With respect to the production model, the two-period approach suggests that production after 1988 became more elastic. An estimated price elasticity of 0.51 before 1988 versus an estimate of 1.04 after 1988. Both coefficients are significant. Figure 12 illustrates the fit of the estimation.

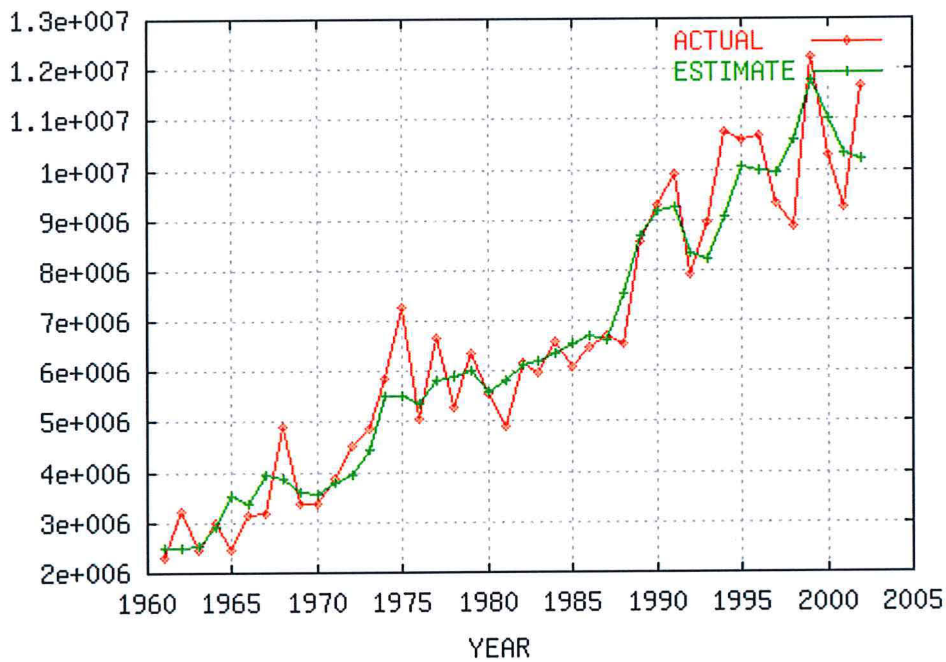

Figure 12: ***Structural break model for processing tomato production in tons.***

Demand

In this section two demand models for processing tomatoes are presented. The first one describes the demand for processing tomatoes at the farm level and the second one illustrates the final demand (at the consumer level) for tomato products.

Demand for processing tomatoes

The demand for processing tomatoes is a function of farmer prices and the price index for tomato paste. The data refer to 21 time periods (from 1982 to 2002). The model describes the industry demand under the assumptions of price taking behavior and market equilibrium. Industry expectations are modeled using lagged prices. The regression model has been estimated with a moving average process of order one. The derived demand equation for processed tomatoes is:

$$\ln Q_t = \beta_0 + \beta_1 \ln PF_{t-1} + \beta_2 PR_{t-1} + \beta_3 t \tag{7}$$

where Q_t represents the quantity demanded of California processing tomatoes , PF_{t-1} denotes the grower price, lagged one time period, PR_{t-1} is the price of tomato paste, lagged one time period, and t is a time trend.

The estimated demand equation is

$$\begin{array}{l} \ln \hat{Q}_t = 15.67 - 0.18 \ln PF_{t-1} + 0.16 \ln PR_{t-1} + 0.03t \\ \qquad\quad (0.06)\ (0.05) \qquad\quad (0.04) \qquad\quad (0.02) \end{array} \tag{8}$$

where $R^2 = 0.815$ and n = 21. Based on the estimates, the demand for California processing tomatoes is inelastic (a statistically significant own-price estimated elasticity of -0.18). The coefficient of tomato paste price is 0.16 and significant. As the price of tomato paste increases the demand for processing tomatoes increases. This is as expected since the demand for processing tomatoes is a derived demand.

Figure 13 shows the fit of the regression.

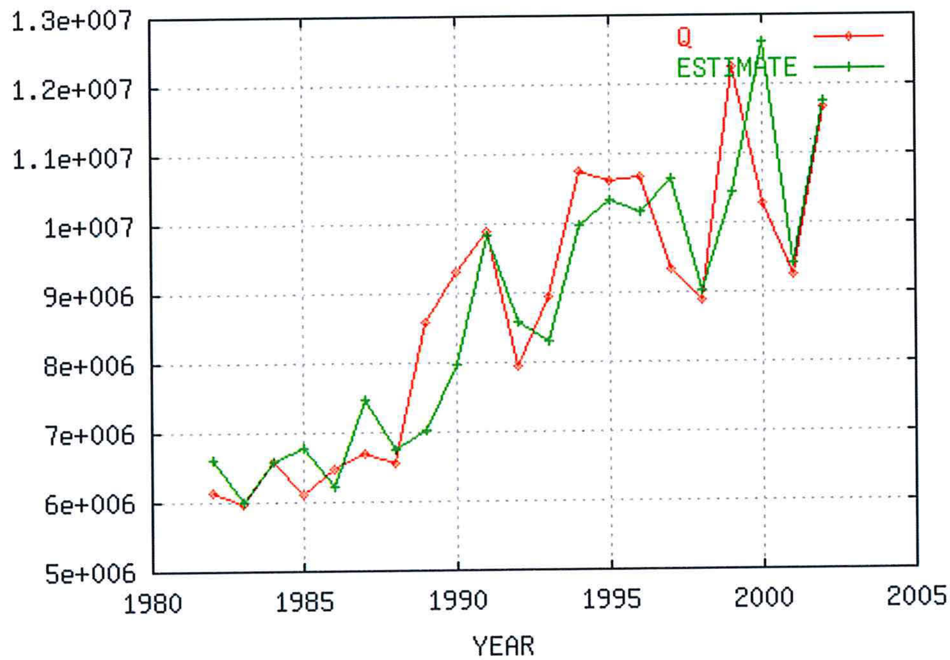

Figure 13: ***Demand for California processing tomatoes (million tons).***

Demand for tomato products

The demand for tomato products was estimated based on quarterly US retail sales data from 1993 to 2004 (Food Institute). Since the data exhibit a strong seasonal pattern, the estimation model is:

$$\ln Q_t = \beta_0 + \beta_1 \ln PT_t + \beta_2 EF_t + \beta_3 PF_t + \beta_4 D_{1t} + \beta_5 D_{2t} + \beta_6 D_{3t} + v_t \tag{9}$$

where PT_t represents the price of tomato products, EF_t denotes the expenditure for food, PF_t represents the price index for food, and D_1, D_2, and D_3 are seasonal dummy variables for the first, second and third quarters.

The model was estimated with a moving average of order four error term (consistent with seasonality). The results are

$$\underset{}{\ln \hat{Q}_t} = \underset{(0.52)}{14.84} - \underset{(0.08)}{0.26} \ln PT_t - \underset{(0.19)}{1.64} EF_t + \underset{(0.22)}{0.86} PF_t + \underset{(0.01)}{0.05} D_{1t} - \underset{(0.01)}{0.33} D_{2t} - \underset{(0.01)}{0.29} D_{3t} \tag{10}$$

where $R^2 = 0.99$ n = 48. The demand for tomato products is inelastic (a significant own-price elasticity estimate of -0.26) and on average is higher during the first and the fourth quarters (since fresh tomatoes are less available). The sign of the food expenditure elasticity is negative which is not as expected.

Figure 14 illustrates the fit of the regression.

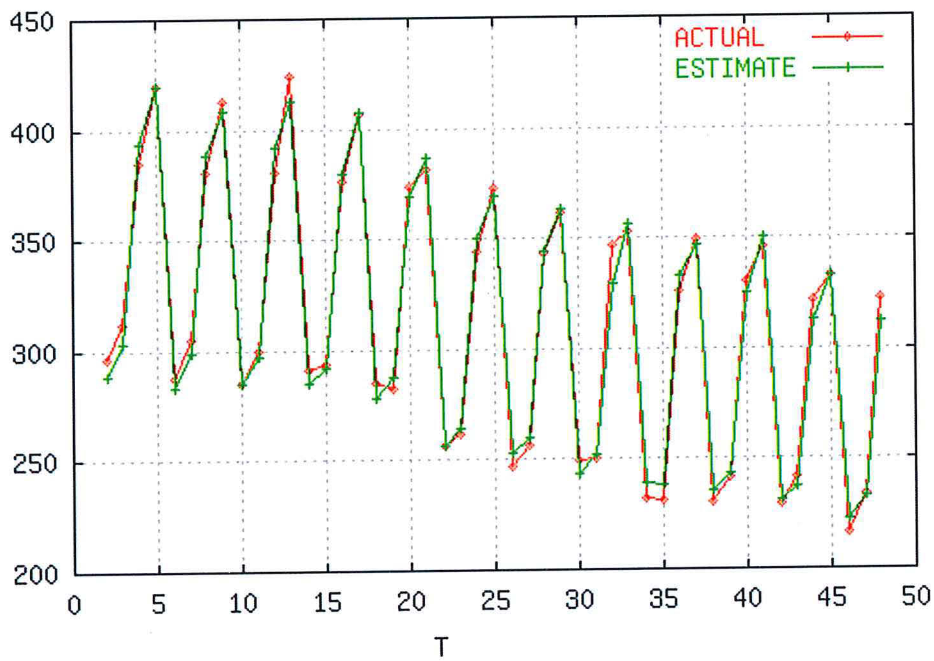

Figure 14: ***Consumers' demand for processing tomato products (1st quarter 1993-4th quarter 2004)***

Fresh Tomatoes

Per capita consumption of fresh tomatoes has been increasing since the '80s (Figure 15). Higher demand triggered a structural adjustment in the industry. Figure 1 shows that, initially, the main acreage adjustment was in Florida, while California increased acreage sharply in the late '80s.

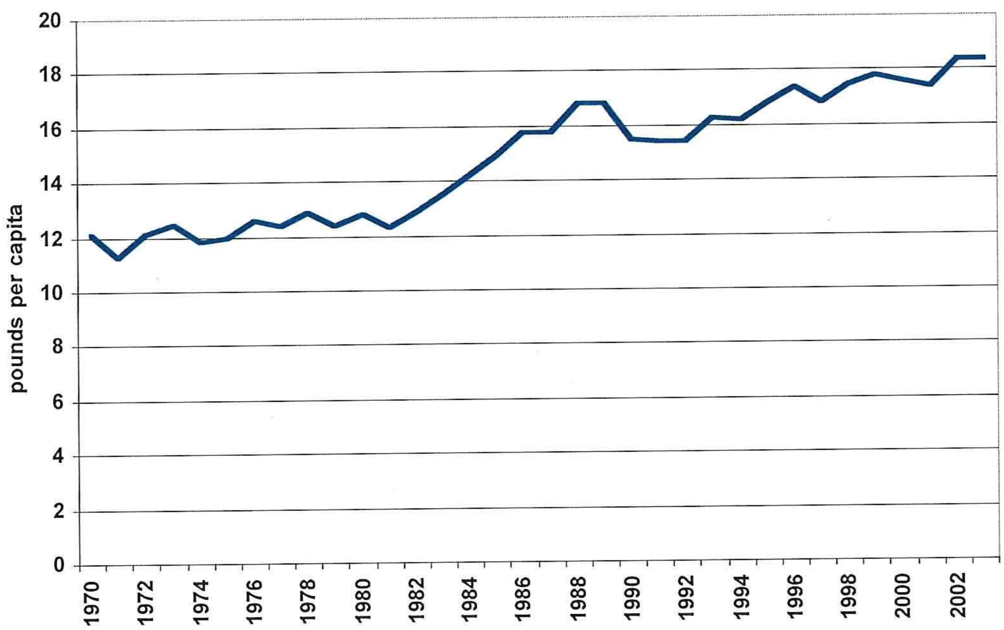

Figure 15: ***US per capita consumption of fresh tomatoes***

Given this trend in the industry the estimations allowed for a structural break. The two periods are 1960-1987 and 1988-2002.

Acreage for Fresh Tomatoes

The acreage model was estimated assuming a partial adjustment process. Price expectations have been modeled using the previous year's price for the period 1960-1987 and a two-year lagged price before the period 1988-2002. This was done because after the structural change, the prices exhibits an alternate pattern, so that the current price is negatively correlated with the previous year, but positively correlated with two periods before. Finally we tested the influence of the processing industry on the fresh tomato acreage, by using the price of processing tomato as a regressor.

What accounts for the structural break in 1987 in fresh tomato acreage? Much of the increase in California acreage can be explained as a response to changes in consumption patterns, according to the USDA. In terms of consumption, tomatoes are the Nation's fourth most popular fresh-market vegetable behind potatoes, lettuce, and onions. Fresh-market tomato consumption has been on the rise due to the enduring popularity of salads, salad bars, and sandwiches such as the BLT (bacon-lettuce-tomato) and subs. Perhaps of greater importance has been the introduction of improved tomato varieties, consumer interest in a wider range of tomatoes (such as hothouse and grape tomatoes), a surge of immigrants with vegetable-intensive diets, and expanding national emphasis on health and nutrition. After remaining flat during the 1960s and 1970s at 12.2 pounds, fresh use increased 19 percent during the 1980s, 13 percent during the 1990s, and has continued to trend higher in the current decade. Although Americans consume three-fourths of their tomatoes in processed form (sauces, catsup, juice), fresh-market use exceeded 5 billion pounds for the first time in 2002 when per capita use also reached a new high at 18.3 pounds. Because of the expansion of the domestic greenhouse/hydroponic tomato industry since the mid-1990s, it is likely per capita use is at least 1 pound higher than currently reported by USDA (the Department does not currently enumerate domestic greenhouse vegetable production). One medium, fresh tomato (about 5.2 ounces) has 35 calories and provides 40 percent of the U.S. Recommended Daily Amount of vitamin C and 20 percent of the vitamin A. University research shows that tomatoes may protect against some cancers.

he partial adjustment acreage function for fresh tomatoes is:

$$\ln A_t = \beta_0 + \beta_1 \ln EP_t + \beta_2 \ln PP_t + \beta_3 t + \beta_4 \ln A_{t-1} + \varepsilon_t \tag{11}$$

where A represents fresh tomato acreage in acres, EP denotes the price expectation in \$/ton (equal to the previous year price for the period 1960-1987 and to the price of two years before for the period 1988-2002), PP denotes the price of processing tomatoes, and t is a time trend.

The estimated fresh tomato acreage function for the period 1960-1987 is:

$$\underset{}{\ln \hat{A}_t} = \underset{(0.96)}{17.43} + \underset{(0.05)}{0.00 \ln EP_t} - \underset{(0.05)}{0.16 \ln PP_t} - \underset{(0.00)}{0.02t} - \underset{(0.07)}{0.67 \ln A_{t-1}} \tag{12}$$

where $R^2 = 0.828$ and n = 27. The estimated coefficient on expected price of fresh tomatoes is positive but insignificant. The results indicate a declining trend in acreage, with disinvestments from the industry regardless of any price expectation. The negative coefficient on lagged acreage (-0.67) and is highly significant and reflects rotation practices.

In the second period (1988-2002), the results of the estimation of fresh tomato acreage function are

$$\ln \hat{A}_t = \underset{(1.24)}{6.81} + \underset{(0.07)}{0.23 \ln EP_t} + \underset{(0.10)}{0.48 \ln PP_t} + \underset{(0.00)}{0.02t} - \underset{(0.12)}{0.04 \ln A_{t-1}} \tag{13}$$

where $R^2 = 0.840$ and n = 15. The estimation suggests a structural change in the second period. The trend is increasing, the coefficient on price expectation is positive and significant (0.23) and the sign on the coefficient of processing tomato price indicates complementarities (0.48).

Figure 16 illustrates the fit of the model for the period 1960-2002.

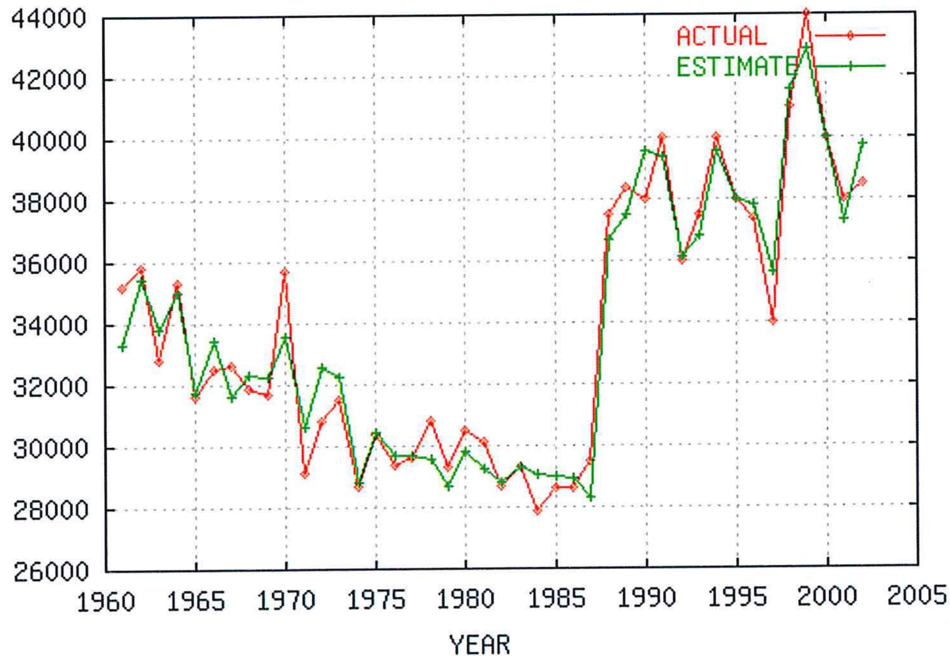

Figure 16: ***California fresh tomato acreage (in acres).***

Production

The partial adjustment model for fresh tomato production is:

$$\ln Q_t = \beta_0 + \beta_1 \ln EP_t + \beta_2 \ln PP_t + \beta_3 t + \beta_4 t^2 + \beta_5 W_t + \beta_6 \ln Q_{t-1} + D_{t,79} + \varepsilon_t \qquad (14)$$

where Q represents annual production in tons, EP denotes the price expectation in \$/ton, PP denotes the price of processing tomatoes[2], also in \$/ton, t is a time trend, W_t represents the water availability (measured by the four river index) and D is a dummy variable identifying the year 1979 which had an exceptional yield. Note that in this equation the time trend including the quadratic trend, captures the effects of technological change. The model was estimated separately for the two time periods, assuming a moving average error process which is consistent with a partial adjustment specification. The results are as follows:

[2] For production, slightly better results can be obtained by using cotton as a competing crop. However, since cotton performs poorly in explaining acreage, we kept processing tomatoes in the estimation for consistency with the acreage equation.

Period 1960-1987:

$$\ln\hat{Q}_t = 10.04 + 0.22\ln EP_t - 0.04PP_t - 0.01t + 0.00t^2 + 0.00W_t + 0.11\ln Q_{t-1} + 0.37D_{t,79} \tag{15}$$
$$\begin{array}{llllllll}(1.51)(0.12) & (0.04) & (0.01) & (0.00) & (0.00) & (0.14) & (0.07)\end{array}$$

where $R^2 = 0.932$ and n = 27.

Period 1988-2002:

$$\ln\hat{Q}_t = 6.82 + 0.27\ln EP_t - 0.05PP_t - 0.02t + 0.00t^2 + 0.00W_t + 0.33\ln Q_{t-1} \tag{16}$$
$$\begin{array}{lllllll}(5.21)\ (0.11) & (0.31) & (0.09) & (0.00) & (0.01) & (0.47)\end{array}$$

where R^2 = 0.789 and n = 15. Based on the estimations, the short run elasticity of fresh tomato production with respect to price expectations was 0.22 before 1987 and 0.27 after 1987. There is no statistical evidence of change in the values of elasticities after the structural break. Given the partial adjustment model, the estimation of long run elasticity is 0.247 (before 1988) and 0.403 (from 1988 on). The trend term coefficients were not significant nor were the coefficienets on the lagged production terms. Figure 17 describes the fit of the regression.

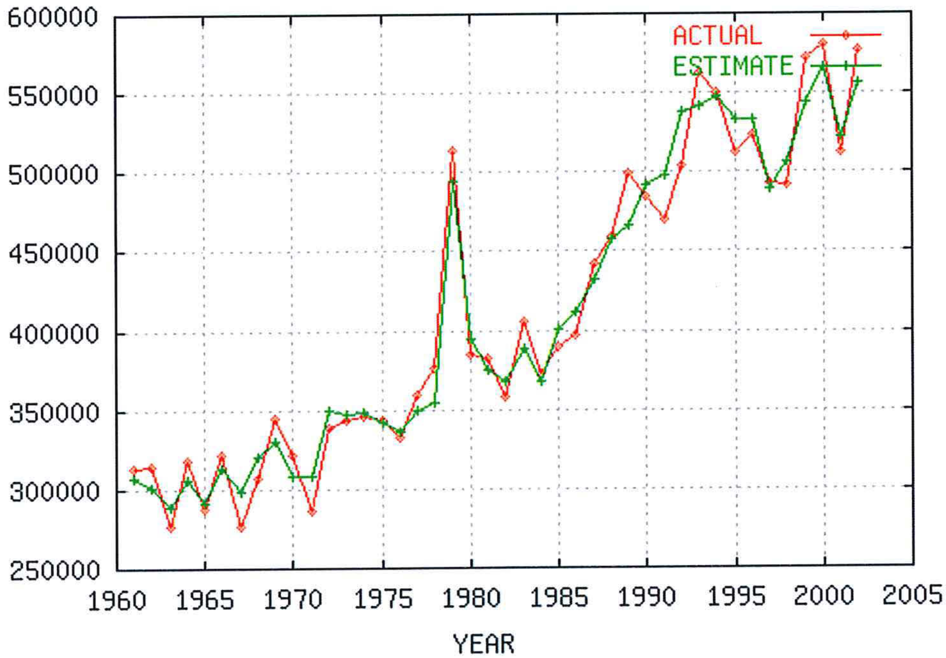

Figure 17: ***California fresh tomato production (in tons).***

Demand

The US demand for fresh tomatoes has been modeled using the Almost Ideal Demand System. The system estimates simultaneously the demand for four of the major vegetables: tomatoes, lettuce, carrots and cabbage. The approach assumes that consumers are price takers and that consumers of the four goods have preferences that are weakly separable. The assumption of weak separability permits the demand for a commodity to be written as a function of its own price, the price of substitutes and complements, and group expenditure.

The almost ideal demand system is

$$w_{it} = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln(\frac{x_i}{P_t^*}) + \varepsilon_{it} \qquad (17)$$

where w_i represents the ith budget share of commodity i, p_j denotes the jth price of the jth good, x_t is group expenditure for the particular set of commodities (fresh tomatoes, carrots, lettuce, and cabbage), and P_t^* is a translog deflator and is given by

$$\ln P_t^* = \alpha_0 + \sum_k \alpha_k \ln p_k + (1/2)\sum_i \sum_j \gamma_{ij} \ln p_i \ln p_j .$$

Adding-up restrictions require that $\sum \alpha_i = 1$, $\sum_i \gamma_{ij} = 0$, and $\sum_i \beta_i = 0$. Homogeneity requires $\sum_j \gamma_{ij} = 0$, and symmetry requires $\gamma_{ij} = \gamma_{ji}$. These conditions hold globally, that is, at every data point.

The demand functions for tomatoes, lettuce and carrots were estimated by maximum likelihood estimation methods, and the results were recovered for the cabbage equation from adding up. The estimated elasticities of demand with respect to prices and income have been calculated from the regression coefficients. The income elasticity is given by

$$\eta_i = 1 + \beta_i / w_i$$

and the price elasticities are given by

$$\varepsilon_{ij} = -\delta_{ij} + [\gamma_{ij} - \beta_i(\alpha_j + \sum_k \gamma_{ik} \ln p_k)] / w_i$$

where $\delta_{ij} = 1$ if $i = j$, zero otherwise.

. The data are for the time period, 1981-2004 and prices are retail prices. The almost ideal demand system was estimated with a first-order autoregressive process ($\hat{\rho} = 0.77$ with an associated asymptotic standard error of 0.08). The estimated elasticities for the fresh vegetable subsystem are given in Table 1.

Table 1: Estimated elasticities calculated using the AIDS estimation.

	Estimated AIDS Elasticities			
	Tomato	Carrots	Lettuce	Cabbage
Tomato	**-0.32***** **(0.10)**	-0.03 (0.09)	-0.07 (0.05)	-0.002 (0.02)
Carrots	-1.51* (0.78)	**-0.53*** **(0.21)**	-0.48 (0.37)	-0.33 (0.45)
Lettuce	-0.19*** (0.05)	-0.09 (0.13)	**-0.71***** **(0.20)**	-0.16 (0.72)
Cabbage	-0.01 (0.04)	-0.17 (0.25)	-0.98 (0.88)	**0.12** **(0.55)**
Income	0.89*** (0.14)	1.44*** (0.24)	0.96*** (0.30)	1.06** (0.41)

a) ***: Significant at the .01 level. **Significant at the .05 level. *Significant at the .10 level.
b) Reported standard errors are bootstrap standard errors computed using a subroutine in SAS written by Dr. Barry Goodwin.

The own price elasticity of tomatoes is estimated to be -0.32, which is highly statistically significant. Therefore demand for fresh tomatoes is relatively inelastic with respect to changes in retail prices. The own-price elasticity of carrots is -0.53 and for lettuce it is -0.71. The estimate of the own-price elasticity of cabbage is positive at 0.12, which is counterintuitive. This finding, however, is not statistically significant. The estimated second-stage expenditure elasticities are all positive and range in values from 0.89 to 1.44. In all cases the expenditure elasticities are statistically significant. All of the cross prices elasticities are negative indicating that the four fresh vegetables are complements. Only the complementarities between tomato quantity with carrot and lettuce prices are statistically significant.

Conclusions

Models for both fresh and processed tomatoes were developed and estimated. An almost ideal demand subsystem was estimated for four fresh vegetables that included tomatoes, carrots, lettuce, and cabbage. The second-stage own-price elasticities were all inelastic except for cabbage which was unexpectedly positive. The conditional expenditure or income elasticites varied from 0.89 for fresh tomatoes to 1.44 for carrots. All of the cross-price elasticities were negative indicating that the four fresh vegetables are gross complements. A plausible explanation for this is that the four commodities are used in salads, especially given that no significant complementarities were found with respect to fresh cabbage.

Ordinary least squares and instrumental variable techniques were used to obtain estimated partial adjustment acreage functions of processing tomatoes. The estimated short-run own-price elasticity estimates were between 0.47 and 0.41. Chow tests confirmed a possible structural break in the acreage function for processed tomatoes around 1988. One possible explanation of the break is the increase use of contracts around this time period.

Estimated own-price elasticities for processed tomatoes in the production function varied between 0.45 and 0.55. Producers respond to prices increases in a positive manner, in accordance with theory.

With respect to demand for processing tomatoes, the own-price elasticity was estimated to be -0.18 and the cross-price estimated elasticity of tomato paste on processing tomatoes was 0.16. Thus, as the price tomato paste increases the derived demand for processed tomatoes increases, as expected.

For the second period the estimated own-price elasticity in the acreage equation was 0.23 indicating that producers respond positively to increases in prices. The short-run elasticity of fresh tomato production with respect to price was 0.22 prior to 1987 and 0.27 after 1987. Thus, through out the sampling period, the own-price elasticity in the fresh tomato production function was found to be inelastic.

References

Baylis, K. and J. Perloff, "End Runs Around Trade Restrictions: The Case of the Mexican Tomato Suspension Agreements", *ARE Update, Giannini Foundation Publication,* 9, No. 2, Dec. 2005.

Deaton, A. and J. Muellbauer, *Economics and Consumer Behavior*, Cambridge University Press, Cambridge, 1980.

Lucier, G. "Tomatoes: A Success Story", *Agricultural Outlook*, USDA, July, 1994, 15-17.

Lucier, G. B-H. Lin, J. Allshouse, and L. Kantor, "Factors Affecting Tomato Consumption in the United States", USDA, ERS, Vegetables and Specialties/VGS-282, Nov. 2000, 26-32.

Plummer, C. "modeling the U.S. Processing tomato Industry", Economic Research Services/USDA, Vegetables and Specialties/VGS-279, Nov. 1999, 21-25.

SUMMARY AND FUTURE RESEARCH

This research project developed acreage, yield, production, and demand models for seven California commodities. Both single and system-of-equations models were developed and estimated. The primary findings are: (1) Domestic own-price and income elasticities of demand for California commodities are predominantly inelastic implying that shocks on the supply side will have large impacts on prices and subsequently on revenues. (2) On the supply side producers are responsive to prices. (3) Estimated supply and demand elasticities are important to policy makers in order to measure welfare gains and losses due to various changes in economic conditions. (4) An almost ideal demand subsystem for four fresh vegetables were estimated. Fresh tomatoes, carrots, lettuce, and cabbage were found to have conditional inelastic own-price elasticities (with the exception of cabbage). All had positive conditional expenditure elasticities. In addition, all four fresh vegetables were gross complements. This result is plausible given that the four vegetables are used in salads. And (5) Better data on prices, acreage, demand, production, yields, and other information would enable better analysis of economic conditions facing California producers and consumers. This report has undated the data on acres, prices and yields in a consistent manner. However, additional updating should be continued in the future.

Estimated own-price, cross-price and income elasticities were obtained for the demand and supply functions for six of the top twenty California commodities according to value of production in 2001 (see, Johnston and McCalla, p. 73). The six commodities are: almonds, walnuts, cotton, alfalfa, rice, and processing tomatoes. The report also includes fresh tomatoes. Fresh tomato per capita consumption is increasing relative to the consumption of processing tomatoes. Future work will include grapes-wine, table, and raisins, citrus fruits, and other commodities.

Future research will examine in more depth the problems of heterogeneity and aggregation. Aggregation across consumers, unless strong conditions hold, results in aggregation biases. These can affect the elasticity estimates. There are different approaches to the problem. The distributional approach incorporates distributional changes in consumer income over time as well as distributional changes in consumer attributes. Future work will also address in more depth the issues involved with the export markets, the role of inventories and stocks, and welfare measures of consumers and producers due to various changes. The role of exports are becoming more important as trade barriers are broken down. Domestic producers find themselves players in global competitive markets.

All of the commodities studied in this report require irrigated water and have exhibited expanded acreage. Processing tomatoes production, for example, has grown to about 300,000 acres currently with 64% grown in the San Joaquin Valley. Acreage of almonds in California rose steadily over the years 1970-2001. In 2001 there were over 500 thousands acres in production. Walnut acreage is about 200,000 acres in California in 2001. Alfalfa hay acreage in California averaged about a million acres per year during the past 30 years. In 2002 there were about 700,000 acres planted to cotton in California. A summary of the harvested acres and the total value of production for the commodities examined in this report is given in Table 1.

Table 1. Harvested Acres and Total Value of Production in 2003

	Harvested Acres	Total Value of Production (in $1000)
Almonds	550,000 (bearing acres)	1,600,144
Walnuts	213,000 (bearing acres)	374,900
Cotton	694,000	753,355
Alfalfa	1,090,000	709,590
Rice	507,000	405,974
Tomatoes		
Processing	274,000	529,214
Fresh	34,000	366,180

Source: California Department of Food and Agriculture.

A concise summary of the models and estimated supply and demand elasticities for each commodity are given Tables 2 and 3 below.

Table 2. Estimated Supply and Demand Elasticities for California Commodities

I. Single-Equation Models [a]

Commodities:	Supply Response (Own-Price)		Domestic Demand	
	Short-Run	Long-Run	Own-Price	Income
Almonds	0.12	12.0	-0.48	0.86
Walnuts	0.02	0.08	-0.26	1.21 (0.43) [b]
Alfalfa	0.35-0.66 [c]	1.06	-0.11	1.74 [d]
Cotton	0.53	0.73	-0.68	NA
Rice	0.23	0.27	-0.08	0.74
Tomatoes				
Fresh	0.27 [e]	0.40	-0.25	0.89
Processing	0.41	0.69	-0.18	0.86

[a] The supply-response elasticities were taken from the estimated acreage equation. Various models were estimated and the reported elasticities represent, in the authors' judgment, the most reasonable estimates based on model specifications and efficient econometric estimators.

b. The value in parenthesis represents the income elasticity post 1983 after structural changes had occurred in the industry.
c. The elasticity varied between 0.35 and 0.66 based on different specifications.
d. The demand for alfalfa hay is a derived demand. The figure reported is the elasticity based on the number of cows in the dairy industry.
e. Post 1988.

Table 3. Estimated Supply and Demand Elasticities for California Commodities

II. System of Equations Models

Commodities	Supply Response (Own-Price)		Domestic Demand	
	Short-Run	Long-Run [a]	Own-Price	Income
Almonds	0.24	0.67	-0.69	1.43
Walnuts	0.15	0.19	-0.48	1.01
Cotton	0.46	15.33	-0.95	-0.05
Rice	0.45	0.72	-0.36	0.33
Tomatoes [b]				
Fresh	NA	NA	-0.25 [c]	0.89
Processing	NA	NA	NA	NA

[a] Based on killing off the lags in a single equation in the system.

[b] The fresh tomato elasticities are based on an AIDS model. NA indicates that a system for these commodities was not estimated.

[c] Based on an almost ideal demand fresh vegetables subsystem.

Positive Mathematical Programming

Richard E. Howitt

A method for calibrating models of agricultural production and resource use using nonlinear yield or cost functions is developed. The nonlinear parameters are shown to be implicit in the observed land allocation decisions at a regional or farm level. The method is implemented in three stages and initiated by a constrained linear program. The procedure automatically calibrates the model in terms of output, input use, objective function values and dual values on model constraints. The resulting nonlinear models show smooth responses to parameterization and satisfy the Hicksian conditions for competitive firms.

Key words: calibration, mathematical programming, nonlinear optimization, production model, sectoral model.

This paper is a methodological paper for practitioners rather than theorists. Instead of a new method that requires additional data, I take a different perspective on mathematical programming using a more flexible specification than traditional linear constraints. Sometimes new methodologies are published, but not implemented. Positive mathematical programming (PMP) is a methodology that has been implemented but not published. Over the past eight years the PMP approach has been used on several policy models at the sectoral, regional and farm level. National sectoral models using PMP for the U.S., Canada, and Turkey include House; Ribaudo, Osborn, and Konyar; Horner et al.; and Kasnakoglu and Bauer. Regional models include Hatchett, Horner, and Howitt; Oamek and Johnson; and Quinby and Leuck. Rosen and Sexton apply PMP to individual farms. The PMP approach uses the farmer's crop allocation in the base year to generate self-calibrating models of agricultural production and resource use, consistent with microeconomic theory, that accomodate heterogeneous quality of land and livestock.

Mathematical programming models are widely used for agricultural economic policy analysis, despite few methodological developments in the past decade. Their popularity stems from several sources. First, they can be constructed from a minimal data set. In many cases, analysts are required to construct models for systems where time-series data are absent or are inapplicable due to structural changes in a developing or shifting economy. Second, the constraint structure inherent in programming models is well suited to characterizing resource, environmental, or policy constraints. In some cases, a set of inequality constraints, such as those found in farm commodity programs, strongly influences crop and resource allocation. Third, the Leontief production technology inherent in most programming models has an intrinsic appeal of input determinism when modeling farm production (Just, Zilberman, and Hochman). In addition, linear programming models are consistent with the Von Liebig production specification, which is preferable for several inputs (Paris and Knapp).

While the PMP approach is unconventional in that it employs both programming constraints and "positive" inferences from the base-year crop allocations, it has one strong attraction for applied analysis: it works. That is to say, the PMP approach automatically calibrates models using minimal data, and without using "flexibility" constraints. The resulting models are more flexible in their response to policy changes, and priors on yield variation or supply elasticities can be specified. With modern algorithms and microcomputers, the resulting quadratic programming problems can be readily solved.

Following a brief overview of past approaches to calibrating programming models of farm production and problems associated with these models, the equivalency of the Kuhn

Richard E. Howitt is professor in the Department of Agricultural Economics at University of California, Davis.

The author would like to acknowledge Stephen Hatchett, Quirino Paris, Phillippe Mean, and an anonymous reviewer for comments that improved the manuscript.

Tucker conditions for the constrained and calibrated models are shown, and three propositions that justify the nonlinearity and dimension of the calibration specification are presented. Formal statement and proofs of the propositions are in the appendix. This is followed by presentation of an empirical calibration method with a simplified graphical and numerical example. The final section of the paper addresses some common empirical policy modeling problems. The ability of PMP models to yield smooth parametric functions and nest LP problems within them is briefly discussed.

Calibration Problems in Programming Models

Programming models should calibrate against a base year or an average over several years. Policy analysis based on normative models that show a wide divergence between base period model outcomes and actual production patterns is generally unacceptable. However, models that are tightly constrained can only produce that subset of normative results that the calibration constraints dictate. The policy conclusions are thus bounded by a set of constraints that are expedient for the base year, but often inappropriate under policy changes. This problem is exacerbated when the model is on a regional basis with very few empirical constraints, but with a wide diversity of crop production.

Brevity only permits a brief overview of some of the past calibration methods in mathematical programming models. A more comprehensive discussion can be found in Hazell and Norton or Bauer and Kasnakoglu. It is worth noting that no one approach has proved satisfactory enough to dominate the applied literature.

Previous researchers (e.g., Day) attempt to provide more realism by imposing upper and lower bounds to production levels as constraints. McCarl advocates a decomposition methodology to reconcile sectoral equilibria and farm-level plans. Both of these approaches require additional micro-level data, and result in calibration constraints influencing policy response.

Meister, Chen, and Heady, in their national quadratic programming model, specify 103 producing regions and aggregate the results to ten market regions. Despite this structure, they note the problem of overspecialization and suggest the use of rotational constraints to curtail the overspecialization. However, it is comparatively rare that agronomic practices are fixed at the margin; more commonly they reflect net revenue maximizing trade-offs between yields, costs of production, and externalities between crops. In the latter case, rotations are functions of relative resource scarcity, output prices, and input costs.

Hazell and Norton suggest six tests to validate a sectoral model. The first is a capacity test for overconstrained models; the second is a marginal cost test to ensure that marginal costs of production, including the implicit opportunity costs of fixed inputs, are equal to the output price; and the third is a comparison of the dual value on land with actual rental values. They also advocate three additional comparisons of input use, production level and product price tests. Hazell and Norton show that the percentage of absolute deviation for production and acreage over five sectoral models ranges from 7% to 14%. The constraint structures needed for this validation are not defined.

In contrast, the PMP approach aims to achieve exact calibration in acreage, production, and price. Bauer and Kanakoglu subsequently applied the PMP approach to one of the sectoral models cited by Hazell and Norton. The results for the Turkish Agricultural Sector model (TASM) showed consistent calibration over seven years.

The calibration problem in farm-level, regional, and sectoral models can be mathematically defined by the common situation in which the number of binding constraints in the optimal solution are less than the number of non-zero activities observed in the base solution. If the modeler has enough data to specify a constraint set to reproduce the optimal base-year solution, then additional model calibration will be redundant. The PMP approach is developed for the majority of modelers who, for lack of an empirical justification, data availability, or cost, find that the empirical constraint set does not reproduce the base-year results. The LP solution is an extreme point of the binding constraints. In contrast, the PMP approach views the optimal farm production as a boundary point, which is a combination of binding constraints and first-order conditions.

Relevant constraints should be based on either economic logic or the technical environment under which the agricultural production is operating. Calibration problems are especially prevalent where the constraints represent allocatable inputs, actual rotational limits, and policy constraints. When the basis matrix has a rank less than the number of observed base-year activities, the resulting optimal solution will suffer from overspecialization of production activities compared to the base year.

A source of these problems is that linear programming was originally used as a normative farm planning method assuming full knowledge of the production technology. Under these conditions, any production technology can be represented as a Leontief technology, subject to resource and stepwise constraints. For aggregate policy models, this normative approach produces a production and cost technology that is too simplified due to inadequate knowledge. In most cases, the only regional production data are average or "representative" values for crop yields and inputs.[1] This common situation means that the analyst is attempting to estimate marginal behavioral reactions to policy changes based on average data observations. The average conditions can be assumed to be equal to the marginal conditions only where the policy range is small enough to admit linear technologies.

Two broad approaches have been used to reduce the specialization errors in optimizing models. The demand-based methods use a range of methods to add risk or endogenize prices. These help resolve the problem, but substantial calibration problems remain in many models (Just).

The other common approach is to constrain the crop supply activities by rotational (or flexibility) constraints, or step functions, over multiple activities (Meister, Chen, and Heady). In regional and sectoral models of farm production, there are few empirically justifiable constraints. Land area and soil type are clearly constraints, as is water in some irrigated regions. Crop contracts and quotas, breeding stock, and perennial crops are others. However, it is harder to justify other constraints such as labor, machinery, or crop rotations on short-run marginal production decisions. These inputs are limiting, but only in the sense that once the normal availability is exceeded, the cost-per-unit output increases due to overtime, increased probability of machinery failure, or disease. If the assumption of linear production (cost) technology is retained, the observed output levels imply that additional binding constraints on the optimal solution should be specified. Comprehensive rotational constraints are a common example of this approach.

An alternative explanation to linear technologies with constraints is that the profit function is nonlinear in land for most crops, and that the observed crop allocations are a result of a mix of unconstrained and constrained optima. The most common reasons for a decreasing gross margin per acre are declining yields due to heterogeneous land quality, risk aversion, or increasing costs due to restricted management or machinery capacity.

Given the exhaustive literature on the addition of risk to LP models, I concentrate on calibrating the supply side by introducing a nonlinear yield (or cost) specification for each production activity. While risk is clearly an important determinant of cropping patterns, as shown below, risk alone usually provides insufficient nonlinear calibration terms to completely calibrate a model.

Behavioral Calibration Theory

Calibrating models to observed outcomes is an integral part of constructing physical and engineering models, but it is rarely formally analyzed for optimization models in agricultural economics. In this section I show that observed behavioral reactions provide a basis for model calibration in a formal manner that is consistent with microeconomic theory. By analogy to econometrics, the calibration approach draws a distinction between the two modeling phases of calibration (estimation) and policy prediction.

On a regional level, information on the output levels produced and the land allocations by farmers is usually more accurate than the estimates of crop marginal production costs. This is particularly true with micro data on land class variability, technology, and risk. This information often features in the farmers' decisions, but is absent in the aggregate cost data available to the model builder. Accordingly, the PMP approach uses the observed acreage allocations and outputs to infer marginal cost conditions for each observed regional crop allocation. This inference is based on those parameters that are accurately observed, and the usual profit-maximizing and concavity assumptions.

Proposition 1 (see appendix A) shows that if the model does not calibrate to observed production activities with the full set of general

[1] The paper is written using cropping activities as examples, but the same procedure can be directly applied to livestock fattening and other activities where the key input is not land but a livestock unit, such as a breeding cow. For an example of PMP applied to a wide range of livestock activities in a national model see Bauer and Kasnakoglu.

linear constraints that are empirically justified by the model, a necessary condition for profit maximization is that the objective function be nonlinear in at least some of the activities.

Many regional models have some nonlinear terms in the objective function reflecting endogenous price formation or risk specifications. Although it is well known that the addition of nonlinear terms improves the diversity of the optimal solution, there are usually an insufficient number of independent nonlinear terms to accurately calibrate the model.

Proposition 2 (appendix A) shows that the ability to calibrate the model with complete accuracy depends on the number of nonlinear terms that can be independently calibrated.

The ability to adjust some nonlinear parameters in the objective function, typically the risk aversion coefficient, can improve model calibration. However, with insufficient independent nonlinear terms the model cannot be calibrated precisely. In technical terms, the number of instruments available for model calibration may not span the set of activities that need to be calibrated.

Consider the following problem where the objective function is specified in a general linear or nonlinear form, $f(\mathbf{x})$. For simplicity, and without loss of generality, activities not observed in the base data are removed from the specification.

$$(1) \quad \max_x f(\mathbf{x})$$

subject to

$$\mathbf{Ax} \leq \mathbf{b}$$
$$(\bar{\mathbf{x}} - \boldsymbol{\varepsilon}_1) \leq \mathbf{x} \leq (\bar{\mathbf{x}} - \boldsymbol{\varepsilon}_2) \quad \mathbf{x} \geq 0, \bar{\mathbf{x}} > 0$$
$$\mathbf{x} \text{ is } k \times 1, \mathbf{A} \text{ is } m \times k,\ m < k$$

where the $\boldsymbol{\varepsilon}_i$ perturbations are defined in appendix B.

Let $\bar{\boldsymbol{\lambda}}_1$ be the $m \times 1$ dual solution vector to problem (1) associated with the set of general constraints. The dual values associated with the set of calibration constraints can be ignored in the analysis of the general constraint duals ($\boldsymbol{\lambda}_1$), since proposition 3 (appendix B) shows that the optimal values for $\boldsymbol{\lambda}_1$ are not changed by the addition of the calibration constraints. Define the $k \times 1$ vector $\bar{\boldsymbol{\gamma}}$ as

$$(2) \quad \bar{\boldsymbol{\gamma}} = \nabla \mathbf{f}(\bar{\mathbf{x}})' - \mathbf{A}'\bar{\boldsymbol{\lambda}}_1$$

where $\nabla \mathbf{f}(\mathbf{x})$ is the $1 \times k$ gradient vector of first derivatives of $\mathbf{f}(\mathbf{x})$. Let $\boldsymbol{\alpha}$ be a $k \times 1$ set of constants such that

$$(3) \quad (\bar{\boldsymbol{\gamma}}_i - \boldsymbol{\alpha}_i) \geq 0\ .$$

Define the $k \times k$ diagonal matrix $\boldsymbol{\Gamma}$ as

$$(4) \quad \boldsymbol{\Gamma} = diag\,[(\bar{\gamma}_1 - \alpha_1)/\bar{x}_1,\ \ldots,\ (\bar{\gamma}_k - \alpha_k)/\bar{x}_k].$$

The matrix $\boldsymbol{\Gamma}$ is positive definite by construction.

Consider the following problem:

$$(5) \quad \max_x\ \mathbf{f}(\mathbf{x}) - \frac{1}{2}\mathbf{x}'\boldsymbol{\Gamma}\mathbf{x} - \boldsymbol{\alpha}'\mathbf{x}$$

subject to

$$\mathbf{Ax} \leq \mathbf{b}$$
$$\mathbf{x} \geq 0.$$

The first-order Kuhn-Tucker conditions for this problem are

$$(6) \quad \nabla \mathbf{f}(\mathbf{x})' - \boldsymbol{\Gamma}\mathbf{x} - \boldsymbol{\alpha} - \mathbf{A}'\boldsymbol{\lambda} = 0.$$

From equation (4) we see that $\boldsymbol{\Gamma}\mathbf{x} = (\bar{\boldsymbol{\gamma}} - \boldsymbol{\alpha})$; therefore, substituting for $\boldsymbol{\Gamma}x$ in (6), we get

$$(7) \quad \nabla \mathbf{f}(\mathbf{x}) - \mathbf{A}'\boldsymbol{\lambda} = \bar{\boldsymbol{\gamma}}.$$

From equation (2) we see that the Kuhn-Tucker condition (6) holds exactly when $\mathbf{x} = \bar{\mathbf{x}}$ and $\boldsymbol{\lambda} = \bar{\boldsymbol{\lambda}}_1$. That is, the calibrated problem (5) will optimize at the values $\bar{\mathbf{x}}$ and $\bar{\boldsymbol{\lambda}}_1$ if the values $\boldsymbol{\Gamma}$ and α are defined by equations (3) and (4).

To summarize, given the three propositions in the appendices, linear and nonlinear optimization problems can be calibrated by the addition of a specific number of nonlinear terms. We use a simple quadratic specification to show that if the quadratic parameters satisfy equations (2), (3), and (4), then the resulting quadratic problem will calibrate exactly in the primal and dual values of the original problem, but without inequality calibration constraints.

In the next section I show how the calibration procedure can be simply implemented in a two-stage process that is initiated with a linear program.

An Empirical Calibration Method

The previous section showed that if the correct nonlinear parameters are calculated for the (k m) unconstrained (independent) activities, the model will exactly calibrate to the base-year values without additional constraints. The problem addressed in this section is to show how the calibrating parameters can be simply and automatically calculated using the minimal data set for a base-year LP.

Because nonlinear terms in the supply side of the profit function are needed to calibrate a production model, the task is to define the simplest specification which is consistent with the technological basis of agriculture, microeconomic theory, and the data base available to the modeler.

A highly probable source of nonlinearity on the primal side is heterogeneous land quality, and declining marginal yields as the proportion of a crop in a specific area is increased. This phenomenon, first formalized by Ricardo (Peach), is widely noted by farmers, agronomists, and soil scientists, but often omitted from quantitative production models.

I use a "Primal" PMP approach which keeps the variable cost/acre constant and has a yield function that decreases the marginal crop yield per acre as a linear function of the acreage planted.[2] This specification is consistent with the large body of evidence from soil science and agronomy that shows variability in soil suitability and consequent crop yield in most agricultural areas, whether on the farm or regional scale. The production function in this paper is Leontief with heterogenous and restricted land inputs.

Obviously this is a considerable simplification of the complete production process. Given the applied goal of this "positive" modeling method, the calibration criteria used is not whether the simple production specification is true, but whether it captures the essential behavioral response of farmers, and can be made to work with available restricted data bases and model structures.[3]

The output from a given cropping activity i under the primal PMP specification with land x_i and two other inputs is

$$(8) \qquad y_i = (\beta_i - \delta_i x_i)\min(x_i, a_{i2}x_i, a_{i3}x_i)$$

where β_i and δ_i are, respectively, the intercept and slope of the marginal yield function for crop i.

The calibrated optimization problem equivalent to equation (5), therefore, becomes

$$(9) \qquad \max \sum_i P_i(\beta_i - \delta_i x_i)x_i - \sum_{j=1}^{3} \omega_j a_{ij} x_i$$

subject to

$$\mathbf{Ax} \le \mathbf{b} \text{ and } \mathbf{x} \ge 0$$

where $a_{i1\,=\,1}$, $\mathbf{A} = \ (m \times n)$ with elements a_{ij}, x_i is the acreage of land allocated to crop i, and ω_j is the cost per unit of the jth input.

The PMP calibration approach uses three stages. In the first stage a constrained LP model is used to generate particular dual values. In the second stage, the dual values are used, along with the data based average yield function, to uniquely derive the calibrating yield function parameters. In the third stage, the yield parameters ($\boldsymbol{\beta}$ and $\boldsymbol{\delta}$) are used with the base-year data to specify the PMP model in equation (9). The resulting model calibrates exactly to the base-year solution and original constraint structure.

Figure 1 shows problem (1) in a diagrammatic form for two activities, with $\mathbf{f}(\mathbf{x})$ simplified to $\mathbf{c}'\mathbf{x}$, one resource constraint and two upper-bound calibration constraints. Note that at the optimum, the calibration constraint will be binding for wheat, the activity with the higher average gross margin, while the resource constraint will restrict the acreage of oats.

Two equations are solved for the two unknown yield parameters ($\boldsymbol{\beta}$ and $\boldsymbol{\delta}$). Defining $\mathbf{f}(\mathbf{x})$ as the quadratic total output function specified in (9), the first equation is the average yield for crop i, $\bar{y}_i$

$$(10) \qquad \bar{y}_i = \beta_i - \delta_i x_i.$$

The second equation uses the value of the dual on the LP calibration constraint (λ_2) which is shown below to be the difference between the value average product (VAP) of the crop and the value marginal product (VMP).

The derivation of the two types of dual value $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$, can be shown for the general case using appendix B. The $\mathbf{A}$ matrix in (1) is partitioned by the optimal solution of (1) into an $m \times m$ matrix $\mathbf{B}$ associated with the variables $\mathbf{x_B}$,

[2] Past working papers on PMP, and most of the applications, have specified the nonlinear part of the profit function as originating from an increase in variable cost per acre with constant yields. Both yield and cost changes are probably present; however, data on yield variability are more easily obtained by an empirical modeler than cost variation.

[3] If more complex specifications of the production function are required, Howitt shows how the calibration principles can be extended to include Cobb-Douglas and nested Constant Elasticity of Substitution (CES) production functions.

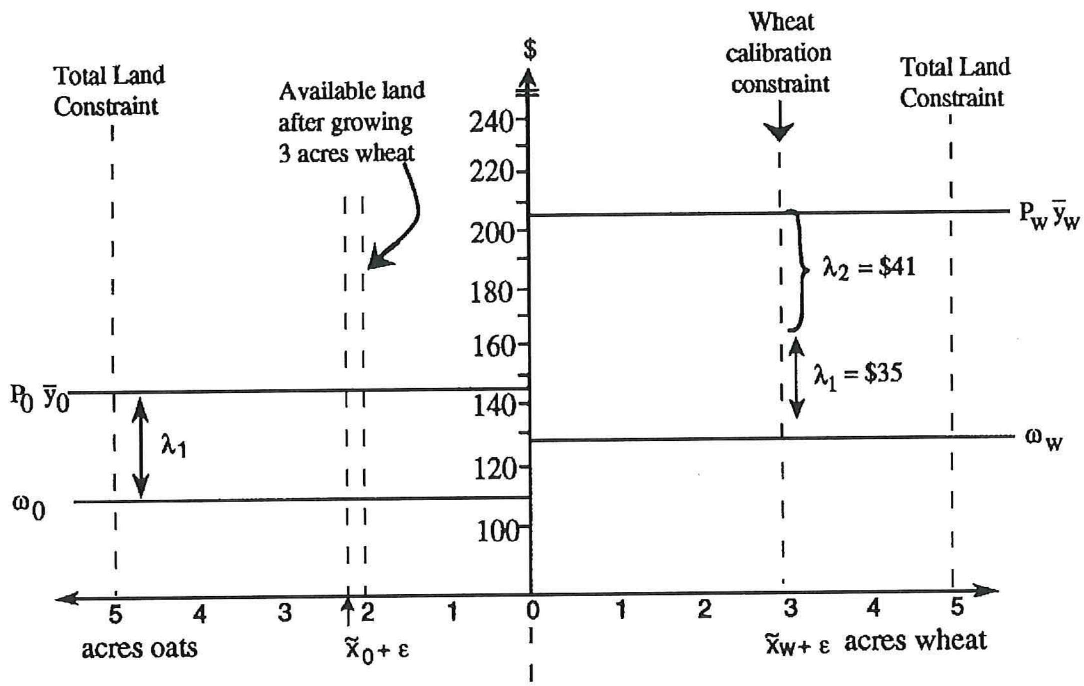

Figure 1. L.P. problem with calibration constraints—two activity/one resource constraint

an $m \times 1$ subset of $\mathbf{x}$ with inactive calibration constraints. The second partition of $\mathbf{A}$ is into an $m \times (k - m)$ matrix $\mathbf{N}$ associated with a $(k - m) \times 1$ partition of $\mathbf{x}$, $\mathbf{x}_N$ of nonzero activities constrained by the calibration constraints. The first partition of equation (B13) in appendix B for $\boldsymbol{\lambda}_1$ is

$$(11) \quad \boldsymbol{\lambda}_1^* = \mathbf{B}'^{-1}\nabla_{\mathbf{x}_B} \mathbf{f}(\mathbf{x}^*)$$

where $\nabla_{\mathbf{x}_B} \mathbf{f}(\mathbf{x}^*)$ is the gradient of value marginal products (VMPs) of the vector $\mathbf{x}_B$ at the optimum value.

The elements of vector x_B are the acreages produced in the crop group limited by the general constraints, and $\boldsymbol{\lambda}_1$ are the dual values associated with the set of $m \times 1$ binding general constraints. Equation (11) states that the value of marginal product of the constraining resources is a function of the revenues from the constrained crops. The more profitable crops ($\mathbf{x}_N$) do not influence the dual value of the resources (proposition 3, appendix B). This is consistent with the principle of opportunity cost in which the marginal net return from a unit increase in the constrained resource determines its opportunity cost. Since the more profitable crops x_N are constrained by the calibration constraints, the less profitable crop group x_B are those that could use the increased resources and, hence, determine the opportunity cost.

The second partition of appendix equation B13 determines the dual values on the upper-bound calibration constraints on the crops

$$(12) \quad \boldsymbol{\lambda}_2 = -\mathbf{N}'\mathbf{B}'^{-1}\nabla_{\mathbf{x}_B} \mathbf{f}(\mathbf{x}^*) + \mathbf{I}\nabla_{\mathbf{x}_N}\mathbf{f}(\mathbf{x}^*)$$

[and substituting equation (11)]

$$\boldsymbol{\lambda}_2 = \nabla_{\mathbf{x}_N}\mathbf{f}(\mathbf{x}^*) - \mathbf{N}'\boldsymbol{\lambda}_1^*$$

Note that the right-hand side of (12) is a $(k - m)$ partition of the right-hand side of (2).

The dual values for the binding calibration constraints are equal to the difference between the marginal revenues for the calibrated crops ($\mathbf{x}_N$) and the marginal opportunity cost of resources used in production of the constrained crops ($\mathbf{x}_B$). Since the stage I problem in figure 1 has a linear objective function, the first term in (12) is the crop average value product of land in activities $\mathbf{x}_N$. The second term in (12) is the marginal value product of land from equation (11). In this PMP specification, the difference between the average and marginal value product of land is attributed to changing land quality. Thus the PMP dual value ($\boldsymbol{\lambda}_2$) is a hedonic measure of the difference between the average and marginal products of land, for the calibrated crops. By analogy to revealed prefer-