the extreme aggregation of the model, but do illustrate how the regional crop-specific calibration can adjust to differing regional technologies and resource endowments.

The linear cost parameters are, for the most part, the same as the base-year data costs in Table 1. Given that there are three binding constraints on allocable resources, two land constraints and one irrigation water limit, the three other crops require nonlinear 'implicit' cost terms for the optimum marginal conditions to hold. For these crops, the linear coefficients on land cost are calibrated so that the marginal and average cost conditions hold. The quadratic cost coefficients for these more profitable crops show wide variation,

Table 2 Parameters for the Calibrated CES Model

CES Scale Parameter				
	CA	*RUS*		
Cotton	153.381	153.588		
Wheat	53.441	69.263		
Rice	17.853	35.825		
CES Share Parameters				
	Land	*Water*	*Capital*	*Chemical*
Cotton (CA)	0.601	0.315	0.054	0.030
Cotton (RUS)	0.937	0.057	0.004	0.002
Wheat (CA)	0.355	0.380	0.170	0.095
Wheat (RUS)	0.847	0.150	0.002	0.001
Rice (CA)	0.141	0.663	0.126	0.071
Rice (RUS)	0.632	0.336	0.021	0.012
Linear Cost Parameters				
	Land	*Water*	*Capital*	*Chemical*
Cotton (CA)	–242.764	25.600	10.000	10.000
Cotton (RUS)	–191.999	28.400	10.000	10.000
Wheat (CA)	33.000	25.600	10.000	10.000
Wheat (RUS)	11.000	28.400	10.000	10.000
Rice (CA)	49.000	25.600	10.000	10.000
Rice (RUS)	–3.570	28.400	10.000	10.000
Quadratic Cost Parameters				
	Land			
Cotton (CA)	414.448			
Cotton (RUS)	76.521			
Wheat (CA)	0.000			
Wheat (RUS)	0.000			
Rice (CA)	0.000			
Rice (RUS)	31.073			

Notes: As Table 1.

as would be expected from the acreage differences. Quadratic cost functions for all cropping activities can be calibrated, if required, but additional information on the yield variability or the elasticity of supply is needed to calibrate these marginal crops.

The prices and resource right-hand side constraints in Table 1 and the parameters in Table 2 are used to define the CES production model shown in equation (5). The resulting CPE model calibrates very closely in terms of output produced, crop input allocations, and dual values on the binding resource constraints. The results of the constrained linear model and the unconstrained calibrated nonlinear model are so similar as to make tabular presentation redundant. The model calibrated and solved for all three stages in under two seconds on a standard 33 MgHz 486 personal computer.

Table 3 shows selected results from a 25 per cent increase in the cost of chemical inputs in both regions. This could be the result of an environmental policy that internalised chemical externalities by a pollution charge.

The theoretical advantages of the CES approach, namely smooth parametric policy responses and the ability to change input use proportions, are shown in the results. The first part of Table 3 shows the percentage change in total input use by crop and region. Cotton production in California – cotton (CA) – is notable in that the 14 per cent reduction in chemical use is more than compensated for by increases in the absolute level of land, water and capital, and

Table 3 Changes in Input and Output Use (25% Increase in Chemical Cost)

Percentage Difference in Total Input Use

	Land	*Water*	*Capital*	*Chemical*
Cotton (CA)	0.296	1.371	0.079	−14.396
Cotton (RUS)	−0.068	−0.146	−0.150	−14.593
Wheat (CA)	0.432	−0.389	−1.654	−15.880
Wheat (RUS)	0.635	0.571	0.557	−13.994
Rice (CAL)	−1.314	−1.845	−3.096	−17.112
Rice (RUS)	−1.365	−1.737	−1.740	−15.952

Percentage Change in Per Acre Input Use

	Water	*Capital*	*Chemical*
Cotton (CA)	1.071	−0.217	−14.648
Cotton (RUS)	−0.078	−0.082	−14.535
Wheat (CA)	−0.817	−2.078	−16.242
Wheat (RUS)	−0.064	−0.078	−14.537
Rice (CA)	−0.539	−1.806	−16.008
Rice (RUS)	−0.377	−0.380	−14.789

Percentage Change in Output

	CA	*RUS*
Cotton	0.080	−0.144
Wheat	−1.653	0.572
Rice	−3.095	−1.737

Notes: As Table 1.

by intensity per acre of water. The output statistics in the last part of Table 3 show that for Californian cotton production, total output increases with chemical costs. This is due to a shift in comparative advantage within California towards cotton production caused by the chemical cost increase. For many crops the trend is to have reductions in total input use for all inputs, and consequent output reductions. Other crops, such as wheat in the rest of the USA, show increases in total output despite large reductions in chemical use. This is due to compensating increases in land area planted, but not in the intensity of capital and water per acre which are reduced slightly. Clearly, even in this very simple model, there is a wide variation in types of substitution stimulated by the increase in chemical cost.

The second important characteristic claimed for CPE models is the smooth response to parametric policy changes. Table 3 shows that the 25 per cent increase in chemical cost produces different percentage changes in input allocation and output. The change produced in total input use across crops and regions ranges from a decrease of 17 per cent in chemicals to an increase of 0.3 per cent in water use. Several inputs and regional outputs are changed very little by the chemical cost increase. Since all crops are still grown in all regions there has been no change of basis; despite this the nonlinear functions are able to show the marginal effects that a cost increase on chemicals will induce.

5. Conclusions

This paper has reviewed model requirements for analysing regional agricultural policy problems and found that, for some policy applications, the conventional empirical approaches available for this task are wanting. Linear programming models have insufficient technical flexibility, while econometric models are often restricted by the data available.

An alternative approach that calibrates more flexible production functions than linear programmes, but uses almost the same minimal data base, is introduced as a compromise between the extremes of linear programming and econometric estimation. The properties of CPE models are shown to meet many of the requirements for modelling regional agricultural policies, while the data requirements are satisfied by the minimal data sets usually available on a regional basis.

While potential difficulties in the nonlinear solution of the many-dimensional nonlinear CPE specification cannot be blithely ignored, initial empirical results indicate that these models are quite tractable. Given the common agricultural policy requirement for modelling regional economic and environmental consequences, the properties of the models seem to justify the additional complexity.

References

Anderson, J. R. (1974). Sparse Data, Climatic Variability and Yield Uncertainty in Response Analysis, *American Journal of Agricultural Economics*, **55**, 77-82.

Antle, J. M. (1983). Sequential Decision Making in Production Models, *American Journal of Agricultural Economics*, **65(2)**, 282-290.

Bauer, S. and Kasnacoglu, H. (1990). Nonlinear Programming Models for Sector Policy Analysis, *Economic Modelling*, July, 275-290.

Brooke, A., Kendrick, D. and Meeraus, A. (1992). *GAMS: A User's Guide*, The Scientific Press, San Francisco.

Dervis, K., de Melo, J. and Robinson, S. (1982). *General Equilibrium Models for Development Policy*. Cambridge University Press, Cambridge.

Gregory, A. W. and Smith, G. W. (1990). Calibration as Estimation, *Econometric Reviews*, **9(1)**, 57-89.

Gregory, A. W. and Smith, G. W. (1993). Statistical Aspects of Calibration in Macroeconomics, in: Maddala, G. S., Rao, C. R. and Vinod, H. D. (eds), *Handbook of Statistics, Econometrics*, Vol. 11, 707-719. North Holland, Amsterdam.

Hoover, K. D. (1995). Facts and Artifacts: Calibration and the Empirical Assessment of Real-Business-Cycle Models, *Oxford Economic Papers*, **47**.

Horner, G. L., Corman, J., Howitt, R. E., Carter, C. A. and MacGregor, R. J. (1992). *The Canadian Regional Agricultural Model: Structure, Operation and Development*. Technical Report 1/92, Agriculture Canada, Ottawa, October.

House, R. M. (1987). *USMP Regional Agricultural Model*. National Economics Division Report, Economic Research Service, USDA, Washington DC, July.

Howitt, R. E. (1995). Positive Mathematical Programming, *American Journal of Agricultural Economics*, forthcoming.

Paris, Q. (1993). *Lecture Notes in Mathematical Programming*. Department of Agricultural Economics University of California, Davis, CA.

Peach, T. (1993). *Interpreting Ricardo*. Cambridge University Press, Cambridge.

Sato, K. (1967). A Two-Level Constant Elasticity of Substitution Production Function, *Review of Economic Studies*, **34**, 201-218.

APPENDIX

Derivation of the Parameters for the CES Production Function

A CES production function with one output, three inputs and constant returns to scale is defined in equation (A1):

$$y = \alpha\,(\beta_1 x_1^{\gamma} + \beta_2 x_2^{\gamma} + \beta_3 x_3^{\gamma})^{\frac{1}{\gamma}} \tag{A1}$$

where $\gamma = \frac{\sigma - 1}{\sigma}$; $\Sigma\, \beta_i = 1$; σ = prior value elasticity of substitution.

Taking the derivative of (A1) with respect to x_1 we obtain

$$\frac{\delta y}{\delta x_1} = \gamma \beta_1 x_1^{(\gamma-1)} \frac{1}{\gamma} \alpha \left(\beta_1 x_1^{\gamma} + \beta_2 x_2^{\gamma} + \beta_3 x_3^{\gamma}\right)^{\left(\frac{1}{\gamma}-1\right)} \tag{A2}$$

since $(\gamma - 1) = -\frac{1}{\sigma}, (\frac{1}{\gamma} - 1) = \frac{1}{\sigma - 1}$.

Simplifying and substituting (A2) can be rewritten as:

$$\frac{\delta y}{\delta x_1} = \beta_1 x_1^{-\frac{1}{\sigma}} \alpha (\beta_1 x_1^{\gamma} + \beta_2 x_2^{\gamma} + \beta_3 x_3^{\gamma})^{\frac{1}{\sigma - 1}}. \tag{A3}$$

Equating $\rho \frac{\delta y}{\delta x_1} = \omega_1$ and $\rho \frac{\delta y}{\delta x_2} = \omega_2$ gives

$$\frac{\omega_1}{\omega_2} = \frac{\beta_1 x_1^{-\frac{1}{\sigma}}}{\beta_2 x_2^{-\frac{1}{\sigma}}}. \tag{A4a}$$

$$\frac{\omega_1}{\omega_3} = \frac{\beta_1 x_1^{-\frac{1}{\sigma}}}{\beta_3 x_3^{-\frac{1}{\sigma}}}. \tag{A4b}$$

From (A4a) we obtain

$$\beta_2 = \beta_1 \frac{\omega_2}{\omega_1} \left(\frac{x_1}{x_2}\right)^{-\frac{1}{\sigma}}. \tag{A5}$$

Likewise from equation (A4b):

$$\beta_3 = \beta_1 \frac{\omega_3}{\omega_1} \left(\frac{x_1}{x_3}\right)^{-\frac{1}{\sigma}}. \tag{A6}$$

But from the constant returns to scale assumption

$$\beta_3 = 1 - \beta_1 - \beta_2. \tag{A7}$$

Substituting (A5) and (A6) into (A7) we obtain:

$$\beta_1 \frac{\omega_3}{\omega_1} \left(\frac{x_1}{x_3}\right)^{-\frac{1}{\sigma}} = 1 - \beta_1 - \beta_1 \frac{\omega_2}{\omega_1} \left(\frac{x_1}{x_2}\right)^{-\frac{1}{\sigma}}. \tag{A8}$$

Dividing through by β_1 and rearranging yields

$$\frac{1}{\beta_1} = 1 + \frac{\omega_2}{\omega_1} \left(\frac{x_1}{x_2}\right)^{-\frac{1}{\sigma}} + \frac{\omega_3}{\omega_1} \left(\frac{x_1}{x_3}\right)^{-\frac{1}{\sigma}}. \tag{A9}$$

Solving (A9) for β_1 and substituting into equation (A5) solves for β_2. Substituting the values into equation (A7) solves for β_3.

The numerical value for the total production, y, in equation (A1) is known from the observed acreage $\bar{x}_1$ and the average yield $\bar{y}$. Using the known values for $\beta_1 \ldots \beta_3$ and equation (A1), we can solve for α as follows:

$$\alpha = \bar{y}\bar{x}_1 / (\beta_1 x_1^{\gamma} + \beta_2 x_2^{\gamma} + \beta_3 x_3^{\gamma})^{\frac{1}{\gamma}}. \tag{A10}$$

The minimal data set needed to specify an LP model are the input allocations and prices, the expected yield, price and any resource or policy constraints. If the elasticity of substitution value and the constant returns to scale assumption are added to this basic data set, the scale and share parameters of the CES production function can be recursively solved for any number of inputs using equations (A9), (A5), (A7) and (A10).

Environmental Modelling & Software 38 (2012) 244–258

Contents lists available at SciVerse ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Calibrating disaggregate economic models of agricultural production and water management

Richard E. Howitt [a,1], Josué Medellín-Azuara [b,*], Duncan MacEwan [a,2], Jay R. Lund [b,3]

[a] Department of Agricultural and Resource Economics, University of California, Davis, One Shields Ave. Davis, CA 95616, USA
[b] Department of Civil and Environmental Engineering, University of California, Davis, One shields Ave. Davis, CA 95616, USA

ARTICLE INFO

Article history:
Received 11 February 2012
Received in revised form
22 June 2012
Accepted 23 June 2012
Available online 31 July 2012

Keywords:
Positive Mathematical Programming
Calibration
California
Hydro-economic models
Water markets
Agricultural production
Drought analysis
Economic impacts

ABSTRACT

This paper describes calibration methods for models of agricultural production and water use in which economic variables can directly interact with hydrologic network models or other biophysical system models. We also describe and demonstrate the use of systematic calibration checks at different stages for efficient debugging of models. The central model is the California Statewide Agricultural Production Model (SWAP), a Positive Mathematical Programming (PMP) model of California irrigated agriculture. We outline the six step calibration procedure and demonstrate the model with an empirical policy analysis. Two new techniques are included compared with most previous PMP-based models: exponential PMP cost functions and Constant Elasticity of Substitution (CES) regional production functions. We then demonstrate the use of this type of disaggregated production model for policy analysis by evaluating potential water transfers under drought conditions. The analysis links regional production functions with a water supply network. The results show that a more flexible water market allocation can reduce revenue losses from drought up to 30%. These results highlight the potential of self-calibrated models in policy analysis. While the empirical application is for a California agricultural and environmental water system, the approach is general and applicable to many other situations and locations.

1. Introduction

The importance of integrating economic and environmental considerations for policy making has fostered the use of hydro-economic models, surveyed by Harou et al. (2009) from a hydrologic perspective and by Booker et al. (2012) from an economic viewpoint. This paper describes in detail methods by which economic models of agricultural production and water use can be calibrated at a scale where the economic variables can directly interact with hydrologic network models. We also develop systematic checks of calibration at different stages, which allows for efficient debugging of models. While the empirical application is for a California agricultural and environmental water system, the approach is general and can be applied to other situations and locations.

Irrigated agriculture is the largest water user and an important part of local economies in arid regions around the world, but it is also a sector which is expected to adapt to changes in urban and environmental water conditions and demands. Production in many of these regions is increasingly constrained by environmental concerns including groundwater overdraft, nitrate runoff, soil erosion, salinity, and balancing water diversions with urban and ecosystem demands. In addition, future population growth and climate change is expected to increase food demand and place additional strain on production, resources, and the environment. Consequently, policymakers seek to design and evaluate agricultural–environmental policies to address these and related issues. Historically policy evaluation is undertaken with aggregate financial and physical data, but these data, and corresponding methods, are being replaced with the influx of micro-level and remote sensing data and improvements in agricultural production models.

We empirically illustrate the ideas in this paper with the example of irrigated agriculture in California, but the methods and insights apply to any agricultural region. The Statewide Agricultural Production Model (SWAP) is a multi-region, multi-input and output model of agricultural production which self-calibrates using the method of Positive Mathematical Programming (PMP) (Howitt,

* Corresponding author. Tel.: +1 530 574 8019; fax: +1 530 752 7872.
E-mail addresses: howitt@primal.ucdavis.edu (R.E. Howitt), jmedellin@ucdavis.edu (J. Medellín-Azuara), macewan@primal.ucdavis.edu (D. MacEwan), jrlund@ucdavis.edu (J.R. Lund).
[1] Tel.: +1 530 752 1521; fax: +1 530 752 5614.
[2] Tel.: +1 760 799 3960; fax: +1 530 752 5614.
[3] Tel.: +1 530 752 5271; fax: +1 530 752 7872.

1364-8152/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envsoft.2012.06.013