Public Comment
Bay Delta Plan Workshop 3
Deadline: 10/26/12 by 12 noon

An Assessment of Models for Measuring the Economic Impact of Changes in Delta Water Supplies

RECEIVED
10-26-12
SWRCB Clerk

David Sunding
Professor, UC Berkeley
Principal, The Brattle Group

Max Auffhammer
Associate Professor, UC Berkeley
Academic Advisor, The Brattle Group

October 24, 2012

This report assesses the extant models for measuring the economic impacts of changes in water supplies from California's Sacramento-San Joaquin Delta. For the urban sector, the principal models are the Least-Cost Planning Simulation Model (LCPSIM), and the more recent Supply-Demand Balance Simulation Model (SDBSIM). The latter model is currently being used to measure the economic benefits of the Bay-Delta Conservation Plan. The principal agricultural impact model is the Statewide Agricultural Production Model (SWAP). This model is the latest in several generations of programming models that measure the change in agricultural resource allocation resulting from changes in surface water supplies in California.

For both urban and agricultural models, the report describes the major features of the model framework, in particular the data and calibration methods used in constructing and applying the models. The report also critically assesses each model and lists areas for further investigation.

On balance, we find that the SDBSIM and SWAP models produce credible estimates of the economic impact of changes in Delta water deliveries. That said, there is room for improvement on many levels. The SWAP model is a structural programming model that relies on a large number of assumptions. It is non-econometric and does not produce standard errors that allow the analyst to assess the statistical significance of results. SWAP is not integrated with a groundwater model, and thus it does not account for changes in groundwater pumping caused by fluctuations in surface water deliveries. As a result, it may underestimate the long-term effects of reductions in Delta deliveries. Further, this report details a number of concerns related to calibration of the SWAP model that may have a large influence on its results.

I. **IMPACT MODELS FOR THE URBAN SECTOR**

This section describes two models for measuring the economic impact of changes in urban water supply reliability. The first model is the Least-Cost Planning Simulation Model (LCPSIM) developed by the California Department of Water Resources. Until recently, this model was the default framework for assessing urban water reliability, especially on the State Water Project. In the past two years, a new framework has been developed by The Brattle Group, working in conjunction with the Metropolitan Water District and the State Water Contractors. The new framework is the Supply-Demand Balance Simulation Model (SDBSIM). It possesses a number of distinct advantages over LCPSIM.

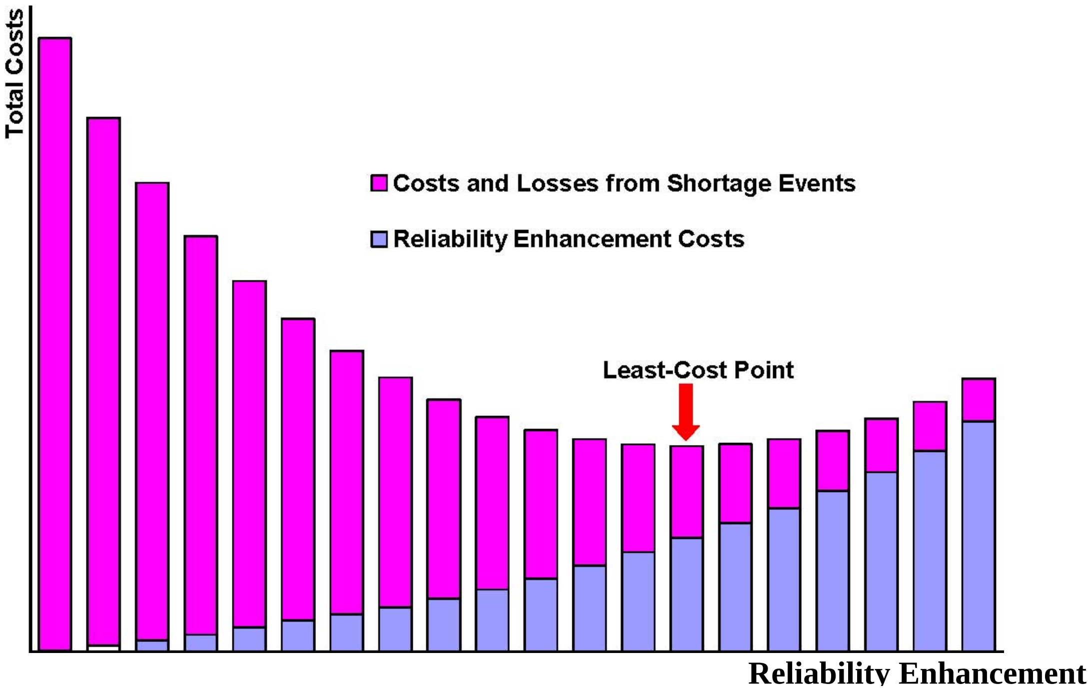
I.A. Least Cost Planning Simulation Model

LCPSIM is a yearly time-step simulation model that was developed and maintained by the California Department of Water Resources and CH2MHill to estimate the economic benefits and costs of improving urban water service reliability at the regional level. The model is similar to load-planning models used in the electricity industry, and simulates a dynamically optimal portfolio of water supplies. LCPSIM has been developed for two regions in California: the South Coast and the San Francisco Bay Area.

Figure 1 LCPSIM Model Regions

The primary objective of the LCPSIM model is to develop an economically efficient regional water management plan by minimizing the total cost of reliability management. Figure 2 illustrates LCPSIM's cost minimization objective. The total cost is the sum of two categories: the cost of reliability enhancement and the cost of unreliable service associated with water shortage events. The cost of reliability enhancement is comprised of three elements: the cost of conservation and recycling, the cost of system operations, and the cost of buying and transferring water. The cost of unreliability is the welfare cost to consumers of a water shortage.[1] This is the

Figure 2 The Effect of Increasing Reliability on Total Cost[2]



value to society of foregone water use that would have otherwise been consumed if not for a shortage. LCPSIM optimizes the degree of reliability over the entire simulation period by determining the portfolio of reliability-enhancing investments that minimizes the cost of these investments plus the cost of shortage in the event that demand cannot be satisfied.

A priority-based objective, mass balance-constrained linear programming solution is used to balance water use with water supply. The model relies on specifically identified demand levels (e.g. year 2020 level) and assesses the water supply portfolio separately given the supply conditions of each hydrologic year between 1922 and 2003.[3] The model keeps track of deliveries to users, deliveries to and from carryover storage, water transfers, and shortage event-related conservation and water allocation programs. Shortages are assumed to be the difference between total demand and the available supply. The LCPSIM allocates the shortages across the

---

1 To access these parameters in LCPSIM: (1) cost of reliability enhancement options: from the RUN/VIEW menu, select VIEW SUMMARY RESULTS and then FULL DISPLAY; (2) system operation costs: from the RUN/VIEW menu, select VIEW LC INCREMENT RESULTS; (3) cost of buying and transferring water: from the RUN/VIEW menu, select VIEW OPERATION TRACE (Excel only) and open the LC Result Report sheet; and (4) cost of a water shortage: from the RUN/VIEW menu, select VIEW OPERATION TRACE (Excel only) and open the LC Result Report sheet.

2 LCPSIM manual- http://www.water.ca.gov/economics/downloads/Models/LCPSIM_Draft_Doc.pdf.

3 LCPSIM manual- http://www.water.ca.gov/economics/downloads/Models/LCPSIM_Draft_Doc.pdf.

different customer classes according to a overall economic cost minimizing rule such that the most shortage is allocated to the sector with the least economic value for water. The highest percentage rate of shortages are allocated to customers who demand water for large landscaping purposes. The next highest shortage is allocated to residential use, followed by commercial, and then industrial. This is because there are less economic impacts from a shortage to the residential sector than there would be to businesses, where a shortage may result in serious adverse economic impacts such as layoffs. Once the shortages are apportioned, the economic losses for all sectors within a region are calculated using a single residential user loss function.

The economic losses are then evaluated against the cost of potential alternative supplies, or reliability enhancement, that could be used to mitigate the water shortages. In the short-run, the model assumes that water transfers and currently existing storage, conservation and recycling programs are available to deal with periodic shortage. In the long-run, the model allows for water supply reliability to be achieved through demand reduction and through supply augmentation from investments including recycling, groundwater storage and recovery, and water transfers. These investments are assumed to reduce frequency, magnitude, and duration of shortage events.[4] They all require capital expenditures to complete, and will take a number of years to implement. The unit costs of these investments vary considerably, and are described in detail in the LCPSIM manual.[5]

The LCPSIM finds the optimal trade off between the cost of alternative supplies and the economic impacts of a water shortage. The alternative supplies are assumed to be employed if the cost of implementation is less than that of the economic impacts from the shortage that results from not having access to these alternative supplies. Alternative supplies are assumed to be realized up until the point that the cost of implementation exceeds the corresponding economic loss given a water shortage. The LCPSIM runs through this process for each of the 82 hydrologic conditions in order to create an expected optimal balance of reliability enhancement and economic loss impacts.

I.B. Supply – Demand Balance Simulation Model

SDBSIM is a probabilistic water portfolio simulation model that apportions and values shortages at the level of the retail agency. The SDBSIM evaluates water shortages in several sectors[6] given demand levels over time and water supply availability for each of the SWP urban agencies. The model runs 83 different trials for each agency by rotating through a historical hydrologic sequence. The shortage and demand outputs are then used to calculate the value of losses to consumers associated with a shortage given a constant elasticity of demand and avoided marginal cost of service.

---

4 http://www.waterplan.water.ca.gov/docs/meeting_materials/ac/01.20.05/LCPSIM-description.pdf.

5 LCPSIM manual- http://www.water.ca.gov/economics/downloads/Models/LCPSIM_Draft_Doc.pdf.

6 All sectors are comprised of single-family residential, multi-family residential, CII (commercial, industrial and institutional), and agriculture.

Forecasting Agency-Level Supply, Demand and Shortages

The water supplies considered in the SDBSIM model consist of the local and imported supplies. Local supplies are comprised of groundwater, local surface water, recycled water, and desalinated water. Imported supplies for Southern California come from the Los Angeles Aqueduct, Colorado River supplies and State Water Project (SWP). Imported supplies for Northern California come from the Hetch Hetchy system, Central Valley Project, and the State Water Project. Individual agencies may have specific import sources; for example, Zone 7 receives imported water from Byron Bethany Irrigation District and the San Diego County Water Authority receives water from Imperial Irrigation District. Estimates of future SWP deliveries from the Delta are forecasted using DWR's CALSIM II model, a generalized water resource simulation that generates hydrological time series forecasts of large, complex river basins.

The projected water demands used in SDBSIM are forecasted using disaggregated econometric models, which capture the impacts of long-term socioeconomic trends on retail demands at the water agency level.[7] These models incorporate projections of demographic and economic covariates that are either forecasted by the agencies themselves or provided by the regional planning agencies Southern California Association of Governments (SCAG) and San Diego Association of Governments (SANDAG).[8] Projections of the covariates are then used to forecast water demand, after which the demand forecasts are adjusted according to expected implementation of conservation programs by individual water agencies. The models forecast demand in five-year intervals for each of the following sectors: single family residential, multifamily residential, commercial/industrial/institutional (CII), and unmetered users. Linear interpolations are generated for the interim years; this results in annual forecasts by sector for each of the SWP urban agencies.

SDBSIM uses an indexed sequential Monte-Carlo simulation method to measure the supply-demand balance outcomes for forecasted years given the pattern of historical hydrologic conditions between years 1922 and 2004. It adjusts the demand and supplies of a forecasted year given a past year of hydrologic conditions, then takes the next sequential forecasted year and adjusts the demand and supplies for that year given the next sequential historical hydrologic year conditions, and so on. For example, SDBSIM would adjust the forecasted demand and supplies for the year 2012 given the hydrologic conditions of the year 1922, and adjust the forecasted demand and supplies of year 2013 given the hydrologic conditions of year 1923, and so on. By preserving the series of climate patterns, or "hydrologic trace", the model is able to capture the operation of storage resources that are drawn upon and refilled over the forecast horizon given a probabilistic sequence of hydrologic conditions. The model then starts over and shifts the hydrologic year by one for each forecasted year. That is, it will adjust the 2012 forecast given the 1923 historical hydrologic conditions, and accordingly will adjust 2013 given 1924 conditions, and so forth. This shifting process is done 83 times such that each forecasted year is evaluated under each hydrologic condition, while still preserving the order of the hydrologic

---

7 The demand for the MWD agencies are forecasted using the MWD – MAIN model. Demand for each of the remaining SWP agencies are forecasted in-house by the agencies.

8 The Metropolitan Water District of Southern California, *Integrated Water Resources Plan: Technical Appendix,* 2010 Update. The underlying figures of the 2010 MWD-MAIN models, with the exception of water rates, rely on the SCAG's 2007 Regional Transportation Plan (TRP-07) and SANDAG's Series 11 Forecast.

conditions, resulting in 83 different reliability outcomes for each forecast year. The model considers the hydrologic conditions of 2004 to be followed by those of year 1922. Thus, when forecasting using a trace that starts with a late hydrologic year, it simply loops back around to the beginning of the climate cycle.

For each year, the SDBSIM model compares the forecasted demand to the sum of available projected local supplies and imported supplies less conservation savings in order to assess the disparity between the amount of water desired and the amount that can be provided. If a shortage exists, the SDBSIM model may release additional supplies from storage or transfer programs until supply and demand are balanced or until these supplies are exhausted. A net shortage for the year results if the gap between supplies and demands is too large to be balanced by storage and transfer programs. If a surplus exists, the SDBSIM model may allocate surplus water to various storage accounts until all storage capacity is used; any remaining surplus supplies are considered unused or "wasted" and are not available for use in subsequent years of the forecast.

There exists considerable uncertainty regarding future hydrological conditions. For example, it is unknown if a major drought like the one experienced in 1924 or 1977 will occur in 2025 or 2050. The timing of such extreme weather patterns may have a significant effect on the value of infrastructure that secures water supply reliability. The advantage of SDBSIM's indexed sequential Monte-Carlo simulation method is that it can account for supply uncertainty by considering 83 different sets of forecasted hydrological time series data and the corresponding supply availability. As suggested earlier, each time series of supply data represents a possible draw from historical hydrological conditions. For example, one SDBSIM simulation uses as input the annual hydrologic conditions from 1922-1960, another SDBSIM simulation uses as input from 1923-1961. In subsequent simulations each year from 1924 to 2004 is considered as the starting year to initialize supply conditions in 2012.[9] In this way, water supply availability between 2012-2050 is computed under a wide range of potential hydrological conditions. Thus, the model produces probabilistic water supply availability given a distribution of potential hydrologic conditions, while also having the ability to predict supply under certain hydrologic conditions.

Valuing Urban Water Supply Reliability

The loss framework utilized in the SDBSIM considers the economic impacts related to water supply interruptions, and emphasizes how water shortages will likely affect ratepayers. The welfare losses during a shortage are determined by the size of the shortage, the forecasted demand, the price elasticity of demand, the utility's pricing structure, and the source of supply unreliability that dictates the avoided marginal maintenance and delivery costs during a shortage.

An important feature of the analysis is that that urban water utilities typically recover capital costs through volumetric prices such that rates are set above marginal cost. The SDBSIM loss

[9] The ordering of years for historical hydrological data is preserved because there is dependence in conditions across years. Hydrological data does not exist beyond 2004. When a simulation requires a time series of hydrologic input data beyond 2004, the time series reverts back to 1922 as the year of hydrological conditions following 2004.

framework recognizes that capital costs are sunk costs, and thus only avoided marginal costs are considered in the loss calculation. Further, it is important to remember that water utilities are public entities so that the impacts of shortages on utilities translate into impacts on ratepayers. In other words, all the welfare loss resulting from a shortage falls on the consumer.

The economic loss calculation in SDBSIM incorporates prevailing water rates in each utility. Water rates combined with observed consumption levels at the prevailing rates provide information about the value of water to households at a single point on the demand curve. Because SDBSIM addresses the economic losses resulting from reducing water consumption below baseline levels, it is necessary to characterize the demand curve at low levels of consumption. The SDBSIM economic loss calculation therefore requires making inferences on consumer willingness to pay for water units at successively higher levels of water rationing, as households are forced to dispense with increasingly higher valued uses of water. To characterize these values, SDBSIM relies on regional water consumption data to estimate demand schedules across households in geographic regions served by individual water purveyors using an econometric model that is capable of explaining water consumption as a function of variables such as rates, income, urban density, and climatic conditions. By comparing agencies over time, and from one place to another, the econometric model traces out more complete demand information than could be gained by looking at a single agency at a single moment in time. As described in subsequent subsections, the results of the statistical analysis are robust and significant at conventional levels used for hypothesis testing.

Theory

The theoretical underpinnings of the SDBSIM loss framework are detailed in the paper by Brozovic, Sunding and Zilberman (2007), who derive an equation for measuring consumer willingness to pay to avoid water service disruptions.[10] Specifically, residential water demand elasticities are estimated for each of *n* retail utilities under a specification of constant elasticity of demand given by:

$$P_i = A_i Q_i^{\frac{1}{\varepsilon_i}}, \qquad i = 1,2,3,\ldots,n, \tag{1}$$

where ε_i is the elasticity of water demand in utility i and A_i is a parameter that scales the magnitude of demand to the price in each agency.

Let P_i^* and Q_i^* respectively denote the retail water price and quantity of water consumed by residential households in utility i under baseline conditions (prior to water rationing). For a given water shortage with an available level of water given by $Q_i(r_i) < Q_i^*$, it is helpful to define the relationship between these quantities in terms of the percentage of water that is rationed in agency i, r_i, as:

$$Q_i(r_i) = (1 - r_i)Q_i^*. \tag{2}$$

[10] Brozovic, N., D. Sunding and D. Zilberman, Estimating Business and Residential Water Supply Interruption Losses from Catastrophic Events, Water Resources Research 43, W08423, 2007 doi:10.1029/2005WR004782.

Making use of equations (1) and (2), consumer willingness to pay to avoid a supply disruption of magnitude $r_i \cdot Q_i^*$ in agency i can be calculated as follows:

$$W_i(r_i) = \int_{Q_i(r_i)}^{Q_i^*} P_i(Q)dQ_i = \int_{Q_i(r_i)}^{Q_i^*} A_i Q_i^{\frac{1}{\varepsilon_i}} dQ_i = \frac{\varepsilon_i}{1+\varepsilon_i} P_i^* Q_i^* [1-(1-r_i)^{\frac{1+\varepsilon_i}{\varepsilon_i}}] \qquad (3)$$

Consumer willingness to pay to avoid a supply disruption in equation (3) can be calculated for each region by constructing an aggregate demand curve to represent the residential water segment (see equation (1)). For regions in which residential customers pay volumetric water rates, P_i^* is the volumetric rate in region i, Q_i^* is the total quantity of water delivered to residences at that price in region i prior to a supply disruption, and ε_i is the elasticity of water demand for region i, which can be estimated from observations of rates and quantities in the region over time along with covariates such as income and weather conditions.

Consumer willingness to pay to avoid a supply disruption in equation (3) depends on the prevailing retail price charged to consumers in each region under baseline supply conditions, P_i^*. Intuitively, the reason for this is that the value of water conservation activities to households in regions with higher water rates is larger than the value of conservation in regions with lower water rates, thus consumers facing higher water rates under baseline supply conditions have greater motivation to engage in conservation activities prior to rationing relative to consumers facing lower baseline water rates. Water conservation is more forthcoming at lower water rates than at higher water rates, and consumer willingness to pay to avoid a given magnitude disruption in water supply is accordingly larger in regions with higher baseline water rates.

The measure of welfare indicated in equation (3) does not account for the avoided costs of service delivery during a shortage. Economic losses that result from water shortage in a given market are mitigated to the extent that delivering a smaller quantity of water reduces the system-wide cost of water service. Because the overall cost of service includes large fixed costs that do not vary with the amount of water delivered through the system (e.g., infrastructure costs, repair and maintenance, administrative expenses, etc.), the avoided cost that results from water shortage is relatively small in relation to total cost. The reduction in the cost of water service that occurs in response to a one-unit reduction in water deliveries is the avoided *marginal* cost of service. Examples of components of avoided marginal cost include the energy and chemical costs of treating water units that are no longer delivered, the reduction in conveyance costs, and the decrease in energy and chemical costs of wastewater treatment that arise from a smaller level of water delivery.

The SDBSIM loss framework assumes that the marginal cost of service delivery is a relatively constant and that it is common across retailers; the delivery cost per unit of water is assumed to be c.[11] This assumption is reasonable—given the lack of data one cannot reject the hypothesis that the cost of service delivery are identical. Once accounting for the avoided cost of service delivery the measure of losses for consumers in retailer i of year t becomes

[11] Avoided marginal cost of service is assumed to be $250 per acre-foot.

$$L_i(r_i) = \frac{\varepsilon_i}{1+\varepsilon_i} \hat{P}_i^* Q_i^* \left[1 - (1 - r_i)^{\frac{1+\varepsilon_i}{\varepsilon_i}} \right] - r_i \cdot Q_i^* \cdot c \quad (4)$$

This framework is easily applied to situations in which shortages may occur in multiple years within a retailer, and across multiple retailers. The framework is also extended in SDBSIM to non-residential sectors including commercial & industrial, agricultural and other uses.

Econometric Model of Urban Water Demand

The econometric model of water demand embedded in SDBSIM is based on a data set covering 119 California water retailers from 1994 to 2011. Although not every retailer is represented in every year, there are over 1,200 price-consumption observation points used to estimate the price elasticity of demand. Data is acquired from direct contact with retailers; price and consumption data are augmented with information obtained from BAWSCA annual reports; consumption data is augmented with data received from the California Department of Water Resources.

The SDBSIM framework links retailer-level price and consumption with demographic variables such as income, household size and lot size (a measure of need for outdoor water use) as well as annual measures of temperature and rainfall. SDBSIM de-means retailer-level consumption to account for shocks common to all retailers within a given year. This modification of the underlying data allows for a comparison of changes in consumption across years due to price changes without confounding changes in statewide hydrological conditions; otherwise, it would be impossible to compare consumption changes in a wet year to consumption changes in a dry year.

After accounting for changes in consumption due to demographic, weather and year-to-year fluctuations in demand, the econometric model of SDBSIM examines the within retailer relationship between any unaccounted for consumption changes and changes in price. In this way it considers a time series of price and consumption data for each retailer to form an overall estimate of consumer willingness to pay to avoid water shortages. Due to the large sample size and consistent relationship between consumption and price, it is possible to perform statistical tests of the price elasticity of demand.

In the literature on residential electricity demand, evidence from consumer electricity consumption in California shows that higher income households have, on average, a lower price elasticity of demand.[12] The SDBSIM econometric model is consistent with similar results in the residential water sector—lower income areas are more responsive to changes in price than higher income areas. As a consequence, higher income areas are more willing to shoulder the burden of avoiding a shortage than relatively lower income areas. The relationship identified by the model between the price elasticity of demand and median income in a service area is statistically significant, which allows the researcher to confidently estimate retailer specific measures of willingness to pay to avoid residential water shortages. As discussed above, accounting for such

[12] Reiss, Peter and Matthew White. 2005. "Household Electricity Demand, Revisited." Review of Economic Studies, 72, p. 853-883.

heterogeneity across retailers is necessary to accurately characterize welfare losses during a shortage. Assuming that all areas value shortage avoidance identically may result in a severe underestimation of welfare losses during a shortage. SDBSIM produces accurate estimates of aggregate welfare losses that can be disaggregated to the agency or retailer level.

Price – Consumption Data Set

The data set used to estimate the SDBSIM econometric model is based on single-family residential Fiscal Year (FY) consumption and prices in 119 California water retailers. The dataset includes 93 retailers in MWDSC service areas and 27 retailers in Northern California (26 agencies belonging to the Bay Area Water Supply and Conservation Association and San Francisco Retail managed by the San Francisco Public Utilities Commission). For the retailers located in the MWDSC service area and San Francisco Retail, historical consumption and rate data from FY 1995-96 through FY 2010-11 were collected directly from retailers with the exception of retailers belonging to MWDOC[13] and SDCWA, for which data was acquired from annual surveys conducted by the wholesale member agencies. For BAWSCA agencies, water consumption and water rates were taken from the BAWSCA Annual Surveys over the period FY 1995-96 through FY 2010-11. The Public Water System Statistics, a survey conducted annually by the Department of Water Resources, is used for retail-level consumption in cases when retailers were not able to provide this data.

SDBSIM uses the sales and accounts data to construct a measure of average monthly household consumption. Average annual single-family household consumption levels for each water agency are calculated by dividing the total single-family residential consumption level for the fiscal year by the number of single-family residential accounts for the given fiscal year. A monthly average consumption level is created by dividing the yearly average by twelve. The construction of the price variable used in SDBSIM requires some additional explanation. In addition to price variation across retailers and time, there is variation in the types of price schedules consumer face. In particular, California retailers use both uniform rate pricing and increasing block pricing. In the former scheme, there is just one uniform rate applied to all water consumed. In the latter, prices depend on how much water has already been consumed in a given month. For example, a retailer using increasing block tiered pricing may charge $1/ccf for the first five ccf in a month, $1.25/ccf for the sixth through twentieth ccf, and $3/ccf for all subsequent units consumed within a month.

Consistent with the conceptual discussion above, SDBSIM uses an equilibrium price of water equal to the price charged ($/ccf) to a residential customer on the median tier in a given year. If the volumetric price ($/ccf) is uniform across all units consumed, then price is set equal to the uniform rate. If there is an increasing block tier structure, then the median rate of all the tiers is assigned as the price. The equilibrium quantity consumed is taken as the monthly average consumption level.

Our price and consumption data set is linked to retailer-specific measures of median income. Retailer-specific measures of median income were constructed based on the 2000 Census using

[13] "Water Systems Operations and Financial Information". April 2011. Municipal Water District of Orange County. Accessible: www.mwdoc.com

area and household density-weighted averages across census tracts comprising the relevant retailer. The first step identifies the area-weighted number of households, n_{ij}, within each Census tract i that intersects with a given retail service area j. In a second step the number of households in intersection ij was used to generate a weighted median income measure for retailer j.

In addition to median income, SDBSIM also uses retailer-specific measures of annual precipitation and summertime maximum temperature. To map weather data, points are geo-referenced at the centroid of each water agency. Based on the resulting set of points, local weather data was extracted from GIS rasters provided by the PRISM Climate Group[14]. In cases when retailer boundaries could not be mapped, a proxy zip code was used to generate the weather data for those retailers.

Model Specification

The regression equation in the SDBSIM econometric model is as follows:

$$\ln(q_{it}) = \beta_1 \cdot \ln(p_{it}) + \beta_2 \cdot \ln(p_{it}) \cdot \ln(income_i) + \beta_3 \cdot W_{it} + \mu_i + \tau_t + \varepsilon_{it} \qquad (8)$$

The subscript i denotes the retailer (i = 1, … , 119), and the subscript t denotes the year (t = 1995, … , 2010). The dependent variable, $\ln(q_{it})$, is the natural log of average monthly household consumption among single-family residential households. The main right-hand side variables of interest are the natural log of price in retailer i of year t, $\ln(p_{it})$, and the natural log of price interacted with the natural log of median household income, $\ln(p_{it}) \cdot \ln(income_i)$. The sum of $\beta_1 + \beta_2 \cdot \ln(income_i)$ is the estimated price elasticity for retailer i. Notice that we obtain heterogeneity in the price elasticities by interacting $\ln(p_{it})$ with the agency-specific measure of $\ln(income_i)$. The regression equation also includes controls for weather with W_{it}, which represents annual precipitation and average summer time max daily temperature in retailer i of year t. The SDBSIM econometric model controls for unobservable factors that may bias the coefficients β_1 and β_2 by including both retailer, μ_i, and year, τ_t, fixed effects. The retailer fixed effects represent a significant advantage of this estimation specification because they control for all time-invariant unobservable characteristics that be correlated with both price and consumption. Any characteristic of a retailer that is time-invariant will be controlled for by SDBSIM.

There may still exist time-varying omitted variables at the retailer-level that bias the SDBSIM coefficients β_1 and β_2. That is, there may exist important unobserved factors that change year to year, and that are correlated with both price and consumption. For example, during a drought there may exist both conservation pricing and intensive conservation campaigns to limit water use. Although the year fixed effects in SDBSIM account for common shocks across all retailers due to drought, there is likely unobserved variation in the intensity of drought and the intensity of conservation campaigns across retail service areas—both may introduce omitted variable bias. The omitted variable, intensity of drought, would likely be positively correlated with water

14 http://www.prism.oregonstate.edu/

consumption and negatively correlated with conservation pricing—such a correlation structure would bias the estimated price elasticities downwards. A second omitted variable, intensity of conservation campaigns, would likely be negatively correlated with water consumption and positively correlated with conservation pricing—such a correlation structure would bias the SDBSIM estimates of the price elasticities upwards. The magnitudes of these biases may be attenuated by the inclusion of the weather variables, strong predictors of drought and conservation campaigns, as retailer-level control variables in the regression. However, if there is residual variation not in these omitted variables which is not captured by weather yet is correlated with both consumption and price, then our point estimates β_1 and β_2 will be biased. SDBSIM breaks the correlation between such omitted variables and price using instrumental variables estimation. Using this estimation strategy, price is first estimated using lagged price according to the following equation:

$$\ln(p_{it}) = \alpha_1 \cdot \ln\left(p_{i,t-1}\right) + \alpha_2 \cdot \ln\left(p_{i,t-1}\right) \cdot \ln(income_i) + \alpha_3 \cdot W_{it} + \theta_i + \rho_t + \eta_{it} \qquad (9)$$

Using the results of the regression in equation (2), the second step is to estimate predicted values of the natural log of p_{it}, $\widetilde{\ln(p_{it})}$, and replace the natural log of price in equation (1) with the predicted values. The final SDBSIM regression equation is[15]:

$$\ln(q_{it}) = \tilde{\beta}_1 \cdot \widetilde{\ln(p_{it})} + \tilde{\beta}_2 \cdot \widetilde{\ln(p_{it})} \cdot \ln(income_i) + \tilde{\beta}_3 \cdot W_{it} + \tilde{\mu}_i + \tilde{\tau}_t + \tilde{\varepsilon}_{it} \qquad (10)$$

where the specification in equation (3) is identical to equation (1) except for the predicted values of log price.

Water Demand Estimation Results

Table 1 presents the estimation results of equation (3). Water rate variables have coefficients significantly different from zero. Importantly, there is a positive and significant coefficient on price interacted with income. This result is evidence that there is statistically significant variation in willingness-to-pay to avoid a shortage according to income levels.

[15] For further explanation of Instrumental Variables estimation see Wooldridge p. 506.

Table 1 Single-Family Residential Demand Estimation Results

	Beta	S.E.	t-stat	p-value	95% C.I.	
Price	-0.415	0.079	-5.26	0	-0.570	-0.260
*Price*Income*	0.108	0.036	3.01	0.003	0.038	0.178
Precipitation	-0.012	0.009	-1.3	0.194	-0.030	0.006
Temperature	0.192	0.114	1.68	0.093	-0.032	0.415
Obs.	1186					
Year FE	Yes					
Retailer FE	Yes					
IV	Yes					

To recover the agency-specific price elasticities, SDBSIM simply takes the sum: $\beta_1 + \beta_2 \cdot \ln(income_i)$, using the agency-specific measures of median income. That is, the price elasticity of agency i equals the sum: -0.415 + .108$\cdot$ $ln(income_i)$. Table 2 shows the range of estimated price elasticities for retailers represented in the SDBSIM model.

Table 2 Estimated Retail-Level Price Elasticities

Retailer	Elasticity
Alameda Co FC & WCD Zone 7	-0.187
Alameda County W.D.	-0.197
Anaheim	-0.241
Antelope Valley East Kern	-0.208
Beverly Hills	-0.198
Burbank	-0.244
Calleguas MWD	-0.198
Castaic Lake WA	-0.198
Central Basin MWD	-0.257
City of Santa Maria	-0.268
Compton	-0.287
Eastern MWD	-0.261
Foothill MWD	-0.202
Fullerton	-0.234
Glendale	-0.251
Inland Empire Utilities Agency	-0.236
Las Virgenes MWD	-0.173
Long Beach	-0.262
Los Angeles	-0.259
MWD of Orange County	-0.210
Mojave WA	-0.261
Palmdale	-0.273
Pasadena	-0.241
San Bernardino Valley MWD	-0.322
San Diego County Water Authority	-0.240
San Fernando	-0.268
San Gorgonio Pass Water Agency	-0.282
San Marino	-0.146
Santa Ana	-0.254
Santa Clara Valley Water District	-0.189
Santa Monica	-0.231
Three Valleys MWD	-0.226
Torrance	-0.230
Upper San Gabriel Valley MWD	-0.247
West Basin MWD	-0.229
Western MWD of Riverside County	-0.241

Estimating Welfare Losses from Water Shortages

SDBSIM uses retailer-specific information on retail prices paid by customers, price elasticities of demand for various sectors, marginal costs of service delivery, and both forecasted demand and shortages for urban agencies receiving Delta water supplies[16.] The median tier price of each agency is collected to construct an agency specific price index. Price elasticities are estimated as previously described. The forecasted demand and shortages are based on the SDBSIM projections over the 83 hydrologic years.

The calculation of losses in SDBSIM is a multi-step process that starts at the level of the retailer. Losses are evaluated separately for each forecasted year for each SWP agency using its own specific economic conditions (baseline price, baseline demand, shortage level, and price elasticity). As suggested, the single-family residential sector receives the majority of shortage allocation; the exact shortage allocation rule is as follows. If an urban agency experiences a shortage in a given year then the first shortage allocation goes to the agricultural sector, which may have its supply reduced by up to 30 percent. Not all urban agencies have an agricultural supply allocation, and if they do, then it is a small sector relative to total agency water demand—as a consequence, the supply reduction in agriculture due to a shortage is small in absolute terms. Hence, the agricultural sector absorbs a relatively small share of a shortage. If there still exists a shortage after reducing the agricultural sector's supply then the next units of shortage are assigned to the single-family residential sector. The single-family residential sector is assigned up to a 30 percent supply reduction before a shortage allocation is made to the commercial and industrial sectors. The rationale for this assumption is that the single-family residential sector has more discretionary water-use, for example, outdoor water use. In a few instances, the projected shortages are so large that a full 30 percent supply reduction occurs in each of the agricultural and single-family residential sectors along with a 20 percent supply reduction in the C&I sector. In these cases, SDBSIM assigns a value of $3,000 per acre-foot to any shortage remaining.

Once an agency-level shortage in a given year has been allocated across the agricultural, single family residential and C&I sectors, SDBSIM calculates the welfare loss in each sector. For the welfare calculations the model uses the price elasticities estimated in the previous section for the single family residential sector; these range from -0.322 to -0.146. For the agricultural sector, SDBSIM employs an elasticity of -0.80, and for the C&I sector, it uses an elasticity of -0.10. These elasticities are consistent with the shortage allocation strategy in which shortage assignments are first made to the agricultural sector which has the lowest value of water, then to the single-family residential sector, and finally to the C&I sector which has the highest valuation of water.

Once the losses have been calculated they are then aggregated across agencies in SDBSIM to generate a measure of total annual losses. The total annual losses are discounted to the present using a 2.275 percent real discount rate. To account for the uncertainty of the timing of shortages this process of loss valuation is conducted for each of 83 unique hydrological trajectories.

---

16 With the exception of Kern County Water Authority.

Comments on Water Rates

The loss function in SDBSIM is dependent on baseline prices for each member agency, therefore, the definition of agency-level water rates will affect the value of water supply reliability we calculate. The index price used to characterize the water rate for households in each region is calculated from the summer rate schedule in cases where water retailers charge seasonal water rates to residential customers. For water retailers that charge volumetric rates for water, the index price used for households in the region is the volumetric price. For water retailers that implement a tiered rate structure, the relevant rate for the economic loss calculation depends on how prices are adjusted across tiers to implement a needed conservation level. SDBSIM assumes that voluntary conservation measures are adopted in proportion to household consumption levels (i.e., that all households respond to a 10 percent conservation need by cutting back water use by 10%), so that conservation is no more likely to occur among customers on any particular tier of the rate structure. The assumption of proportional adjustment of water use on all rate tiers leads to a conservative measure of index prices in the sense that conservation may be more forthcoming among households on higher pricing tiers and because agencies implementing conservation through price changes may raise water rates to a greater degree on higher rate tiers than on lower rate tiers (or alternatively reduce the quantity of water that qualifies for the lower rates), facilitating a disproportionate level of conservation on higher tiers of the rate structure than on lower tiers of the rate structure.

Under proportional rate adjustment, the relevant water rate for the economic loss calculation in equation (7) is a weighted average of the prices paid by each household in the service area for the last unit of water consumed. For many water retailers, water rates involve an inclining tiered structure, and the price index in equation (6) depends on the distribution of individual households across the pricing tiers, with the relevant rate for each household comprised of the rate paid for the last unit of water (i.e., the highest tier on which consumption takes place). That is, the price index is an index of marginal rates, which exceeds the average rate paid by households (total sales revenue divided by total water deliveries) because households on higher pricing tiers also pay lower rates on a portion of water purchased on lower tiers.

Put differently, the price index in equation (6) would accord with the average rate paid by households (total sales revenue divided by total water deliveries) for each water retailer in the case where all water units consumed by a household are priced at the rate on the highest tier of consumption. Because information is not available to construct such a price index, the index price for each water retailer is taken to be the price on the median tier of the inclining block rate structure. For most water retailers, the typical rate paid by single-family households for the last unit of consumption in summer months turns out to align with the median tier in the rate structure (frequently the second tier in a three-tiered rate structure), which is consistent with our choice of price index. The rates used in SDBSIM are net of any additional surcharges charged to customers at higher elevation zones, as cost premiums to higher elevation zones are assumed to be offset by the higher costs of pumping to these zones.

I.C. Comparing LCPSIM and SDBSIM

The SDBSIM is preferable to the LCPSIM approach as it does not just evaluate the water supply portfolio in one year of demand, but relies on demand forecasts out to 2050 in order to evaluate the water need over time. Concurrently, the SDBSIM does not only gauge the supply availability given one particular hydrologic condition at a time, as in the LCPSIM, but runs through a historical sequence of hydrologic conditions given an evolving demand. This approach captures an additional dimension of the water supply portfolio that accounts for the previous years supplies that may affect the current state through potential storage supply availability. In addition, the SDBSIM runs through 83 different possible hydrologic sequences that are matched up to forecasted water demand, creating a full picture of possible portfolio outcomes.

SDBSIM's agency level analysis, as opposed to the LCPSIM's regional level analysis, is another benefit of the model because it reveals the agency-specific value of a water shortage. This granular analysis is sought after because it more accurately assesses water supply availability throughout the model over time. The LCPSIM assumes that there are facilities and institutional agreements in place to move water as needed within a region to minimize the impact of shortage events. These assumptions are generally true for the regions assessed in the LCPSIM, however they are not appropriate for all areas and could lead to biased calculations of the benefits from additional reliability supplies. SDBSIM's agency level analysis avoids these biases, as well as makes it possible to better compute the value of agency-specific economic impacts of the resulting shortages. Since each agency within the region may have drastically different value for a unit of water, a disaggregated approach leads to a more precise representation of water value rather than valuing all agencies within a region under the same water demand estimation. Furthermore, the SDBSIM has separate loss functions for each sector within the agency while the LCPSIM uses a single residential loss function across all sectors. As a result, the SDBSIM is better able to capture the varying economic impacts across sectors.

Finally, the SDBSIM is a more desirable approach because it is a simulation model and does not endogenize water supply alternatives. The LCPSIM assumes certain levels of alternative supplies and associated costs that are integrated into the optimization of the model. This approach is problematic because it is difficult to accurately map the supply curve for alternative supplies due to the typically location-specific nature of their costs. Since LCPSIM hinges on these assumptions, the resulting welfare impact analysis will likely be biased as a result. Moreover, the implementation of new alternative supplies would most likely result in increased water prices in an attempt to recover the cost of the project. The LCPSIM does not take into account the likely decline in water demand in response to the price hike. This may also lead to biased results. Conversely, the SDBSIM treats alternative supplies as exogenous inputs. The model is flexible such that current and future alternative supplies can be selected for the model and no assumptions on costs are necessary.

II. IMPACT MODELS FOR THE AGRICULTURAL SECTOR

The Statewide Agricultural Production Model (SWAP) is an optimization model of California's agricultural economy, developed for use as a policy analysis and planning tool. The model is developed and maintained by researchers at UC Davis. It has been applied in numerous studies of California agriculture, including analyses conducted for the Bureau of Reclamation and the California Department of Water Resources.

II. A. Description of SWAP

SWAP is calibrated using the technique of Positive Mathematical Programming (PMP), which relies on observed data to deduce the marginal impacts of future policy changes on cropping patterns, water use, and economic performance (Howitt 1995). As a multi-input, multi-output model, SWAP determines the optimal crop mix, water supplies, and other farm inputs necessary to maximize profit subject to heterogenous agricultural yields, prices, and costs. SWAP's outcomes reflect the impacts of environmental constraints on land and water availability, and can be adapted to reflect any number of additional policy or technological constraints on farm production.

The PMP approach taken by SWAP allows for calibration of parameters that exactly match base-year conditions, using observed data on land use, farmer behavior, and other exogenous information. Under the fundamental assumption of profit-maximizing behavior by farmers, the model uses a non-linear objective function to derive parameters that satisfy first-order conditions for optimization under the base year's observed input and output data. While aggregate data on variables such as crop yield and acreage is often available, it is much more difficult to estimate a crop's marginal production costs. In lieu of relying on these often inaccurate estimates, the PMP technique uses the more reliable aggregate data to infer the marginal costs of production for each crop in a given region.

Aggregate data used in SWAP comes from a variety of sources. Crops are aggregated into 20 categories defined in collaboration with the California Department of Water Resources (DWR), with a proxy crop identified to represent production costs and returns for each category. Input costs and yields for the proxy crops are derived from the regional cost and return studies from the UC Extension Crop Budgets (UCCE 2011). Base applied water requirements are derived from DWR estimates (DWR 2010). Commodity prices from the model's base year are obtained from the California County Agricultural Commissioner's reports. County-level data are aggregated to a total of 37 agricultural subregions, based off of DWR Detailed Analysis Units (DAU). The SWAP regions aggregate one or more DAUs, which are chosen based on similar microclimate, water availability, and production conditions.

The SWAP model specifically accounts for both surface and groundwater supplies. In total, the SWAP model considers a number of types of surface water: State Water Project (SWP) delivery, Central Valley Project (CVP) delivery, and local deliveries or direct diversions. Where applicable, water costs include both the SWP and CVP charge as well as a district's charge. For groundwater, the model includes both the fixed costs of pumping as well as variable costs based off O&M and energy costs. For more detailed estimation of costs associated with long-run depth to groundwater changes, the SWAP model can be further linked to a separate groundwater model.

Using the input data sources described above, the SWAP model solves a PMP calibration function specified as follows for agricultural regions *g*, crop types *i*, production inputs *j*, and water sources *w*:

$$\underset{xl_{gi,land},watl_{gw}}{\text{Max}} \prod = \sum_{g}\sum_{i}(v_{gi}yld_{gi} - \sum_{j\neq water}\omega_{gij}\,\alpha_{gij})xl_{gi,land} - \sum_{g}\sum_{w}(watl_{gw}\varpi_{gw})$$

The terms $xl_{gi,land}$and $watl_{gw}$ signify land and water use. Region-specific crop prices and yield are represented by v_{gi} and yld_{gi}, while ω_{gi} and $\varpi_{gi}$are input and water costs. α_{gij} are regional Leontief coefficients, depicting the observed level of input use for each crop in each region. Farm production is constrained by the availability of land and water, which are separated in the calibration given that any individual region may be constrained by either one of the two. The land and water constraints are defined as

$$\sum_{i} xl_{gi} \leq b_{g,land} \quad \forall\, g$$

and

$$\sum_{w} watl_{gw} \leq \sum_{w} watcons_{gw} \quad \forall\, g$$

where $b_{g,land}$ and $watcons_{gw}$are land and water availability constraints in each region. The PMP approach calculates imputed "shadow values" the constraining inputs, which reflect the true value of an additional unit of land or water in the region. Each additional unit or land or water allows for additional agricultural output, which will be dependent on the crop produced and the price for that crop in the regional market. The imputed shadow values are thus a function of the revenues from constrained crops, and reveal each region's willingness-to-pay for additional units of constrained inputs as a function of their productive opportunities.

In addition to the resource shadow values for land and water, the addition of a calibration constraint forces the program to optimize according to observed base year cropping patterns. As detailed in Howitt (1995), an arbitrarily small number is included as a perturbation term (ε) to decouple the resource and calibration constraints:

$$x_{gi,land} \leq \tilde{x}_{gi,land} + \varepsilon \quad \forall\, g, i \qquad \varepsilon = 0.0001$$

The more profitable crops in the model will end up limited by the calibration constraints. The less profitable crops are not constrained by the calibration value and therefore determine the shadow values of the constrained input resources, in this case those of land and water. The shadow values on land and water are thus set equal to the marginal net return of a unit increase in those resources, which is a function of revenues from the constrained crops.

The imputed values from PMP calibration are next used to parameterize regional production functions for each crop. The production functions are specified using a constant elasticity of substitution (CES) and have constant returns to scale, as the total value of production is allocated exactly among the different inputs. The use of the CES production function allows for substitution of inputs at a specified substitution elasticities. For example, applied water rates could be partially substituted for by improving irrigation efficiency through capital expenditures on improved irrigation technology (although care should be taken to account for improvements in irrigation technology that have already occurred). The CES functions are defined as

$$y_{gi} = \tau_{gi}[\beta_{gi1}{x^{\rho_i}}_{gi11}\beta_{gi2}{x^{\rho_i}}_{gi2} + \cdots + \beta_{gij}{x^{\rho_i}}_{gij}]^{1/\rho_i}$$

Where y_{gi} represents output of crop i in region g based on the combined inputs j. The relative use of different production factors is depicted by the share parameters β_{gij}. Scale parameters are given by τ_{gi}, and x_{gij} represents production factor usage. If data is available, specific substitution elasticities can be estimated and applied. Alternatively, a fixed value for all inputs can be used. Assuming a constant elasticity of substitution σ, $\rho_i = \frac{\sigma-1}{\sigma}$.

Optimal input allocation is determined by the first order condition which sets the value of marginal product from each input equal to the marginal cash cost plus opportunity cost for that input. Using the shadow values calculated in the PMP calibration step, this value will be equal to the base input price plus the shadow values on the constrained resources. For crops bound by the calibration constraint, the calibration shadow value is additionally added. This process can be generalized for any number of regions and crops. Under the constraint of constant returns to scale, one can algebraically solve for the share values β_{gij}. Since the value of total production y is known, substituting in the calculated share values allows for final calculation of the scale parameter τ.

The next step in the SWAP model is estimation of an exponential land cost function, using information on acreage response elasticities and the calibration constraint shadow value. The use of an exponential cost function avoids problems associated with quadratic cost functions that estimate a linear marginal cost for land. Namely, linear estimates can result in negative marginal costs over a range of low land areas, forcing a modeler to adopt unrealistic marginal production costs near the lower bound in order to fit a desired supply elasticity. The use of an exponential cost function, on the other hand, bounds marginal costs above zero and thus avoids this problem. The total land cost function is defined as

$$TC_{gi} = \delta_{gi}e^{\gamma_{gi}x_{gi.,land}}$$

where δ is the minimum fixed cost of producing crop i in region g, and γ is the response function's elasticity parameter. These parameters are calculated by regressing the calibration shadow value of land against the observed base level of land use and the elasticity of supply for each crop group.

Agricultural prices in the SWAP model are treated as endogenous by calculating individual demand functions for each crop group. First, a statewide demand function for each crop is calculated using crop demand elasticities estimated by Green et al. (2008). The specified downward-sloping demand curves represent consumers' willingness to pay for each individual crop. All else equal, as production of a given crop increases its price is expected to decrease. While the statewide price is assumed to be constant across all modeled regions, regional prices are allowed to deviate due to region specific differences in production levels, crop quality, climate, and other factors.

The individual crop demand functions are specified as

$$p_i = \xi \alpha_i^1 - \alpha_i^2 \left(\sum_g \sum_j y_{gij} \right)$$

where p_i is crop price, α_i^1 and α_i^2 represent the intercept and slope of the demand curve, and ξ allows for a shift in demand due to further exogenous factors. To calculate the statewide California crop price, observed prices are weighted by the relative proportion of statewide production in each region *g*. Subtracting the statewide price from regional observed prices yields the regional marketing cost rmc_{gi}, reflecting differences in prices due to region-specific factors.
At this point, the calibrated functions are aggregated into a nonlinear profit maximization program which considers farm production optimization and considers the previously specified CES production functions, crop- and region-specific exponential land cost functions, and crop demand functions specified above. Accounting for endogenous crop prices, the program maximizes the sum of producer and consumer surplus as follows:

$$\begin{aligned} \underset{x_{gij}, wat_{gw}}{\text{Max}} \; PS + CS = & \sum_i \left(\xi \alpha_i^1 \left(\sum_g y_{gi} \right) + \frac{1}{2} \alpha_i^2 \left(\sum_g y_{gi} \right)^2 \right) \\ & + \sum_g \sum_i \left(rm_{gi} \left(\sum_j y_{gi} \right) \right) \\ & - \sum_g \sum_i \left(\delta_{gi} \exp(\gamma_{gi} x_{gi,land}) \right) \\ & - \sum_g \sum_i \left(\omega_{gi,supply} x_{gi,supply} + \omega_{gi,labor} x_{gi,labor} \right) \\ & - \sum_g \sum_i \left(\varpi_{gw} wat_{gw} \right) \end{aligned}$$

The program optimizes for each region *g*, crop *i*, and water source *w*. The four production inputs are written out separately, as land cost is estimated by the exponential cost function, and water costs vary by source. The first term in the above equation is equal to the sum of gross revenue plus consumer surplus for each crop in each region. The second term represents region-specific additional revenue from regional crop prices higher than the statewide base price. The third term represents total land costs, the fourth represents total labor and supply costs, and the fifth and final term represents total water costs.

The authors of the SWAP model apply additional constraints to ensure the estimation of realistic outcomes (Howitt 2012). Simple input and water constraints limit model output according to the total input availability in each region. While the CES production function allows for substitution between inputs, the model is further constrained to prevent the model from reducing applied water rates below those normally observed. This ensures that applied water levels under stress irrigation are not unreasonably low.

Further constraints include limiting the amount of perennial crops which can be retired, as farmers would be expected to devote resources in the short run to preserving established perennial stands that have large investment costs. Limiting the amount of perennial retirement assumes that only older stands near retirement would be taken out of production (an assumption that may not be realistic). Additionally, a silage constraint is added to ensure that produced crops continue to meet the regional feed requirements of California dairy herds.

The model is extensible in that any number of additional constraints can be added to more accurately depict agronomic, environmental, or political conditions in an applied setting. However, some constraints may need to be relaxed in order for the model to calibrate properly. A final overall test of calibration for the model examines the difference in input allocation and production outputs between the base data and the modeled outcome, which should be nearly identical.

At this point, if the calibration test is specified the model is ready for use in policy application and sensitivity analysis. There are three fundamental assumptions that are important to note. First, the model assumes water is interchangeable among all crops in a region. Second, farmers are expected to act in a way that maximizes annual profits, by equating the marginal revenue of water to its marginal cost. Finally, it is assumed that each region adopts a crop mix that will maximize regional profits.

II.B. An Assessment of the SWAP Model

Input Data

For each crop group the modelers choose a representative crop, which reflects the variable input costs for each crop group. The table below taken from the SWAP documentation lists the crop croups and the chosen representative crop:

SWAP Crop Group	Proxy Crop	Other Crops
Almonds and Pistachios	Almonds	Pistachios
Alfalfa	Alfalfa Hay	
Corn	Grain Corn	Corn Silage
Cotton	Pima Cotton	Upland Cotton
Cucurbits	Summer Squash	Melons, Cucumbers, Pumpkins
Dry Beans	Dry Beans	Lima Beans
Fresh Tomatoes	Fresh Tomatoes	
Grain	Wheat	Oats, Sorghum, Barley
Onions and Garlic	Dry Onions	Fresh Onions, Garlic
Other Deciduous	Walnuts	Peaches, Plums, Apples
Other Field	Sudan Grass Hay	Other Silage
Other Truck	Broccoli	Carrots, Peppers, Lettuce, Other Vegetables
Pasture	Irrigated Pasture	
Potatoes	White Potatoes	
Rice	Rice	Wild Rice
Safflower	Safflower	
Sugar Beet	Sugar Beets	
Subtropical	Oranges	Lemons, Misc. Citrus, Olives
Vine	Wine Grapes	Table Grapes, Raisins

For the group grain, for example, wheat is chosen as the proxy crop for wheat, sorghum, oats and barley. The model obtains input costs of land, labor and other supplies from the University of California Cooperative Extension (2011) cost and return studies. These studies are location and crop specific detailed studies, which aid farmers in obtaining best practice estimates of input costs for a given farming technique and location for a given crop. For wheat, for example, there are two studies (2008 and 2009) which provide such estimates for irrigated wheat and wheat “for grain”. The irrigated wheat study is for the Sacramento Valley and the “wheat for grain” study is for the lower San Joaquin Valley. There are no other studies available for wheat in the database the modelers cite in the model documentation.

The notes significant year to year variation in yields and extrapolates out of sample: “Reported average wheat yields in Sacramento Valley over the past ten years ranged from 1.56 to 2.58 tons per acre. [...] In this study 3.0 tons per acre is used.” The 3.0 tons per acre figure is outside the range of what is actually observed in the data for this location. For irrigated wheat, the report indicates per acre operating costs of $351 per acre.

The wheat for grain study on the other hand uses yields of 3.5 tons per acre and notes a range of 2-4.5 tons per acre. The per acre operating costs are approximately $490 per ton. These are the only two data points available for wheat in all of California. These two studies are used to proxy for oats, sorghum and barley throughout California - in areas where wheat is currently grown and areas where it is not.

One could check these numbers against an available report for grain sorghum, which is available for the South San Joaquin Valley. Yields are assumed to be 4 tons per acre with a range of 2-5 tons. Total operating costs are $464 per acre. These are significant differences from the numbers for wheat which proxies for Sorghum. To illustrate the differences within group one need only look at labor requirements. The labor requirements differ quite a bit across these crops and locations. The irrigated wheat San Joaquin South study estimates 1.57 hours of labor per acre. The Sorghum study 2.17 hour and the wheat for grain study estimates 3.17 hours of labor per acre. This is 100% difference across grains within a group. The difference between sorghum and wheat in the same region is still significant (39% difference). A 39% difference in labor intensity is anything but marginal. Further the irrigation requirements by crop vary significantly The reports state that sorghum requires 30 acre inches, while wheat requires 20 acre inches, which is a 50% difference. This puts in to question the representativeness of the parameters of these costs studies as a location specific estimate for a given crop group.

Further, there is less than complete coverage for these crops across space and there is significant variation in input requirements across space as the above listed examples illustrate – even within crop group.

Water Requirement Data

While the UC Extension studies do provide estimates of water requirements for the studied crops, which resulted in the estimated yields further used by SWAP, the model does not use these, but instead uses the DWR obtained estimates, which have spatially broader coverage. The “Annual Land & Water Use Estimates” provide estimates of applied water by crop and DAU. While there is broad coverage of crops, these data are only publicly available through 2001. For any given region the number in the database for a given crop is zero as it only appears if there was observed (irrigated) area for a crop. This is problematic for a model like SWAP, which of course allows for switching into crops, which were previously not grown in a certain area. The location specific water requirements are an important determinant of this decision and appear not to be observed. For the 277 DAUs and 20 crops in 1999 there should be 5540 irrigation intensities. For 4232 out of these 5540 a zero is recorded, which indicates a missing value rate of 76%. This is of course not by intent, but rather by the fact that crops are grown in very specific areas and we do not observe most crops grown in most areas. From a practical perspective, this feature is partially offset by the fact that SWAP aggregates DAUs into SWAP regions, but the fundamental issue of missing values remains.

We have obtained only the aggregated SWAP region specific numbers for land use and applied water for 2005. Even after aggregation there remains a significant number of missing values. The issue of within region and crop group heterogeneity cannot be addressed by using these data.

Land Use Data

The land use data come from the same source as the water requirement data. Land use is aggregated to the SWAP regions for the crop groups for a single year (2005). This is clearly problematic, as one has no idea as to what the available land for crop production by crop group is. The SWAP documentation is silent on this matter. In a single year, some land will lie fallow. Further, in some years previously unused marginal lands may be converted to farmland.

Finally, not all land can produce any crop. A good example of this phenomenon is the area of Westlands Water District that is impacted by shallow groundwater. SWAP does not utilize readily available GIS layers of soil characteristics to determine the available amount of land for a given crop and the potentially available land for irrigated and non-irrigated agricultural production. There are other land use data layers available (e.g., the USDA's NASS layers) that could inform or help verify the DWR provided land use data and verify the model's fit for more than a single period.

Aggregation

SWAP aggregates production into production regions. Some of these regions follow natural boundaries such as water districts with similar water supplies (e.g., Westlands Water District). Other SWAP regions are not well defined, however. For example, SWAP Region 19 includes state and federal districts, and areas without any groundwater availability. These districts are heterogeneous enough that they should not be aggregated. Further, assuming unrestricted water trading among these districts is not realistic.

Groundwater Extraction

SWAP is not integrated with any groundwater model, and treats groundwater availability as exogenous. Thus, it does not capture the fact that if there is significant groundwater extraction the water table may fall and pumping costs may rise. In reality, groundwater extraction costs are endogenous in the long run, and will influence the shadow value of surface water.

We would add that to the extent that the SWAP does not account for variability in the quality of groundwater and its ability to serve as a suitable replacement supply for all types of crops, then the model would tend to underestimate the impacts of reductions in surface supplies. Further, it is uncertain if the model accounts for the availability of groundwater, the installed capacity to pump, and/or the ability to transport groundwater to places without availability/capacity.

Concerns about the Linear Calibration Program

In step II SWAP maximizes farmer profits and uses observed 2005 land values as the calibration constraints. The model for each region chooses land use for crop i in region g as well as water use for crop i and region g. What is known here are region specific crop prices, yields per acre for each crop and region, input and water prices as well as observed input use for all inputs except for water. The profits are maximized subject to a region specific total land constraint and source specific water consumption constraints. The model is solved using a numerical optimization routine and designed to reproduce 2005 land use patterns almost exactly.

There are a few points of concern with this approach. The model takes as given and correct the 2005 land distribution, input prices, water prices, and total land constraints. All of these values are taken as given known exactly here in order to get the all important shadow values for land

and water. If these values are observed with error, the shadow values may change radically. Howitt eloquently discusses the sensitivity of these models to changes in the input data in a recent *ARE Update* article.

Concerns about the Production Function Parameter Calibration

Step III uses Howitt's PMP approach to derive parameters for a Constant Returns to Scale Constant Elasticity of Substitution Production Function for each region and crop. While one must make some parametric assumptions in any empirical study, the choice of functional form and parameterization has significant consequences for model prediction. This is especially true in non-econometric models where the majority of parameters are not estimated based on real world data but assumed by the modeler. Below we outline the most significant assumptions made here:

- **Constant returns to scale.** The CRS assumption in most basic terms assumes that doubling all inputs exactly doubles output. In the long run a CRS production function has constant average and marginal costs of production. This has to be and should be verified by crop empirically. Whether average costs are increasing or constant has significant ramifications for optimal output choice. A decreasing returns to scale assumption is very likely a much more realistic assumption.

- **Elasticity of substitution.** SWAP assumes an elasticity of substitution between any two inputs of 0.15. This is not empirically verified and has maybe the most significant ramifications for the optimal input choice. In the most basic terms what this assumes is a very limited ability by farmers to substitute between land, labor, water and other supplies. Assuming that farmers can maintain a given output level with the same substitution flexibility between water and land as between land as between labor and other supplies is unrealistic. Further assuming that this elasticity is identical for all crops is certainly not true. They justify this assumption with "experience from previous analyses", yet given the importance of this assumption, a Monte Carlo analysis should at least be conducted to demonstrate the importance of these assumptions. It is important to note that figure 3 in the model documentation is an artificial construct and bears no connection to a specific crop or farming reality.

- **First order conditions.** The optimization program solves by setting the marginal value product equal to the marginal cash cost plus opportunity cost for the input. If there are no estimates on the marginal productivity of a crop, the modelers make the significant assumption that "marginal productivity decreases 25% over the base condition productivity and thus use 25% of the land resource shadow value". It is important again to note that this is necessary in order to make the model solve, but an assumption that will have significant consequences for the solution of the model and which is not based on farming reality but rather a somewhat arbitrary assumption on part of the modeler.

- **Numerical scaling issues.** As the model documentation notes, there are scaling issues affecting the ability of the model to arrive at a solution. The modelers provide a scaling and rescaling solution. As Knittel and his coauthors recently pointed out, these numerical solution techniques are very sensitive to starting points and solution algorithms. The model documentation does not show whether this is an issue here.

Some Additional Observations

- **Demand Functions.** The model uses linear demand functions with elasticities based on a single study (Green et al 2008). It is somewhat nonstandard to assume an elasticity at the starting point for a linear demand function as the demand elasticity is not constant along this demand function. Great care has to be paid that the model accurately reflects the demand functions estimated by Greene and their staring point relative to the 2005 number used by SWAP.
- **Drought years.** 2005 was not a drought year. It is not clear that the model can be used to represent what happens during drought years as it is parameterized based on a non-drought year. One could conduct an exercise and see how well the model predicts drought and non-drought year land use and water use.

II.C. Recommendations Regarding the SWAP Model

The SWAP model is a useful tool that has provided important insights into the impact of changes in water availability. It has a long track record of use in program evaluation, cost-benefit analysis and academic research. State and federal agency staff members are familiar with its workings, and with its results. Nonetheless, we are concerned that SWAP is built on a very large number of relatively untested assumptions. We also have concerns about the underlying data, and about the calibration procedures used to fit the model to the data. We would like to recommend that DWR undertake the following steps to improve the suite of models available to analyze agricultural impacts of changes in Delta water supplies:

- The state should conduct a systematic peer review of SWAP, focusing on the large number of assumptions underlying the model (only some of which have been described in this report).

- We recommend that the predictions of the SWAP model be tested against real-world changes in land allocation. Such a test could be undertaken via a "backcasting" exercise where SWAP is calibrated to historical conditions, ideally by an independent research team, and then used to predict the impacts of actual past changes in water availability.

- DWR should work to integrate SWAP with a groundwater model. This project takes on additional importance given the potential for large changes in water availability associated with future State Board actions.

- The UC Davis researchers should consider reconfiguring the SWAP regions to better correspond to actual water rights, project service areas, and groundwater conditions.

- DWR should develop an econometric model for the agricultural sector in the San Joaquin Valley. There is a large amount of land use data that has become available in recent years that could be used as the basis for such a model. Other models could be developed for other quantities of interest, including farm employment and its relation to water deliveries. A key advantage of an econometric model is that it would produce standard errors around forecasts, a key omission of the SWAP model.

References

DWR. California Department of Water Resources, Land and Water Use Data. Available at: http://www.water.ca.gov/landwateruse/ (Last accessed October 2012).

Green, R., et al. "Estimation of Supply and Demand Elasticities of California Commodities." Working Paper. Department of Agricultural and Resources Economics, University of California, Davis (2008).

Howitt, Richard E. "Positive Mathematical Programming." American Journal of Agricultural Economics, 77: 329-342 (1995).

Howitt, Richard E. "A Calibration Method for Agricultural Economic Production Models." Journal of Agricultural Economics, 46(2): 147-159 (1995).

Howitt, Richard E., et al. "Calibrating Disaggregate Economic Models of Agricultural Production and Water Management". Environmental Modelling & Software, 38: 244-258 (2012).

UCCE. University of California Cooperative Extension, Crop Budgets. Available at: http://coststudies.ucdavis.edu (Last accessed October 2012).

Department of Agricultural and Resource Economics
University of California, Davis

Estimation of Supply and Demand Elasticities of California Commodities

by

Carlo Russo, Richard Green, and Richard Howitt

Working Paper No. 08-001

June 2008

Giannini Foundation of Agricultural Economics

ESTIMATION OF SUPPLY AND DEMAND ELASTICITIES OF CALIFORNIA COMMODITIES

by

Carlo Russo, Richard Green, and Richard Howitt

ABSTRACT

The primary purpose of this paper is to provide updated estimates of domestic own-price, cross-price and income elasticities of demand and estimated price elasticities of supply for various California commodities. Flexible functional forms including the Box-Cox specification and the nonlinear almost ideal demand system are estimated and bootstrap standard errors obtained. Partial adjustment models are used to model the supply side. These models provide good approximations in which to obtain elasticity estimates.

The six commodities selected represent some of the highest valued crops in California. The commodities are: almonds, walnuts, alfalfa, cotton, rice, and tomatoes (fresh and processed). All of the estimated own-price demand elasticities are inelastic and, in general, the income elasticities are all less than one. On the supply side, all the short-run price elasticities are inelastic. The long-run price elasticities are all greater than their short-run counterparts. The long-run price supply elasticities for cotton, almonds, and alfalfa are elastic, i.e., greater than one.

Policy makers can use these estimates to measure the changes in welfare of consumers and producers with respect to changes in policies and economic variables.

Keywords: Consumer Economics: Empirical Analysis (D120); Agricultural Markets and Marketing (Q130); Agriculture: Aggregate Supply and Demand Analysis; Prices (Q110)

This research project was funded by the California Department of Food and Agriculture. The authors very much appreciate helpful comments on earlier revisions by Steven Shaffer and Charles Goodman.

Carlo Russo is a Ph.D. student in the Department of Agricultural Economics at University of California, Davis. He can be contacted by e-mail at russocar@primal.ucdavis.edu. Richard Green is a professor and Richard Howitt is a professor and department chair, both in agricultural and resource economics at UC Davis. They can be contacted by e-mail at green@primal.ucdavis.edu and howitt@primal.ucdavis.edu, respectively.

ESTIMATION OF SUPPLY AND DEMAND ELASTICITIES OF CALIFORNIA COMMODITIES

by

Carlo Russo, Richard Green, and Richard Howitt

June 2008

Introduction[1]

California's agricultural sector can be characterized as being in a constant state of flux. On the consumer side of the market there have many changes in recent decades. Demographically, the proportion of married women in the labor force over the past four decades has doubled. In addition, demand patterns have been influenced by health and diet concerns. For example, there has been a 350% increase in sales of organic foods during the past decade. Demands for specialized and niche products are also on the increase.

The structure of fresh vegetable sales are more concentrated with fewer and larger retail buyers, and environmental regulations are being imposed to ensure better food safety. Competition from foreign suppliers is increasing. Technological changes have occurred in the processing of agricultural materials. Morrison-Paul and MacDonald noted that food prices today often appear less responsive to farm price shocks than in the past. Their research, however, found improving quality and falling relative prices for agricultural inputs, in combination with increasing factor substitution, has counteracted these forces to encourage greater usage of agricultural inputs in food processing.

---

[1] For an excellent discussion of the changes in California's agricultural sector, see Johnson and McCalla.

On the production side, global markets and trade liberalization has greatly impacted domestic markets. Land lost to urban expansion and an ever-growing pressure on water available impact California producers. The number of farms in California is decreasing while the sizes of farms are getting larger. While the price for California's fruits, nuts and vegetables is determined in domestic and export markets, the profitability of competing field, fiber and fodder crops is influenced by federal subsidies and state regulations. These impacts on California agriculture occur as both demand and supply side policies change.

In order to better understand and evaluate the consequences of these changes on consumer and producer welfare, it is essential to obtain reliable estimates of supply and demand elasticities of California commodities. To the best of our knowledge, there is no current comprehensive study that provides accurate up-to-date supply and demand elasticity estimates of California's major crops. Frequently cited works reporting demand elasticities are Carole Nuckton's Giannini Foundation publications (1978, 1980), "Demand Relationships for California Tree Fruits, Grapes, and Nuts: A Review of Past Studies" and "Demand Relationships for Vegetables: A Review of Past Studies". However, given the significant structural changes noted above, there are many causal factors that need to be updated to generate current supply and demand elasticities.

A more recent article, "Demand for California Agricultural Commodities" by Richard Green in the winter 1999 issue of *Update* reports estimates of own-price elasticities for selected commodities. The commodities included food (in general), almonds, California iceberg lettuce, California table grapes, California prunes, dried fruits (figs, raisins, prunes), California avocados, California fresh lemons, California residential water, and meats (beef, pork, poultry, and fish). All of the elasticity estimates are reported in research publications by faculty of the Department of Agricultural and Resource Economics at the University of California at Davis. Individual sources for the commodities are given in the reference section.

The primary purpose of this research project is to obtain updated supply and demand elasticity estimates of major California commodities. That is, short and long-run own-price elasticities of supply and own price, cross-price and income elasticities of demand. In this study sophisticatedly simple models are used (Zellner). The models focus on California agriculture. As a consequence, we tried to emphasize the specificity of California supply, contrasting it when possible, with aggregate US or the most relevant competing states' supply. Modeling the demand for California commodities was a challenging task, considering that markets are integrated and often statistics about retail prices do not discriminate products by origin. Also, for most crops we focused on the demand at the wholesale level. Thus, farm gate price may be based on a standard "mark down" of the price paid by the buyers. The modeling of wholesale demand was also convenient for those products (for example nuts) that are consumed mostly as ingredients of final goods. Exceptions to this approach relate to alfalfa and tomatoes. The former commodity is a major input for the California dairy industry so we estimated a derived demand. For fresh tomatoes we estimated the consumer demand at the US level.

Each crop presented specific modeling issues which are described in detail in the following sections. A brief discussion of the theoretical foundations of the models will be given, but detailed theoretical underpinnings of the models can be found in standard microeconomic textbooks.

The analysis will start with some of the most highly valued crops in California: almonds and walnuts, alfalfa hay, cotton, rice, and fresh and processing tomatoes. Future research will examine grapes (including raisin, table, and wine); lettuce (head and leaf); citrus (grapefruit, lemons, and oranges), stone fruits (apricots, nectarines, preaches, plums, and prunes); and broccoli.

Before a discussion of the theoretical models, data sources, econometric techniques, and the empirical results a brief literature review is provided.

Literature Review

1. Some Estimated Demand and Supply Elasticites from Previous Studies

One of the first attempts to compile a table of demand elasticity estimates for California crops was Nuckton (1978). She reported own-price elasticity of demand estimates for several California commodities including apples, cherries, apricots, peaches and nectarines, pears, plums and prunes, grapes, grapefruit, lemons, oranges, almonds, walnuts, avocados, and olives. Table 1 is a compilation of the empirical estimates that Nuckton reported. Estimates for the different studies varied widely, but Table 1 attempts to summarize the results from the main studies.

In 1999 Green published more recent elasticity estimates of California commodities from various sources. The table of elasticity estimates is repeated below in Table 2.

Table 1. Selected Elasticity Estimates of California Commodities[1]

Commodity	Own-Price Elasticity of Demand	Comments
Apples	-0.458 to –0.81	Fresh; some estimates were elastic
Cherries	-4.27	Sweet; retail; based on 20 cities
Apricots	-1.345	Fresh, farm level
Peaches & Nectarines	-0.898	Fresh
Pears	elastic	Based on reciprocal price flexibilities
Plums and Prunes	-0.630	Fresh, farm level
Grapes	-0.327 (-0.267) –0.160	Fresh; table grapes (raisin) wine
Grapefruit	-1.25	Fresh, retail level
Lemons	-0.210 (-0.38)	Fresh (processing)
Oranges	-0.72 (-2.76)	Fresh farm (fresh retail)
Almonds	-1.74 (-14.164)	Domestic shelled (export shelled)
Walnuts	-0.464	Shelled; wholesale
Avocados	elastic	Based on reciprocal price flexibilities
Olives	elastic	Based on reciprocal price flexibilities

Source: Nuckton, C., "Demand Relationship for California Tree Fruits, Grapes, and Nuts: A Review of Past Studies." Giannini Foundation, August 1978.

Table 2. Estimates of Own-Price Elasticities for Selected Commodities[1]

Commodity	Own-Price Elasticity
Food (in general)	-0.42
Almonds	-0.83
California Iceberg Lettuce	-0.16
California Table Grapes	-0.28
California Prunes	-0.44
Dried Fruits (Second Stage or Conditional)	
Figs	-0.23
Raisins	-0.67
Prunes	-0.35
California Avocados	-0.86
California Fresh Lemons	-0.34
Meats(Second Stage or Conditional)	
Beef	-0.84
Pork	-0.79
Poultry	-0.58
Fish	-0.57
California Residential Water	-0.16

Source: Green, R., “Demand for California Agricultural Commodities”, *Update*, Department of Agricultural and Resource Economics, Vol. 2, No. 2, Winter, 1999.

Some sources for the entries in Table 2 are as follows: food (Blanciforti, Green, and King); California iceberg lettuce (Sexton and Zhang); dried fruits (Green, Carman, and McManus); California avocados (Carman and Green); California fresh lemons (Kinney, Carman, Green, and O'Connell); and California residential water (Renwick and Green).

2. Examples of Market Conditions for Selected Commodities

A brief review of some recent selected articles illustrates the complexities of the market conditions facing California producers and consumers. In addition, a discussion of some economic factors that influence the supply and demand for certain products is given. The market situation for different crops varies dramatically. For some commodities, export and import markets are important. Other crops are perennial and have to be model differently than annual crops. Expectations of producers have to be incorporated in the supply response functions for these crops and a dynamic rather than a static approach has to be used. Rotation patterns can affect the supply response for certain crops such as alfalfa and cotton. A model for each crop has to incorporate these unique market characteristics associated with that particular crop. A few examples of the characteristic of the markets for a selected number of commodities are given below.

Alston, et al (1995) found an elasticity of demand for California almonds of –1.05. The demand for almonds in the United States is more elastic than almond demand in major importing countries. From a policy viewpoint, the inelastic demand for California almonds in export markets suggest that the industry can raise prices and profits in the short run by restricting the flow of almonds to these markets. In the long, however, this approach would lead to a decline in the almonds industry's share of the world market as competitors respond to higher prices with increased rates of almond plantings. They found little evidence for good substitutes for almonds among other nuts. Filberts in some European markets are an important exception to this rule. On the supply side, Alston et al (1995) found that almond yields in California are highly volatile, but yields can be predicted with good accuracy as a function of past yields, February rainfall, and the age distribution of almond trees. The major competitor to the California almond industry is the Spanish almond industry. Spanish almonds are a close substitute for California almonds in several European markets. This implies that changes in Spanish almond production have important effects on the California industry. Thus, a model of the almond industry must include both domestic and export markets on the demand side and the perennial nature of almond production (including alternate bearing years) on the supply side. Since there is little evidence of substitutes for almonds in the domestic market, a single-equation demand function can be estimated in order to obtain own-price and income elasticities for almonds.

With respect to table grapes, Alston et al (1997) obtained an estimated domestic own-price elasticity of demand for table grapes of –0.51, an income elasticity of demand of 0.51, and an elasticity of demand with respect to promotion of 0.16. Alston et al's (1997) study was primarily concerned with the effectiveness of promotion of table grapes. Their econometric results provided strong evidence that promotion by the California Table Grape Commission had significantly expanded the demand for California table grapes both domestically and in international markets. They evaluated the costs and benefits of a promotional campaign for various supply elasticity values. The policy implications were that the benefits from promotion were many times greater than either the total costs or the producer incidence of costs of a check-off program for table grapes. The own-price elasticity of –0.51 is inelastic implying that consumers are not very responsive to changes in prices of table grapes.

Almonds and grapes are two commodities for which international markets exist for the products. Thus, in order to properly model the supply and demand functions for these goods, exports and imports must be taken into account in addition to the domestic markets.

The own-price elasticity of demand for prunes, evaluated at the means, was found to be –0.4 by Alston et al (1998). The corresponding elasticity of demand with respect to income is 1.6, which, as they report, is larger than expected. Their study concludes that results from their analysis of the monthly, retail data support strongly the proposition that prune advertising and promotion has been an effective mechanism for increasing the demand for prunes and returns to producers of prunes. Based on their empirical results, they recommended that the prune industry could have profitably invested even more in promotion during the period of their investigation (September 1992 to July 1996).

Another perennial crop is alfalfa. Knapp and Konyar estimated the perennial crop supply response for California alfalfa. They employed a state-space model and the Kalman filter in order to generate parameter estimates as well as estimates of new plantings, removals, and existing acreage by age group. The estimated price elasticities for California alfalfa supply under quasi-rational expectations were –0.25 for the short run (one year) and –0.29 for the long run (10-20 years). The magnitudes of these supply elasticites appear reasonable with the longer-run elasticity a bit larger, as expected, in absolute value, than its short-run counterpart. In addition, Knap and Konyar found positive cross-price elasticity estimates for competing crops. Thus, producers react to prices of substitutes and act accordingly. Alfalfa is typically planted for three to four years and then removed from production. Frequently, cotton and alfalfa involve a rotation pattern. To our knowledge no one has attempted to model the rotation phenomena that exists between alfalfa and cotton. One of the models to be developed and estimated in this report incorporates this rotation pattern into the supply response models estimated for cotton and alfalfa.

ALMONDS

Figures 1A-6A in Appendix A provide a graphic overview of the domestic and foreign markets for California almonds for the years 1970-2001 (USDA). The figures contain information on marketable almond production, domestic per capita consumption, export and import of almonds, acreage in California, yield per acre, and grower price (nominal and real). A brief description of the almond industry will be given before the empirical results are presented.

Production of almonds exhibit a well-known alternate bearing-year phenomenon, that is, a high production year is followed immediately by a lower crop year and this pattern continues. Exports of almonds over the years 1970-2001 have continued to increase from less than 100 million pounds in 1970 to over 500 million pounds in 2001. Per capita consumption of almonds has also continued to increase over the same time period (Figure 2A). In 1970 per capita consumption of almonds were less than 0.4 pounds per capita and they increased to over 1 pound per capita in 2001. Acreage of almonds in California rose steadily over the years 1970-2001 from less than 200 thousand acres in 1970 to over 500 thousands acres in 2001. Per acre yield of almonds in

California exhibit a “see-saw” pattern, but the trend from 1970 has been increasing. Nominal grower prices for almonds have been volatile over the 30-year period from 1970 to 2001 reaching a peak in 1995 of $2.50 per pound. The major policy implication from Figure 6A; however, is that the real grower price, adjusted for inflation, has been steadily decreasing over the 1970-2001 period. The 2001 real grower price of almonds was barely over 50 cents per pound down from the peak real price of about $3.00 per pound in 1973. A causal glance at Figures 1A-6A in Appendix A indicates that the almond market is continually changing and a lot of world marketing forces affect California’s production and sales of almonds. Supply and demand models are developed and estimated for almonds and the results are given in the next section.

Some theoretical and data issues must be addressed before the models and estimations are presented. First, should a researcher use a singe-equation approach or a system approach? In this report both approaches are presented, although single equation estimations are usually considered to be less efficient. There are several reasons for considering this model. Based on previous research work by the authors, alternative nuts were found to be weak substitutes for almonds in the United States domestic market. Similar results were also found by Alston et al (1995). Thus, the advantages of imposing theoretical restrictions such as Slutsky symmetry conditions may be of little value in a demand system or subsystem for nuts. In addition, retail prices for almonds do not exist since they are used as ingredients in confectionaries. This has two important implications. First, are the demand functions retail or farm-level demands? Wohlgenant and Haidacher developed the theoretical relationships for the retail to farm linkages for a complete food demand system. Their approach, however, assumes that both retail and farm-level prices exist. In our case retail prices do not exist so we cannot employ their approach. This limitation of the demand models needs to be considered when interpreting the elasticity estimates. For example, farm-level own-price elasticities are generally more elastic than retail own-price elasticities for food commodities. Second, this may imply that nuts are not weakly separable from other food commodities.[2] This would rule out estimating a nut demand subsystem. The model that we employed uses CPI to account for the prices of other food items and commodities.

Given the alternate bearing phenomenon of almonds, there is a demand for consumption and a demand for storage. Alston et al (1995) did not find evidence of a stockholding effect. Thus, we followed their approach and assume that the demand function reflects consumption responses and not storage effects.

Finally, there is a calendar year versus a crop year problem involved with data collection. Alston et al (1995), when they estimated the domestic demand for almonds, used total availability (harvest received by handlers) minus US calendar year net exports minus stocks carried out plus carryins as their dependent variable.

Single equation estimation: demand

Based on standard microeconomic theory, it is assumed that an individual (representative) consumer behaves in such a way so as to maximize a well defined quasiconcave utility function subject to a budget constraint (see, e.g., Deaton and Muellbauer). The domestic aggregate demand for almonds can be written as

---

[2] A reviewer questioned this assumption. Nuts appear to be <u>not</u> weakly separable from other food commodities since they are used as ingredients in other food products. One implication of weak separability is that demands for the weakly separable goods can be expressed as a function of prices within the group and group expenditure. In theory, for example, if the price of cakes decreases, then one would expect that the quantity demanded of cakes would increase and consequently the demand for nuts would increase violating one of the implications of weak separability. Weak separability of nuts could be tested in a demand system if data were available and thus, in principle, is a refutable hypothesis.

$$Q_t = f(AP_t, WP_t, CPI_t, PCIN_t) \quad (3)$$

where Q_t represents per capita almond consumption, AP_t represents the price of almonds, WP_t denotes the price of walnuts, a possible substitute for almonds, CPI_t represents the consumer price index and captures the price of all other goods, and $PCIN_t$ denotes per capita income.[3]

With respect to functional forms for the almond demand equation, Box-Cox flexible functional forms

$$\frac{Q_t^{\lambda}}{\lambda} = \beta_0 + \beta_1 \frac{X_{t1}^{\lambda}}{\lambda} + \cdots + \beta_K \frac{X_{tK}^{\lambda}}{\lambda} + \varepsilon_t \quad (4)$$

were estimated by maximum likelihood procedures where λ can take on any value. All of the estimations in the report are carried out using SHAZAM, version 10. The linear and double logarithmic forms are special cases of the Box-Cox specification. The linear and double-log functional forms in the almond demand equation were tested against the more flexible Box-Cox functional form and in both cases the linear and double-log specifications were strongly rejected. The values of the likelihood ratio statistics were 43.7 for the linear and 14.85 for double-log model. The chi-squared critical value with one degree of freedom is 3.841 at the five percent significance level. Table 1 presents the estimations. The homogeneity condition of degree zero in all prices and income (HOD) does not hold globally in the Box-Cox specification unless the functional form is double

[3] Demand theory describes the behavior of individual consumers. The estimations, however, use aggregate data over all consumers. This can result in aggregation biases. If the observations are time series of cross-section data on randomly selected households, then it can be shown that the aggregate coefficients converge, as the number of households (N) goes to infinity, in probability to the micro coefficients (Theil). The disturbance terms are heteroskedastic, however. White's heteroskedastic-consistent standard errors for the estimated coefficients must be used. A recent excellent and thorough treatment of the conditions needed to avoid aggregation bias including exact aggregation and the distributional approach is given in Blundell and Stoker. They consider heterogeneity of consumers and distribution of income over time.

log.[4] The linear, double-log, and Box-Cox estimated functional forms for almond demand equations are presented in Table 3. In order to make the different models comparable, homogeneity was imposed in the double-log models and the other models were deflated by CPI.

[4] The homogeneity condition is $\lambda = 0 \; and \; \Sigma\beta_j = 0$ where the $\beta's$ are price and income coefficients; see Pope, *et al*. Linear specifications cannot be HOD by construction.

Table 3. Almond Demand Functions[1]

	Linear	**D. Log**	**D. Log-A**[2]	**Box-Cox**	**Box-Cox-A**[3]
AP[4]	-0.0016	-0.480	-0.377	-0.2671	-2.386
p-value	(0.0036)	(0.0004)	(0.0010)	(0.0004)	(0.0007)
elasticity	-0.351	-0.480	-0.377	-0.477	-0.378
WP[5]	0.0001	0.103	0.002	0.0436	-0.0267
p-value	(0.3898)	(0.5895)	(0.9912)	(0.5948)	(0.9891)
elasticity	0.465	0.103	0.002	0.097	-0.002
PCIN	0.00001	0.870	0.973	0.2911	29.404
p-value	(0.000)	(0.0038)	(0.0120)	(0.0036)	(0.0251)
elasticity	0.465	0.870	0.973	0.864	0.928
Const	-0.403	-5.14	-5.429	-4.270	-78.211
p-value	(0.000)	(0.0042)	(0.0068)	(0.0319)	(0.0394)
R^2	0.62	0.74	0.80	0.74	0.82
lnL	28.484	14.051	17.66	35.91	40.239
λ				0.107	-0.340
ρ			0.49		0.56

[1] *Q* is in pounds per capita, *AP* and *WP* are in cents per pound, and *PCIN* is in dollars.

[2,3] "A" denotes autocorrelated correction models.

[4,5] These are grower prices since retail prices do not exist.

The models were estimated using annual data from 1970 to 2001, a total of 32 observations. The Durbin-Watson values were 1.23 and 1.12 in the linear and double-log functional forms. The critical values are 1.244 and 1.650 at the five percent significance level, thus in the double-log and Box-Cox specifications the models were also estimated with an AR(1) error process. The estimated autocorrelation coefficients were 0.49 (double-log) and 0.56 (Box-Cox) with an estimated asymptotic standard error of 0.15 (double-log) and 0.14 (Box-Cox). The estimated own-price elasticity of domestic demand for almonds ranged from –0.48 to -0.35. The estimated elasticity was -0.38 in the Box-Cox functional form with an AR(1) error process. The estimates were highly significant with small p-values. Also, the estimated cross-price elasticity with walnuts was positive in four of the five models, but none of the coefficients were statistically significant; the smallest p-value being 0.39. The results confirm the absence of gross substitution effects between almond and walnuts. All of the estimated income coefficients were positive and ranged from 0.46 to 0.97 with small p-values. A sequential Chow and Goldfeld-Quandt test was conducted to determine if any structural changes had taken place during this period. No evidence was found of any structural changes.

Additional models were estimated using the dependent variable, US total consumption of almonds plus California exports minus US imports. The dependent variable captures the international demand for US almonds as well as the domestic demand. The ordinary least squares estimated double-log regression had an R^2 of 0.92. The estimated own-price elasticity of demand for almonds was -0.270 with an associated p-value of 0.022. The estimated model had a positive time trend coefficient of 0.05 (p-value =0.03) income elasticity was 2.10 with a p-value of 0.07.

Single equation estimation: supply

On the supply side, estimated almond acreage, yield, and marketable production functions were estimated for the period 1970 to 2001. The almond acreage was estimated using a partial adjustment model of the form:

$$\begin{aligned} A_t^* &= \alpha + \beta P_t \\ A_t - A_{t-1} &= (1-\gamma)\left(A_t^* - A_{t-1}\right) + \varepsilon_t \end{aligned} \tag{5}$$

where equations (5) are the desired almond acreage and equation (6) is the actual acreage; respectively. By substitution and some simplifications, the model can be estimated as:

$$A_t = (1-\gamma)\alpha + (1-\gamma)\beta P_t + \gamma A_{t-1} + \varepsilon_t \tag{6}$$

where A_t is the almond acreage (in acres), P_t is the average real almond grower price per pound over the previous eight years and, and ε_t is an error term included to capture all omitted factors that affect almond acreage.

This specification was chosen because it incorporates the behavior of producers whom adjust their acreage when they realize that the desired acreage (A_t^*) differs from the actual acreage the previous year (A_{t-1}). The adjustment coefficient, $1-\gamma$, indicates the rate of adjustment of actual acreage to desired acreage. The partial adjustment model is a model that captures producers' behavior (see, e.g., Kmenta). Almond trees take between five and six years to be fully productive. The acreage equation assumes a long-run planning process based on past prices, which are considered a proxy of the farmers' expectations about future prices.

The estimated acreage equation, with all variables expressed in logarithm form and based on 1979-2001 annual observations, is:

$$\ln \hat{A}_t = -0.32 + 0.12 \ln P_t + 0.97 \ln A_{t-1} . \tag{7}$$
$$\quad (0.31) \; (0.03) \qquad (0.04)$$

The values in parentheses are standard errors. The coefficient of determination of the regression is R^2=0.97. The Durbin-h statistic is 1.40 which is asymptotically not significant, thus there is no evidence of autocorrelation. The estimated short-run price elasticity is 0.12 with an associated p-value of 0.0016. The estimated coefficient on lagged acreage is 0.97 with an associated p-value of 0.0000. The estimated acreage response equation provides empirical evidence that almond producers respond positively to anticipated price increases in almonds.

The yield equation for almonds is:

$$\ln Y_t = \beta_1 + \beta_2 \ln P_{t-1} + \beta_3 Rain_t + \beta_4 T_t + \beta_5 T_t^2 + \varepsilon_t \tag{8}$$

where Y_t is almond yield in pounds per acre, P_{t-1} is the real grower price of almonds in cents per pound in the previous year, $Rain_t$ is rainfall in inches in March, and T_t is a time trend that is a proxy for technological change.

The ordinary least squares estimated yield equation for almonds for the years 1971-2001 is (equation (9))

$$\ln \hat{Y}_t = 6.39 + 0.07 \ln P_{t-1} - 0.20 \ln Rain_t + 0.05 T_t - 0.001 T_t^2 \tag{9}$$
$$(0.48) \; (0.09) \qquad (0.05) \qquad (0.01) \quad (0.0003)$$

where the values in parentheses are standard errors. The estimated R^2 is 0.68 which indicates an adequate fit of the model with the data. All of the p-values for the estimated coefficients are less than 0.10 except for one associated with lagged price. The coefficient on lagged price is positive (0.07) but not significant. The coefficient on

March rainfall is negative (-0.20) reflecting the effect of rain on increased brown rot disease and decreased pollination. The coefficient on the time trend is positive (0.05) and significant indicating that, conditioned on all the other variables, yields are increasing over the time period, 1971-2001. The coefficient on time squared is negative (-0.001) and significant reflecting that the time trend is increasing at a decreasing rate. The increasing trend can be due to technology and improvement of production practices. The almond yield equation exhibits an alternate bearing phenomenon since the autocorrelation was negative ($\hat{\rho} = 0.38$) with an asymptotic t-value of 2.26.[5] The model was estimated using the autocorrelation method of Pagan in SHAZAM. The other autocorrelation methods, ML and Cochrane–Orcutt gave similar results.

Finally, a production function for almonds was developed and estimated. The model is:

$$\ln Q_t = \beta_1 + \beta_2 \ln P_{t-1} + \beta_3 \ln Rain_t + \beta_4 \ln Q_{t-1} + \varepsilon_t \tag{10}$$

where Q_t is California almond production in millions of pounds, P_{t-1} represents the lagged price of almonds in cents per pound, $Rain_t$ represents March rainfall in inches, and Q_{t-1} denotes lagged production. The model is a partial adjustment model and includes the effect of alternate crop years and weather. As in the yield equation, the alternative bearing phenomenon is captured by a negative autocorrelation coefficient.

The estimation of the model, correcting for autocorrelation, is

[5] Several methods were used to capture the alternate-year yield phenomenon. For example, a dummy variable was added to the function with zero values for low-yield years and ones for high-yield years. Due to weather conditions and new varieties of trees that started bearing, the data exhibits a high-low pattern for a number of years followed by two high-yield years in a row or two low-yield years in a row. The high-low pattern continues for a few years but the pattern may be reversed. History then repeats itself. It is difficult to capture these phenomena with a dummy variable in the systematic part of the equation. This

$$\ln \hat{Q}_t = -0.44 + 0.19 \ln P_{t-1} - 0.20 \ln Rain_t + 0.97 \ln Q_{t-1} \quad (11)$$
$$(1.24)(0.15) \qquad (0.07) \qquad (0.11)$$

where the numbers in parentheses are estimated standard errors. The R^2 of the model is 0.71. The elasticity of production with respect to the lagged own price (for given values of the production in the previous year, the weather conditions and the alternate crop years) is 0.19 but not significant (p-value= 0.20). The coefficient on March rainfall is negative as explained above and the estimated coefficient on lagged production is positive and highly significant. The alternate crop pattern was capture by a negative autocorrelation coefficient of -0.55 with an associated asymptotic t-value of 3.74.[6]

WALNUTS

Data for the years 1970-2001 are presented in Appendix B for walnuts. California marketable production, total domestic consumption, exports and imports, per capita consumption, acreage, yield, and grower prices, both nominal and real for walnuts are given in Figures 1B-6B in Appendix B. An overview of the walnut industry can be seen by an examination of the Figures. Marketable production of walnuts has slowly increased from just below 100 million pounds in 1970 to over 250 million pounds in 2001. Exports of walnuts exhibit a similar pattern of that to production (see Figure 1B in Appendix B). Per capita consumption of walnuts has remained relatively stable at 0.4 pounds over the period 1970-2001 (Figure 2B). Acreage has slowly increased over the period starting with about 150 thousand acres in 1970 to about 200 thousand in 2001.

---

was not the case with walnuts where the alternate pattern was consistent throughout the sample period. See the patterns in the data for almond yields, walnut yields, and walnut production in Appendix C.

[6] Alternative functional forms of the production function were estimated including a Box-Cox specification, models with moving average error schemes, etc. The Box-Cox functional form yielded a price elasticity of 0.29 and a model estimated with a moving average error term yielded a slightly lower price elasticity estimate of 0.23.

Yields of walnuts are more volatile over the period than acreage but with a steady trend upward over the period 1970-2001 (Figure 4B). Real grower prices have decreased over the period from 1970 to 2001 (Figure 6B). Real grower prices reached a peak in about 1978 of $2.00 per pound and have declined ever since to about 60 cents per pound in 2001.

Demand, acreage, yield, and production equations were estimated for walnuts using annual data from 1970 to 2001. The United States domestic demand for walnuts is estimated and reported first.

The model for US per capita consumption of walnuts is

$$Q_t = f(AP_t, WP_t, CPI_t, PCIN_t) \tag{13}$$

where Q_t represents per capita walnut consumption in pounds, AP_t represents the price of almonds in cents per pound where almonds are a possible substitute for walnuts, WP_t denotes the price of walnuts in cents per pound, CPI_t represents the consumer price index and captures the price of all other goods, and $PCIN_t$ denotes per capita income in dollars.

The restriction of homogeneity of degree zero in all prices and income was imposed. When the model for all the years, 1970 to 2001, was estimated by ordinary least squares, the Durbin-Watson value was small (0.796) indicating a possible misspecified model. Consequently, sequential Chow and Goldfeld-Quandt tests were performed and they indicated a structural break in 1983. Two demand functions were estimated, one using data from 1971 to 1983 and one employing data from 1983 to 2001. The estimated models, double-log and Box-Cox functional forms, are presented in Table 4.

Table 4. Walnut Demand Functions

	Pre 1983		Post 1983	
	Double Log	**Box-Cox**	**Double Log**	**Box-Cox**
AP	-0.210	-0.449	-0.082	-0.19E-06
p-value	(0.039)	(0.136)	(0.325)	(0.667)
elasticity	-0.210	-0.197	-0.082	-0.023
WP	-0.284	-0.825	-0.267	-0.26E-07
p-value	(0.068)	(0.113)	(0.063)	(0.051)
elasticity	-0.284	-0.266	-0.267	-0.251
CPI	-1.039	-1.435	-0.633	-0.61E-05
p-value	(0.029)	(0.612)	(0.414)	(0.307)
elasticity	-1.039	-0.677)	-0.633	-0.807
PCIN	1.534	5.349	-0.983	0.10E-09
p-value	(0.007)	(0.339)	(0.201)	(0.398)
elasticity	1.039	1.207	-0.983	0.427
Constant	-7.361	-17.519	-4.50	-0.333
p-value	(0.005)	(0.304)	(0.207)	(0.000)
R^2	0.759	0.763	0.705	0.726
DW	2.563	2,43	2.069	2.507
lnL	15.988	26.029	25.492	44.217
λ	0	-0.15	0	2.06

The R^2 values range from 0.71 to 0.76. The fit of the models to the data was not as good as for the almond demand equations. The Durbin-Watson statistics did not indicate any problems with autocorrelation. The estimated own-price elasticity of demand for walnuts ranged from –0.266 to -0.284 for the time period prior to 1983 and from -0.251 to -0.267 after the year 1983. The p-values were 0.068 (pre 1983) and 0.63 (post 1983) for the double-log models and 0.113 (pre 1983) to 0.051 (post 1983) for the Box-Cox functional forms. The Box-Cox equation post 1983 was estimated with a time trend. Its estimated coefficient was -0.03 with an associated p-value of 0.014. Three of the four estimated income elasticities were positive with only the post 1983 for the double-log specification negative (-0.983). Only one of the estimated almond cross-price elasticities was significant at any reasonable level. Thus, the sample evidence finds little substitution effects between almonds and walnuts. Based on the sample evidence the estimated own-price elasticity of demand for walnuts is inelastic.

What are some economic factors that can explain the structural break around 1982-83? From Figure 6B, real walnut prices dropped dramatically in 1983. There was a large supply of walnuts that year and inventory levels increased significantly. In addition, the United States imposed a tariff on pasta and Italy, one of the largest importers of U.S. walnuts, retaliated by placing an embargo on U.S. walnuts. Exports dropped causing increases in inventory levels.

Another model was estimated where the dependent variable was US total consumption of walnuts plus California exports minus US imports. The dependent variable captures domestic plus net export demand. Again, sequential structural tests indicated a structural break around 1983. The results from this estimated equation yielded a total own-price elasticity of demand for walnuts of –0.354 prior to 1983 and an estimated value of –0.061 after 1983. The estimated coefficient of determination for this equation was 0.923. The wide difference between the estimated own-price elasticities of demand between the two time periods may be due, in part, to structural changes mentioned above. The primary policy implications are that the demand for walnuts is inelastic with little evidence that almonds are an important substitute for walnuts.

On the supply side, acreage, yield, and production equations were estimated for walnuts, using a partial adjustment model. The estimated acreage equation is

$$\underset{(1.16)}{\ln \hat{A}_t = 2.90} + \underset{(0.01)}{0.02} \ln P_t + \underset{(0.00)}{0.00T_t} + \underset{(0.00)}{0.00T_t^2} + \underset{(0.10)}{0.74} \ln A_{t-1} \qquad (14)$$

where A_t represents acreage of walnuts in acres, P denotes walnuts grower prices of walnuts in cents per pound and T is a time trend. Values in parentheses represent standard errors. The estimated coefficient of determination, R^2, was 0.953. The estimated short-run elasticity of acreage with respect to price is 0.02, which implies that acreage is inelastic with respect to the current price. The estimated lagged acreage coefficient was 0.74 and highly significant indicating a partial adjustment by producers of walnut acreage over time. Figure 6 charts the actual acreage of walnuts to the predicted values.

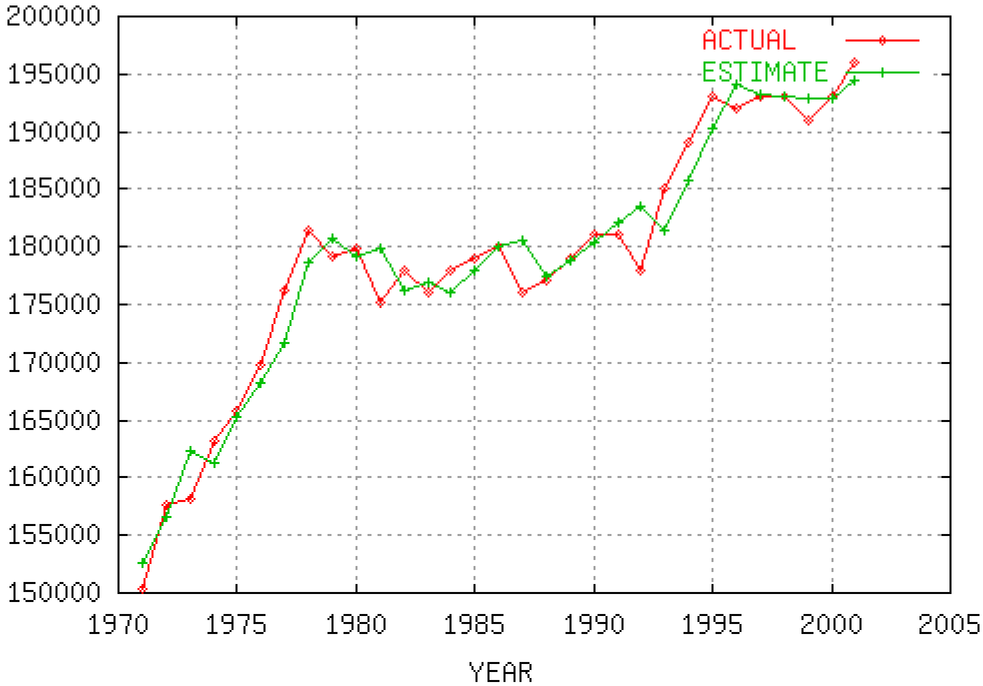


Figure 1: Walnuts acreage. Actual and estimated (in acres).

The value of the Durbin h statistic (-0.37) indicates that autocorrelation is not a problem.

The ordinary least squares estimated yield equation for walnuts, based on the years 1972-2001, is

$$\ln \hat{Y}_t = -0.01 - 0.03 \ln P_{t-1} + 0.03 TAM_t + 0.14 D_t + 0.01 T_t - 0.0002 T_t^2 \qquad (15)$$
$$\quad (0.60) \; (0.08) \qquad (0.03) \qquad (0.03) \quad (0.01) \quad (0.0003)$$

where Y_t, the dependent variable is yield of walnuts in pounds per acre, P_{t-1} is lagged real grower price of walnuts in cents per pound, T_t is a time trend, TAM_t is the average temperature in March, D_t is a dummy variable that is equal to one in high-yield years and zero for low-yield years (more specifically, D=1 in 1970 and alternates from 1 to 0 throughout the sampling period) and is included to capture the alternate yield-year phenomenon. The coefficient of determination is 0.72. The Durbin-Watson calculated value of 1.78 does not support evidence of negative correlation. The “see-saw” pattern exhibited by walnut yields is more consistent that than for almond yields and thus the dummy variable included in the systematic part of the equation picks up the alternative bearing phenomenon (see Appendix C). The estimated coefficient on D is positive and highly significant as expected and the coefficient on March temperature is positive as expected but not significant. There is a little evidence of a positive time trend. The lagged price coefficient is unexpectedly negative but not significant.

The final estimation for walnuts consists of estimating a production function for the years 1971-2001. The estimated production function, corrected for autocorrelation, is:

$$\underset{}{\ln \hat{PR}_t} = \underset{(1.84)}{3.52} + \underset{(0.06)}{0.003} \ln P_{t-1} + \underset{(0.02)}{0.03} TAM_t + \underset{(0.07)}{0.23} D_t + \underset{(0.13)}{0.69} \ln PR_{t-1} \qquad (16)$$

where the dependent variable, PR_t, is walnut production in millions of pounds, P_{t-1} is walnut price in cents per pound, TAM_t is the March temperature, and D_t is a dummy variable that takes on the values of 1 and 0 and accounts for the alternate year production phenomenon. The R^2 of the regression is 0.82. The estimated autocorrelation coefficient is -0.47 with as asymptotic t-value of 2.60. The alternate year dummy coefficient is positive and highly significant as expected is picking up all the alternate production year effect. The estimated coefficient on lagged walnut price is positive but insignificant and the estimated coefficient on lagged production is positive and significant. The positive sign on March temperature is as expected.

SUR Estimation

The results of the estimations suggest that walnuts and almonds cannot be considered as close substitutes or complements because the cross-price elasticities were not significantly different from zero. However, the possible relations across the two markets can be explored using a demand system of *seemingly unrelated equations* (SUR). In this system, correlation in the errors across equations is assumed. Some of the same omitted factors may influence both almond and walnut demands.

The equations are estimated using an iterative SUR procedure to achieve efficiency. Also the properties of symmetry and zero homogeneity were imposed. The estimation of the system (eq. 17) is:

$$\ln PC_t^W = -4.17 - 0.14 \ln P_t^W - 0.20 \ln P_t^A - 0.48 \ln CPI_t + 0.82 \ln PCIN_t - 0.19 T_t - 0.07 D_t$$

(3.57) (0.14) (0.08) (0.07) (0.78) (0.01) (0.08)

$$\ln PC_t^A = -5.45 - 0.20 \ln P_t^A - 0.18 \ln P_t^W - 0.67 \ln CPI_t + 1.05 \ln PCIN_t$$

(1.64) (0.08) (0.17) (0.40) (0.29)

where numbers in parentheses are standard errors, PC^W and PC^A are the per-capita consumption of walnuts and almond, respectively. PC^w and PC^A are grower nominal prices of walnuts and almonds, respectively, D_t is a dummy variable that takes on the value of zero prior to 1983 and the value of one after 1983. The remaining variables are as defined above except per capita income is also expressed in nominal terms. The system R^2 is equal to 0.81. The estimated own-price elasticity of walnuts is -0.14 and that of almonds -0.20; with only the estimated own-price elasticity of almonds being highly significant. The estimated income elasticity for walnuts is 0.82 and that of almonds is 1.05.

Some Policy Implications

Based on the models estimated for almonds and walnuts the own-price elasticity of US domestic demand for almonds was found to be between –0.35 and -0.48.. These estimates are inelastic and imply that almond producers are vulnerable to large swings in prices of almonds due to supply shifts. Similar estimates of the own-price elasticity of US domestic demand for walnuts were obtained. The estimated own-price elasticities for walnuts ranged from –0.25 to –0.28. Walnut producers face the same marketing situation as almond producers, that is, prices of walnuts fluctuate widely due to shifts in the supply function of walnuts.

The estimated acreage response equation for almonds indicated that producers respond positively to lag prices. The estimated short-run price elasticity of acreage for almonds was 0.12 and significant. This is relatively small but does indicate that producers are responsive to increases in prices over time. For walnuts the estimated short-run price elasticity of acreage was 0.02 and significant. Again, the value is small but positive.

The estimated yield equations for both almonds and walnuts reflected a significant alternate-year phenomenon. For almonds the phenomenon was capture by a significant and negative autocorrelation coefficient. For walnuts it was captured by a dummy variable. Yields for almonds are significantly affected by a time trend. Yields of almonds are increasing over the time period 1979-2001, based on the estimated yield equation. For walnuts, yields were positively affected by temperature in March and a time trend, but neither coefficient was significant.

A SUR demand system was estimated for walnuts and almonds. The domestic own-price elasticity of demand for walnuts was estimated to -0.14 and that of almonds -0.20 with almonds being significant. The estimated income elasticity of demand for walnuts was 0.82 and that for almonds was 1.05 with the estimated income elasticity in the almond equation being significant. The evidence does not support gross substitution between almonds and walnuts.

The primary policy implication based on these results is that almond and walnut producers are facing an inelastic domestic demand for their products. Combine this with the volatility of the supply function due to temperature and rainfall changes, wide variations in prices exist which lead to wide variations in profits from year to year. Storage, improved technology, and an expanding export market are factors that may mitigate the volatile market conditions facing US producers of almonds and walnuts.

References

Alston, J., H. Carman, J. Christian, J. Dorfman, J-R. Murua, and R. Sexton, "Optimal Reserve and Export Policies for the California Almond Industry: Theory, Econometrics and Simulations." *Giannini Foundation Monograph*, No. 42, February 1995.

Alston, J., J. Chalfant, J. Christian, E. Meng, and N. Piggott, "The California Table Grape Commission's Promotion Program: An Evaluation." *Giannini Foundation Monograph*, No. 43, November 1997.

Alston, J., H. Carman, J. Chalfant, J. Crespi, R. Sexton, and R. Venner, "The California Prune Board's Promotion Program: An Evaluation." *Giannini Foundation Research Report*, No. 344, March 1998.

Blanciforti, L., R. Green, and G. King, "U.S. Consumer Behavior Over the Postwar Period: An Almost Ideal Demand System Analysis." *Giannini Foundation Monograph*, No. 40, August 1986.

Blundell, R. and T. Stoker, "Heterogeneity and Aggregation", *Journal of Economic Literature*, 43(June 2005):347-391.

Carman, H. and R. Green, "Commodity Supply Response to a Producer Financed Advertising Program: The California Avocado Industry" *Agribusiness: An International Journal*, 1993, 605-621.

Goodwin, B. and G. Brester, "Structural Change in Factor Demand Relationships in the U.S. Food and Kindred Products Industry.' *American Journal of Agricultural Economics 77(1995):69-79.*

Green, R., H. Carman, and K. McManus, "Some Empirical Methods of Estimating Advertising Effects in Demand Systems: An Application to Dried Fruits." *Western Journal of Agricultural Economics 16(1):63-71.*

Green, R., "Demand for California Agricultural Commodities." *Update* Vol. 2, No. 2, Winter 1999.

Johnston, W. and A. McCalla. "Whither California Agriculture: UP, Down, or Out? Some Thoughts about the Future." *Giannini Foundation Special Report* 04-1, August 2004.

Kinney, W., H. Carman, R. Green, and J. O'Connell, "An Analysis of Economic Adjustments in the California-Arizona Lemon Industry." *Giannini Foundation Research Report*, No. 337, April 1987.

Kmenta, J. *Elements of Econometrics*, 2nd ed., Macmillan Pub. Co., New York, 1986.

Knapp, K. and K. Konyar, "Perennial Crop Supply Response: A Kalman Filter Approach." *American Journal of Agricultural Economics*, 73(1991):841-849.

Morrison Paul, C. and J. MacDonald, "Tracing the Effects of Agricultural Commodity Prices and Food Costs." *American Journal of Agricultural Economics* 85(2003):633-646.

Nuckton, C., "Demand Relationships for California Tree Fruits, Grapes, and Nuts: A Review of Past Studies", *Giannini Foundation Special Publication* No. 3247, August 1978.

Nuckton, C., "Demand Relationships for Vegetables: A Review of Past Studies", *Giannini Foundation Special Report 80-1*, 1980.

Pope, R., R. Green, and J. Eales, “Testing for Homogeneity and Habit Formation in a Flexible Demand Specification of U.S. Meat consumption” *American Journal of Agricultural Economics* 62(1980): 778-784.

Renwick, M. and R. Green, “Do Residential Water Demand Side Management Policies Measure Up? An Analysis of Eight California Water Agencies” *Journal of Environmental Economics and Management* 40(2000):37-55.

Sexton, R. and M. Zhang, “A Model of Price Determination for Fresh Produce with Application to California Iceberg Lettuce.” *American Journal of Agricultural Economics* 78(1996):924-34.

Wohlgenant, M. and R. Haidacher, “Retail to Farm Linkage for a Complete Demand System of Food Commodities”, USDA, ERS, Technical Bulletin No. 1775, 1989.

Zellner, A. “Philosophy and Objectives of Econometrics”, *Journal of Econometrics*, 2006.

Appendix A: Almonds

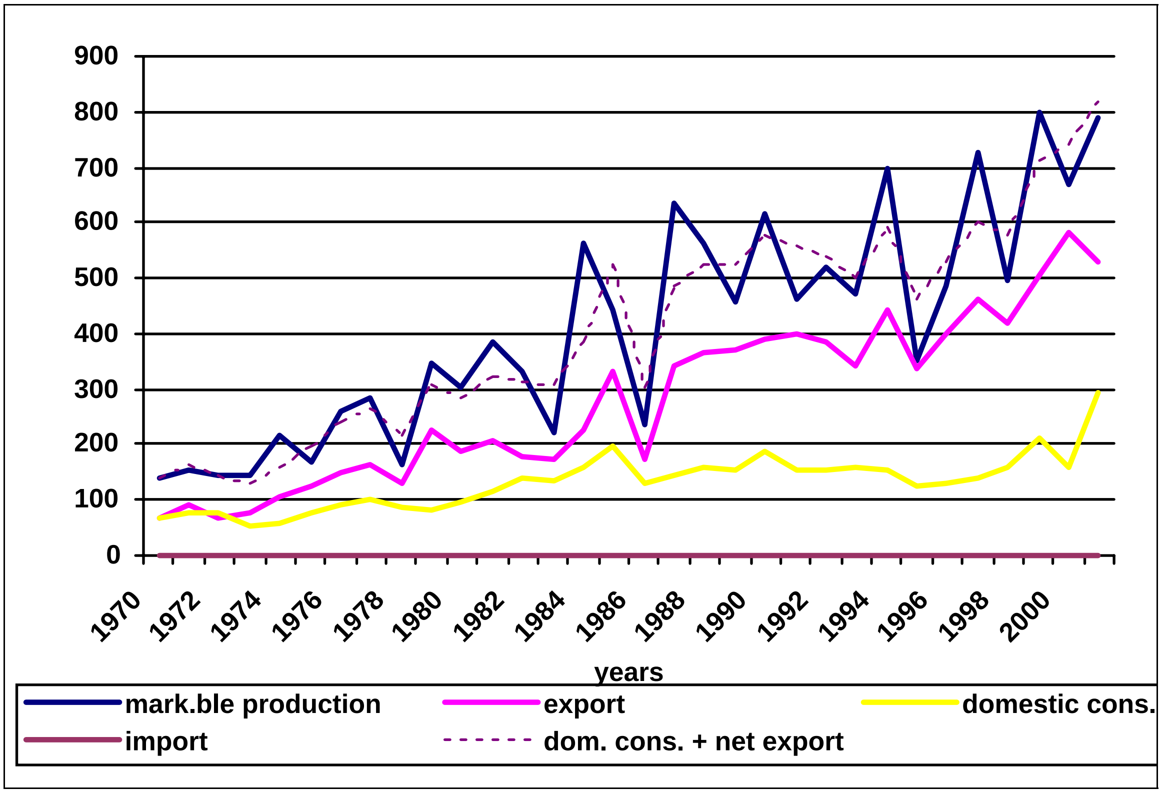


Figure 1A: California marketable production, US domestic consumption, export and import of Almonds. Years 1970-2001(millions of lbs).

Source: USDA

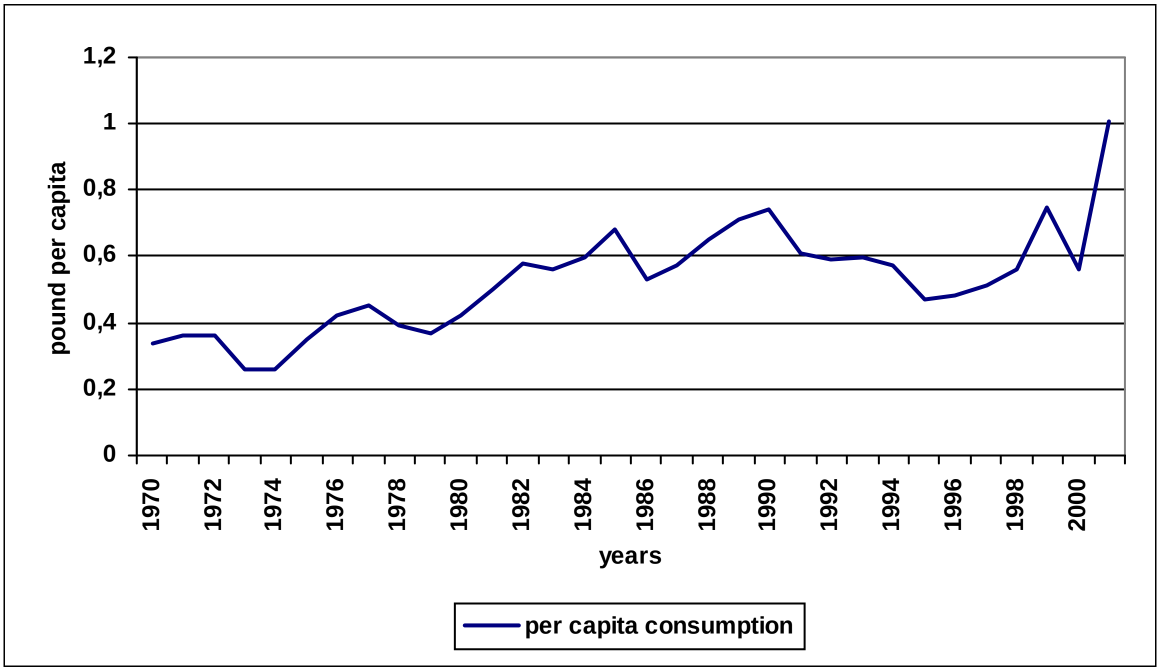


Figure 2A: US per capita consumption of Almonds. Years 1970-2001

Source: USDA

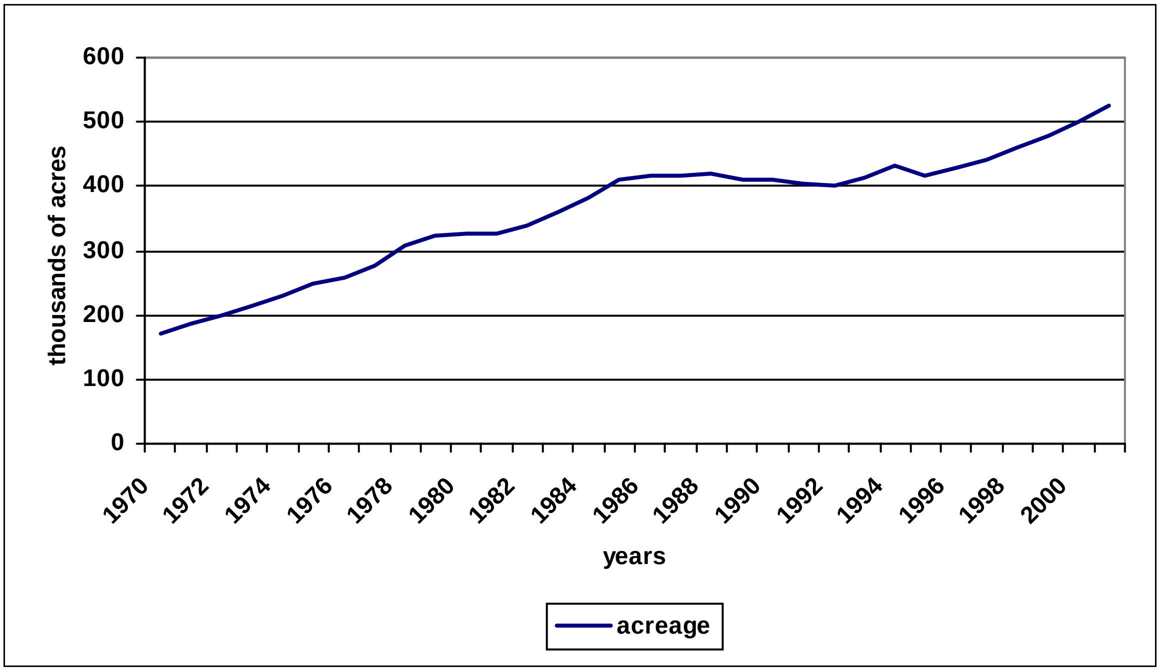


Figure 3A: Acreage of almonds in California. Years 1970-2001

Source: USDA

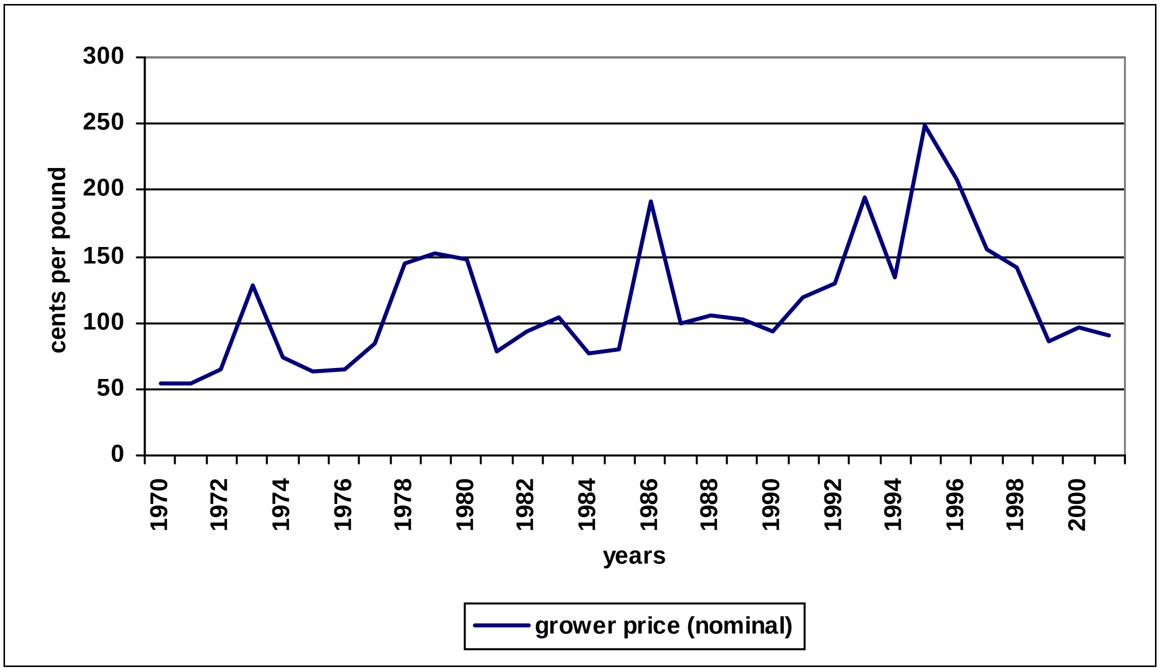


Figure 4A: Grower price for almonds in California (nominal values). Years 1970-2001

Source: USDA

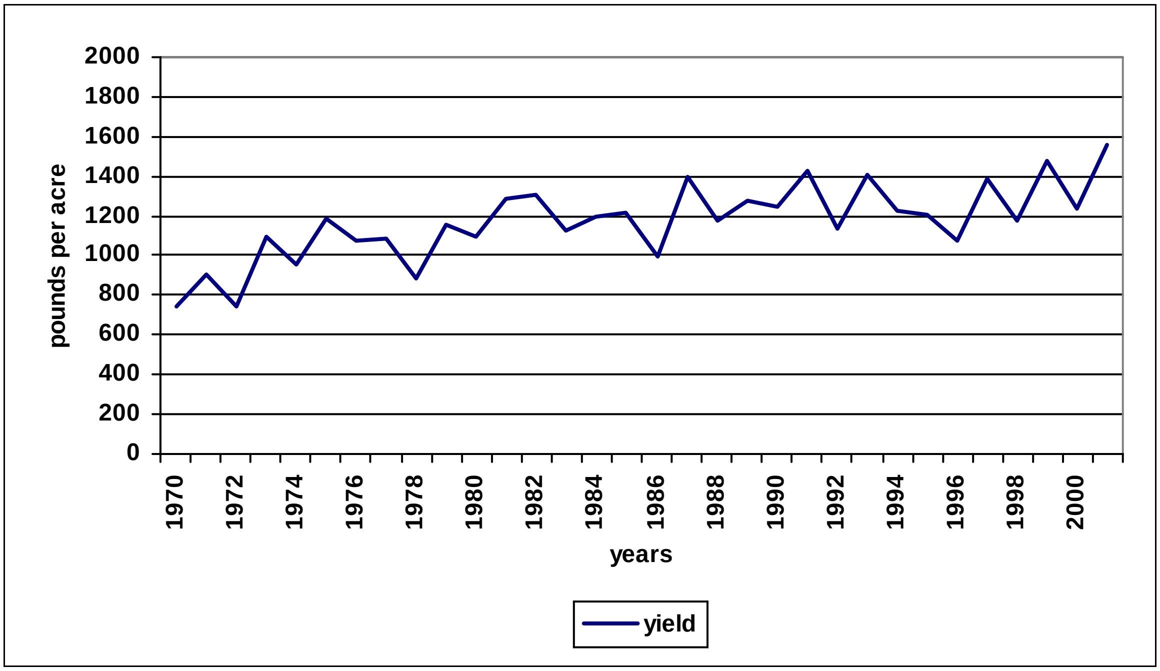


Figure 5A: Yield of almonds in California. Years 1970-2001

Source: USDA

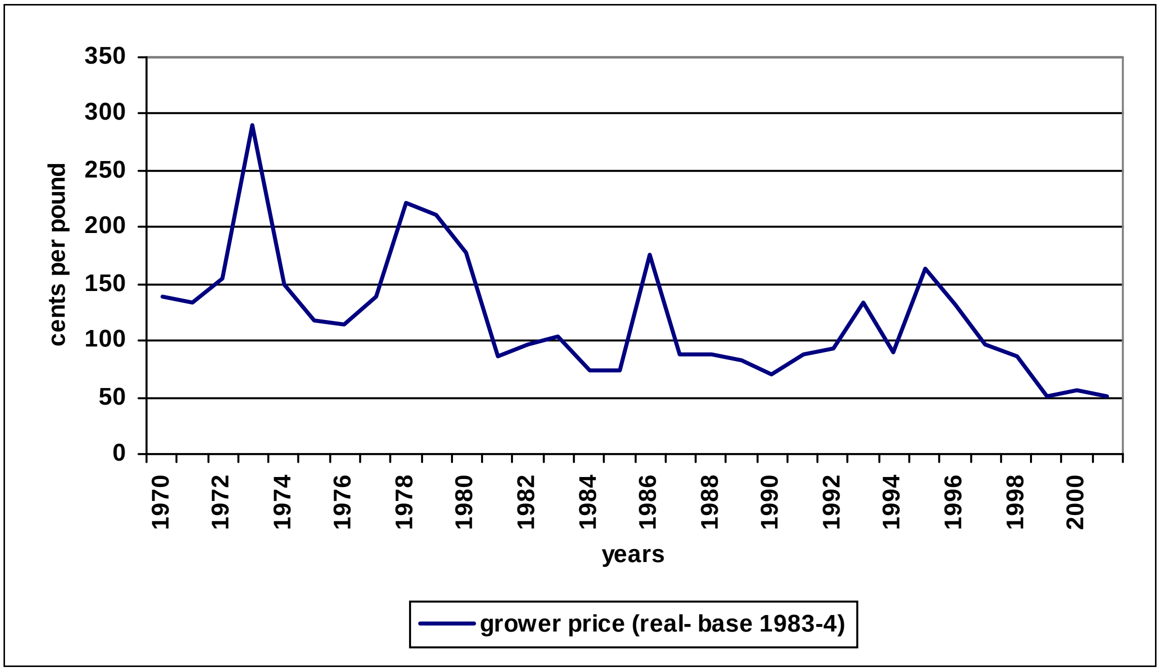


Figure 6A. Real grower price for almonds in California (real values). Years 1970-2001.

Source: USDA

Appendix B: Walnuts

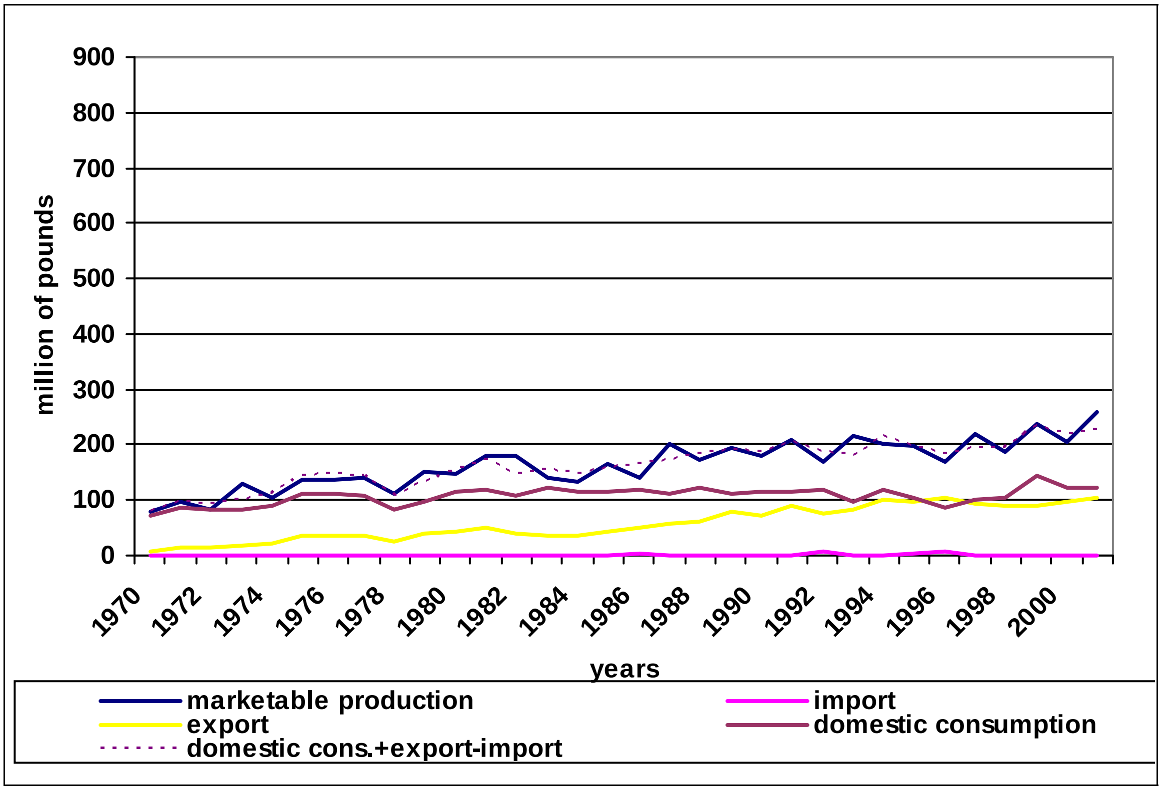


Figure 1B: ***California marketable production, US domestic consumption, export and import of Walnuts. Years 1970-2001***

Source: USDA

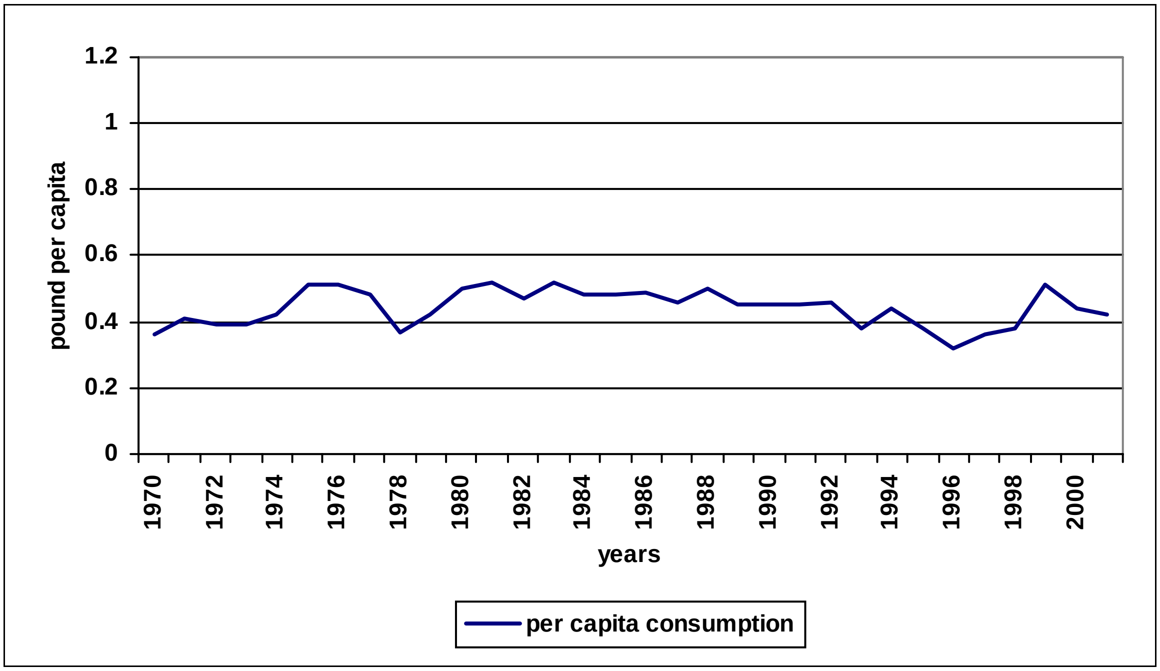


Figure 2B: ***US per capita consumption of Walnuts. Years 1970-2001***

Source: USDA

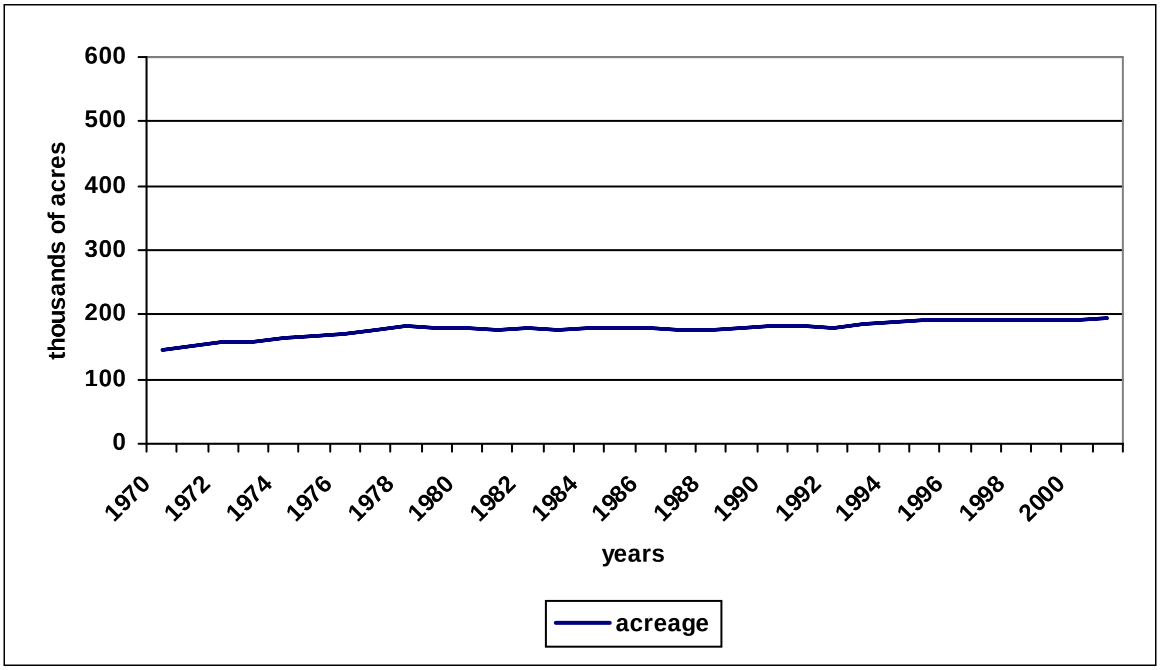


Figure 3B: ***Walnut acreage in California. Years 1970-2001***

Source: USDA

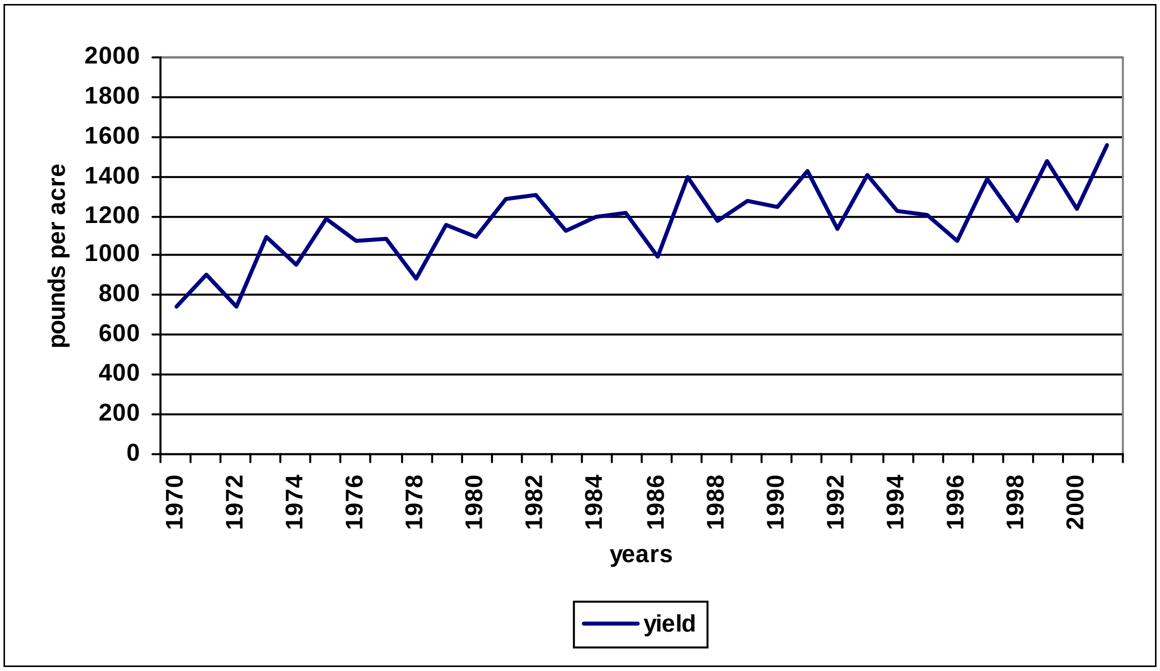


Figure 4B: ***Per acre yield of Walnuts in California. Years 1970-2001***

Source: USDA

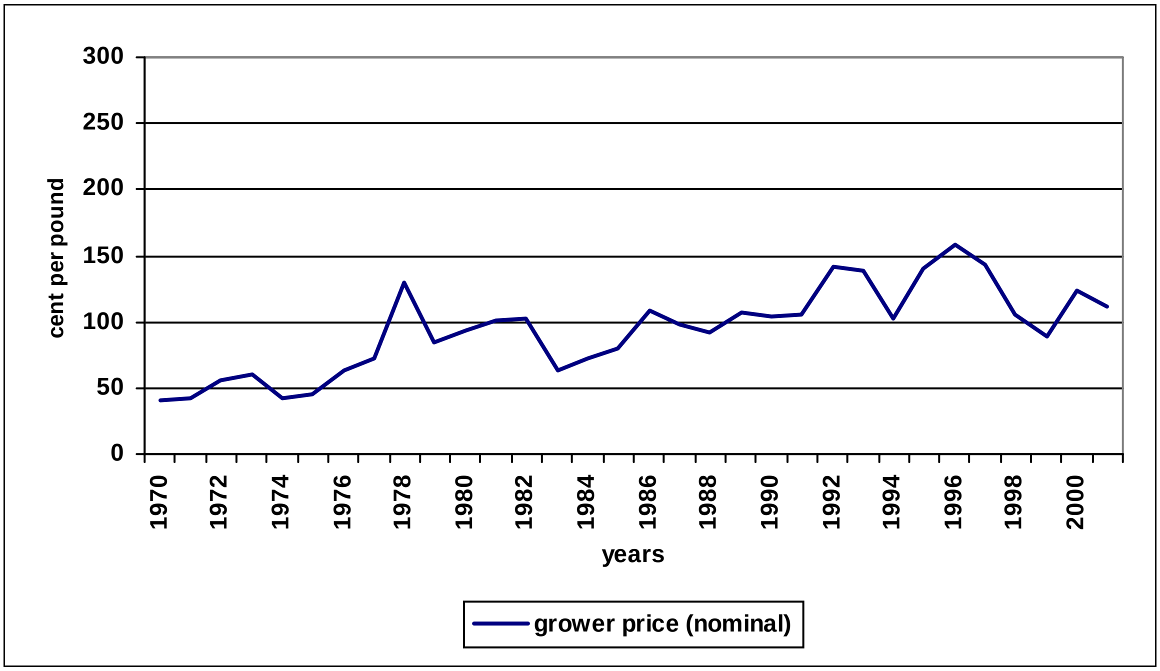


Figure 5B: ***Grower price for walnuts in California (nominal values). Years 1970-2001***

Source: USDA

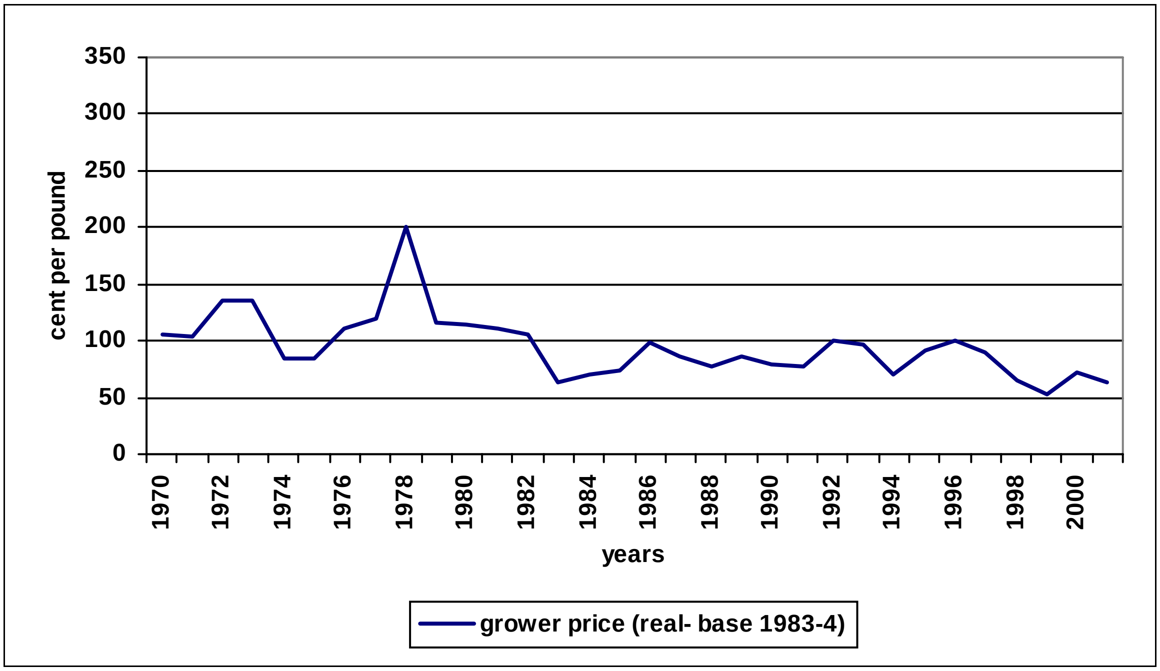


Figure 6B: ***Real grower price for walnuts in California (real values). Years 1970-2001***

Source: USDA

APPENDIX C: Almond Yields, Walnut Yields, and Walnut Production, 1970-2001

Year	Almond Yields (Pounds/Acre)	Walnut Yields (Pounds/Acre)	Walnut Production (Millions Lbs)
1970	877	740	108000
1971	863	900	135000
1972	759	740	116000
1973	726	1100	174000
1974	995	950	155000
1975	748	1190	198000
1976	1100	1080	183000
1977	1130	1090	192000
1978	588	880	160000
1979	1160	1160	208000
1980	985	1100	197000
1981	1250	1290	225000
1982	1250	1310	234000
1983	1020	1130	199000
1984	672	1200	213000
1985	1550	1220	219000
1986	1140	1000	180000
1987	601	1400	247000
1988	1580	1180	209000
1989	1410	1280	229000
1990	1190	1250	227000
1991	1610	1430	259000
1992	1210	1140	203000
1993	1370	1410	260000
1994	1190	1230	232000
1995	1700	1210	234000
1996	885	1080	208000
1997	1190	1390	269000
1998	1720	1180	227000
1999	1130	1480	283000
2000	1130	1240	239000
2001	1740	1560	305000

Source: USDA.

ALFALFA AND COTTON

Introduction

Historically, from 1950-2002, alfalfa and cotton have been among California's top commodities in terms of total value (Johnston and McCalla). In 1950 cotton was ranked third in terms of value of production in California with a value of $202 million. By 2001, cotton had slipped to the eighth most valuable commodity in California in value of production. The trend has been downward during the period 1950-2002. Hay (85% alfalfa) was ranked fifth in 1950 in California with a value of production of $121 million. In 2001, hay was ranked seventh in value of production just ahead of cotton.

Models are developed for California alfalfa and cotton acreage, production, and consumption. Both single equation and systems of equations are estimated. The data consist of 33 annual observations from 1970 to 2002. In some models, there were slightly fewer observations due to lags in the specifications. A brief description of the alfalfa market is given prior to reporting the estimations of the models. In addition, some issues related to the nature of the data are discussed.

Alfalfa

Alfalfa hay acreage in California has averaged about a million acres per year during the past 30 years (Figure 1A). Alfalfa contributes about 85 percent of the value of all hay production in California. Alfalfa is influenced by profitability of alternative annual crops such as cotton, tomatoes, trees, and vines. The demand for alfalfa hay is determined to a large degree by the size of the state's dairy herd, which consumes about 70 percent of the supply. Horses consume about 20 percent. Alfalfa is a perennial crop with a three to five-year economic life. Since it is a water intensive crop, its profitability is strongly influenced by water and water costs. In addition, alfalfa is important in crop rotations because of its beneficial effects on the soil (Johnston, p. 87).

Alfalfa production in California has been increasing annually since the mid nineties (Figure 2A). It reached a peak in 2002 at 8.1 million tons. The increase in production has been primarily due to the upward trend in yields (Figure 3A) and not to increases in acreage. Alfalfa real grower price in California, using a 1983/84 base, has exhibited a downward trend since the early eighties (Figure 6A). In 2002 the real grower price was about $60 per ton.

Model for Alfalfa Acreage

A partial adjustment model of alfalfa acreage is based on the following equation:

$$\begin{aligned} \ln A_t = \beta_0 + \beta_1 \ln A_{t-1} + \beta_2 \ln P_t + \beta_3 \ln risk_t + \beta_4 crit_t + \beta_5 crit_t * \ln A_{t-1} \\ + \beta_6 crit_t * \ln P_t + \beta_7 crit_t * \ln risk_t + \varepsilon_t \end{aligned} \quad (1)$$

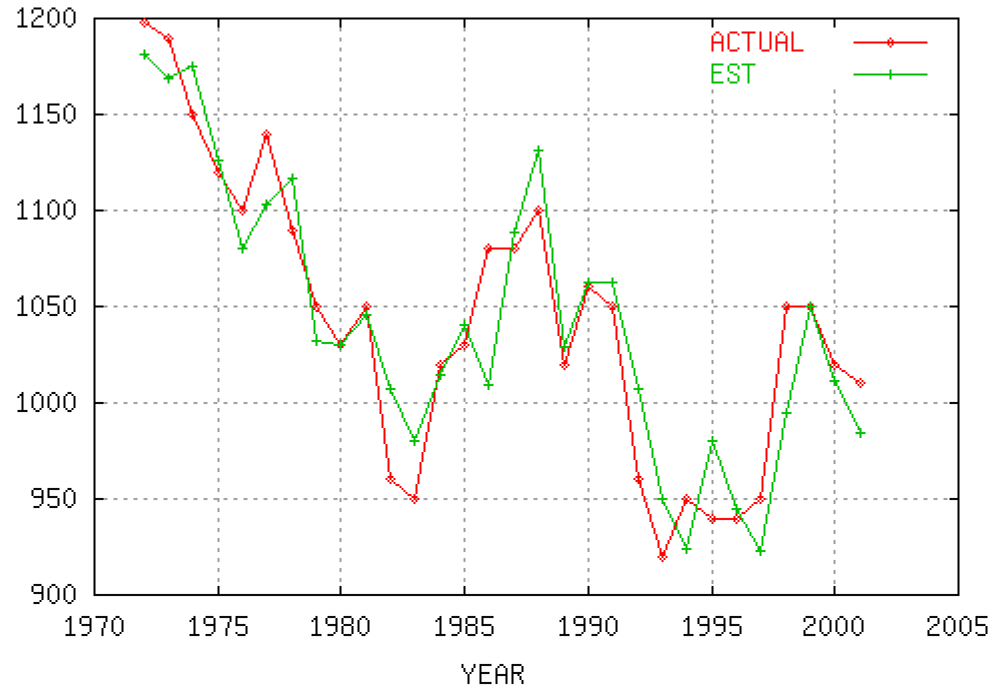
where A_t represents planted alfalfa acreage in thousands of acres, P_t is alfalfa price per ton, $risk_t$ is the variability in alfalfa price (measured by the standard deviation), and $crit_t$ is a dummy variable identifying the *critical years* for water scarcity (i.e., the year when the *Four river index* fell below the value of 5.4). The *Four river index* is an index to measure the water availability in California based on four river flows. The higher the value the more water available. Two interaction terms are also included in the model to capture the effects of water scarcity on prices and risk.

The results of the estimation are (equation 2):

$$\begin{aligned} \ln \hat{A}_t = \underset{(1.66)}{4.08} + \underset{(0.17)}{0.67} \ln A_{t-1} + \underset{(0.16)}{0.35} \ln P_t - \underset{(0.27)}{0.61} \ln risk_t - \underset{(10.95)}{23.80} crit_t + \underset{(1.26)}{2.56} crit_t * \ln A_{t-1} \\ + \underset{(0.59)}{0.31} crit_t * \ln P_t + \underset{(0.58)}{0.67} crit_t * \ln risk_t \end{aligned} \quad (2)$$

where the numbers in parentheses are estimated standard errors. The estimation supports the hypothesis that alfalfa acreage is influenced by prices, ceteris paribus. The short-run price elasticity of acreage is 0.35 and significant when ample water is available and 0.66 when there is a shortage of water. Acreage increases with price expectations and decreases with increases in perceived risk, as anticipated. Also the availability of water has a significant impact on acreage. An F-test on the joint significance of the variable "crit" and its cross products allows us to reject the null hypothesis of no impact at a 90% confidence level (p-value: 0.0787). The signs of the coefficients are consistent with a reduction of planting of new crop acreage during critical years of water scarcity. Furthermore, the estimated coefficient on lagged acreage is 0.67 and significant supporting the partial adjustment framework.

The regression R^2 is 0.847, indicating a good fit. The Durbin h test indicates that there is no autocorrelation in the disturbance terms. Graph 1 depicts the actual and estimated values for alfalfa acreage:



Graph 1: ***Actual and estimated values of alfalfa acreage (in thousands of acres).***

Model for Alfalfa Yield

Alfalfa yield is modeled by the following equation:

$$\ln Y_t = \beta_0 + \beta_1 \ln P_{t-1} + \beta_3 \ln CP_{t-1} + \beta_4 FRI_t + \beta_5 D_t + \varepsilon_t \quad (3)$$

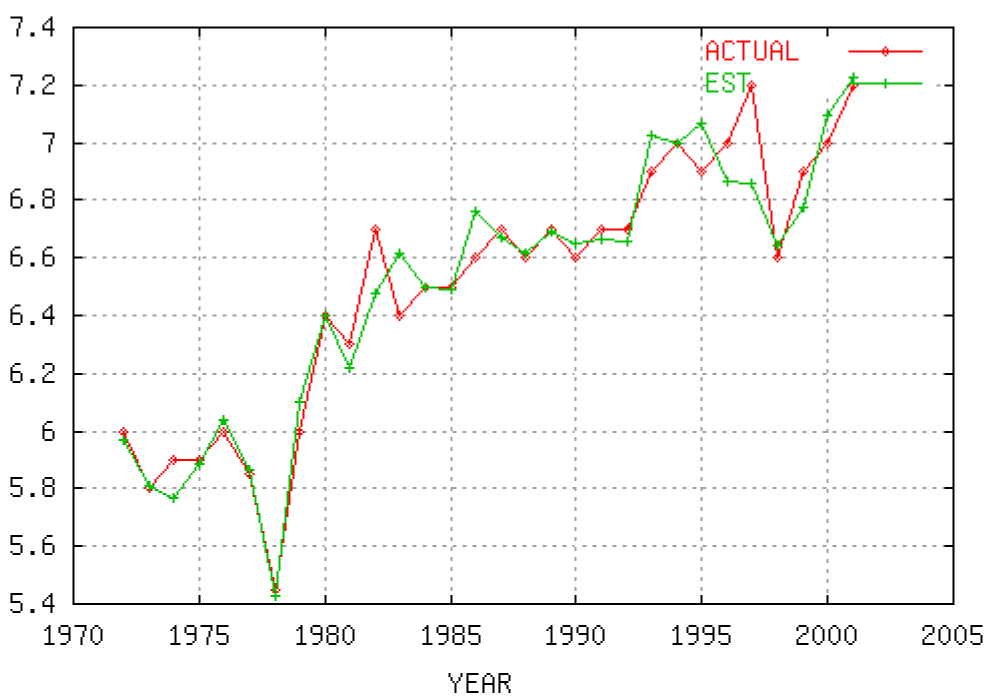
where Y_t is alfalfa yield in tons, P_{t-1} is lagged alfalfa price per ton, CP_{t-1} is lagged cotton price \$/lb.(the rotation crop), FRI_t is the value of the *Four River Index* (approximating the availability of water) and D_t is a dummy variable identifying the year 1978 as an outlier. The model includes a moving average component of order two.

The estimated yield equation is:

$$\ln \hat{Y}_t = \underset{(0.02)}{1.31} + \underset{(0.00)}{0.08} \ln P_{t-1} - \underset{(0.01)}{0.14} \ln CP_{t-1} + \underset{(0.00)}{0.01} FRI_t - \underset{(0.03)}{0.12} D_t \quad (4)$$

where numbers in parentheses are standard errors. The estimated equation indicates that yields respond positively to changes in prices and water availability. Both of these estimated coefficients are highly significant. Alfalfa yields are negatively related to last year's cotton price since they compete for the same irrigated land.. The estimated coefficient is also highly significant. The 1978 dummy coefficient is negative and significant as expected as it was a major drought year. Including a dummy variable for one year is equivalent to eliminating the 1978 observation.

The regression exhibits a good fit (R^2 is 0.93) and the tests ruled out autocorrelation (the Durbin-Watson statistics is 2.00) in the disturbance terms. . Graph 2 describes the actual and estimated alfalfa yields.



Graph 2: ***Actual and estimated values of alfalfa yield (tons/acre).***

Production

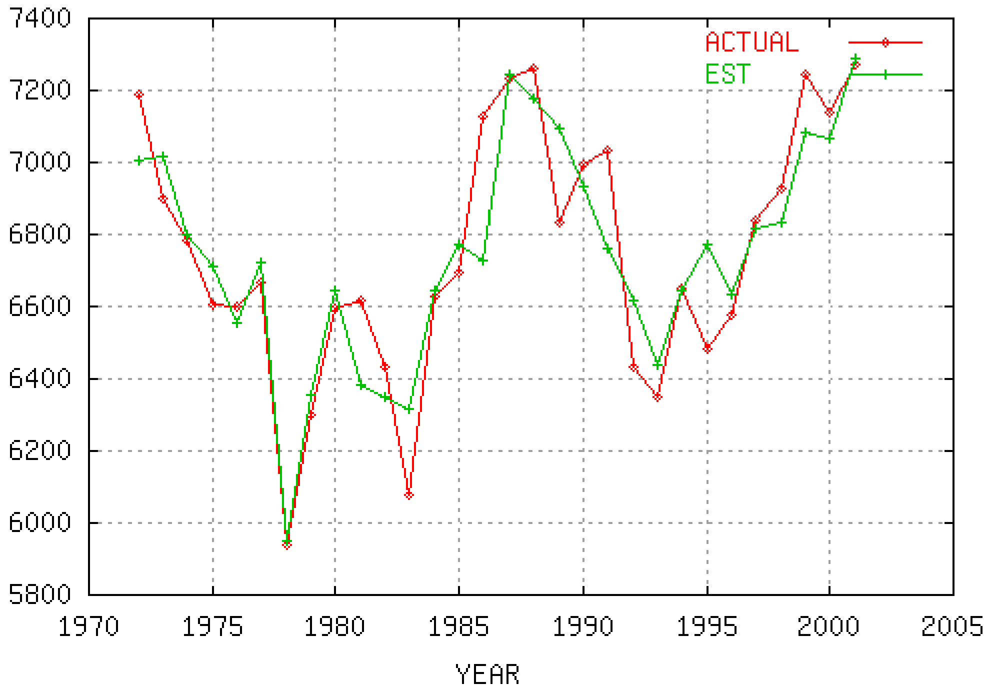
The estimated alfalfa production equation (Table 1) is presented in tabular form in order to better facilitate interpretations of estimated coefficients:

Table 1. Alfalfa Production Equation

Variable	Coefficients	Standard errors
Constant	4.87	1.98
Lag of Log Production	0.69	0.21
Lag of Log Alfalfa Price	0.44	0.17
Lag of Log Alfalfa Risk	-0.75	0.28
Lag of Log Cotton Price	-0.07	0.03
Dummy for *critical years*	-12.07	5.77
Crit*Lag of Log Production	1.33	0.74
Crit*Lag of Log Alfalfa Price	-3.87	1.27
Crit*Lag of Log Alfalfa Risk	3.61	1.02
Crit*Lag of Log Cotton Price	0.13	0.08
Dummy for outlier (1978)	-0.08	0.03

The estimated own-price elasticity is 0.44 and significant at the usual 5% significance level which suggests that alfalfa production is relatively inelastic. Alfalfa production is negatively related to risk (price volatility) and cotton prices. Both estimated coefficients are significant. Water shortages have a negative impact on alfalfa production (see the estimated coefficient of -12.07 on the dummy variable for critical years and is significant).

The regression R^2 is 0.817. The Durbin h statistics (-0.62) indicates that there is no problem with autocorrelation in the errors. Graph 3 plots actual and estimated values of alfalfa production.



Graph 3: ***Actual and estimated values of alfalfa production (thousands of tons).***

Demand

The estimated demand function for alfalfa is a derived demand. Dairies and horse enterprises demand about 90 percent of alfalfa. The assumption made in the estimations is that the market for alfalfa is in equilibrium, that is, that quantity demanded is equal to quantity supplied given the ease of storage this is expected.

The estimated demand equation for alfalfa is given by

$$\hat{Q}_t = -5.904 - 0.107\,price_t + 0.243 milkps_t + 1.736 cows_t + 0.105\,prmix_t - 0.606\,prmilk_t \quad (6)$$

(2.626) (0.107) (0.042) (0.288) (0.039) (0.113)

where Q_t is the quantity demanded of alfalfa in tons, $price_t$ is the real grower price of alfalfa in \$/ton, $milkps_t$ is the milk price support, cow_t is the number of cows, $prmix_t$ is the price of a combination of corn and soybeans, and $prmilk_t$ is the real price of milk. All variables are expressed in logarithmic form.

The coefficient of determination, $R^2 = 0.888$, indicates a good fit of the model with the data. The own-price elasticity of demand is -0.107 which is inelastic, but not statistically significant. The estimated coefficient of milk support price is 0.243 implying that the quantity demanded of alfalfa increases as the support price of milk increases. The estimated coefficient on real price of milk is negative. The coefficient on the number of cows is positive and statistically significant. This is reasonable given that about 70% of the demand for alfalfa is from dairies. All of the coefficients in the demand equation are statistically significant at the five percent level of significance except for own price.

System for Alfalfa

A three-equation system for alfalfa was developed and estimated. Iterative three-stage least squares are used to estimate a model consisting of acreage, production, and demand relationships for alfalfa. We assume that the market for alfalfa is in equilibrium, that is, that quantity demanded is equal to production. We further assume that stocks are included in the demand for alfalfa. Thus, the three endogenous variables are: acreage, production, and alfalfa price. The estimators will be asymptotically efficient given that the model is specified correctly. The gain in efficiency is due to taking into account the correlation across equations. And three-stage least squares will purge (asymptotically) the correlation that exist between endogenous variables on the right hand side of the equations in the model with the error terms.

The estimated alfalfa system is given by

$$\hat{A}_t = 4.210 + 0.133 price_{t-1} - 0.277 risk_{t-1} + 0.532 A_{t-1}$$
(0.097) (0.159) (0.159) (0.111)

$$\hat{Y}_t = 2.630 + 0.601 A_t + 0.037 price_{t-1} - 0.088 prcot_{t-1} + 0.199 Y_{t-1} - 0.109 D_t \quad (7)$$
(0.834) (0.150) (0.015) (0.021) (0.128) (0.109)

$$\hat{Q}_t = 3.962 - 0.020 price_t - 0.061 prcorn_t + 0.037 prsoy_t + 0.475 Q_{t-1} - 0.114 D_t + 0.091 cow_t$$
(1.227) (0.015) (0.036) (0.037) (0.108) (0.027) (0.101)

where A_t represents acreage of alfalfa, Y_t denotes production of alfalfa, Q_t is the quantity demanded of alfalfa, and the remaining variables are defined above. The own-price elasticity is 0.133 in the acreage response equation but is not statistically significant at the five percent level of significance. Acreage response decreases as risk increases as measured by the standard deviation of alfalfa monthly prices. Production of alfalfa is positively related to alfalfa price, is negatively related to cotton prices, and positively correlated to past acreage and production. Alfalfa demand has a very low own-price elasticity of demand of -0.020. Alfalfa demand is negatively related to price of corn but positively related to soybean prices. Demand is positively related to the number of cows. Recall that about 70% of the demand for alfalfa is from dairies. The majority of the estimated coefficients are statistically significant at the 5% level.

Cotton

Cotton is the most important field crop gown in California. Growers in California grow two types of cotton: Upland, or Acala and Pima. Upland cotton makes up about 70 to 75 percent of the California cotton market and is the higher-quality cotton. Upland has a worldwide reputation as the premium medium staple cotton, with consistently high fiber strength useful in many apparel fabric applications. Export markets are important, attracting as much as 80 percent of California's annual cotton production in some years making it California's second highest export crop (Johnston, p. 84). Historically, California cotton, in terms of value of production, was the third highest ranking crop in California in 1950 below cattle and calves and dairy products. In 2001 cotton was ranked the eighth highest valued crop below milk and cream, grapes, nursery products, cattle and calves, lettuce, oranges, and hay (McCalla and Johnston).

There has been a downward trend in cotton acreage and production in California since 1979. California growers produced 3.4 million bales of cotton on l.6 million acres in 1979. In 2002 they produced about 2 million bales of cotton on 700,000 acres (Figures 10A and 12A). Cotton yields have experienced an upward trend since 1979 (Figure 11A). Nominal producers' prices in California for cotton exhibit an upward trend since the 1970s, but real producers' prices in California has exhibit a downward trend since the mid seventies (Figures 13A and 14A).

Recently the World Trade Organization (WTO) ruled against U.S. cotton subsidies. U.S. cotton subsidies totaled about $10 billion in 2002 and the WTO ruled that the subsidies created an unfair competition for Brazil, which filed the complaint. California producers received about $1.2 billion in subsidies in 2002. California cotton is not as subsidized as cotton in other states, such as Texas, because subsidies are based on price and California's higher-quality cotton is more expensive (Evans, May 3, 2004).

Acreage, production, and demand equations are estimated for California cotton. Single equation and system of equations models are developed and estimated. In this report we aggregated the different cotton varieties. Disaggregated models of cotton were also estimated because of changes in the cotton industry and to allow for different impacts for subsidized and unsubsidized varieties. The number of observations in the disaggregated models present in the next section are limited due to the relatively recent introduction of Pima in California.

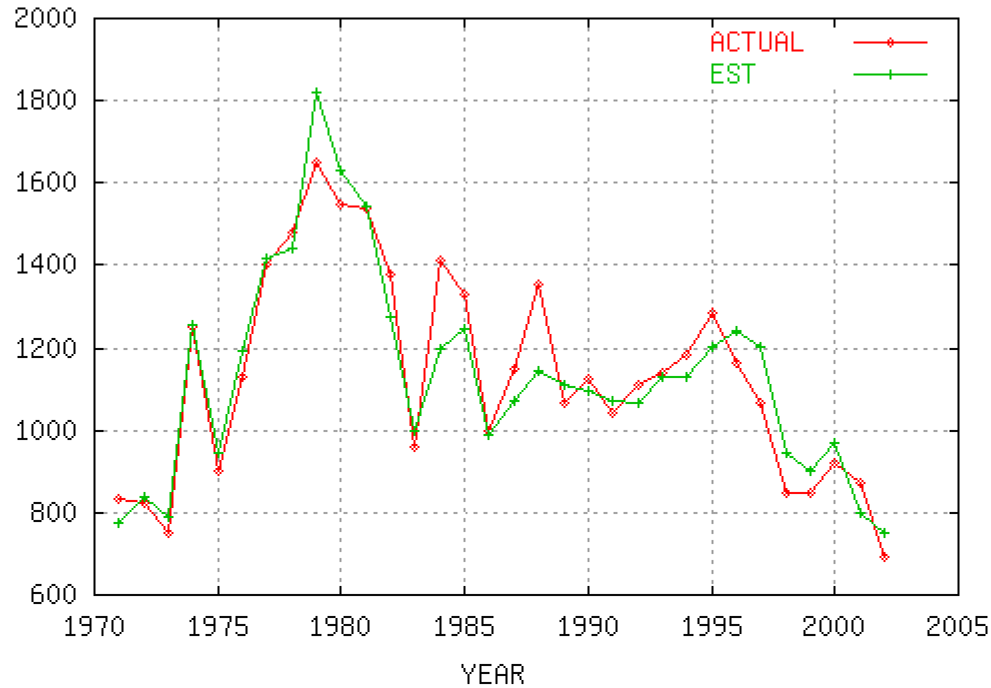
Acreage

The estimated planted acreage relationship, a partial adjustment model, for California cotton is

$$\ln \hat{A}_t = -4.19 + 0.53 \ln price_t - 0.05 \ln riskc_t - 1.47 \ln pricealf_t + 2.87 \ln riska_t + 0.27 \ln A_{t-1}$$
$$(1.26)(0.06) \quad (0.03) \quad (0.26) \quad (0.42) \quad (0.07) \tag{8}$$

where A_t is cotton acreage in thousands of acres, $price_t$ is real cotton price in \$/lb., $risk_t$ is the standard deviation of monthly cotton prices and is a measure of risk, $pricealf_t$ denotes real alfalfa price in \$/ton, and $riska_t$ represents the standard deviation of monthly alfalfa price and is a measure of risk of growing alfalfa. All variables are expressed in logarithmic form.

The estimated coefficient of determination is R^2 =0.899. The short-run own-price acreage elasticity of cotton is 0.53 and is highly significant. Cotton acreage decreases with an increase in risk in growing cotton and as price of alfalfa increases. All of the estimated coefficients are statistically significant at the 1% level except for the risk coefficient associated with cotton which is significant at the 10% level. A graph depicting the estimated acreage equation with the actual cotton acreage is given in Graph 4.



Graph 4: *Actual and estimated cotton acreage (thousands of acres).*

The Durbin h statistics (1.12) fails to reject the null hypothesis of no autocorrelation in the disturbances.

Production

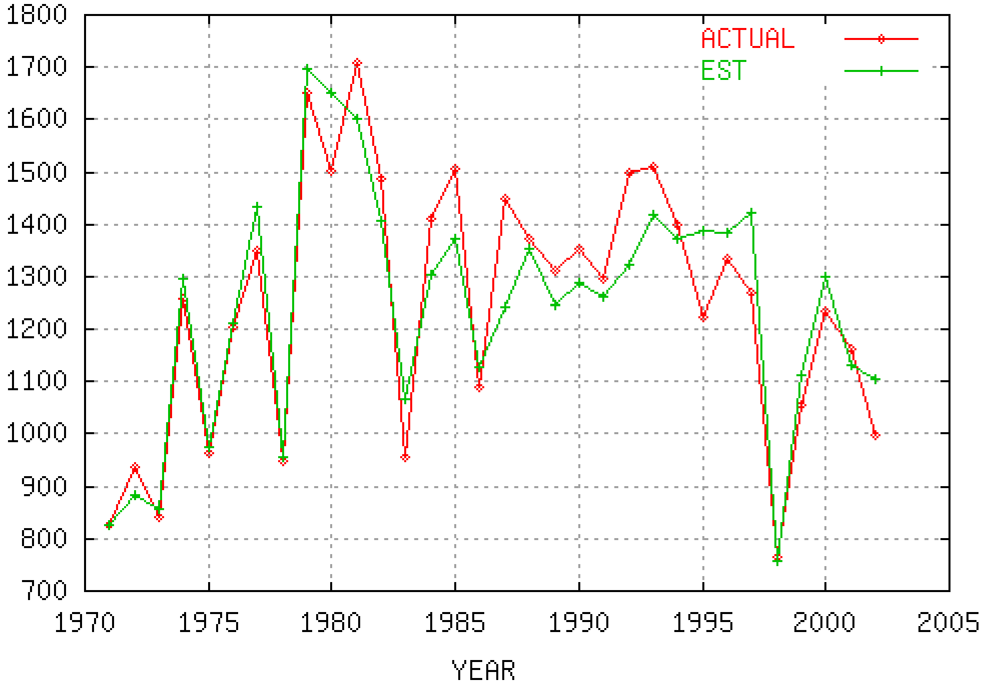
The estimated production relationship for cotton, an adaptive expectations model, is (eq. 9)

$$\hat{Y}_t = -7.066 \ + 0.497\,pricec_t - 0.499 riskc_t - 1.844\,pricea_t + 4.067 riska_t + 0.011 Y_{t-1} - 0.313 D_t$$

(2.444) (0.115) (0.036) (0.543) (0.880) (0.009) (0.081)

where Y_t denotes cotton production in 1000 bales, and D_t denotes a dummy variable for the drought year, 1978. The remaining variables are as defined above. An adaptive expectations models implies a moving average error process of order one and the production function was estimated with a MA(1) error scheme.

The goodness of fit yields an $R^2 = 0.878$. All of the estimated coefficients are statistically significant from zero at the 5% level except for the risk measure for cotton and lagged cotton production. The short-run price elasticity is 0.497 and the long-run price elasticity is 0.503 [0.497/(1-0.011)]. The estimate coefficients on risk and the dummy variable are negative as anticipated. A plot of the estimated production of cotton with the actual production of cotton is given in Graph 5.



Graph 5: *Actual and estimated cotton production (1000 bales).*

Demand

The estimated demand function for cotton is given by (eq. 10)

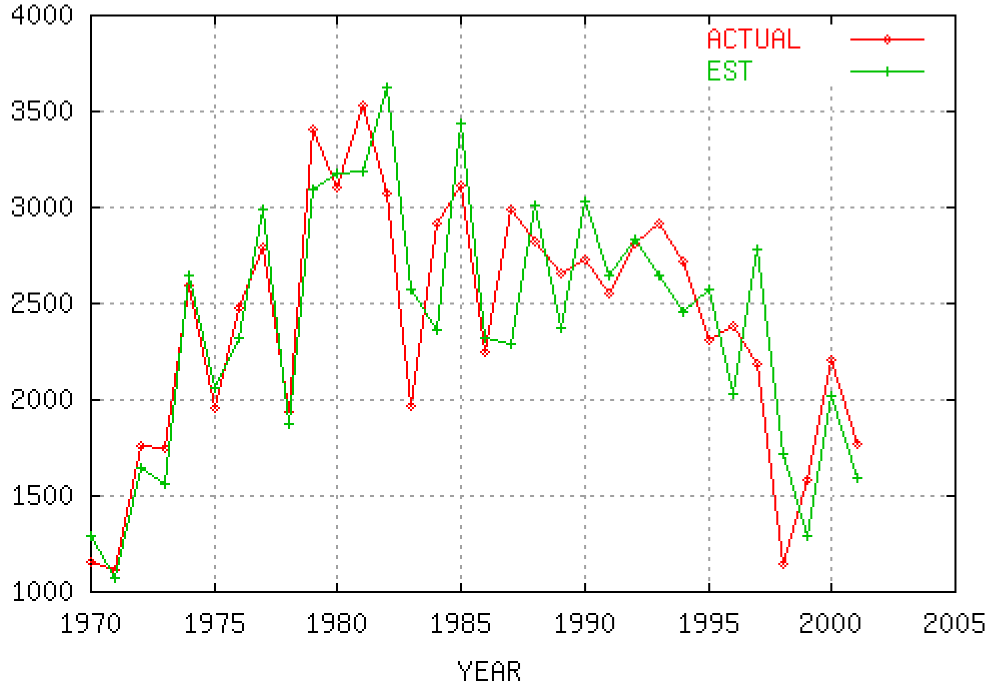
$$\hat{Q}_t = -12.631 - 0.684\,prc_t + 0.360\,prus_t + 0.827\,prray_t - 0.064\,prpol_t + 0.000\,pop_t$$
(13.490) (0.228) (0.293) (0.493) (0.544) (0.000)

$$-0.217pop_t - 0.070t - 0.004t^2$$
(0.100) (0.117) (0.002)

where Q_t denotes the US disappearance plus US imports of cotton, prc_t denotes the real grower price of California cotton, $prus_t$ represents the United States price of cotton, $prray_t$ denotes the price of rayon, a substitute for cotton, $prpol_t$ denotes the price of polyester, a substitute for cotton, pop_t represents US population, D_t is a dummy variable for the drought year in 1978, and t denotes a time trend. All variables, except the time trend and dummy variable, are expressed in logarithmic form.

The overall goodness of fit was 0.756. The estimated own-price elasticity of California cotton is -0.684 and significant. The positive coefficient on rayon indicates that it is a gross substitute for cotton while the negative sign on polyester indicates a gross complement. There is a negative sign associated with the time trend indicating that the demand for cotton has been decreasing over the sample period

A plot of the estimated and actual demand series for cotton is depicted in Graph 6.



Graph 6: Actual and estimated cotton demand (thousands of bales).

System for Cotton

A two-equation system for cotton was developed and estimated by iterated three-stage least squares (3SLS). The estimated cotton production and demand system (eq.11) is

$$\ln \hat{Y}_t = -1.13 + 0.46 \ln Pc_t - 0.49 \ln Riskc_t - 0.82 \ln Pa_t + 2.14 \ln Riska_t$$
$$\quad (2.49)\ (0.12) \qquad (0.05) \qquad (0.52) \qquad (0.87)$$
$$+0.03 \ln \mathrm{Y}_{t\text{-}1} - 0.41 D_t$$
$$(0.03) \qquad (0.20)$$

and

$$\ln \hat{Q}_t = 6.89 - 0.95 \ln Pc_t + 1.24 \ln \mathrm{Prus}_t + 0.23 \ln \mathrm{Prray}_t - 0.00 \ln \mathrm{Prpol}_t - 0.05 \ln Pcin_t$$
$$(1.96)\ (0.99) \qquad (0.78) \qquad (0.41) \qquad (0.37) \qquad (0.04)$$
$$-0.24 \mathrm{D}_t \ + 0.07 \mathrm{t} - 0.03 \mathrm{t}^2$$
$$(0.23) \qquad (0.02)\ (0.00)$$

where $pcin_t$ denotes per capita income and the remaining variables are defined above. The first equation represents the production equation for cotton and the second equation is the demand function for cotton. All variables are expressed in logarithmic form. The own-price elasticity is 0.46 for the production of cotton and the own-price elasticity of demand for cotton is -0.95. Both elasticities are inelastic and of the correct sign. The signs on the risk variables are as expected. The cross-price elasticity estimates of rayon and polyester indicate that they are both gross substitutes for cotton. The estimated coefficients on time and time squared indicates that the demand for cotton is trending upward at a decreasing rate. The sign on per capita income coefficient is unexpectedly negative, but not significant.

Modeling Variety Substitution

In California, currently two major varieties of cotton are grown: Upland (Acala) and Pima. Variety differentiation is a phenomenon that is relatively recent, because until late 1980s the so-called “law of one variety” allowed California farmers to grow only Upland (Acala). The bill was revised in 1988 and again in 1991 introducing a broader set of choices for farmers. In 2004, 550 thousand acres of Upland and 220 thousand acres of Pima were planted. Figures 7 and 8 summarize the acreage and production trends from 1970 to 2002.

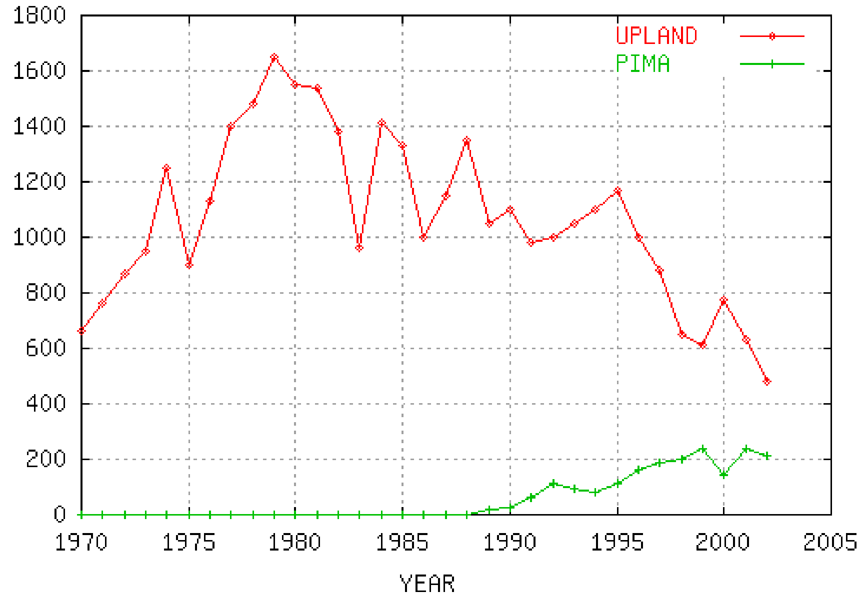


Figure 7: Upland and Pima planted acreage in California (thousand of acres).

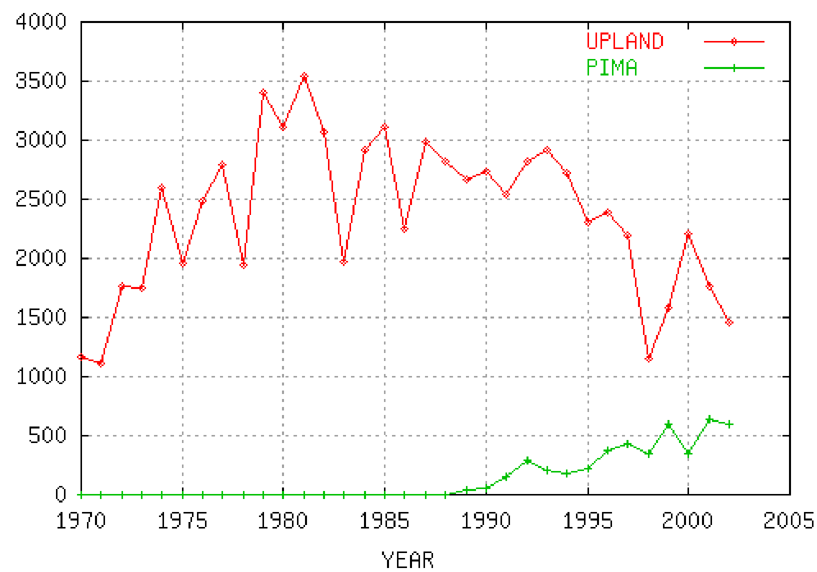


Figure 8: Upland and Pima production in California (thousand of bales).

The graphs show that pima acreage and production are gradually increasing over time. Farmers are gradually adopting the new variety. Since the abolishment of the law of one variety is relatively recent, we have no way to assess if the process has reached a steady state. However, pima cotton is more sensitive to rainfall conditions, and experts expect that the final crop pattern in California will be a mixture of pima and upland, depending on local weather conditions.

The rationale for the adoption of the new variety can be found, in part, in Figure 9, that reports the real grower prices for pima and upland.

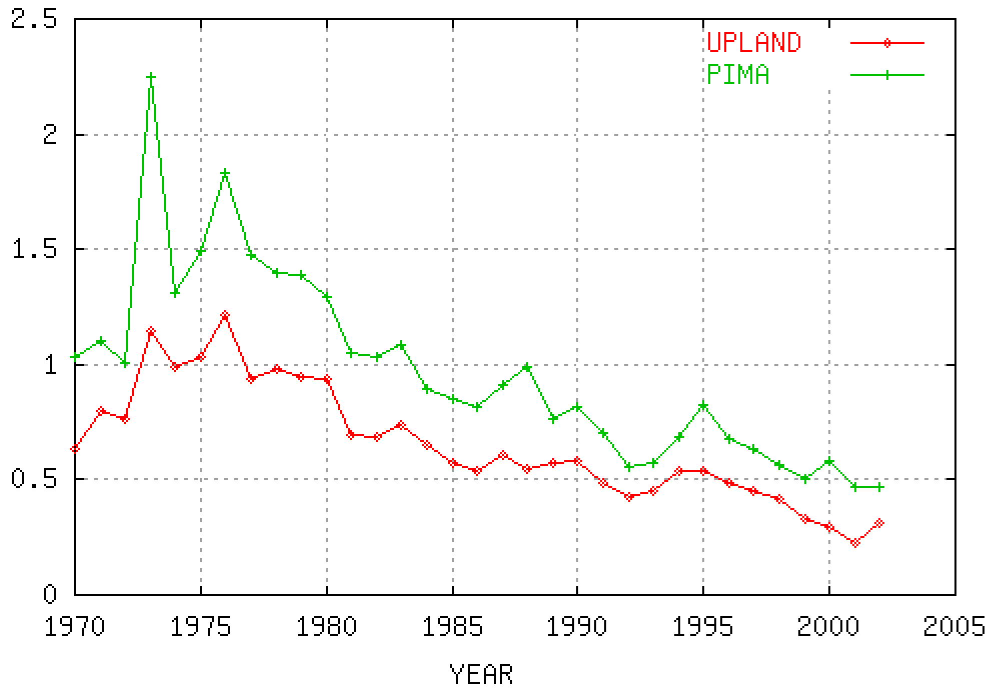


Figure 9: Real prices for Pima and Upland (dollars/lb.).

The graph shows that pima growers benefit from a price premium relative to upland producers. If weather conditions are favorable, pima is considered more profitable. The time trends also show that the price of upland and pima are cointegrated, suggesting a strong theoretical argument for modeling aggregate cotton production regardless of variety (as we did in the previous section).

In this section we adopted a partial adjustment model of the new variety based on relative prices. Given the relevance of the pima production, the model can provide useful indications, however it must be pointed out that: (i) the phenomenon is still too recent to allow reliable statistical analyses based on a time series approach, and (ii) the short time series poses a strong constraint in the number of explanatory variables that can be incorporated into the model.

We designed a model based on an equation for pima acreage and an equation for upland acreage. In both cases we assumed that farmers follow a behavior pattern based on partial adjustments of acreage.

The equations are

$$\begin{aligned} AP_t &= \beta_0 + \beta_1 AP_{t-1} + \beta_2 P_t^P + \beta_3 P_t^U + \varepsilon_t \\ AU_t &= \alpha_0 + \alpha_1 UP_{t-1} + \alpha_2 P_t^U + \alpha_3 P_t^P + u_t \end{aligned} \qquad (12)$$

where AP and AU are pima and upland acreage, respectively, P^P and P^U are pima and upland real prices and ε and u are error terms. All the variables are in logarithm form. The model was estimated both as single equations and as a SUR system. The results of the estimation are the following.

Single-equation estimations:

Upland estimation:

$$\hat{AU}_t = -0.66 + 0.91 AU_{t-1} + 1.76 P_t^U - 0.86 P_t^P \qquad R^2 = 0.81$$
$$(2.19)\ (0.32) \qquad (0.75) \qquad (0.39) \qquad (13)$$

Pima estimation:

$$\hat{AP}_t = 4.49 + 0.74 AP_{t-1} + 2.98 P_t^P - 3.86 P_t^U \qquad R^2 = 0.96$$
$$(0.72)\ (0.08) \qquad (0.78) \qquad (1.14) \qquad (14)$$

where the number in parentheses are standard errors. The test statistics for a single coefficient possess a t distribution with 10 degrees of freedom.

Upland cotton prices have a positive impact on acres planted to Upland. When prices of Pima increase, the acres planted to Upland decrease. Thus, Upland and Pima are gross substitutes. Both price coefficients are significant. With respect to the Pima acreage equation, Pima prices have a positive effect on acres planted to Pima. Upland prices have a negative relationship, as expected, with Pima planted acres.

SUR estimation

Upland

$$\hat{AU}_t = -0.74 + 0.71 AU_{t-1} + 1.92 P_t^U - 0.57 P_t^P \qquad R^2 = 0.78$$
$$(2.16)\ (0.32) \qquad (0.82) \qquad (0.40) \qquad (15)$$

Pima

$$A\hat{P}_t = 4.42 + 0.78AP_{t-1} + 2.89P_t^P - 4.26P_t^U \qquad R^2 = 0.96$$
$$\quad (1.11)\ (0.05) \qquad (1.00) \qquad (1.72) \tag{16}$$

The two procedures (single-equation approach and SUR) give similar estimations. In the SUR results the coefficient of Pima prices is insignificant in determining Upland acreage. However, it must be noted that the explanatory variables have a high degree of multicollinearity.

The model confirms the hypothesis that the relative prices of Upland and Pima are driving forces in the adoption process at the state level in California.

Conclusions

The estimated models indicate that the short-run own-price elasticity of alfalfa acreage is inelastic (0.35) but more elastic (0.66) when ample water is available. By applying water marginally through out the growing period, a producer can obtain more cuttings of alfalfa. Alfalfa yields are also responsive to increases in prices. The own-price elasticity of yields is 0.08 and highly significant. Alfalfa yields are negatively related to the previous year's cotton price. Production is positively related to own price with an estimated elasticity of 0.44 and significant. Production was negatively related to risk with an elasticity of risk equal to -0.75. Demand for alfalfa is a derived demand and is positively related to the number of cows and milk price support and negatively related to its own price.

The estimated own-price elasticity of cotton acreage is 0.53 and highly significant. Cotton acreage decreases with an increase in risk in growing cotton and as price of alfalfa increases. The short-run own-price elasticity of cotton production is

0.497 and the long-run estimate is 0.503. The own-price elasticity of cotton demand is -0.684. Rayon is a substitute for cotton. The empirical results support the fact that alfalfa and cotton are rotating crops in California.

In recent years there has been an increase in Pima acreage relative to the traditional Upland variety in California. Upland cotton prices have a positive impact on acres planted to Upland. When Pima prices increase, the acres planted to Upland decrease. A similar situation applies to Pima acreage. That is, an increase in Upland prices causes a decrease in Pima acreage. Thus, the empirical results support that hypothesis that relative prices of Upland and Pima have a significant impact on the adoption of the two varieties.

Future research needs to focus on the collection of more data related to the consumption of California cotton and alfalfa, stocks and inventories, and interstate trade of alfalfa between California and Oregon and Nevada.

References

Davidson, R. and J. MacKinnon. *Econometric Theory and Methods*, Oxford University Press, Oxford, 2004.

Evans, J., "Cloud Over Cotton Farmers", *Sacramento Bee*, May 3, 2004.

Hamilton, J., *Time Series Analysis,* Princeton University Press, Princeton, NJ, 1994.

Johnston, W., "Cross Sections of a Diverse Agriculture: Profiles of California's Production Regions and Principal Commodities" in *California Agriculture Issues and Challenges*, J. Siebert, ed., University of California, Giannini Foundation, August 1997.

Johnston, W. and A. McCalla, "Whither California Agriculture: Up, Down or Out? Some Thoughts About the Future", Giannini Foundation, forthcoming 2004.

Stock, J. and M. Watson. *Introduction to Econometrics,* Addison-Wesley, Boston, MA, 2003.

White, K., *SHAZAM, User's Reference Manual*, Version 9.0, Irwin/McGraw-Hill, Vancouver, B.C., Canada, 2000.

ALFALFA HAY

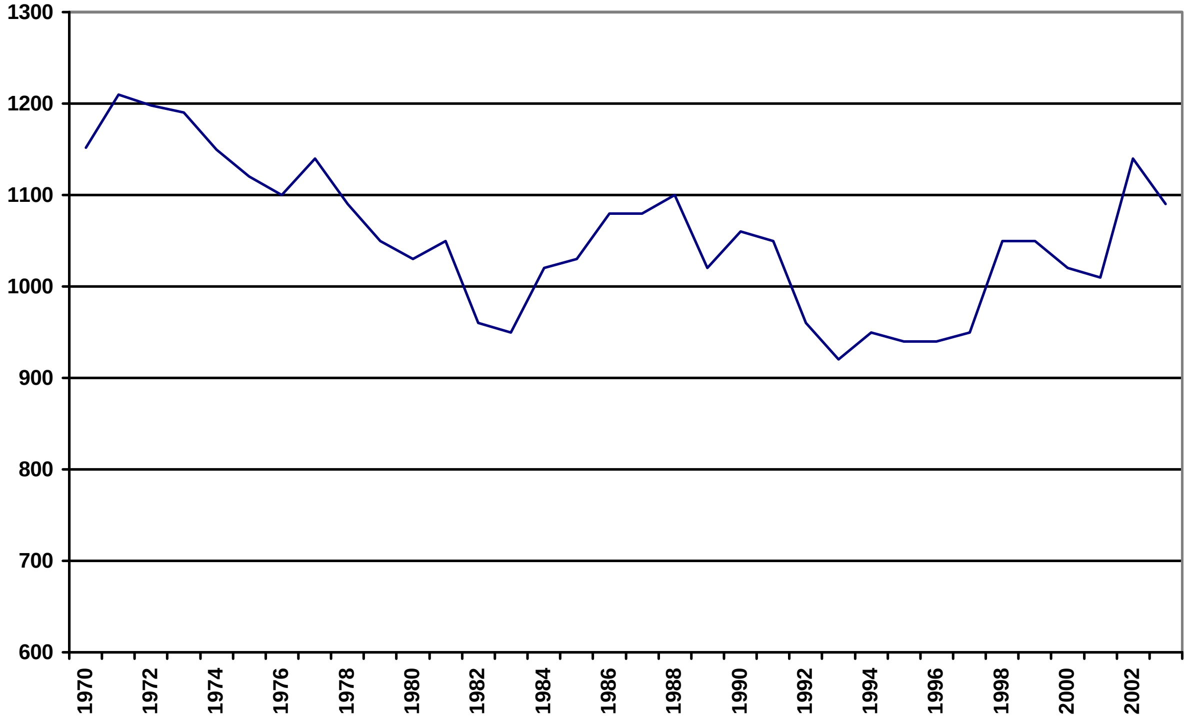


Figure 1A: Harvested Acreage for Alfalfa in California (Thousands of Acres).

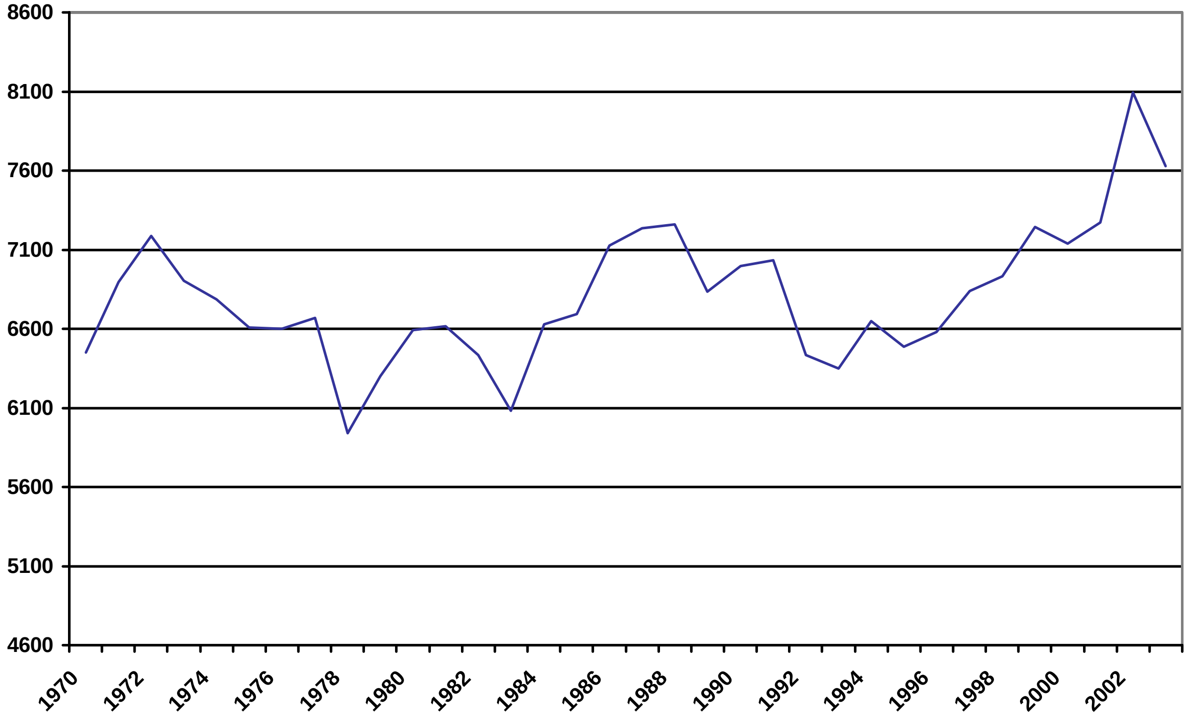


Figure 2A: Alfalfa production in California (Thousands of tons)

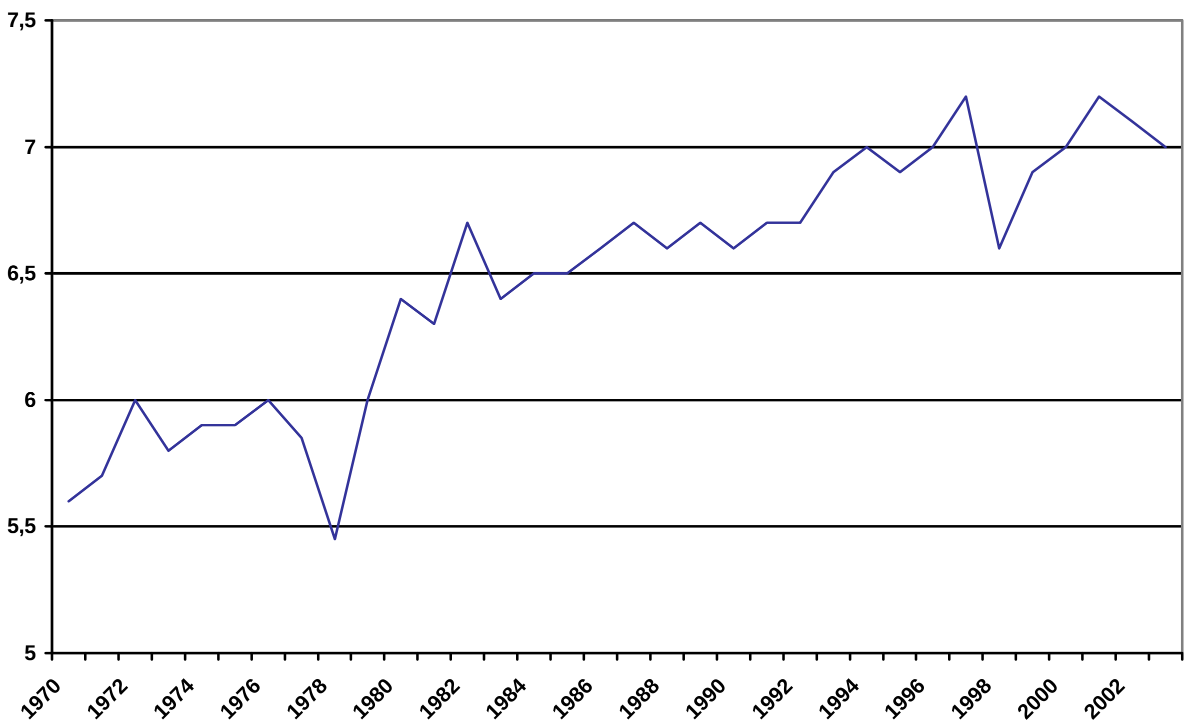


Figure 3A: Alfalfa Yield in California (tons)

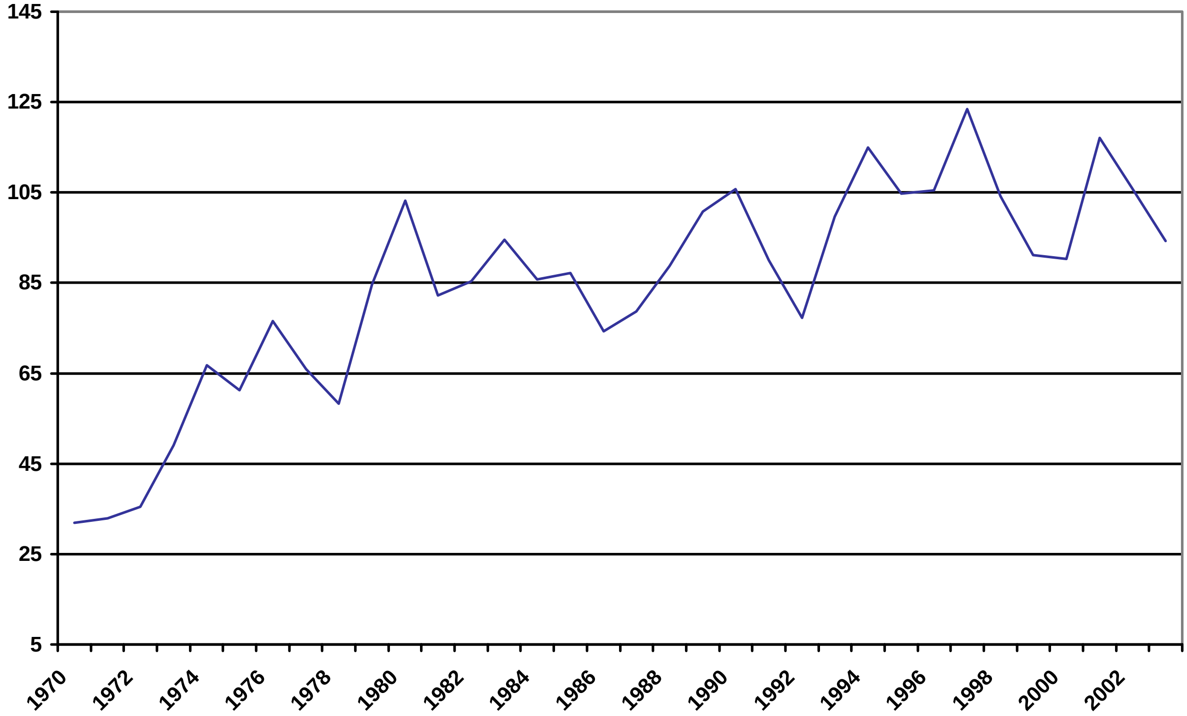


Figure 4A: Alfalfa Nominal Grower Price in California (12 month average)

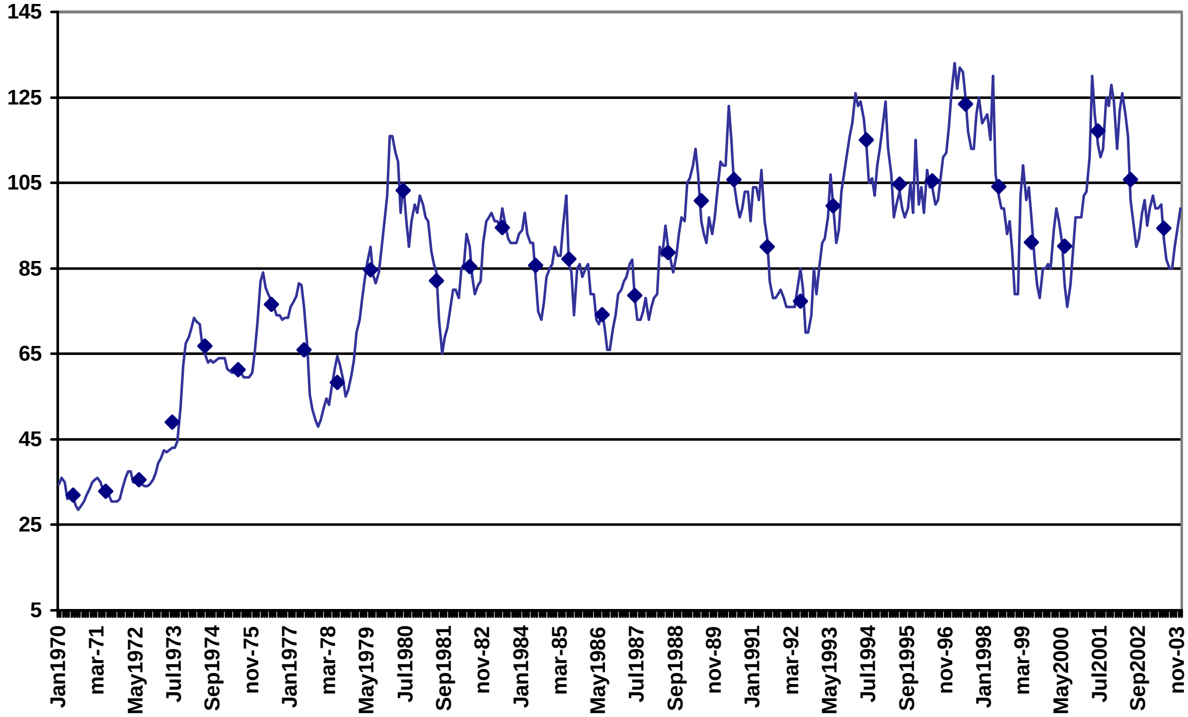


Figure 5A: Alfalfa Nominal Grower Price (Monthly- dollars per ton)

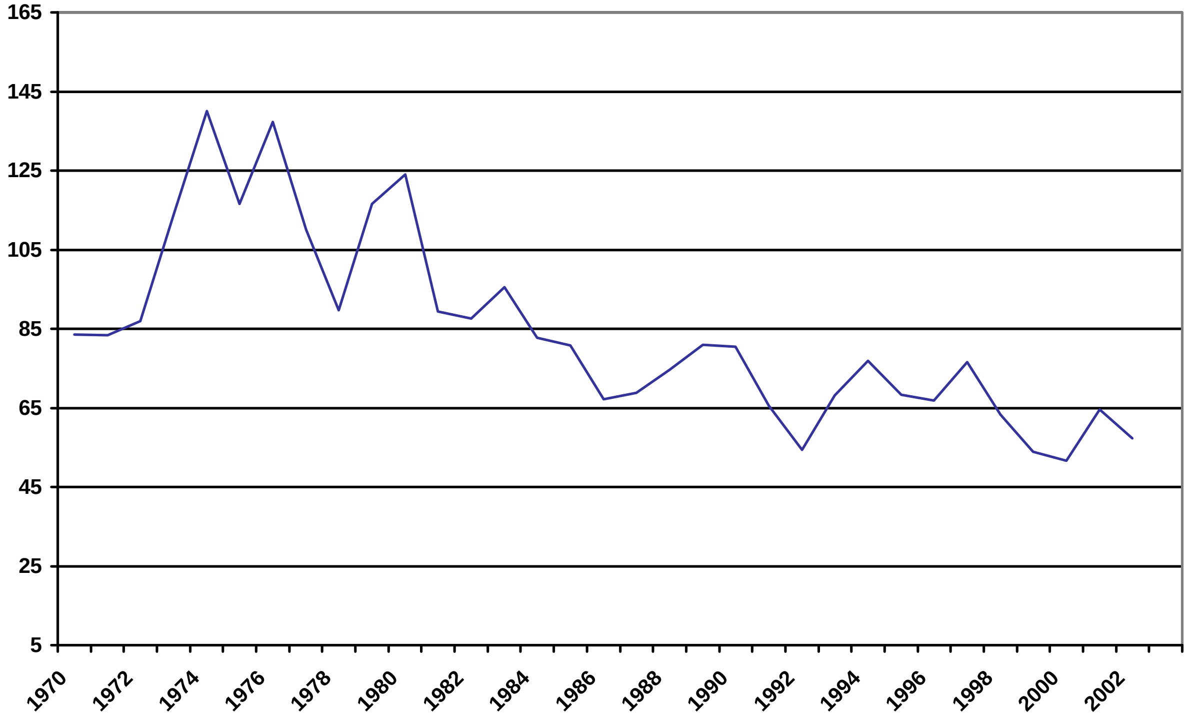


Figure 6A: Alfalfa Real Grower Price in California (12 month average, dollars per tons – base 1983/4)

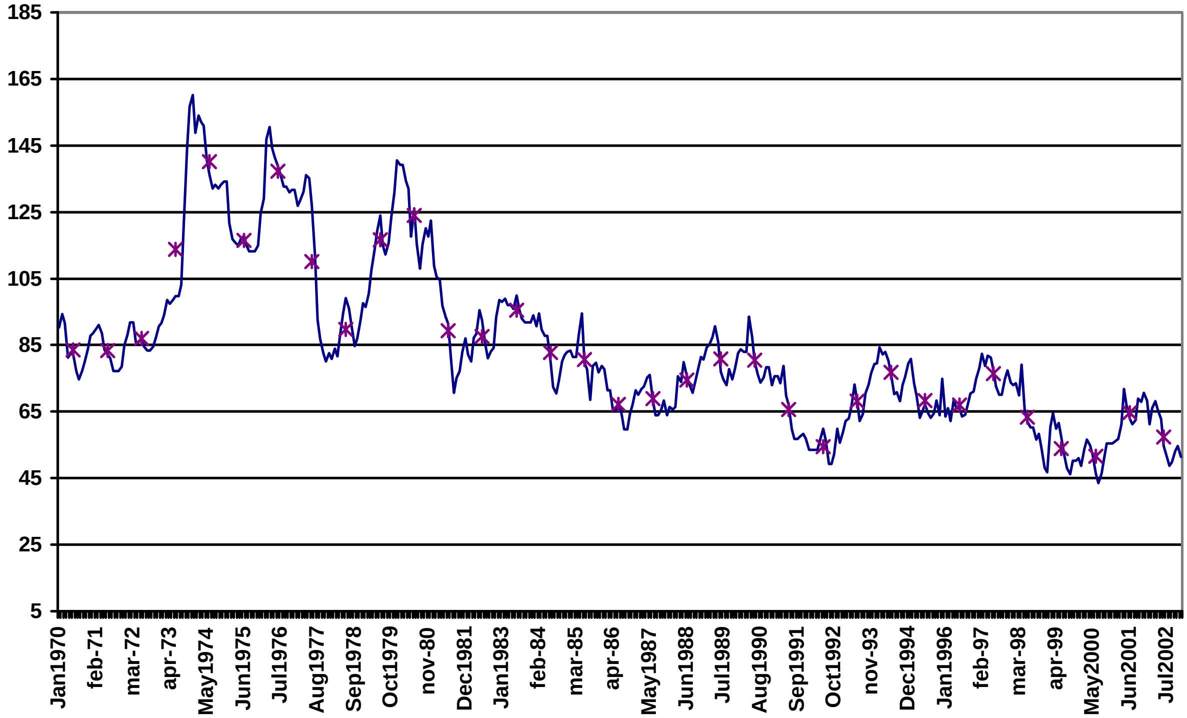


Figure 7A: Alfalfa Real Grower Price in California (monthly, dollars per tons– base 1983/4)

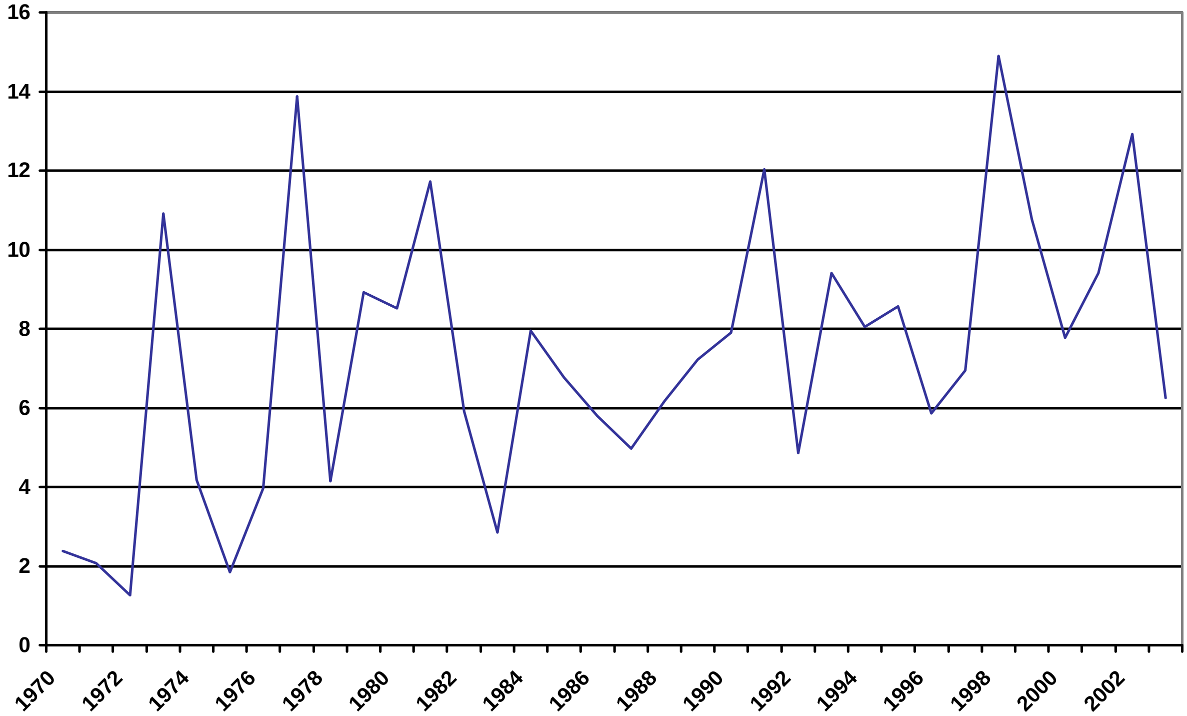


Figure 8A: Standard Deviation of Monthly Alfalfa Price (Nominal, $/month).

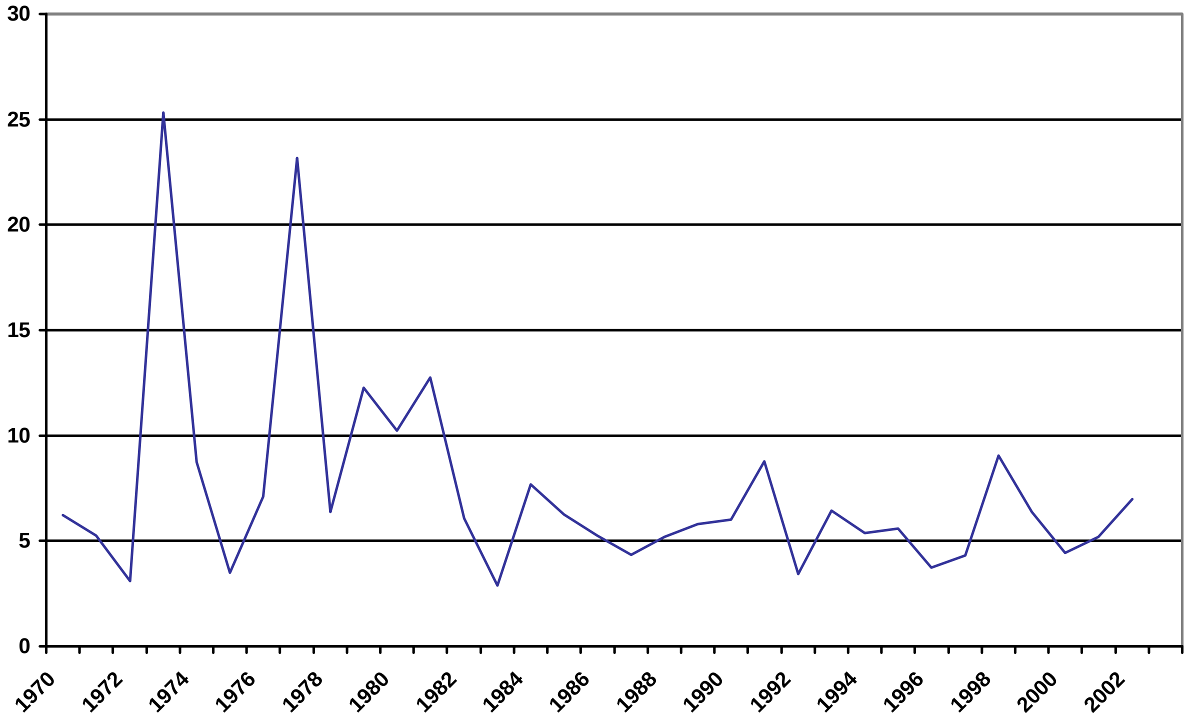


Figure 9A: Standard Deviation of Monthly Alfalfa Price (Real dollars per month, base 1983/4).

COTTON

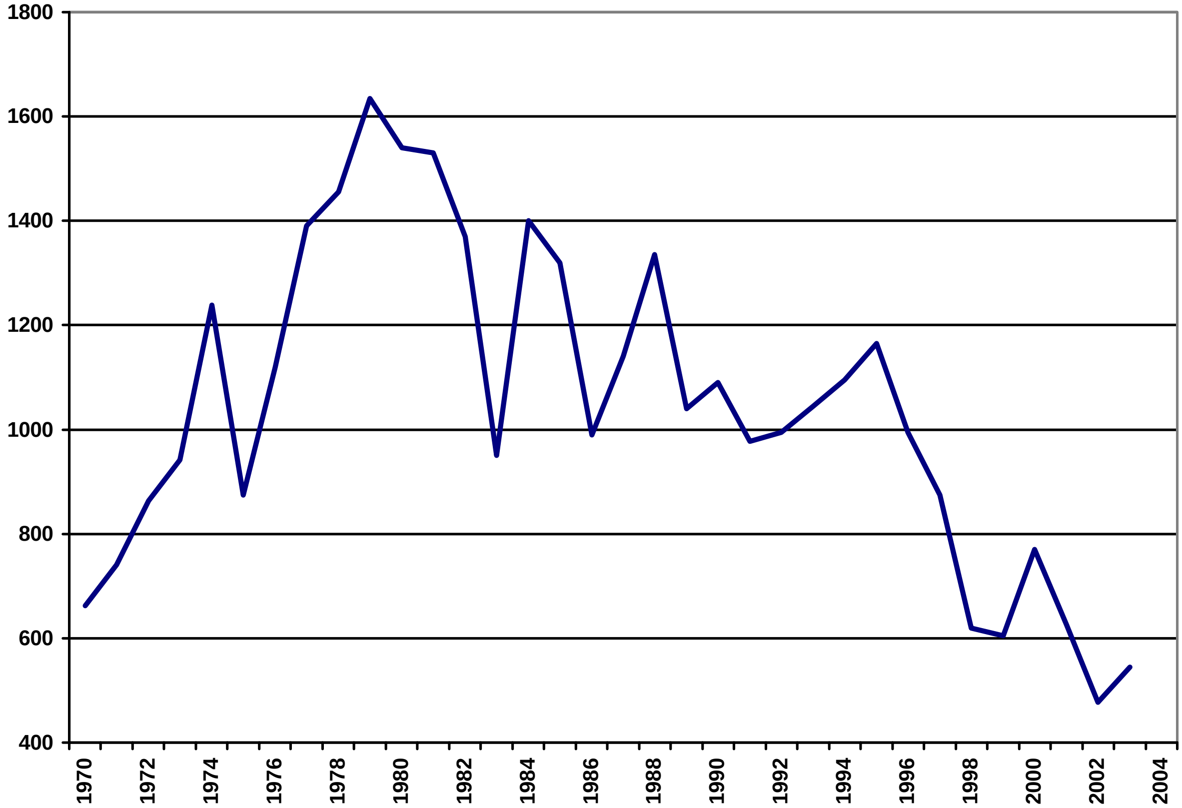


Figure 10A: Cotton Acreage in California (acres/ in thousands).

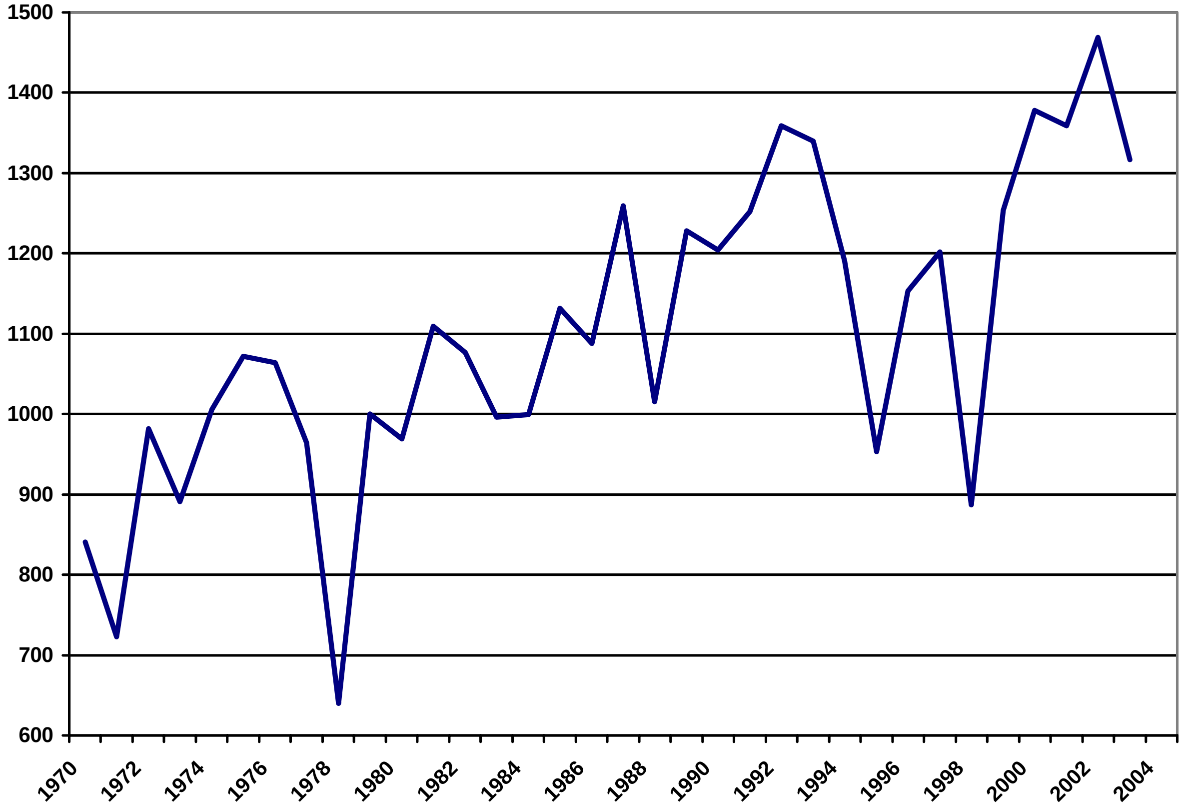


Figure 11A: Cotton Yield in California (pounds).

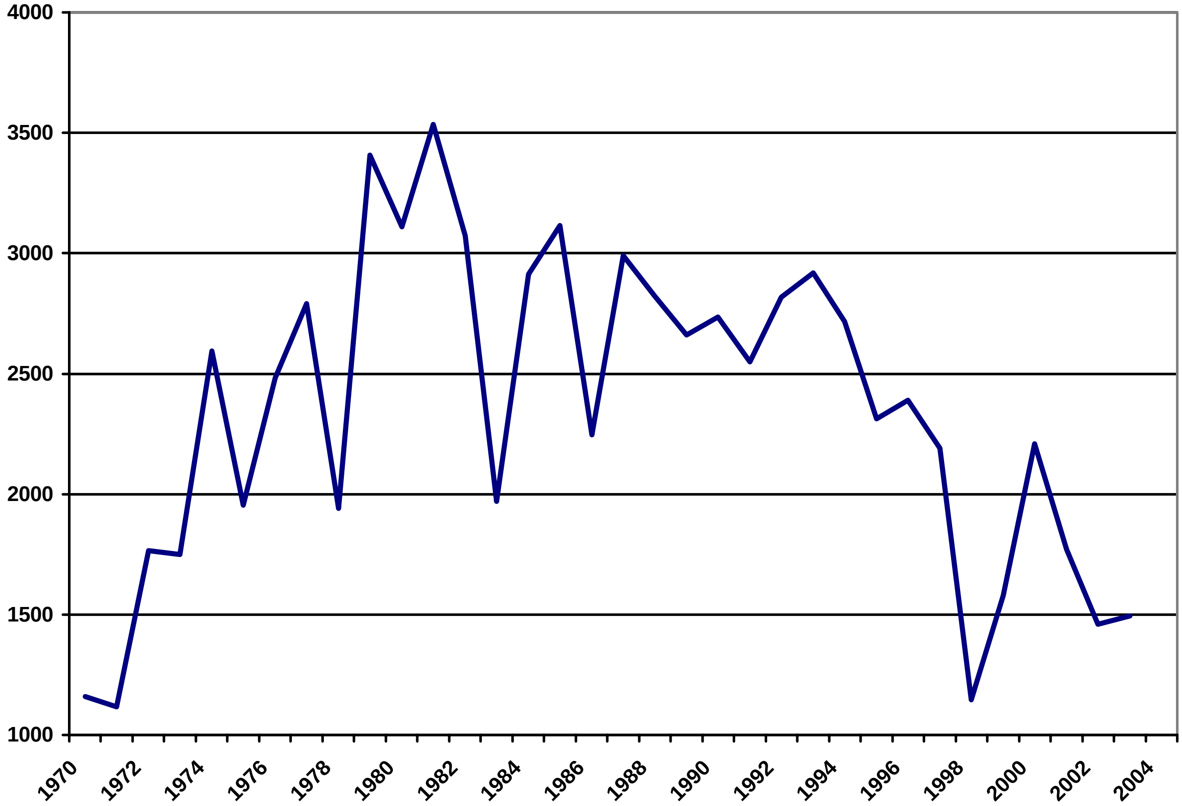


Figure 12A: Cotton Production in California (1000 bales).

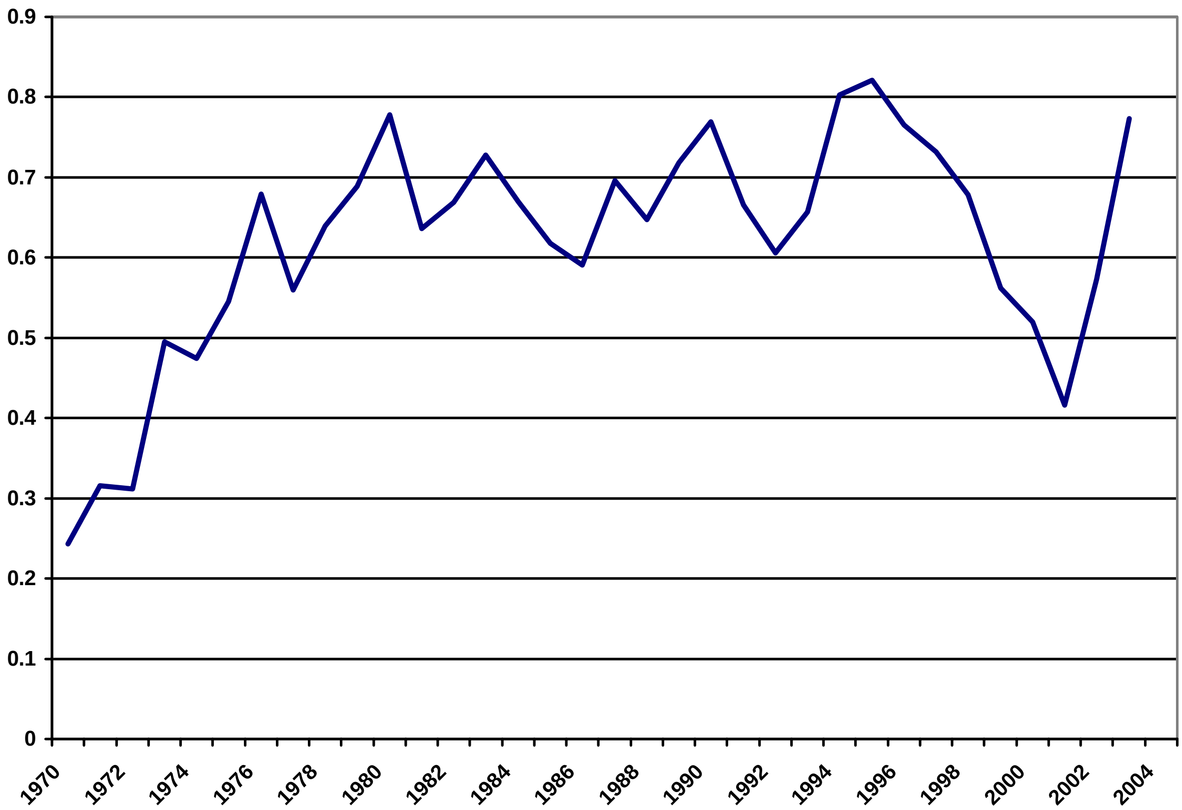


Figure 13A: Nominal Producers' Price in California for Cotton ($/lb).

Figure 14A: Real Producers' Price in California for Cotton ($/lb).

RICE

National vs. State Model

California is one of the major producers of rice in the US. The other most important states are Arkansas, Louisiana, Mississippi, Missouri and Texas. The market in California appears to be fully integrated with the southern states, as suggested by an empirical check of the law of one price. This conclusion is hardly surprising, given that rice is a storable and easily transportable commodity. Figure 1 illustrates the law of one price between California and Arkansas.

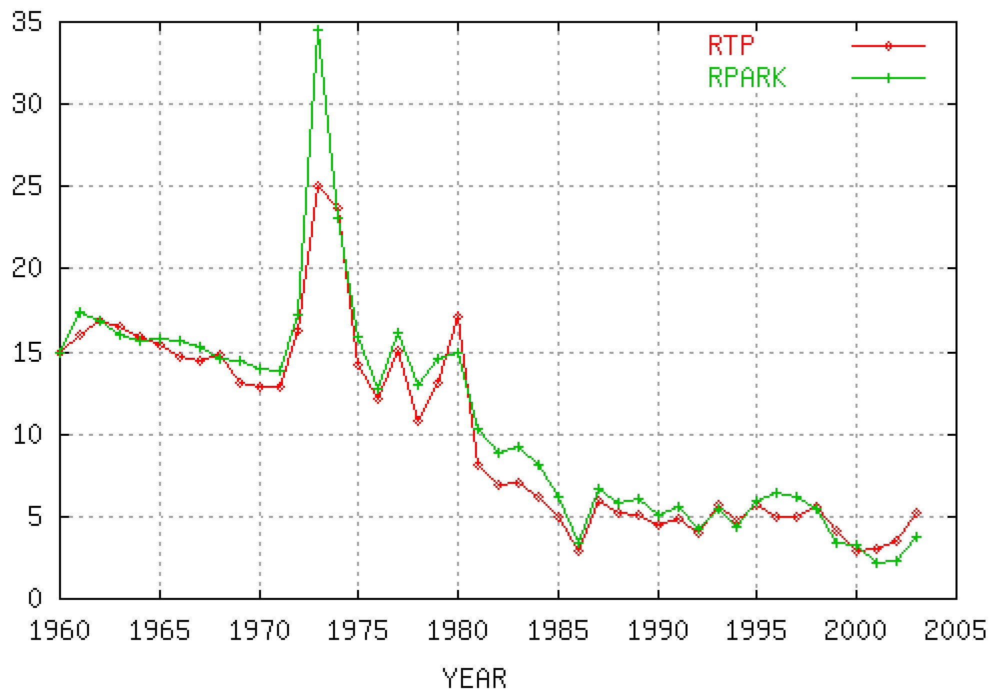


Figure 1: ***Rice grower price (real) in California (RTP) and Arkansas (RPARK) (in real dollars per cwt.)***

A simple ordinary least squares regression of California rice price on Arkansas price gives an R^2 of 0.939 and an estimated slope coefficient between 0.80 and 0.94 (with a 95% confidence level). Moreover, a simple cointegration test suggests the absence of unit roots in the disturbances. Thus, California price and Arkansas rice prices move together over the long run. Market integration suggests that a US level model can be useful to describe California rice production. In this study, however, we present both national and state models.

The US Market

We estimated two models for the US rice industry. The first one is based on a longer time series, but does not account for policy distortions or trade. The second model considers the influence of policy and trade but data limitations constrain the length of the available time series.

A simplified model

A simplified production model is

$$\ln Q_t = \beta_0 + \beta_1 \ln P_{t-1} + \beta_2 \ln P_{t-2} + \beta_3 t + \beta_4 \ln Q_{t-1} + \beta_5 D_t + \varepsilon_t \quad (1)$$

where Q is the quantity of rice production in tons, P_t is rice price per ton , t is a time trend and D is a binary variable identifying the years 1977 and 1983 (outliers).

Prior to reporting the estimated production function for rice, a brief discussion of some aberrations of the rice market will be explained. Around 1976-77 there was a price collapse that caused producers to rotate to other crops or not plant rice at all. This lead to decreases in rice production. In the early eighties rice prices collapsed again and this caused many growers to forfeit their crop to the government because the price was below the value of the government loan. This was not only the case with rice, but other program crops such as wheat and corn. In an attempt to reduce acreage and sell off the rice that the government had claimed, the government implemented the 50/92 plan. Subsides were directly linked to production. Thus, if a grower did not produce he was not paid. The 50/92 program allowed the grower to produce on 50% of his acreage and receive 92% of the subsidies that he would receive if he had produced on 100% of his land. This reduced production allowed the government to reduce the stocks of commodities that they had to claim in 1981-82. The 50/92 program ran until about 1988. Since then subsidies have been decoupled from production to prevent problems like this from happening again. The 50/92 program was popular in the south, especially in Texas where their production was lower and they had low fixed costs of land, but in California it was only widely used for a few years. Policy variables are incorporated into some of the models below.

The estimated partial adjustment production model for rice, for the time period 1972-2004, is:

$$\ln \hat{Q}_t = \underset{(0.68)}{2.32} + \underset{(0.07)}{0.23} \ln P_{t-1} - \underset{(0.08)}{0.07} \ln P_{t-2} + \underset{(0.00)}{0.02} t + \underset{(0.16)}{0.41} \ln Q_{t-1} - \underset{(0.07)}{0.26} D_t \quad (2)$$

with R^2 = 0.896 and n = 33. The Durbin *h* test did not indicate problems with autocorrelation. The coefficient on lagged production is positive and significant. This indicates that there is some adjustment each year in the production of rice. By removing the lags, i.e., by assuming $Q_t = Q_{t-1}$, the long-run price elasticity of production is 0.27 which is inelastic and significant, but indicates that rice producers do respond to price changes. The estimated coefficient on the time trend variable is 0.02 and significant indicating a positive trend over time. The estimated coefficient on the dummy variable is negative (coefficient = -0.26) and significant for the outlier years as expected.

Figure 2 describes the fit of the regression (in logarithmic scale--the original series is in 000 cwt).

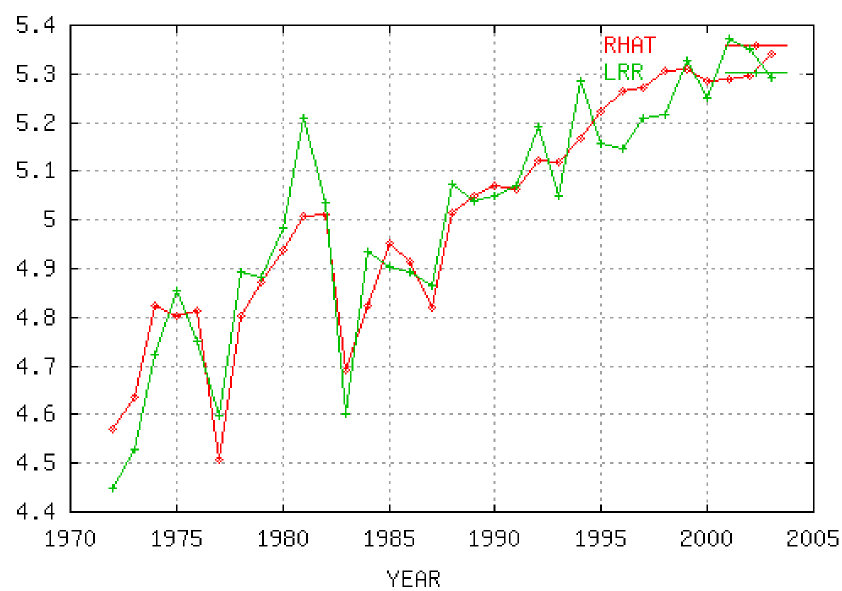


Figure 2: ***US rice production actual (LRR) and estimated (RHAT)***

(logarithmic scale)

Domestic Demand for Rice

The US domestic demand equation for rice is:

$$\ln PC_t = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln PCINC_t + \beta_3 \ln CPI_t + \varepsilon_t \quad (3)$$

where PC represents domestic consumption in pounds per capita, P denotes rice price per cwt, $PCINC$ represents per capita income in dollars per capita, and CPI is the consumer price index.

The estimated domestic demand function for rice, corrected for first-order autocorrelation, is:

$$\underset{(1.16)}{\ln \hat{PC}_t = -4.51} - \underset{(0.05)}{0.08} \ln P_t + \underset{(0.25)}{0.74} \ln PCINC_t - \underset{(0.29)}{1.47} \ln CPI_t \quad (4)$$

with R^2 = 0.93 and n = 34. The results of the simple model suggest that rice consumption is price inelastic (estimated own-price elasticity of -0.08), however, it is not significantly different from zero (p-value = 0.09). Domestic consumption of rice is positively related to income with a statistically significant (p-value = 0.0000) estimated income elasticity of 1.56. The estimated autocorrelation coefficient was 0.57 with an asymptotic t-ratio of 4.07 and after the correlation the Durbin-Watson statistic did not indicate any problems with autocorrelation.

The single equation estimates may be inefficient, given that errors may be correlated across equations. To overcome this problem we estimated a seemingly unrelated regression (SUR) production-consumption system for rice based on the simplified model specification. The results are:

$$\ln \hat{Q}_t = \underset{(0.07)}{11.21} + \underset{(0.03)}{0.14} \ln EP_t + \underset{(0.002)}{0.02}t - \underset{(0.07)}{0.19}D_t \quad \text{(production equation)} \quad (5)$$

$$\ln \hat{PC}_t = \underset{(2.85)}{1.15} - \underset{(0.05)}{0.03} \ln P_t + \underset{(0.61)}{0.36} \ln PCINC_t - \underset{(0.61)}{0.32}CPI_t + \underset{(0.002)}{0.02}t \quad \text{(demand equation)}$$

(6)

where EP represents the expected price of rice (price lagged one time period). The individual equation R^2s are high (0.93 and 0.89; respectively). The estimated own-price elasticity of production is

0.14 and but not significant. The own-price elasticity of demand for rice is -0.03, but it is not significant either. The income elasticity of demand for domestic rice is 0.36 and is also not significant. The explanatory variables were highly collinear which accounts for some of the estimated coefficients being insignificant.

An Alternative Model

An alternative model considers policy and exports. However, due to the short time series (1986-2003), the model must be parsimonious. For a comprehensive and disaggregated treatment of the influence of commodity programs on the rice acreage response to market prices, see McDonald and Sumner.[1]

The least squares estimated production equation is:

$$\begin{aligned}\ln \hat{Q}_t &= 10.843 + 0.176 \ln P_{t-1} + 0.003 PSE_t + 0.034t \\ &\quad (0.236)(0.067) \qquad (0.001) \qquad (0.033)\end{aligned} \tag{7}$$

with $R^2 = 0.87$ and n = 18. The policy variable, PSE_t, is the OECD percentage producer support estimate for the U.S. that is a comprehensive or aggregate measure of total policy support. The other explanatory variables are as defined above. The estimated policy coefficient is positive with a value of 0.003 and almost significant (p-value = 0.079). The estimated expected price elasticity of production is 0.176 and is significant (p-value = 0.02). The estimated coefficient on the time trend indicates that production has been increasing over time.

Overall the results suggest that public support has a significant and positive effect on production. The fit of the regression is depicted in Figure 3.

[1] McDonald and Sumner incorporate detailed rice commodity programs into their approach. Their approach is based on an econometric estimation of a marginal cost curve, some assumptions about the distribution of parameters of their cost function combined with a simulation methodology. Their main policy results indicate that models that do not take into account all the programs' rules produce smaller structural parameters. They cite previous studies that find the acreage elasticites for rice vary from 0.09 to 0.34 which their results indicate are too small.

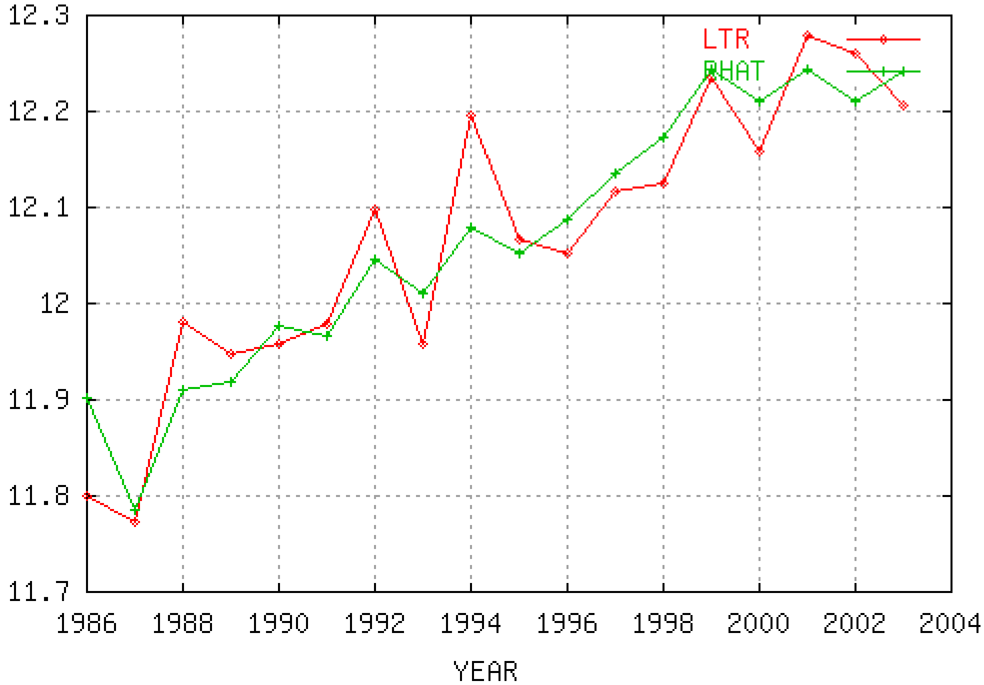


Figure 3: ***US rice production actual (LTR) and estimated (RHAT)***

(in billions of lbs)

Export Demand for Rice

The estimated export equation for rice is:

$$\ln \hat{E}X_t = \underset{(6.52)}{31.81} - \underset{(0.19)}{0.49} \ln P_{us,t} + \underset{(0.31)}{0.91} \ln P_{Thai,t} - \underset{(0.62)}{1.99} \ln Inc_{Japan,t} + \underset{(0.01)}{0.04} t \qquad (8)$$

with R^2 = 0.78, n = 18, and where EX_t represents US exports of rice in 000 cwt, P_{US} represents the grower price for US rice in \$/cwt, P_{Thai} denotes the price for rice in Thailand (the major competitor in the world market) and Inc_{Japan} represents per capita income in Japan (the major importer of US rice). The estimated results indicate that US exports decrease with US price increases (US price elasticity of exports is -0.49), increase with increases in Thailand rice prices, and have been increasing over time, conditioned on the other variables. The negative sign on per capita income in Japan was not expected.

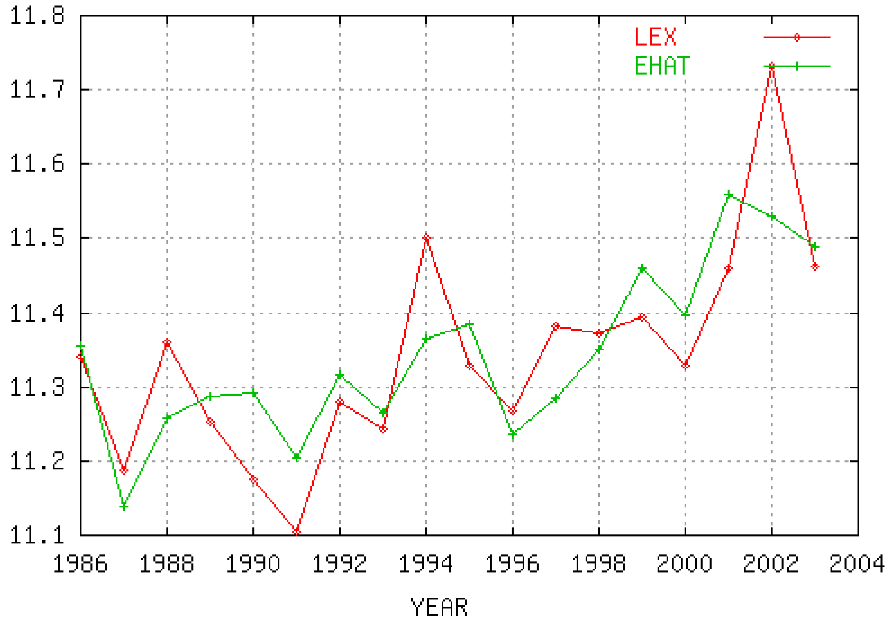


Figure 4: ***US rice export actual (LEX) and estimated (EHAT)***

In order to account for price endogeneity, correlated errors across equations, and to obtain more efficient estimates, we estimated a system of two equations for US rice, under the market clearing assumption. Lagged price was used as the instrumental variable for current price to account for endogeneity of prices. The system was estimated by iterative three stage least squares (3SLS). The estimators have the same asymptotic properties as maximum likelihood estimators. That is, they are consistent, asymptotically normally distributed and efficient. Iterative 3SLS converge to the same value as MLE, but are not equivalent because of a Jacobian term in the likelihood function. The first equation is a production function and the second equation is a demand function. The system results are:

$$\ln \hat{Q}_t = 8.92 + 0.45 \ln P_{US,t} + 0.27 \ln P_{Thai,t} + 0.01 PSE_t + 0.06t \quad (9)$$
$$(1.74)\ (0.39)\qquad (0.14)\qquad (0.01)\qquad (0.02)$$

$$\ln \hat{Q}_t = 2.68 - 0.36 \ln P_{US,t} + 0.39 \ln P_{Thai,t} + 0.33 Inc_t + 0.34 Inc_{Japan,t} \quad (10)$$
$$(3.77)\ (0.17)\qquad (0.25)\qquad (0.21)\qquad (0.49)$$

The fit of the system is depicted graphically in Figure 5 (the R^2 for the first equation is 0.718 and for the second is 0.888).

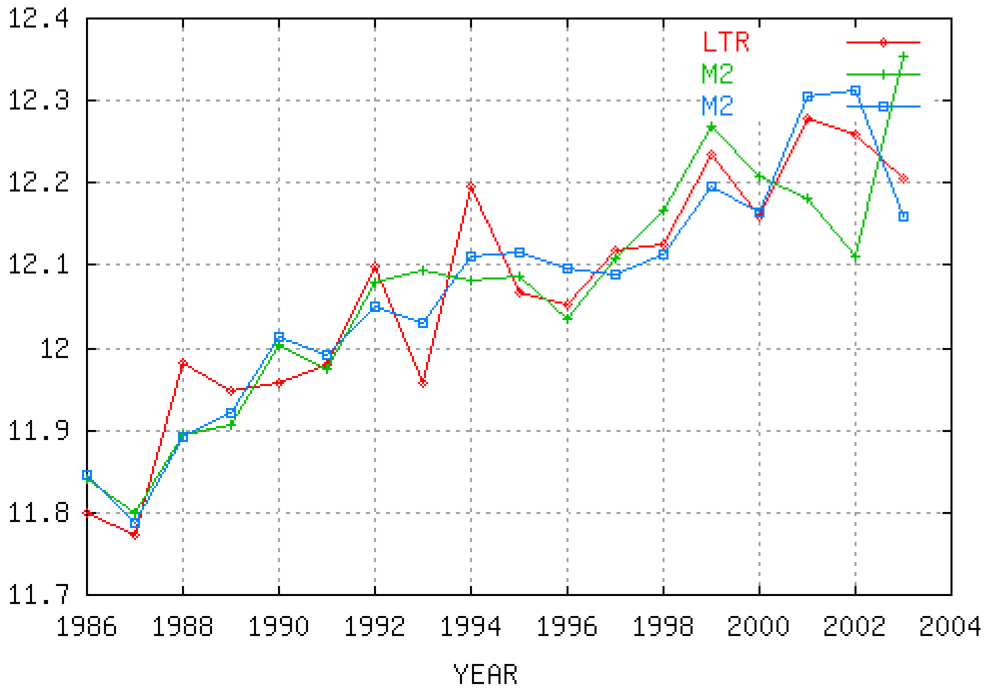


Figure 5: ***Estimation of supply and demand for US rice, under market equilibrium assumption***

The estimated US price expectation (the lag price) elasticity of supply is 0.45 which is also inelastic but is not significant. The estimated coefficient of Thailand price of rice is 0.27 with a t-ratio of about two. The index for price support is positive but not significant. There is also a positive (0.06) and significant time trend in the supply of rice. According to the estimated price coefficient (-0.36), the elasticity of demand of US rice implies that an increase of 1% in price results in a decrease of 0.36% change in the quantity demanded. As the price of Thailand rice increases, the demand for US rice increases, but again the estimated coefficient is not significant. The income elasticity is 0.33 and the estimated coefficient of Japanese income is 0.34 as expected. Both coefficients are not significant, however.

California Market

The estimated production function of California rice is:

$$\ln \hat{Q}_t = \underset{(0.72)}{7.56} + \underset{(0.16)}{0.48} \ln P_{t-1} + \underset{(0.05)}{0.11} Pay_t - \underset{(0.04)}{0.005} Loan_t + \underset{(0.01)}{0.04} t + \underset{(0.52)}{1.21} D_t - \underset{(0.48)}{1.23} D_t * Pay_t \qquad (11)$$

with $R^2 = 0.816$, n = 21, and where Q denotes California production, P denotes grower price, Pay represents direct payments, $Loans$ are the interest rate on marketing loans and D is a dummy variable identifying the years 1996 and after to account for policy changes.

The estimated own-price elasticity is 0.48 (and significant) which is higher than the corresponding estimated value for US production. Producers respond positively to increases in direct payments and to policy changes occurring after 1996. There is also a positive time trend. Interest rates on marketing loans did not have a significant impact on California production. Figure 6 depicts the fit of the California production model.

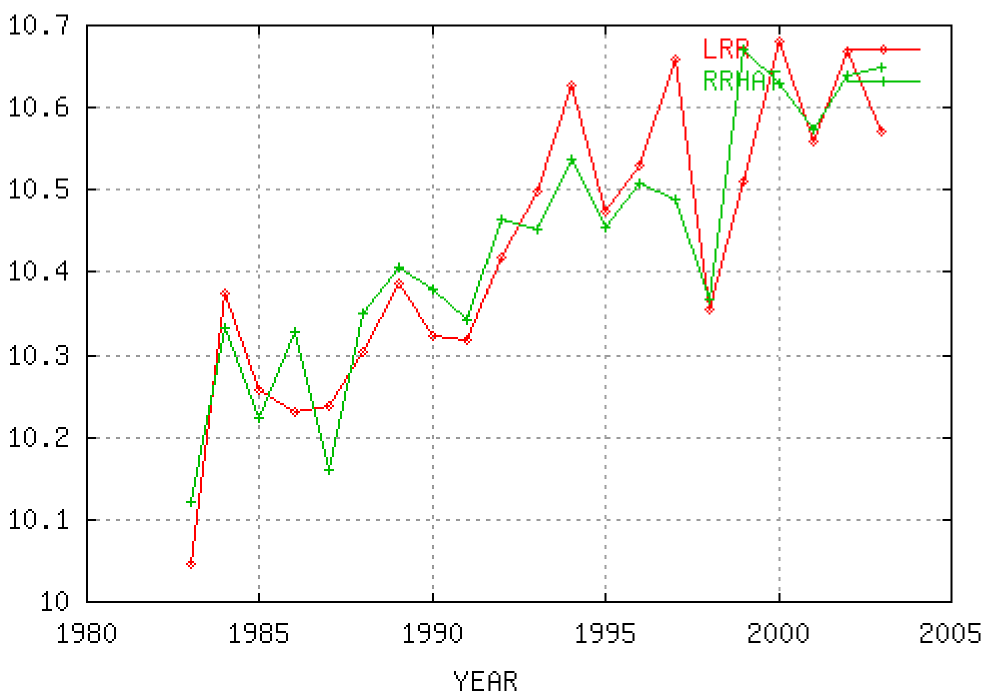


Figure 6: ***California rice production actual (LRR) and estimated (RRHAT)***

Conclusions

Rice producers in California and throughout the United States respond positively to increases in rice prices. The short-run price elasticity of production, based on a partial adjustment model, for the US was estimated to be 0.23. When policy variables were included in the production equation the price elasticity dropped to 0.18 (see eq. 7). Rice producers respond positively to support programs. The production equation was an aggregated one. For a disaggregated approach that estimates how rice producers respond to different support programs, see McDonald and Sumner.

The estimated own-price elasticity of demand for rice was found to be inelastic (-0.140) for a SUR system. The income elasticity for rice was estimated to be 0.74 in a single-equation demand function (eq. 4).

US rice producers export less when the US price increases (estimated elasticity =-0.49). They export more when the Thailand rice price increases (estimated Thailand price elasticity of 0.91) since Thailand is a major competitor in the world market.

References

McDonald, J. and D. Sumner, "The Influence of Commodity Programs on Acreage Response to Market Price: with an Illustration Concerning Rice Policy in the United States", *American Journal of Agricultural Economics*, 85(4): 2003: 857-871.

TOMATOES

Background

The United States is the world's second leading producer of tomatoes, after China. Fresh and processed tomatoes combined accounted for almost $2 billion in cash receipts during the early 2000s. Mexico and Canada are important suppliers of fresh market tomatoes to the United States and Canada is the leading importer of U.S. fresh and processed tomatoes.

The characteristics of tomato consumption are changing. Fresh tomatoes consumption increased by 15% between the early '90s and early 2000s, while the use of processed products declined 9%. Currently, the per capita consumption is 18 pounds per person of fresh tomatoes, and 68 pounds for processed tomatoes (fresh-weight basis).

The U.S. fresh and processing tomato industries consist of separate markets. According to ERS (website) four basic characteristics distinguish the two industries. Tomato varieties are bred specifically to serve the requirements of either the fresh or the processing markets. Processing requires varieties that contain a higher percentage of soluble solids (averaging 5-9 percent) to efficiently make tomato paste, for example.

- Most tomatoes grown for processing are produced under contract between growers and processing firms. Fresh tomatoes are largely produced and sold on the open market.
- Processing tomatoes are machine-harvested while all fresh-market tomatoes are hand-picked.
- Fresh-market tomato prices are higher and more variable than processing tomatoes due to larger production costs and greater market uncertainty

Policy

Tomato production is not covered by price or income support. However, tomato producers may benefit from general, non crop specific-programs such as federal crop insurance, disaster assistance, and western irrigation subsidies. The only federal marketing order in force for tomatoes covers the majority of fresh-market tomatoes produced in Florida between October and June.

With respect to imports, the United Stated negotiated a voluntary price restraint on fresh tomato imports from Mexico starting in 1996. Mexico agreed to set a floor price of $0.21 per pound of tomatoes exported to the United States The effect of the policy was to reduce Mexican exports to the U.S. and there were sizeable fresh tomato diversions (to other importing countries) and diversions into processing; see Baylis and Perloff for more details of this policy.

California Production

California is the second leading producer of fresh tomatoes in the US, after Florida. Figures 1-3 compares fresh tomatoes planted acreage, production and nominal price for US, Florida and California.

California accounts for about 95 percent of the area harvested for processing tomatoes in the United States—up from 79 percent in 1980 and 87 percent in 1990. The other major producers are Texas, Utah, Illinois, Virginia, and Delaware and Florida. In Figure 1, total U.S. fresh tomato acreage has declined over the period 1960 to 2002, but acreage in California and Florida has remained steady. The declined in acreage has come from the states of Texas, Utah, Illinois, Virginia, and Delaware (Lucier). Figures 4-6 illustrates the trends for California and US planted acreage, production and nominal prices for processed tomatoes.

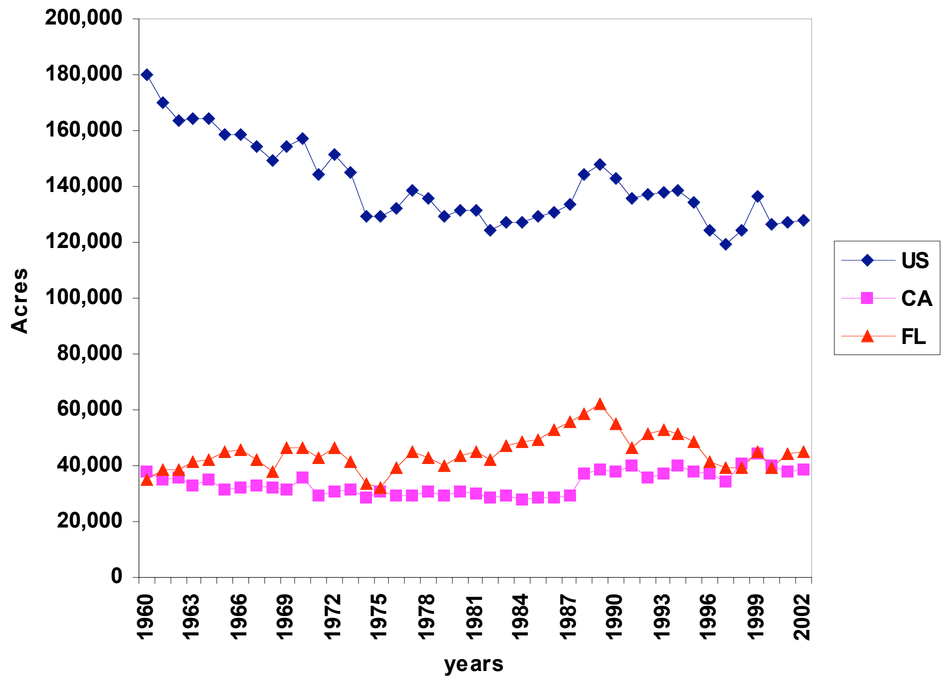


Figure 1: ***Fresh tomato acreage 1960-2002 – (source ERS)***

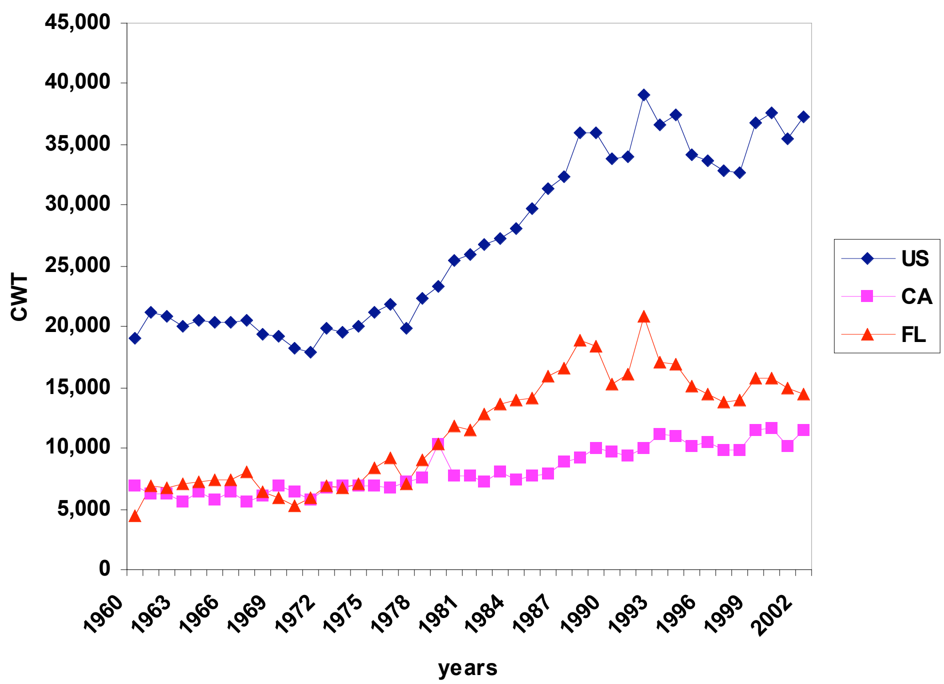


Figure 2: ***Fresh tomato production 1960-2002 – (source ERS)***

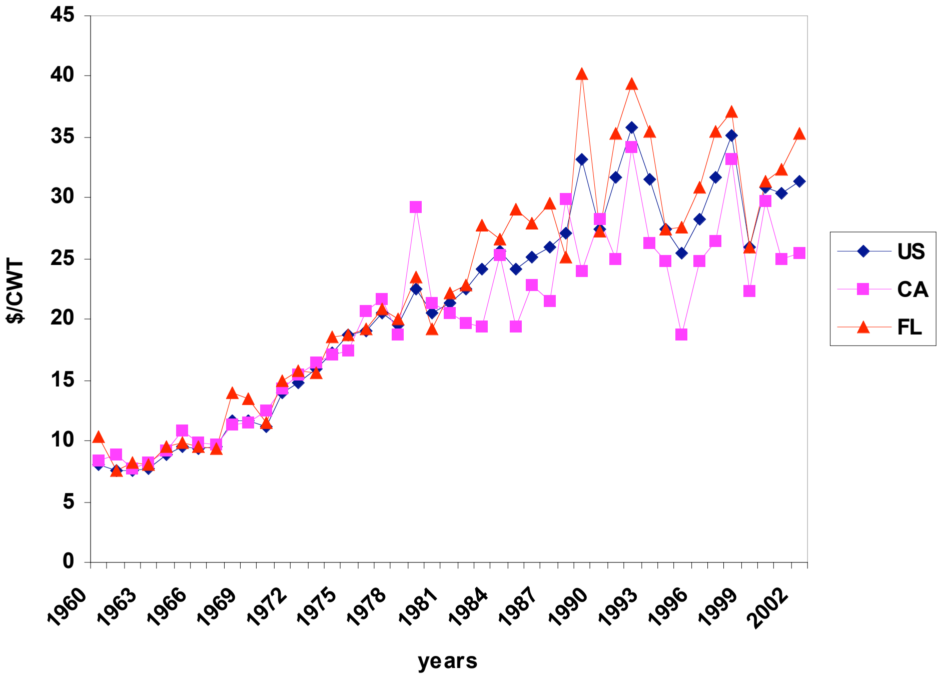


Figure 3a: ***Fresh tomato nominal prices 1960-2002 –(source ERS)***

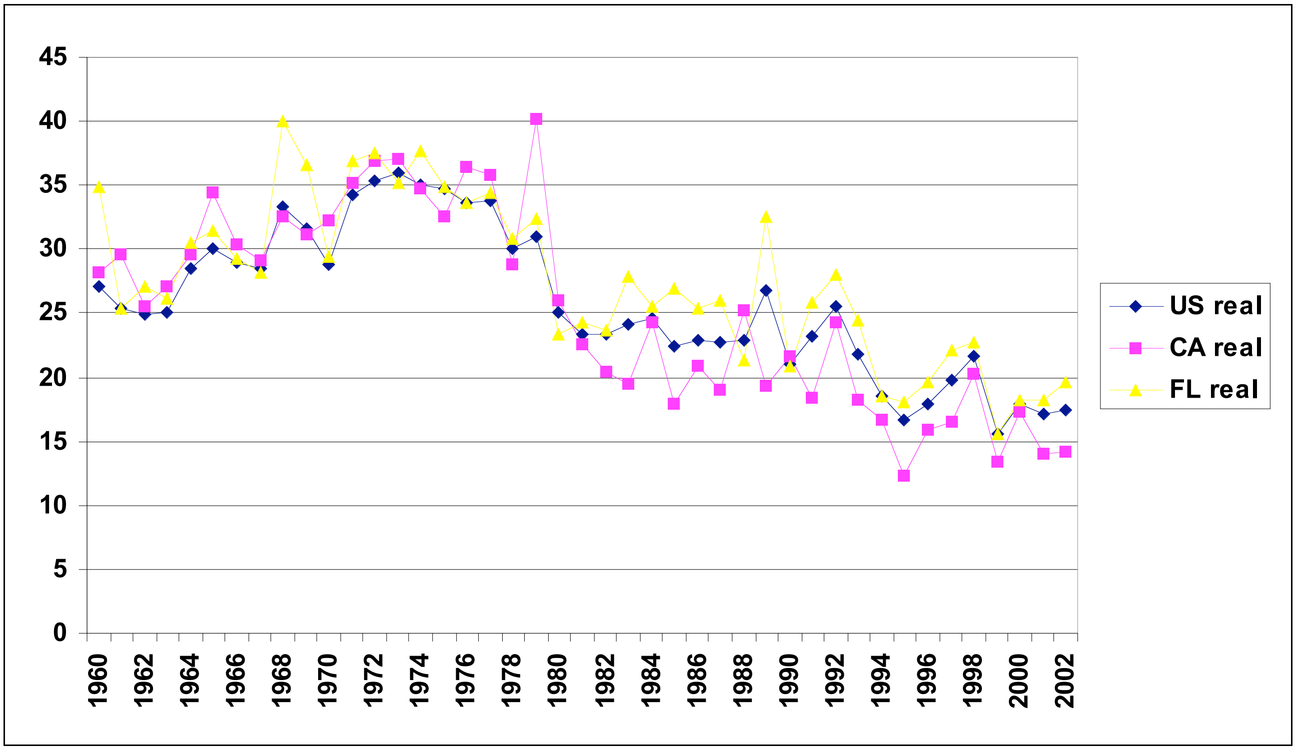


Figure 3b*: Fresh tomato real price 1960-2002 (base 1983-84)*

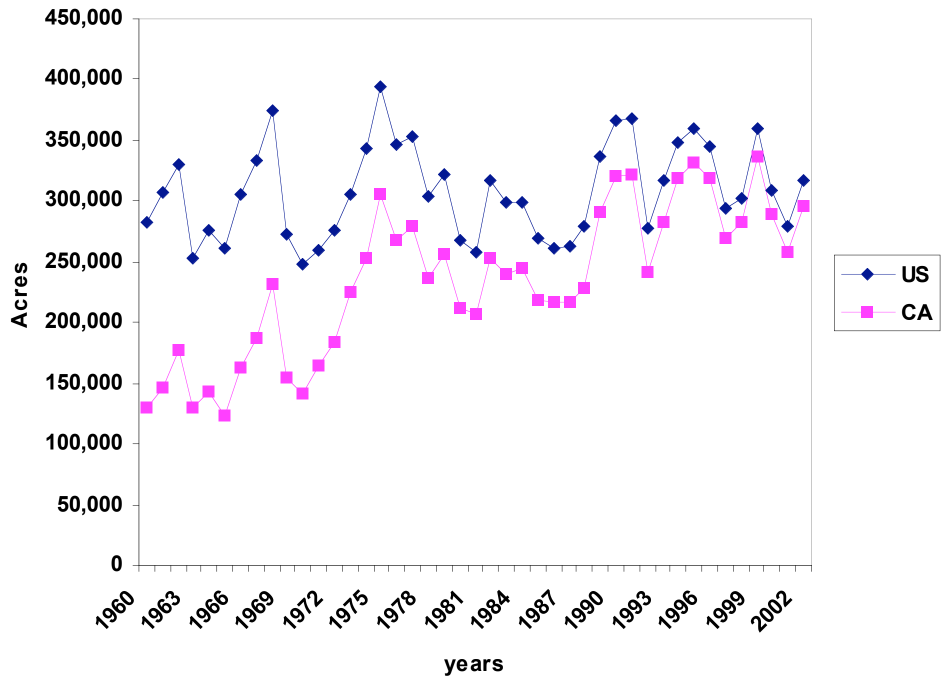


Figure 4: ***Processing tomato acreage 1960-2002 – (source ERS)***

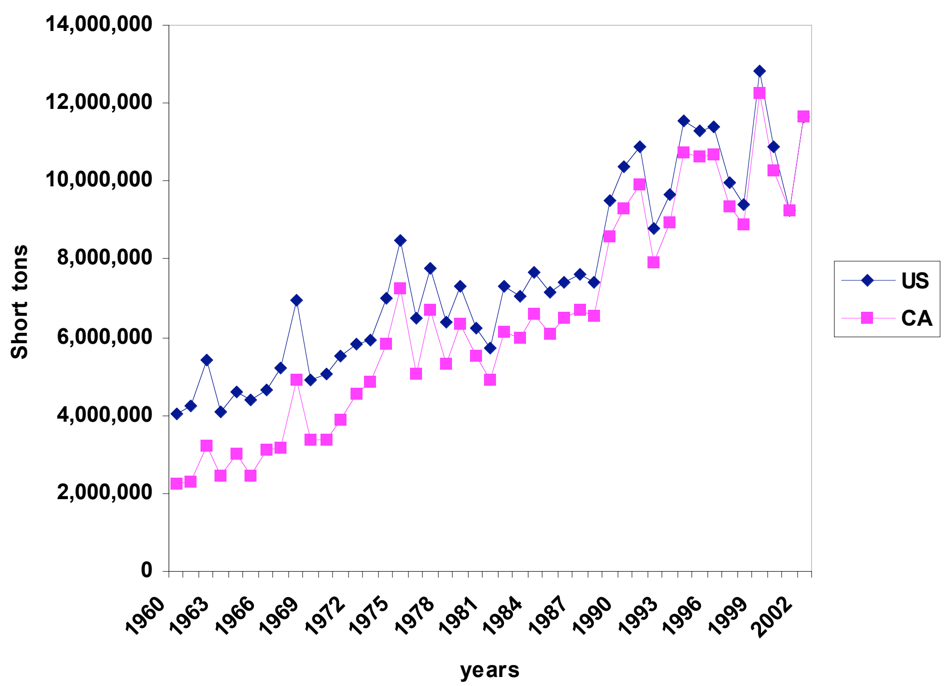


Figure 5: ***Processing tomato production 1960-2002 – (source ERS)***

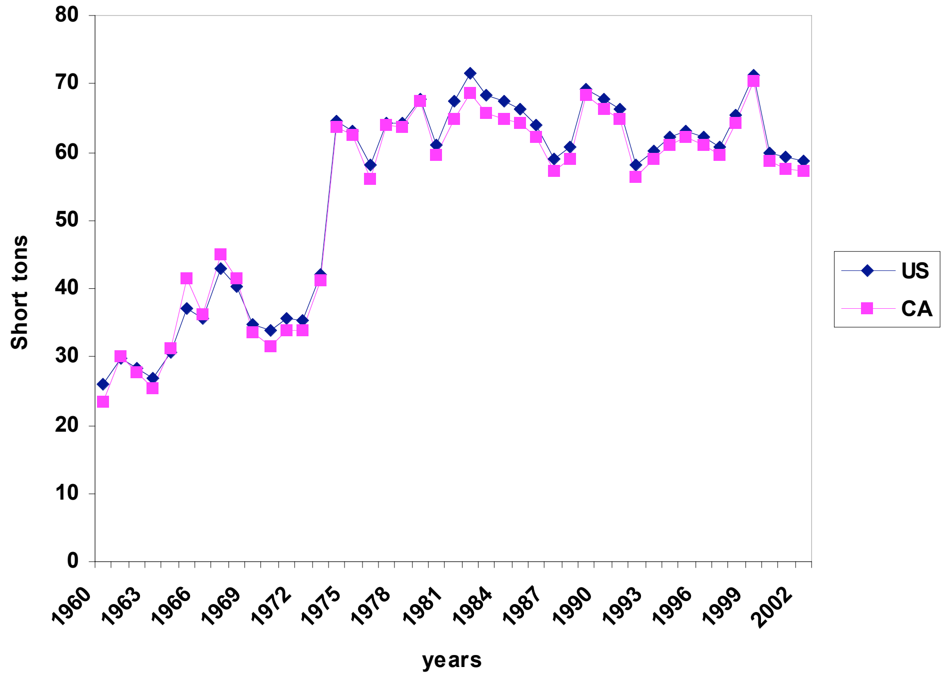


Figure 6: ***Processing tomato nominal prices 1960-2002 ($/ton) – (source ERS)***

Processing Tomatoes

Tomato growing is based on grower-processor contract agreements. The majority of production is traded this way with the spot market playing a marginal role. Most initial processing is by firms that manufacture tomato paste, a raw ingredient. Tomato paste is storable up to 18 months. Downstream firms transform the paste in final consumer products. According to the Food Institute, at the end of the process, raw material (tomatoes and fees) account for 39%-45% of total production cost.

According to the ERS, there was a radical structural change in the processing industry in the late 1980's and early 1990's. A period of relatively high prices in the late 1980s triggered new investments. This finally resulted in excess supply and decreasing prices. As a consequence, many processors went bankrupt and the whole industry was restructured. The current structure is the result of such adjustments.

Estimation

A brief industry description highlights two key points prior to the estimations.

Price expectations. The majority of production is sold under contract. This has two implications: i) producers know (with good approximation) prices when planning production, so we do not need to model expectations; rather we assume perfect information, ii) the actual contract price is unobservable, being industry private information. It is reasonable to assume that the spot market price is correlated with contract price according to the additive error formula:

spot price = contract price + error .

We use the spot price as a proxy for the real contract price. However, since the measurement error is likely to be correlated with the error terms in the production equations (for example in case of unexpected shortage, we expect higher spot prices) we use an instrumental variable (IV) approach. The instrument is the previous year's spot price, which is correlated with the current spot price, but uncorrelated with random shocks in current production.

Structural change. The industry underwent structural changes from the late '80s until the early '90s. Much of the change is likely due to continued expansion in food-service demand, especially for pizza, taco, and other Italian and Mexican foods (Lucier). Increased immigration and changes in America's tastes and preferences have contributed to rising per capita tomato use (Lucier, *et al)*. Commercial varieties were developed to expedite packing, shipping, and retailing in the processing market. Mechanical harvesting and bulk handling systems replaced hand harvest of processing tomatoes in the California in the 1960's as the new varieties were introduced. Increases in yields are due to the development of higher yielding hybrid varieties and improved cultural practices such as increases in use of transplanting (Plummer). The hypothesis of structural change was tested on both the supply and demand side.

Acreage

The acreage equation is based on a partial adjustment model:

$$\ln A_t = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln A_{t-1} + \beta_3 t + \varepsilon_t \tag{1}$$

where A_t represents acreage at time t in actual acres, P_t represents the spot price of processing tomatoes in \$/cwt, and t is a time trend.

The OLS estimated acreage function for the years 1960-2002 is:

$$\begin{array}{l} \ln \hat{A}_t = 5.67 + 0.47 \ln P_t + 0.32 \ln A_{t-1} + 0.03t \\ \qquad\quad (1.38)\ (0.12) \qquad (0.13) \qquad\quad (0.01) \end{array} \tag{2}$$

with $R^2 = 0.815$, n = 42 and where the numbers in parentheses are standard errors.

The instrumental variable estimated acreage equation is:

$$\begin{array}{l} \ln \hat{A}_t = 5.67 + 0.41 \ln P_t + 0.36 \ln A_{t-1} + 0.02t \\ \qquad\quad (1.39)\ (0.18) \qquad (0.14) \qquad\quad (0.01) \end{array} \tag{3}$$

with $R^2 = 0.814$ and n = 42.

Figures 7 and 8 compare the fits of the two regressions.

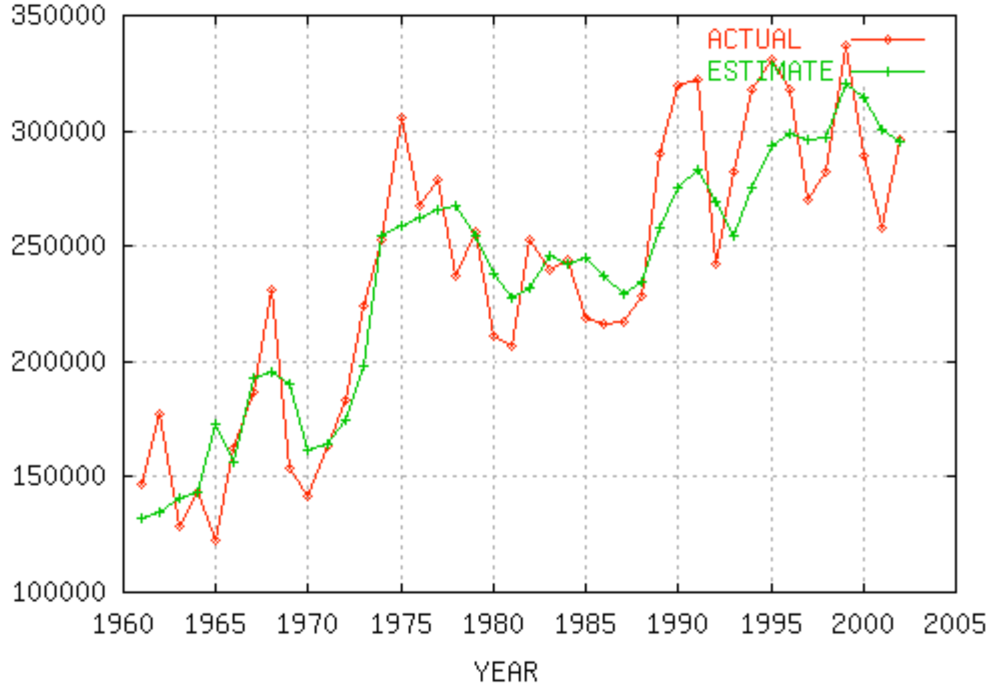


Figure 7: ***OLS estimation of processing tomato acreage (in acres).***

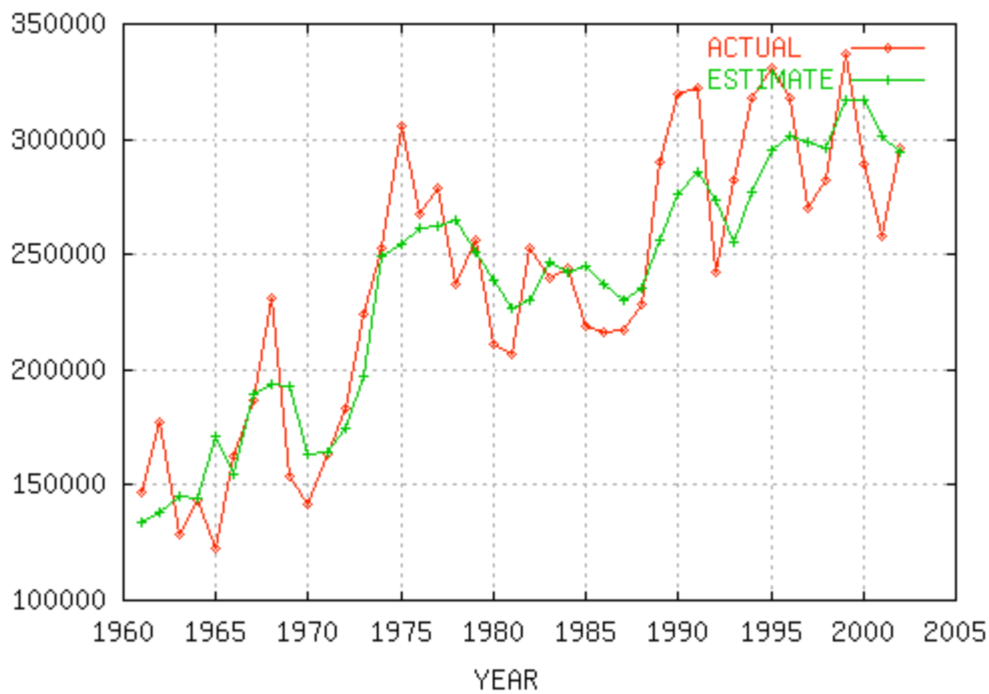


Figure 8: ***IV estimation of processing tomato acreage (in acres).***

The two estimation procedures –OLS and IV- give similar results. According to the partial adjustment model, the IV estimate of the short-run elasticity of acreage with respect to a change in price is 0.41 compared to the OLS estimate of 0.47. The estimate of the long-run price elasticity is 0.64. The coefficients on lagged acreage and the time trend are both positive. All the coefficients are statistically significant from zero.

Structural change

The Chow test confirmed the possibility of a structural break in the late '80s. The estimation of the model for the two periods (before and after 1988) gave the following results:

Dep. Variable: Tomato Acreage	Before 1988		After 1988	
Variable	estimate	std. dev.	estimate	std. dev.
Constant	5.62	1.73	2.20	3.21
Price	0.51	0.15	1.09	0.36
Lag Acreage	0.32	0.16	0.40	0.19
Time Trend	0.02	0.01	0.03	0.01

Table 1. Chow test results for processing tomato acreage function.

By observing the results prior to 1988 and past 1988, almost all of the coefficients are significantly different from zero. Most of the estimated coefficients differ little in magnitudes between the two periods. However, the short-run elasticity of acreage with respect to price is 0.51 before 1988 and 1.09 after 1988. Producers are much more responsive to prices after 1988 regarding their acreage. What explains this difference? Producers are, apparently, more responsive to price changes with the increased use of contracts and other structural changes mentioned above.

Figure 9 depicts the fit of the estimated structural-break model.

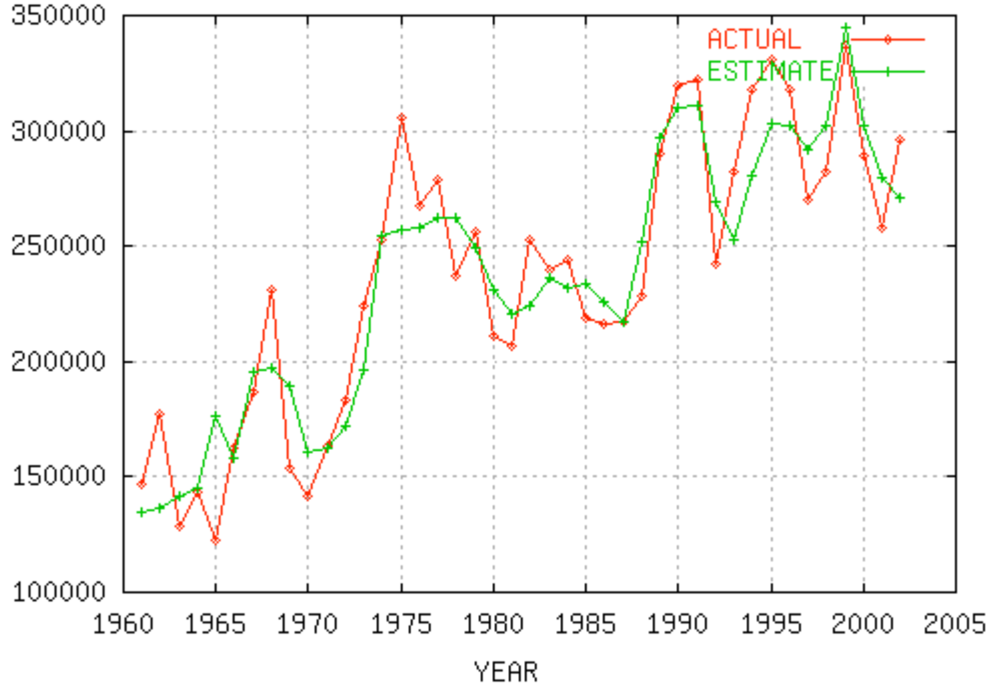


Figure 9: ***Structural break model for processing tomato acreage (in acres)***

Production

The partial adjustment model for processed tomato production is

$$\ln Q_t = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln Q_{t-1} + \beta_3 t + \varepsilon_t \tag{4}$$

where Q_t denotes production at time t in tons, P_t represents real price of processing tomatoes in \$/cwt, and t is a time trend. The OLS estimated production function is

$$\begin{aligned} \ln \hat{Q}_t &= \underset{(1.98)}{11.00} + \underset{(0.13)}{0.45} \ln P_t + \underset{(0.14)}{0.10} \ln Q_{t-1} + \underset{(0.01)}{0.04} t \end{aligned} \tag{5}$$

with $R^2 = 0.92$ and n = 42.

The same model, estimated by using lagged prices as instrumental variables, gave comparable results:

$$\begin{aligned} \ln \hat{Q}_t &= \underset{(1.99)}{11.03} + \underset{(0.19)}{0.55} \ln P_t + \underset{(0.15)}{0.07} \ln Q_{t-1} + \underset{(0.01)}{0.05} t \end{aligned} \tag{6}$$

with $R^2 = 0.91$ and n = 42. The OLS estimate of the own-price elasticity is 0.45 compared to that of 0.55 for the instrumental variables estimate. Both coefficients are significant. Coefficients of lagged acreage are both positive but not significant. And both coefficients on the time trends are positive and significant.

Figures 10 and 11 compare the fit of the two estimations.

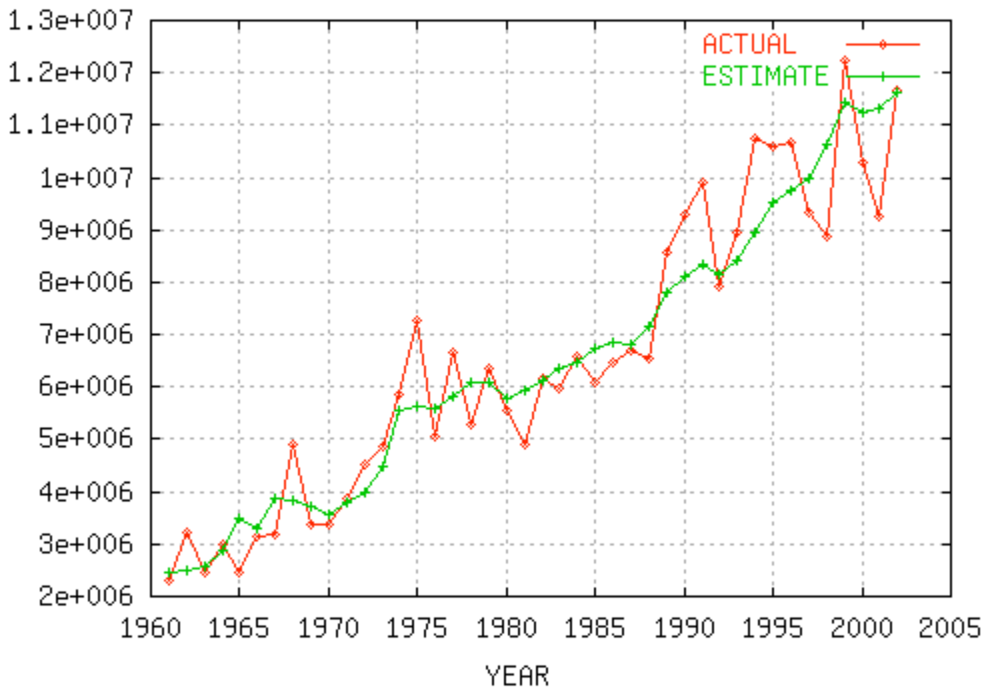


Figure 10: ***Production estimation for processing tomato (OLS) in tons.***

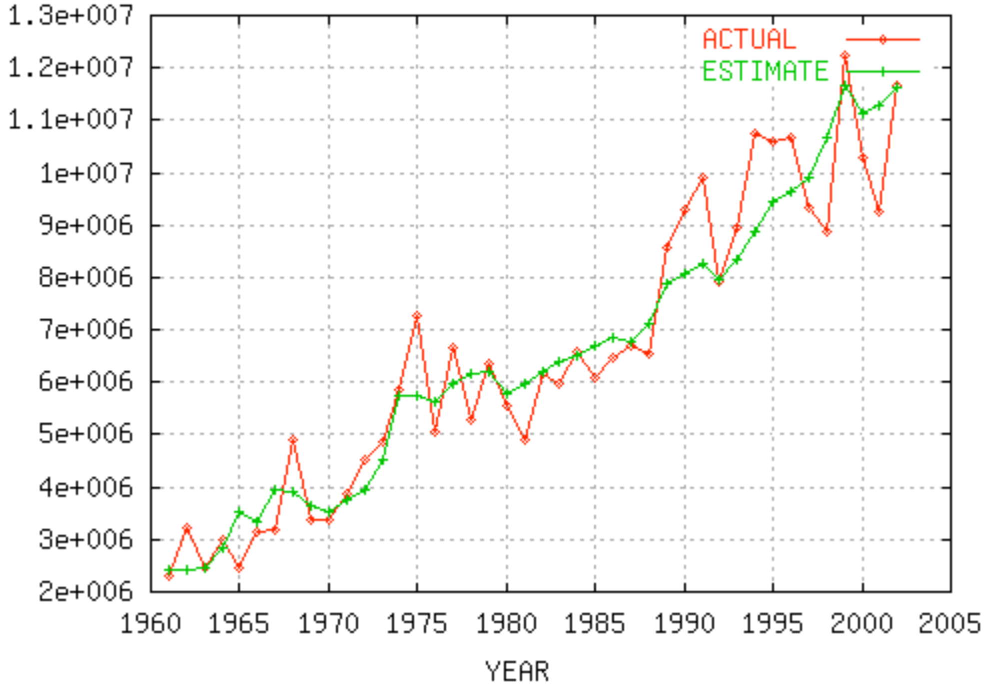


Figure 11: ***Production estimation for processing tomato (IV)in tons.***

Although a Chow test did not reject the null hypothesis of no structural change[1], we present the estimates for the two-period model, to provide a comparison with the acreage model.

Dep. Variable Production in tons	Before 1988		After 1988	
Variable	estimate	std. dev.	estimate	std. dev.
Constant	12.11	2.52	4.89	5.29
Price	0.51	0.17	1.04	0.47
Lag Acreage	0.01	0.19	0.35	0.25
Time Trend	0.05	0.01	0.04	0.01

Table 2. Chow test results for processing tomato production function

[1] The test has a p-value of 0.117.

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.

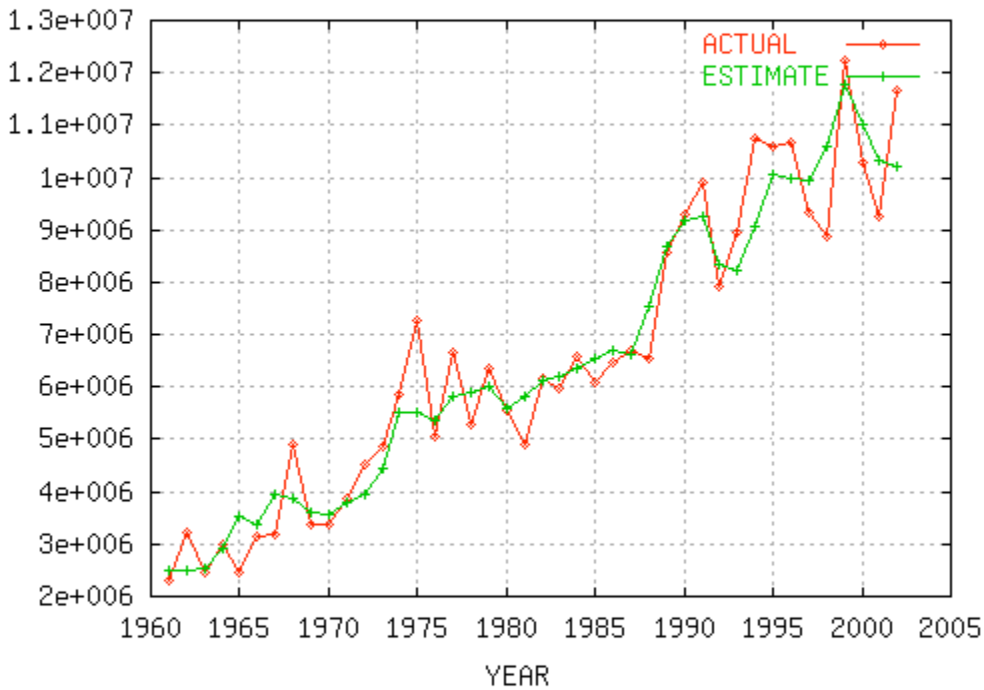


Figure 12: ***Structural break model for processing tomato production in tons.***

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.

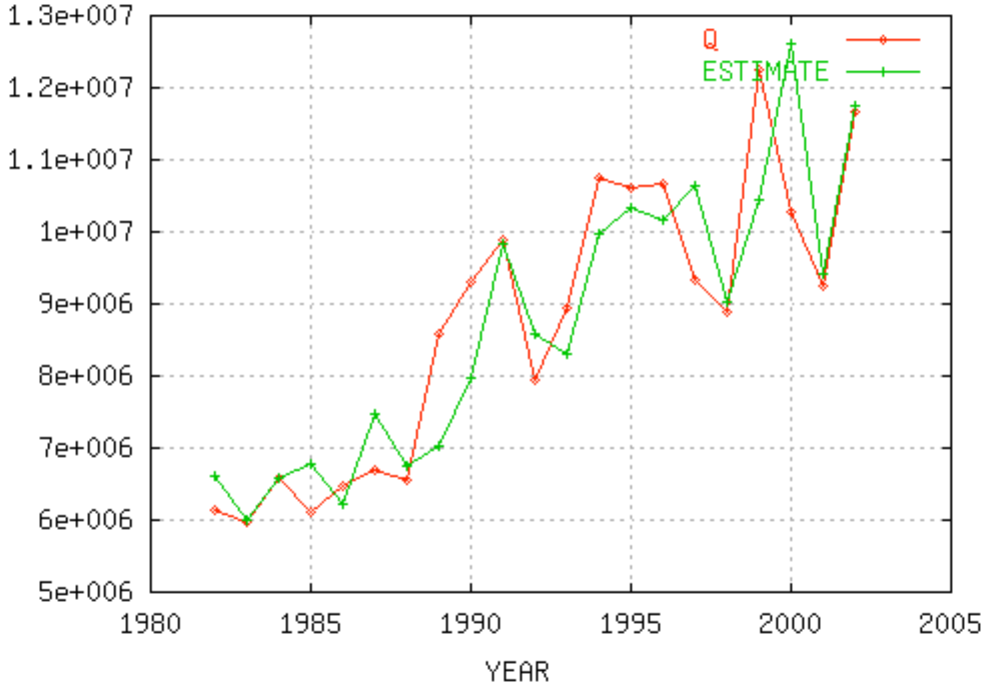


Figure 13: ***Demand for California processing tomatoes (million tons).***

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:

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

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

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

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.

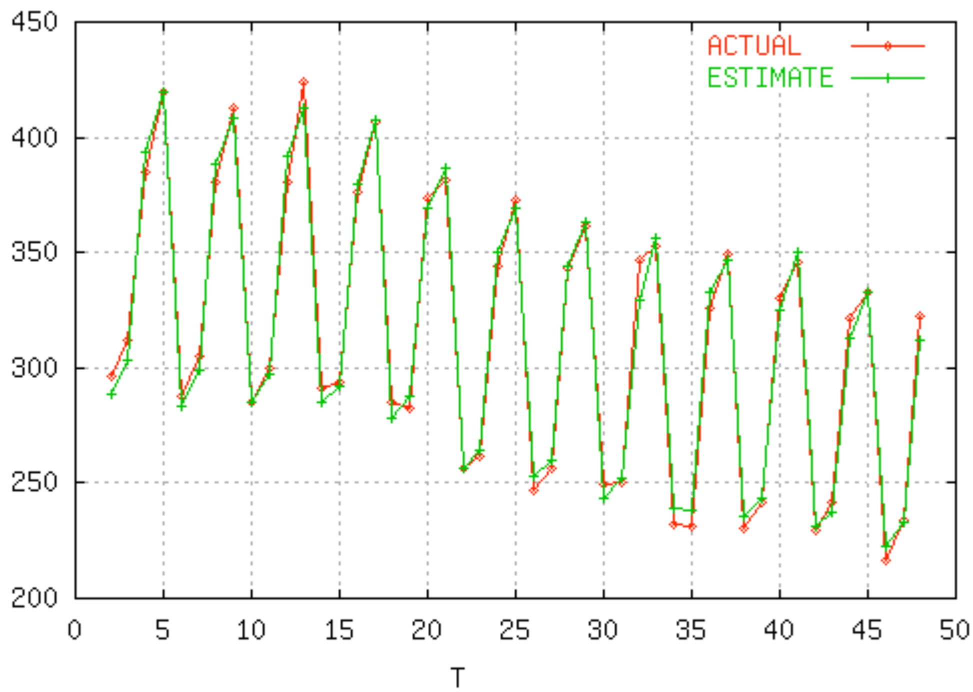


Figure 14: ***Consumers' demand for processing tomato products (1st quarter 1993-4th quarter 2004)***

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.

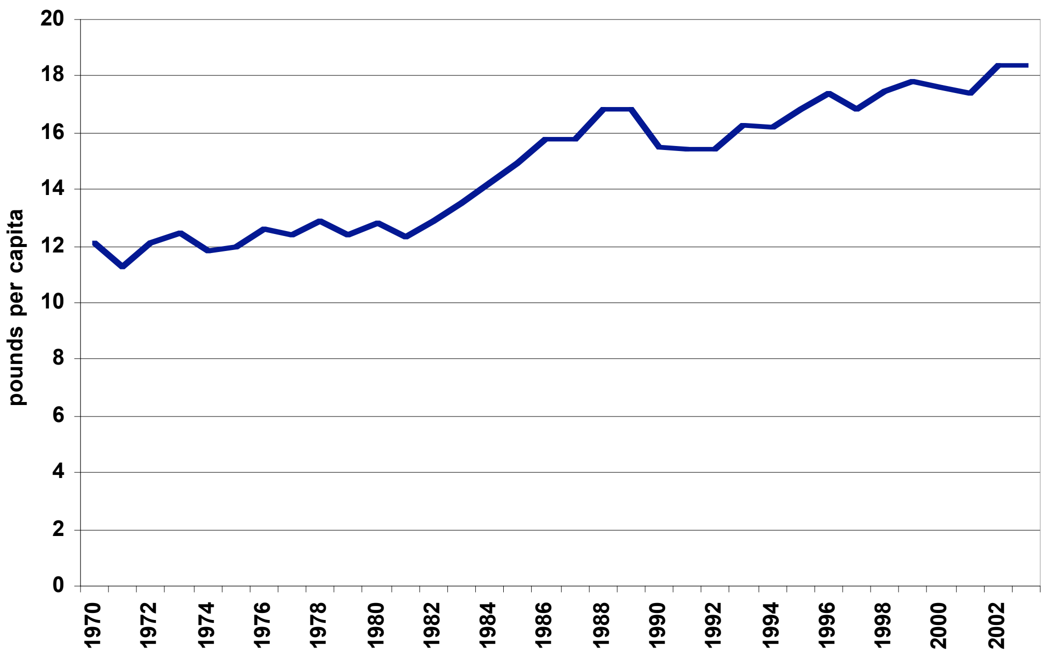


Figure 15: ***US per capita consumption of fresh tomatoes***

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.

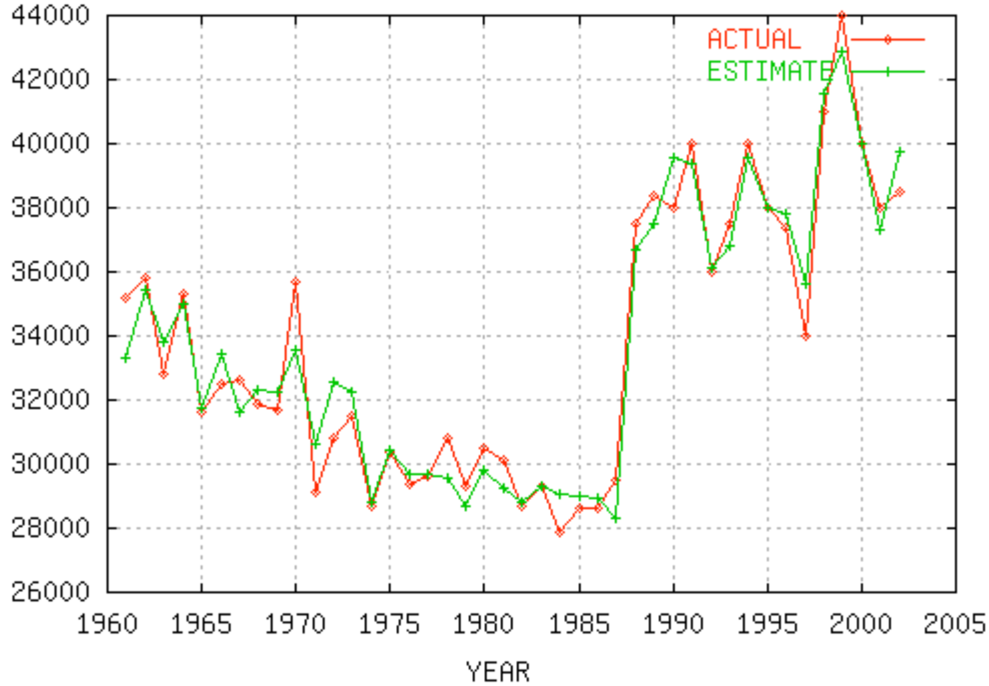


Figure 16: ***California fresh tomato acreage (in acres).***

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:

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

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

Period 1988-2002:

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

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.

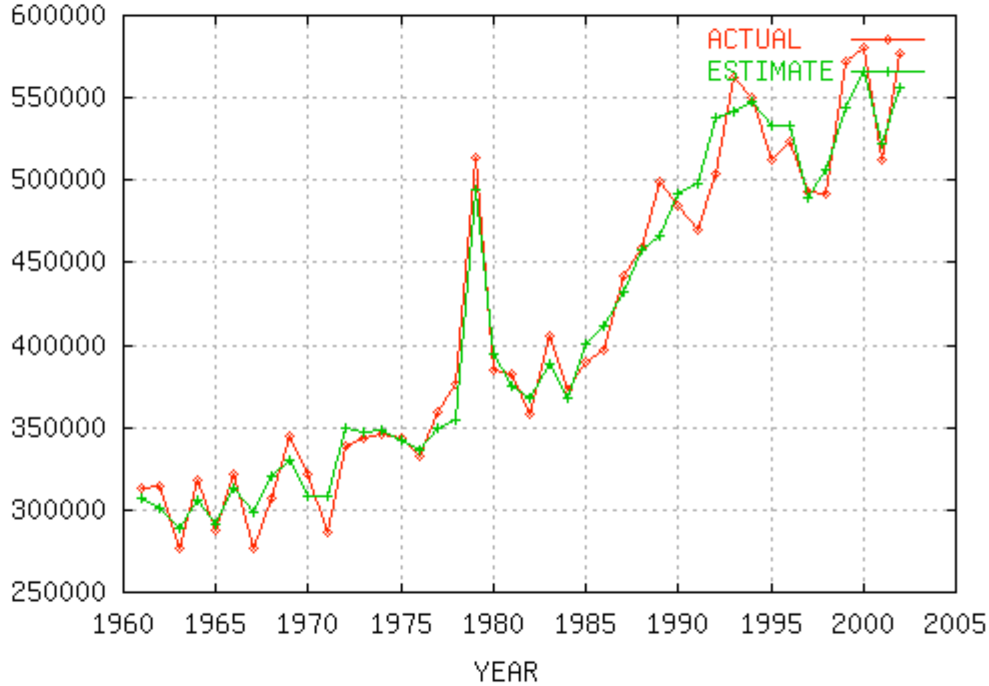


Figure 17: ***California fresh tomato production (in tons).***

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 nonzero 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.

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

subject to

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

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

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

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

$$(3)\qquad (\bar{\boldsymbol{\gamma}}_i - \boldsymbol{\alpha}_i) \ge 0\,.$$

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

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

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

Consider the following problem:

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

subject to

$$\mathbf{A}\mathbf{x} \le \mathbf{b}$$
$$\mathbf{x} \ge 0.$$

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

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

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

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

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

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

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

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

subject to

$$\mathbf{Ax} \leq \mathbf{b} \text{ and } \mathbf{x} \geq 0$$

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 $\mathbf{f(x)}$ simplified to $\mathbf{c'x}$, 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 $\mathbf{f(x)}$ as the quadratic total output function specified in (9), the first equation is the average yield for crop i, $\bar{y}_i$

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

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.

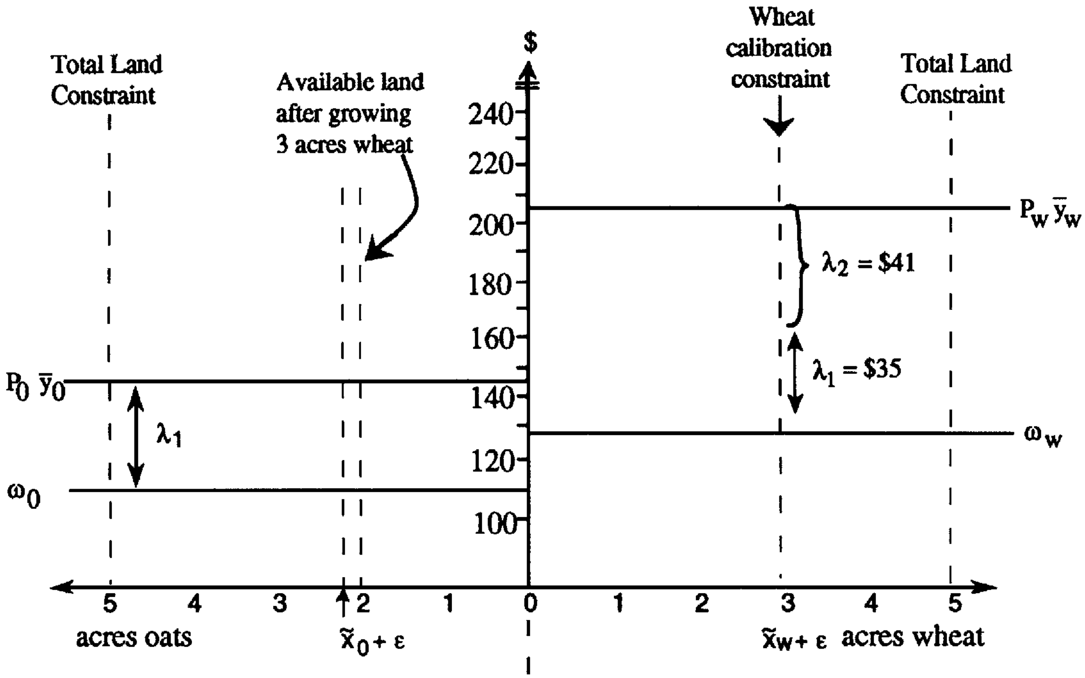


Figure 1. L.P. problem with calibration constraints—two activity/one resource constraint

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

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

where $\nabla_{x_B} \mathbf{f}(\mathbf{x}^*)$ 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

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

[and substituting equation (11)]

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

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-

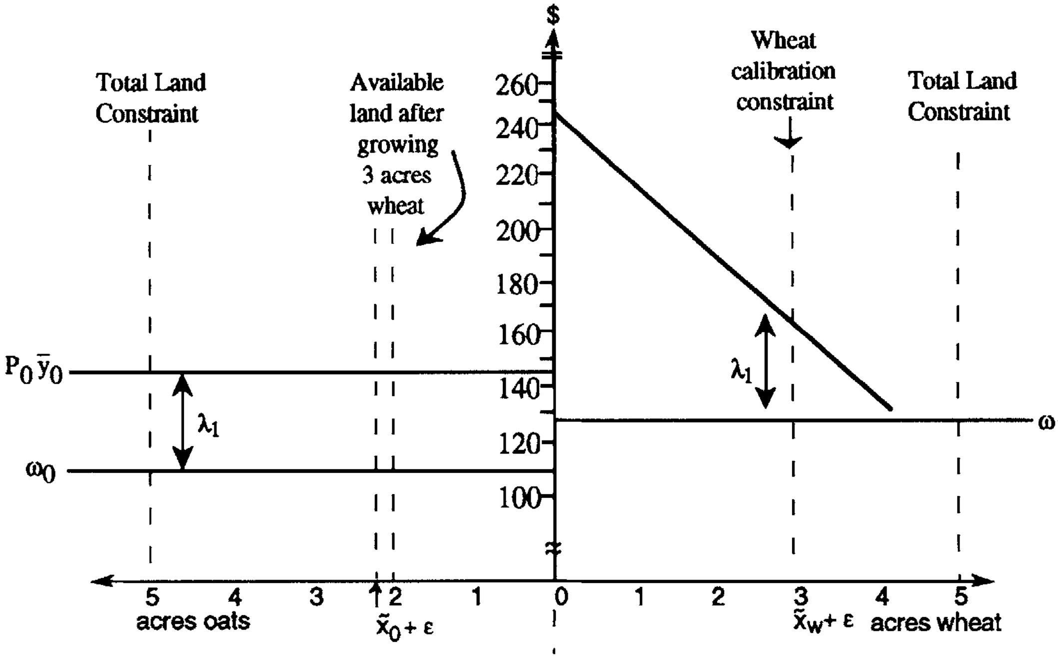


Figure 2. PMP yield function on wheat

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

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

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

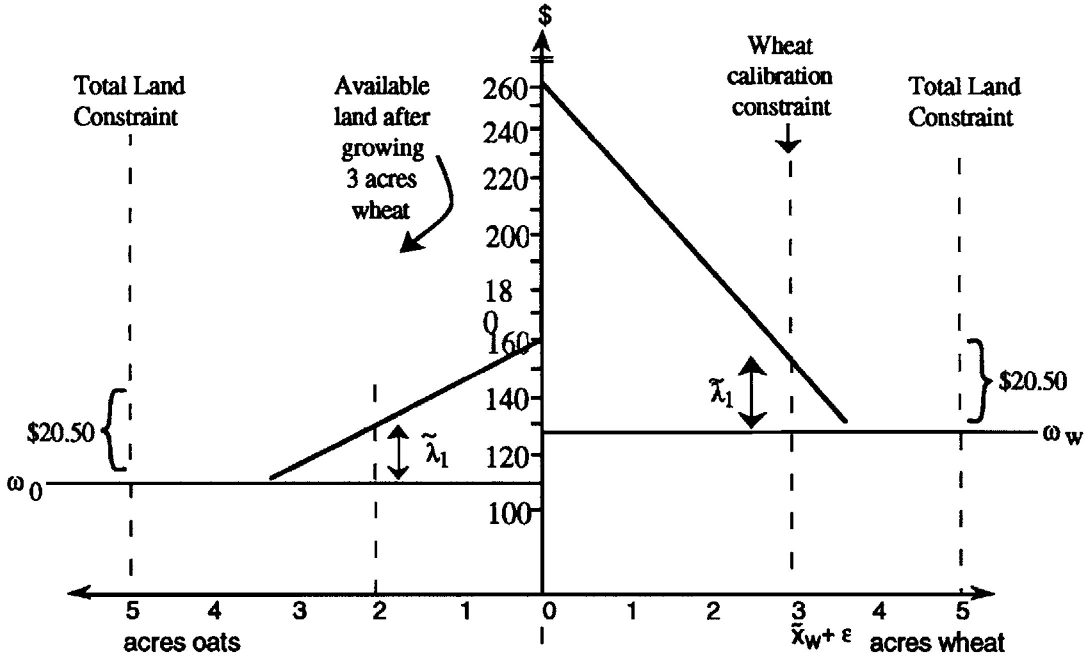


Figure 3. PMP model—quadratic yields on all crops

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 ($\boldsymbol{\lambda}_1$) 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 λ_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

$$x_w + x_0 \leq 500.$$

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 the 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 **x**, **f**(**x**).

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

subject to

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

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 **A** as the $(m \times m)$ optimal solution basis matrix **B** and the remaining partition of **A** as **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

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

or

(a3) $\max \mathbf{f}(\mathbf{x}_B, \mathbf{x}_N) \qquad \text{subject to } \mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}$

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

(a4) $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N$.

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

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

and symmetric. Writing (a7) as

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

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.

(b1) maximize $\mathbf{f(x)}$
subject to
$$\mathbf{Ax} = \mathbf{b} \quad \text{(I)}$$
$$\hat{\mathbf{A}}\mathbf{x} < \hat{\mathbf{b}} \quad \text{(II)}$$
$$\mathbf{Ix} = \tilde{\mathbf{x}} \quad \text{(III)}$$
$$\mathbf{x} = k \times 1,\ \mathbf{A} = m \times k\quad \hat{\mathbf{A}} = (l - m) \times k$$
$$\tilde{\mathbf{x}} = k \times 1\quad k > m\quad \mathbf{b} = m \times 1\quad \hat{\mathbf{b}} = (l - m) \times 1.$$

$\tilde{\mathbf{x}}$ 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.

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

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.

(b5) maximize $\mathbf{f(x)}$
subject to
$$\mathbf{Ax} = \mathbf{b}$$
$$\mathbf{Ix} \le \tilde{\mathbf{x}} + \boldsymbol{\varepsilon}$$

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

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

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

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

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

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

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

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

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 $\bar{x}_{il}$, 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.

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

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

$$Ix \leqslant \tilde{x} + \epsilon. \tag{1c}$$

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):

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

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

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

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

1995a). SWAP covers over 93% of irrigated agriculture in California, most of which is in the Central Valley, and calibrates exactly to an observed base year of land use and input allocation data through use of exogenous elasticities and assumed profit-maximization behavior by farmers. This paper, (i) documents SWAP and motivates application to other regions, (ii) discusses SWAP construction and emphasizes the sequential calibration diagnostic checks used in the model, (iii) extends the applied PMP literature with more flexible production and cost functions, and (iv) links the SWAP model to the infrastructure of a hydro-economic network model for water supply in California (CALVIN). We conclude with an empirical example and estimate the value of water markets for California's San Joaquin Valley.

The next section highlights the importance of micro-level policy analysis, in various geographic regions, with models similar to SWAP. We place SWAP in the context of the existing literature of optimization models and PMP. In the subsequent section we construct SWAP with particular emphasis on the sequential calibration routine and improvements over previous PMP models. The calibration routine has six steps with model consistency checks at each stage. Improvements over similar PMP models include Constant Elasticity of Substitution (CES) production functions, exponential PMP land cost functions, and endogenous crop prices. Finally, we demonstrate an application of SWAP for evaluation of water markets in the San Joaquin Valley. We conclude the paper with a discussion of extensions, limitations, and future work on SWAP.

1.1. Micro-level analysis of agricultural policies

In the U.S. and other agricultural economies the demand for micro-level analysis of agricultural policies that reflect the effects on local agricultural and environmental resources is growing for several reasons. National agricultural policies are increasingly driven or constrained by environmental criteria. Furthermore, there are an increasing number of regional (state) level policies that proscribe the use of agricultural inputs (land, water, labor and supplies) and resources. The era of unfettered commodity price support programs whose impact could be measured by aggregate financial or physical outcomes is waning, as are the aggregate demand and supply methods used to measure such outcomes. Finally, the complex physical and economic interaction between the environment and agricultural policies is difficult to accurately capture using standard econometric techniques based on aggregate data.

Calibrated optimization models for micro-level analysis, such as SWAP, focus on spatially heterogeneous commodity, resource, and input specific policies. Instead of using data from the outcome of economic optimization to estimate aggregate elasticities, calibrated optimization models use prior estimates of elasticities of demand, supply, and substitution coupled with observed micro-input data on regional production to calibrate the model. In the SWAP model we additionally assume that profit-maximizing behavior and short run equilibrium conditions led to the observed base year resource allocation. Since these models use an explicit primal specification of agricultural production, they can model policies defined in terms of physical resource limits rather than financial outcomes.

1.2. Optimization models and Positive Mathematical Programming

Moore and Hedges (1963) first introduced models of irrigated agriculture as a way to estimate irrigation water demand. They, and later studies, used mathematical (typically linear) programming models to estimate irrigation water demand elasticities. Gardner (1983) reviewed studies on irrigation water demand, with emphasis on California, completed during the 1960s and 1970s. This literature has since evolved to focus on large-scale regional optimization models. Today optimization models are used to analyze water demand and agricultural–environmental policies, since these models work better with a multitude of resource constraints and complex interactions between agriculture and the environment (Griffin, 2006).

A major problem that initially plagued optimization models was a tendency to overspecialize in crop production (Howitt, 1995a). In response, the 1980's saw the first models based on the technique of Positive Mathematical Programming (PMP). PMP is a deductive approach to simulating the effects of policy changes on cropping patterns at the extensive and intensive margins. The term "positive" implies the use of observed data as part of the model calibration process. PMP has several advantages over traditional optimization models. First, the PMP cost function calibrates the model exactly to observed values of production output and factor usage. Second, PMP adds flexibility to the profit function by relaxing the restrictive linear cost assumption. A third advantage is that PMP does not require large datasets. Heckelei and Britz (2005) note that PMP models can be viewed as a bridge between econometric models, with substantial data requirements, and more limited traditional optimization models. Finally, programming models including the subset of PMP models such as SWAP are more responsive to policy changes than statistical (inductive) models of agricultural production (Scheierling et al., 2006).

Calibration of production models by PMP has been reviewed extensively in the literature and variations on the base method have been developed. Buysse et al. (2007) and Heckelei and Wolff (2003) argue that shadow values from calibration and resource constraints are an arbitrary source of information for model calibration. Subsequent research suggests the use of exogenous information such as land rents instead of shadow values (Heckelei and Britz, 2005; Kanellopoulos et al., 2010). Heckelei and Britz (2005) and Paris and Howitt (1998) propose a generalized maximum entropy (GME) formulation to estimate resource and calibration constraint shadow values. However, the GME procedure has seen little use in applied research. Merel and Bucaram (2010) and Merel et al. (2011) propose calibration against exogenous, and potentially regionally-disaggregate, supply elasticity estimates.

Research on linked hydrologic and economic models has evolved parallel to research on PMP with a focus on improved policy simulations and analysis. Economic models typically omit a hydrologic representation and hydrologic models lack the ability to economically allocate water. Hybrid hydrologic–economic models can be holistic (one model) or compartmental (sequential iteration between different models) (Cai, 2008; Braat and vanLierop, 1987). Compartmental hydrologic–economic models are frequently a hydrologic model linked with an economic model calibrated by PMP. Gomann et al. (2005) link the RAUMIS economic model, calibrated using PMP, to GROWA98 and WEKU hydrologic models to model the effects of Nitrogen tax relative to a quota on dairy herds to increase water quality in Germany. In an example of work in California, Quinn et al. (2004) adopt a compartmental approach and develop the PMP APSIDE economic model which is linked to the CALSIM II water model. They also include climate simulations, in a third model, to evaluate climate change impacts in California. vanWalsum et al. (2008) introduce the bio-economic model Waterwise which is linked to the DRAM PMP model. They use the model to evaluate European Union water quality policies in the Netherlands.

Despite the many papers employing PMP models to infer economic values for water and environmental resources, we cannot find any publication that focuses on the calibration procedure for PMP economic models and formal diagnostic tests for each calibration stage. The calibration-diagnostic iterative procedure applies to standalone economic models and linked hydrologic–economic models in diverse geographic regions.

2. The Statewide Agricultural Production Model (SWAP)

2.1. SWAP modeling framework

A model is, by definition, a simplified representation of a real system. In the process of abstracting and simplifying a real system, a model loses some information; thus even with theoretically consistent structure it is unlikely that a model will calibrate closely to observed (base year) data. The problem is well documented in agricultural production modeling (Hazell and Norton, 1986). One solution is to use observed farmer behavior, in the form of observed land use patterns, and additional exogenous information to calibrate parameters of the structural model that exactly reproduce observed base-year conditions. Positive Mathematical Programming is a common calibration method for structural agricultural production models.

The SWAP model is a regional model of irrigated agriculture in California, calibrated using PMP. PMP can derive model parameters so that first-order conditions for economic optimization are satisfied at an observed base year of input and output data. This is accomplished by assumed profit maximizing behavior by farmers and a non-linear objective function. SWAP offers three key improvements over traditional PMP models. First, SWAP includes regional exponential PMP land cost functions, which corrects the inability of previous models, with quadratic functions, to handle large policy shocks. Second, SWAP includes regional Constant Elasticity of Substitution (CES) crop production functions which allow limited substitution between inputs. Leontief production functions were common in most previous models. Finally, regional crop prices are endogenously determined based on a statewide demand function.

SWAP was originally developed to be the agricultural economic component for the CALVIN model of the California water system (Draper et al., 2003). It has subsequently been used in a wide range of policy analyses in California. SWAP has been used to estimate economic losses due to salinity in the Central Valley (Howitt et al., 2009), economic losses to agriculture due to alternative conveyance in the Sacramento-San Joaquin Delta (Appendix to Lund et al., 2007), economic losses to agriculture and confined animal operations in California's Southern Central Valley (Medellin-Azuara et al., 2008), and economic effects of water shortage on Central Valley agriculture (Howitt et al., 2011). The model has also been linked to agronomic yield models in order to estimate effects of climate change on irrigated agriculture in California (Medellin-Azuara et al., 2012). Variations of SWAP also have been applied in other regions such as the US-Mexico border basins (Howitt and Medellín-Azuara, 2008; Medellin-Azuara et al., 2009). The model is used for policy analysis by the California Department of Water Resources (DWR, 2009) and the United States (U.S.) Department of the Interior (Interior), Bureau of Reclamation (Reclamation, 2011).

SWAP is defined over homogenous agricultural regions and assumes that farmers maximize profits subject to resource, technical, and market constraints. Farmers sell and buy in competitive markets where any one farmer cannot affect the price of any commodity. The model selects crops, water supplies, and other inputs that maximize profit subject to constraints on water and land, and subject to economic conditions regarding prices, yields, and costs. The model incorporates water supplies from state and federal projects, local water supplies, and groundwater. As conditions change within a SWAP region (e.g. the quantity of available project water supply increases or the cost of groundwater pumping increases) the model optimizes production at both the extensive and intensive margins by adjusting the crop mix, water sources and quantities used, and other inputs. It will also fallow land in response to resource conditions.

The SWAP model is written in GAMS (General Algebraic Modeling System) and solved using the non-linear solver CONOPT-3. The objective is to maximize the sum of producer (regional profits) and consumer surplus.

2.2. Model development and calibration

Development of the SWAP model is divided into calibration and policy analysis phases. Calibration is analogous to parameter estimation in econometric models or calibration in Computable General Equilibrium (CGE) models. Policy analysis estimates the effects of changing prices, costs, resources, or institutions given the calibrated parameter values.

We detail the calibration procedure for SWAP and emphasize model improvements and diagnostic checks in the process. The calibration procedure for SWAP reflects most of the ten steps discussed in Jakeman et al. (2006) with particular emphasis on sequential calibration and a parallel set of diagnostic tests to check model performance. Stepwise model development procedures have been applied for many modeling problems, including neural networks (Piuleac et al., 2010), and computational fluid dynamics (Blocken and Gualtieri, 2012). The stepwise tests specified in Fig. 1 are ordered in a logical sequence. For example, the first test for positive net returns is a necessary condition for an optimal solution in the calibrated linear program. Likewise, the equality of the input marginal value products to their opportunity costs is a necessary

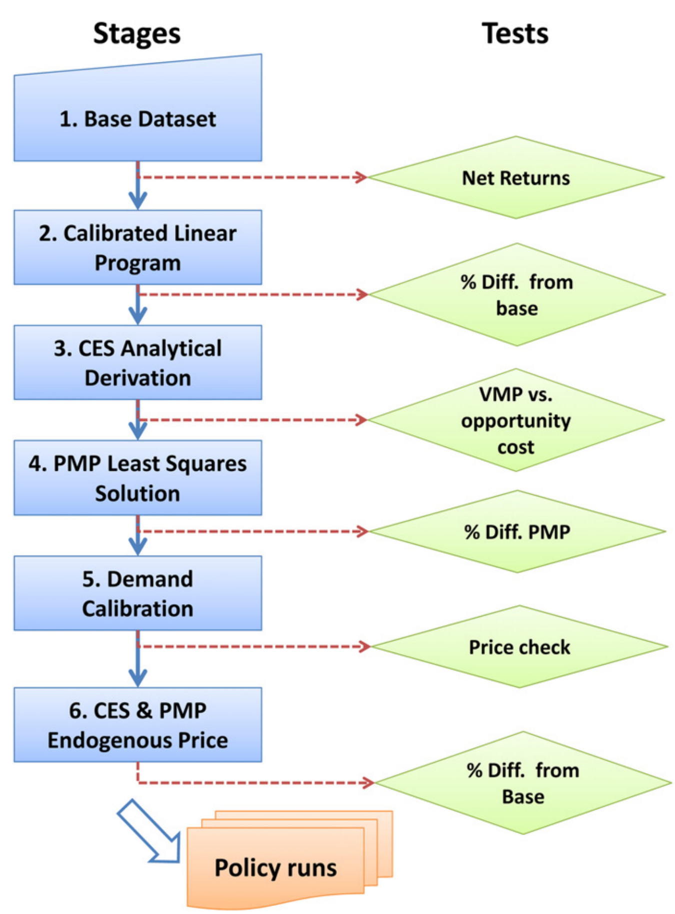


Fig. 1. SWAP calibration stages and tests.

condition for optimal input calibration in the nonlinear CES model. The sequential tests defined in Fig. 1 are a blueprint for model validation and identification of potential problems.

The calibration phase of the SWAP model uses a sequential six-step process outlined in Fig. 1. The six steps are (i) assemble input, output and elasticity data, (ii) solve a linear program subject to fixed resource and calibration constraints, (iii) derive the CES production function parameters using input opportunity costs from step two, (iv) estimate the crop and region-specific PMP cost functions using a least squares method, (v) calibrate the aggregate demand functions and regional adjustment costs using prior demand elasticity estimates, and (vi) optimize and simulate the calibrated SWAP model which includes tests for adequate calibration in terms of input and output prices and quantities.

Model calibration data should be representative of "normal" production conditions in the relevant region. We take 2005 as the base year in the SWAP model because it represents the most recent data available for an average water and price year in California. The model calibrates to the base year in terms of the following parameters: crop output quantities, output prices, input quantities, input value marginal products, variable costs, and imputed costs to fixed inputs.

2.2.1. Step I: data assembly

The level of spatial aggregation is important for defining the scope and method of analysis. Disaggregated production models typically require more data but tend to be effective in policy analysis in rural economies (Taylor et al., 2005). When agricultural production is homogeneous and production conditions are relatively stable, there is less information gained from disaggregation. SWAP aggregates agricultural production data to the level of representative regions. The SWAP regions are based on the California Department of Water Resources (DWR) Detailed Analysis Units (DAU). Each SWAP region is composed of one or more DAU with homogenous microclimate, water availability, and production conditions. This scale is more suitable for statewide hydro-economic models that require marginal economic values of water for competing agricultural and urban demand locations (Draper et al., 2003). The SWAP model has 27 base regions in the Central Valley plus the Central Coast, the Colorado River region that includes Coachella, Palo Verde and the Imperial Valley and San Diego, Santa Ana and Ventura, and the South Coast. The model has a total of 37 agricultural regions, only 27 regions in the Central Valley are considered for the analysis in this paper. Fig. 2 shows California agricultural area covered in SWAP.

We aggregate crops into 20 representative crop groups. A single crop group can represent several individual crops. Irrigated land use represents the area of all crops within the group, production costs and returns are represented by a single proxy crop for each group. The current 20 crop groups were defined in collaboration with DWR (DWR, 2010). For each group we choose the representative (proxy) crop based on four criteria: (i) availability of

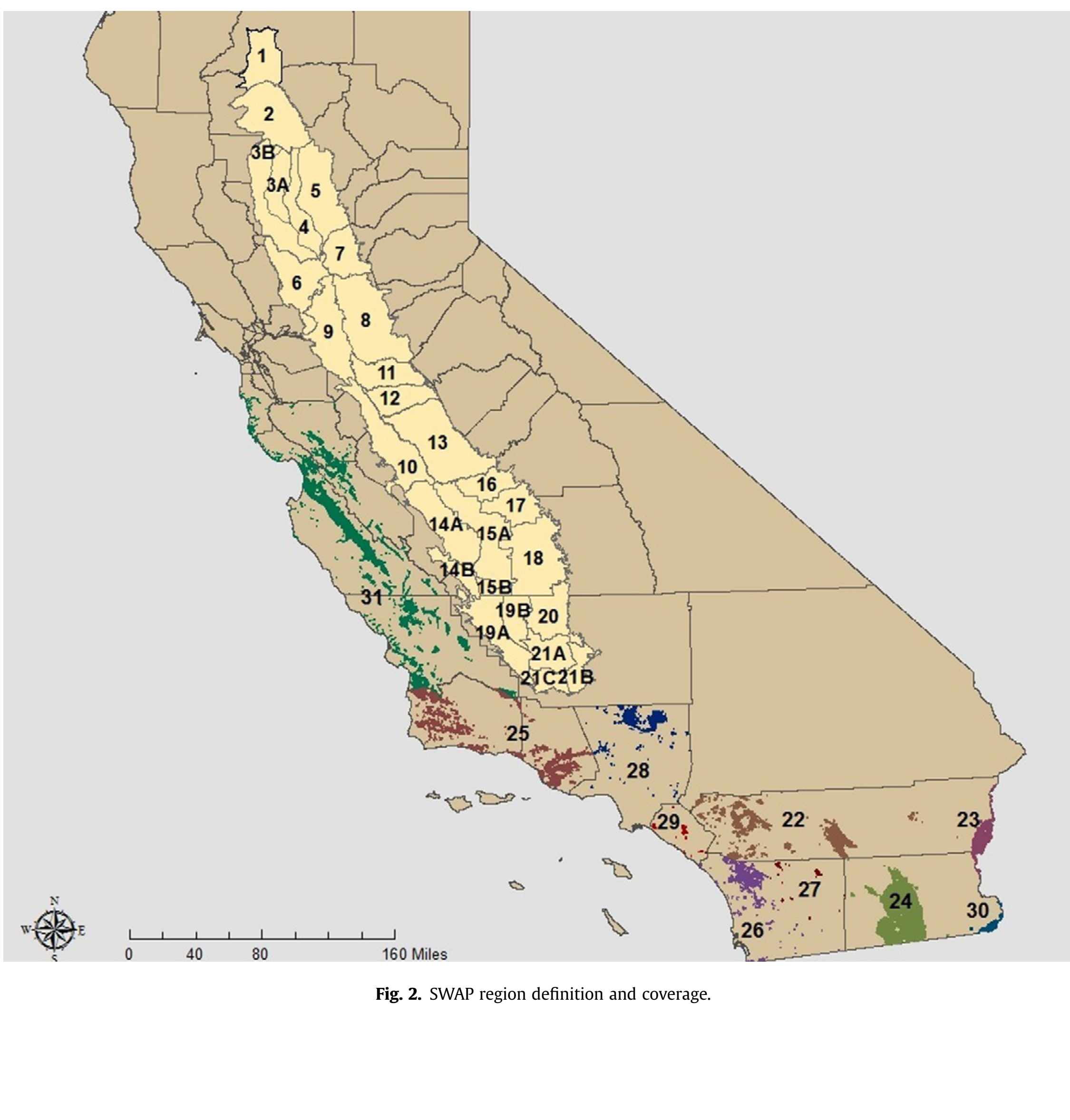


Fig. 2. SWAP region definition and coverage.

a detailed production budget, (ii) representative of the largest land use within a group, (iii) representative of water use (applied water) of all crops in the group, and (iv) having similar gross and net returns as other crops in the group. The relative importance of these criteria varies by crop. The 20 crop groups include almonds and pistachios, alfalfa, corn, cotton, cucurbits, dry beans, fresh tomatoes, grains, onions and garlic, other deciduous, other field, other truck, irrigated pasture, potatoes, processing tomatoes, rice, safflower, sugar beets, subtropical, and vines.

Variable input costs for the crop groups are derived from the regional cost and return studies from the University of California Cooperative Extension (UCCE, 2011). There are four aggregate inputs to production, (i) land, (ii), labor, (iii) water, and (iv) other supplies. All inputs except water are derived from the UCCE Budgets. Since cost budgets represent best management practices, SWAP also uses the corresponding yields from the budgets. Commodity prices for the base year in the model are from the California County Agricultural Commissioner's reports published by the U.S. Department of Agriculture (USDA, 2011).

We derive applied water per hectare (base) requirements for crops in SWAP from DWR estimates (DWR, 2010). DWR estimates are based on Detailed Analysis Units (DAU). An average of DAU's within a SWAP region is used to generate a SWAP region specific estimate of applied water per hectare for SWAP crops.

The SWAP model includes five types of surface water: State Water Project (SWP) delivery, three categories of Central Valley Project (CVP) delivery, and local surface water delivery or direct diversion (LOC). The three categories of CVP deliveries represent water service contract and include Friant Class 1 (CVP1), Friant Class 2 (CL2), and water rights settlement and exchange delivery (CVPS). CVP and SWP water costs have two components, a project charge and a district charge. The sum of these components is the region-specific cost of the individual water source.

Groundwater pumping costs are calculated as two components, the fixed cost per cubic meter based on typical well designs and costs within the region, plus the variable cost per cubic meter. The variable cost per cubic meter is O&M plus energy costs based on average total dynamic lift within the region. In our example application we consider a short run drought analysis and hold dynamic lift and groundwater pumping costs constant. Long run policy analysis may link the SWAP model to a groundwater model such as the Central Valley Hydrologic Model (CVHM) to simultaneously estimate changes in regional depth to groundwater (Reclamation, 2011).

The model calibration approach, discussed in the following section, is driven by the first order conditions and fixed resource constraints. Since the underlying objective is to maximize profits, subject to inequality constraints on the fixed inputs, each regional crop production activity must have a positive gross margin at the base calibration values. As such, the essential test at this stage is to ensure that the gross margin over variable costs is positive for those crops actually grown. If the net returns to land and management are negative after checking the data, there are several ways of addressing the problem. The simplest approach may be to use a lower bound calibration constraint in Step II to calculate the needed reduction in the land opportunity cost from the lower bound constraint shadow value. More generally, the researcher should consult extension agents and other experts to identify potential inconsistencies in the crop budgets or other input data.

2.2.2. Step II: linear calibration program

In this step we solve a linear program of farm profit maximization with calibration constraints set to observed values of land use. All other production inputs are normalized to land. The Lagrangian multipliers on the calibration and resource constraints are used in steps three and four to parameterize regional CES production functions and exponential PMP cost functions. We define sub-index g for (27) agricultural (SWAP) regions, i for (20) crop groups, j for (4) production inputs, and w for (6) individual water sources.

We solve a linear program to obtain marginal values on calibration and resource constraints. The linear program objective function is to maximize the sum of regional profits across all crops by optimizing land use $xl_{gi,land}$ and water use $watl_{gw}$. Equation (1) defines the objective function,

$$\max_{xl_{gi,land},watl_{gw}} \Pi = \sum_g \sum_i \left(v_{gi} yld_{gi} - \sum_{j \neq water} \omega_{gij} a_{gij} \right) xl_{gi,land} - \sum_g \sum_w \left(watl_{gw} \varpi_{gw} \right), \tag{1}$$

where v_{gi} are region-specific crop prices (marginal revenue per tonne of output), yld_{gi} are the base yields for crop i in region g, ω_{gij} are input costs, ϖ_{gw} are water costs, and a_{gij} are regional Leontief coefficients defined in Equation (2). $\tilde{x}_{gij}$ represents the *observed* level of input use.

$$a_{gij} = \frac{\tilde{x}_{gij}}{\tilde{x}_{gi,land}} \tag{2}$$

Production is constrained by resource availability of binding inputs including land and water. These are treated separately in the calibration program, since regions may be binding in land, water, or both. The land resource constraints are defined as

$$\sum_i xl_{gi} \le b_{g,land} \quad \forall g, \tag{3}$$

where $b_{g,land}$ are region-specific land availability constraints. The water constraints are defined by region and water source,

$$\sum_i aw_{gi} xl_{gi} \le \sum_w watl_{g,w} \quad \forall g, \tag{4}$$

and

$$\sum_w wat_{gw} \le \sum_w \text{watcons}_{gw} \quad \forall g \tag{5}$$

where watcons_{gw} are region and water source-specific constraints, and aw_{gi} are crop water requirements (applied water per hectare) and may reflect regional difference in average irrigation efficiency or consumptive use. Define λ_g^L and λ_g^W as the shadow values for Equations (3) and (4), respectively.

A calibration constraint forces the program to reproduce base year observed cropping patterns. We include a perturbation ($\varepsilon = 0.0001$) to decouple the resource and calibration constraints as detailed in Howitt (1995a),

$$x_{gi,land} \le \tilde{x}_{gi,land} + \varepsilon \quad \forall\ g, i. \tag{6}$$

We add the calibration constraint to land only, and use the shadow value of land λ_{gi}^C as the marginal price needed to calibrate optimal land allocation in Equation (6). The other inputs are calibrated by using the first order conditions for the CES production function defined later in the process.

Two tests are applied to the output of the Step II model. The first test measures any deviation in regional crop input allocation by the model. Percentage deviations in input use by crop and region of less than 1% are permissible given the small perturbations in the calibration constraints, but any input deviation greater than this implies negative gross margins, or unduly restrictive fixed input constraints. The second calibration test verifies that the number of non-zero dual values on calibration constraints plus the number of non-zero shadow values on binding resource constraints equal the number of non-zero production activities in each region. If this test does not hold, the model will not have sufficient cost information to calibrate the full set of non-zero activities as some crops should have interior solutions, but do not have calibration shadow values to derive them.

2.2.3. Step III: production function parameter calibration

In this step we sequentially derive the parameters for the Constant Returns to Scale (CRS) CES production function for each region and crop following the procedure developed in Howitt (1995b). The CES is a flexible functional form which allows for a constant rate of substitution between production inputs and nests Leontief (fixed proportions) and Cobb–Douglas (unit substitution) production technologies. Researchers use various types of quadratic functions in agricultural optimization models (Cai, 2008). The model which preceded SWAP in California, the Central Valley Production Model, modeled production along the water use-irrigation efficiency isoquant (Reclamation, 1997). SWAP improves previous methods and calibrates a CES production function for each crop and region. One key property of the CES production function is that it defines the rates at which inputs can be substituted for each other, for example, applied water used in irrigation can be partly substituted for by increased irrigation efficiency which requires additional labor and capital.

The Constant Returns to Scale (CRS) CES production functions for every region and first-order conditions for an optimum input allocation yield a sequential set of conditions to solve for the parameters of the CES. The theoretical properties may be found in Beattie and Taylor (1985). We define the CES functions as

$$y_{gi} = \tau_{gi}\left[\beta_{gi1}x_{gi1}^{\rho_i} + \beta_{gi2}x_{gi2}^{\rho_i} + \ldots + \beta_{gij}x_{gij}^{\rho_i}\right]^{v/\rho_i}, \quad (7)$$

where y_{gi} represents output of crop i in tonnes for region g, by combining aggregate inputs j. The scale parameters are (τ_{gi}) and the relative use of production factors is represented by the share parameters β_{gij}. Production factor use is given by x_{gij}. The returns to scale coefficient is v and CRS requires that the coefficient is set at 1.

The SWAP model uses a non-nested CES production function with the same elasticity of substitution between any two inputs. The SWAP model is also able to handle a nested-CES production function with two or more sub-nests and corresponding versions of the model have been developed. If data are available the substitution elasticity should be estimated. If substitution elasticities are available from existing studies those can be used. Currently there are insufficient data to estimate the elasticity of substitution, thus the value is fixed at $\sigma = 0.17$ for all inputs. We assume this value to allow for limited substitution between inputs based on experience from previous analyses.

Limited substitution between inputs is consistent with observed farmer production practices. Namely, we observe that farmers can, over a limited range, substitute among inputs in order to achieve the same level of production. Fig. 3 shows an example of a CES production surface. To show the CES function as a 3-dimensional surface two inputs (supplies and land) are held constant. The vertical axis shows total production of alfalfa in Region 15 given different combinations of water and labor which are shown on the horizontal axes. Fig. 3 illustrates two important aspects of the CES production function. First, substitution between inputs can be seen by holding production constant (the vertical axis) and sliding around the production surface. There is limited substitution between water and labor, as shown by the "sharp" corners to the production surface. Second, Fig. 3 demonstrates the ability of SWAP to model deficit (stress) irrigation by farmers or, more generally, the marginal product of a given input. Faced with a water shortage we expect that farmers may deficit-irrigate some crops. Holding labor constant and sliding along the production surface, as water is decreased production (yield) decreases as well. Additional restrictions can be imposed to incorporate exogenous agronomic data.

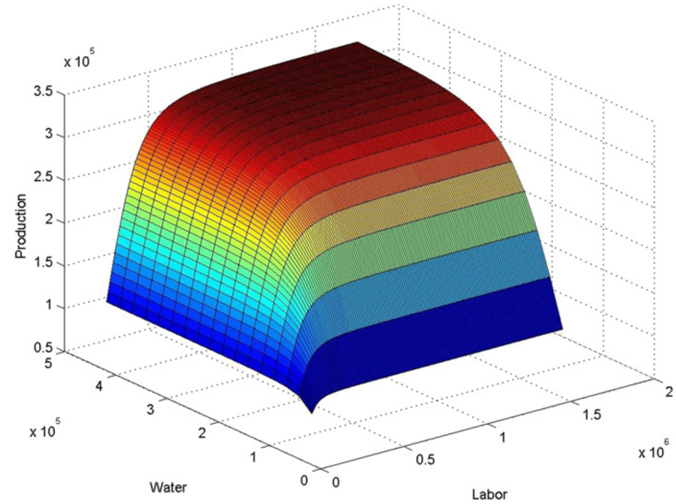


Fig. 3. Simplified CES production function surface for alfalfa in region 15.

The first order condition for optimal input allocation is that the value marginal product (output price times the marginal product) of each input for each crop and region is equal to the marginal cash cost plus opportunity cost of the input. This is equal to the base input price plus the dual value on the resource constraints, λ_g^L and λ_g^W, and, when binding, the dual value on the calibration constraint, λ_{gi}^C. The linear program in Step II will not have calibration shadow values for activities associated with the binding resource constraints. In the absence of prior estimates of the marginal productivity of these crops, we impose the assumption that marginal productivity decreases 25% over the base condition productivity and thus use 25% of the land resource shadow value as a proxy for the calibration shadow value, and adjust the other calibration values accordingly. While this is a general assumption over different regions and crops, it provides a robust method for full calibration of all the observed crops without inducing infeasibilities from more arbitrary exogenous restrictions.

Let the cost per unit of each input, inclusive of marginal cash cost and opportunity cost of input j be ω_j. To simplify notation, consider a single crop and region and normalize the price per unit output to 1. Define

$$\rho = \frac{\sigma - 1}{\sigma}, \quad (8)$$

and the corresponding farm profit maximization problem, optimizing over input use X_j, is written as,

$$\max_{x_j} \pi = \tau\left[\sum_j \beta_j x_j^{\rho}\right]^{v/\rho} - \sum_j \omega_j x_j. \quad (9)$$

Constant returns to scale requires that $v = 1$ and

$$\sum_j \beta_j = 1. \quad (10)$$

We use the restrictions imposed by constant returns to scale and take ratios of any two first order conditions to derive the familiar optimality condition that marginal rate of technical substitution equals the ratio of input costs. Let l correspond to all $j \neq 1$ and by rearranging and using the restriction in Equation (10) we can explicitly solve for the first (or any arbitrary) coefficient,

$$\beta_1 = \frac{1}{1 + \frac{x_1^{(-1/\sigma)}}{\omega_1}\left(\sum_l \frac{\omega_l}{x_l^{(-1/\sigma)}}\right)}. \tag{11}$$

We use the same procedure as above for all other β_l where $l \neq 1$, thus

$$\beta_l = \frac{1}{1 + \frac{x_1^{(-1/\sigma)}}{\omega_1}\left(\sum_l \frac{\omega_l}{x_l^{(-1/\sigma)}}\right)} \frac{\omega_l}{\omega_1} \frac{x_1^{-1/\sigma}}{x_l^{-1/\sigma}}. \tag{12}$$

We calculate the scale parameter, for each region and crop, from the definition of the CES production function, evaluated at the base level. The scale parameter is

$$\tau = \frac{\left(yld/\tilde{x}_{land}\right)\cdot \tilde{x}_{land}}{\left[\sum_j \beta_j x_j^{\rho}\right]^{\upsilon/\rho_i}}. \tag{13}$$

The process generalizes to any number of regions and crops. In SWAP this process is automatically performed for all crops and regions and the production functions are fully calibrated.

2.2.3.1. Numerical scaling issues in optimization models. From the first order conditions we see that

$$\beta_l = \frac{\omega_l}{\omega_1} \frac{\beta_1 x_1^{(-1/\sigma)}}{x_l^{(-1/\sigma)}}, \tag{14}$$

for any given input l. If input costs (marginal cash cost plus opportunity cost) of two inputs are of a different order of magnitude this can cause the β_j coefficients to become unbalanced and lead to numerical issues with model calibration. Specifically, an ill-conditioned calibration routine will tend to set $\beta_l \approx 1$ and all other $\beta_j \approx 0$. In turn, the model will not calibrate with a low elasticity of substitution (large value in the exponent). This type of data issue is common with large-scale regional production models since inputs are aggregated into coarse categories. For example, other supplies have a much larger cost per unit land than labor costs for many crops causing ill-conditioned matrices that impede numerical convergence to an optimal solution.

There are many sophisticated scaling approaches but a simple solution used in SWAP is to numerically scale input costs into units of the same order of magnitude. We use land costs as the reference scale and convert input costs, except for land, into land units. We calculate the ratio of input use to total hectares, for each crop and region, and normalize the costs of production into the corresponding unit. This scaling is used throughout the SWAP program. At the end of the program we use a de-scaling routine which simply reverses this process to convert input use and costs back into standard units.

2.2.4. Step IV: estimating an exponential PMP cost function

The SWAP model posits that farmers cultivate the best land first for any given crop so additional land put into production will be of lower quality. The effect will vary over space and will depend on several additional factors including management skills, field-specific physical capital, and the dynamic effects of crop rotation. In general, additional land into production requires a higher cost to prepare and cultivate. We combine this unobservable (directly) information with average production costs to calibrate exponential land cost functions in the model.

PMP land cost functions are calibrated using information from acreage response elasticities and shadow values (implied values) on calibration constraints. Merel and Bucaram (2010) derive conditions for the exact calibration to elasticities for the Leontief and CES model with a quadratic PMP cost function. They show that the approach used here can be defined as myopic calibration, since it does not account for the effect of crop interdependency on the marginal elasticity. However they do show that under so-called "number of crops" and "dominant response" conditions, the myopic approach can be an adequate approximation. With 20 representative crops, the SWAP model is likely to satisfy both conditions, though we have not numerically tested the conditions since they are derived for a quadratic PMP cost function. In another more general formulation, Merel et al. (2011) show that a decreasing returns to scale CES function can calibrate exactly to a wider set of elasticities. They also propose that for multiple regions such as in SWAP, the individual region elasticities be allowed to vary as long as the weighted aggregate crop elasticity calibrates to the prior value. This modification will be incorporated in future versions of the SWAP model.

Previous PMP models, such as CVPM, were specified with quadratic PMP land cost functions. Fig. 4 shows a comparison of the exponential PMP cost function and the more frequently used quadratic PMP cost function that implies a linear marginal cost on land. Calibrating a quadratic total cost function subject to a supply elasticity constraint can result in negative marginal costs over a range of low hectares for a specific crop and region. This is inconsistent with basic production theory and can result in numerical difficulties both in the calibration phase and with policy analysis. The exponential cost function is always bounded above zero, by definition, which is consistent with observed costs of production. The marginal factor cost of land has the required first and second order conditions for calibration and minimizes the difference from the prior elasticity value. A second practical advantage is that the exponential cost function often can fit a desired elasticity of supply without forcing the marginal cost of production at low hectares to have unrealistic values. A quadratic PMP cost function, often forces the modeler two choose between an

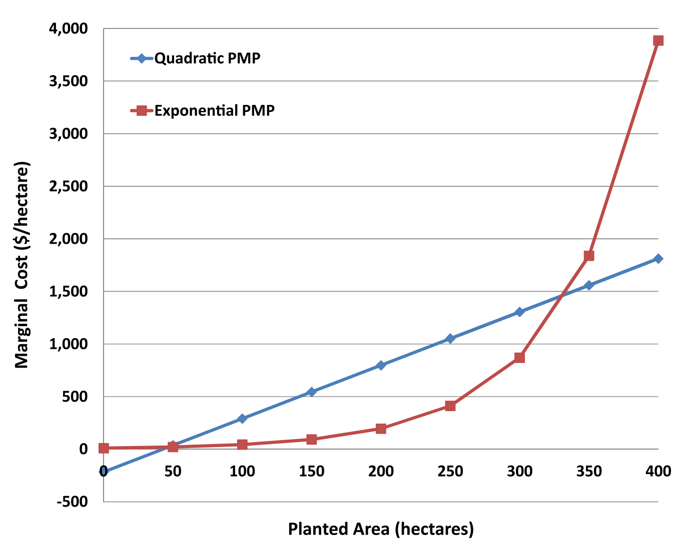


Fig. 4. Comparison of quadratic and exponential PMP land cost functions (adapted from Medellín-Azuara et al., 2010).

unrealistic elasticity, which influences policy response, or an unrealistic initial marginal cost of production. Researchers considering using a quadratic total function should beware of the potential for negative marginal costs.

Formally, in Step IV of PMP calibration we estimate parameters for the exponential cost function. We define the total land cost function as

$$TC(x_{land}) = \delta e^{\gamma x_{land}} \tag{15}$$

where δ and γ are the intercept and the elasticity parameter for the exponential land area response function, respectively. These parameters are from a regression of the calibration shadow values on the observed quantities, restricted by the first order conditions, and elasticity of supply for each crop group from previous studies. For clarity, consider a single PMP cost function within a single region for a specific crop, defined as

$$MC = \frac{\partial TC}{\partial x_{land}} = \delta\gamma e^{\gamma x_{land}}, \tag{16}$$

where marginal cost equals cash cost plus marginal opportunity cost. The acreage supply elasticity, η, is

$$\eta = \frac{\partial x_{land}}{\partial TC}\frac{TC}{x_{land}}, \tag{17}$$

where

$$\frac{\partial x_{land}}{\partial TC} = \frac{1}{\delta\gamma e^{\gamma x_{land}}}. \tag{18}$$

Simplifying and noting that the logarithmic version of the equation is linear,

$$\ln(\eta\delta\gamma x_{land}) + \gamma x_{land} = \ln(R). \tag{19}$$

Thus, two conditions, Equations (16) and (19), must be satisfied at the calibrated (observed) base level of land use. The former is the PMP condition and holds with equality, the latter is the elasticity condition which we fit by least-squares.

The test at this stage of calibration is to calculate the deviation of the marginal PMP cost at the base land allocation from the shadow value of the corresponding calibration constraint, λ^{C}_{land}, derived in Step II. If deviations are more than a few percentage points in this test, the model does not accurately calibrate, usually due to a non-optimal solution in the least squares fit for the parameters, or an unduly restrictive elasticity constraint on the estimation.

2.2.5. Step V: calibrating demands for endogenous prices

We include endogenous prices through downward sloping demand functions for all crops in SWAP. This represents the consumer side of the market and provides a mechanism for calculation of consumer surplus in the model. As such, the objective function is to maximize the sum of producer and consumer surplus.

We define a subroutine to estimate a statewide demand function for each crop based on the California crop demand elasticity as estimated by Green et al. (2006). We specify the model with linear California-specific crop demand functions. The demand curve represents consumer's willingness-to-pay for a given level of crop production. All else constant, as production of a crop increases, the willingness-to-pay for additional production is expected to fall and to clear the market the price must also fall. The extent of the price decrease depends on the elasticity of demand or, equivalently, the price flexibility. The latter refers to the percentage change in crop price due to a percent change in production given a perfectly competitive market.

We account for regional price differences in the California statewide demand functions. Crop demand includes both in-state and out of state demands for California crops. The statewide demand functions are defined using a base price and regional prices may include deviations from that base price. The state-wide market price of each crop is assumed constant across regions in the state. Regional deviations from the base reflect variations in distance from markets, production contracts, crop quality, variety, harvest season, and other factors.

Production shares by region and price flexibilities of demand are the relevant data needed to calibrate the demand functions. The price flexibilities are based on earlier work for the CVPM model (Reclamation, 1997). We specify a linear inverse-demand function with two parameters, for crop i in region g, defined as

$$p_i = \xi\alpha_i^1 - \alpha_i^2\left(\sum_g\sum_j y_{gij}\right). \tag{20}$$

The crop price is p_i and parameters α_i^1 and α_i^2 represent the intercept and slope of the crop-specific inverse demand curve, respectively. The parameter ξ is a potential parallel shift in demand due to exogenous factors. We calculate the California price for crop i by weighting the regional observed prices v_{gi} by the fraction of region g in the statewide production. Proportion of production (pp_{gi}) is defined as

$$pp_{gi} = \frac{\tilde{y}_{gi}}{\sum_g \tilde{y}_{gi}}, \tag{21}$$

where $\tilde{y}_{gi}$ is the base production. The weighted California price is consequently defined as

$$wp_i = \sum_g v_{gi}pp_{gi}. \tag{22}$$

The regional marketing cost is the difference between the observed regional price (base) and the calculated California crop price. This reflects differences in price which can be attributed to various region-specific differences discussed above and is defined as

$$rmc_{gi} = v_{gi} - wp_{gi}. \tag{23}$$

Given the above definitions, we can calculate the parameters of the inverse demand functions. For a given price flexibilities (χ_i), the slope parameter is

$$\alpha_i^2 = \frac{\chi_i wp_{gi}}{\sum_g \tilde{y}_{gi}}. \tag{24}$$

Consequently, the intercept is

$$\alpha_i^1 = wp_i - \alpha_i^2\sum_g \tilde{y}_{gi}. \tag{25}$$

The test at this stage is to substitute the regional production quantities into Equation (20) and check to see if the equilibrium price adjusted by the regional marketing cost calibrates closely, within a few percentage points, to the regional price.

2.2.6. Step VI: a calibrated non-linear optimization program

The last step in SWAP calibration combines the calibrated functions into a non-linear optimization program. This base program does not include a policy shock and is used to ensure that the calibrated model reproduces observed base year conditions. We include endogenous price determination, agronomic constraints, and resource constraints in the program. With endogenous prices, the objective function is to maximize the sum of producer and consumer surplus.

$$\begin{aligned}\underset{x_{gij},wat_{gw}}{\text{Max}}\ PS+CS = &\sum_i \left(\xi\alpha_i^1\left(\sum_g y_{gi}\right)+\frac{1}{2}\alpha_i^2\left(\sum_g y_{gi}\right)^2\right)\\ &+\sum_g\sum_i\left(rm_{gi}\left(\sum_j y_{gi}\right)\right)\\ &-\sum_g\sum_i\left(\delta_{gi}\exp\left(\gamma_{gi}x_{gi,land}\right)\right)\\ &-\sum_g\sum_i\left(\omega_{gi,\text{supply}}x_{gi,\text{supply}}+\omega_{gi,labor}x_{gi,labor}\right)\\ &-\sum_g\sum_w\left(\varpi_{gw}wat_{gw}\right).\end{aligned} \tag{26}$$

The choice variables are inputs (land, labor, water, and other supplies) for each region g and crop i, in addition to total regional water use by source. The first term of the objective function, Equation (26), is the sum of gross revenue plus consumer surplus for all crops and regions, measured relative to the base crop prices. The second term captures the region-specific gross revenue associated from deviations in regional prices from the base prices (these are denoted regional marketing costs). The third term is the region and crop specific PMP land costs. These include both the direct costs of land reported in the base data and the marginal costs inferred from the shadow values on the resource and calibration constraints. The fourth term accounts for labor and other supply costs across all regions and crops. Finally, the fifth term of the objective function is the sum of irrigation water costs by region, crop, and water source. This term is written separately to emphasize that SWAP includes water costs that vary by source.

We define a convex constraint set with resource, agronomic, and other policy constraints. First, the production technology generates the regional crop production y_{gi} as defined in Equation (7). Resource constraints include regional input constraints,

$$\sum_i x_{gij} \le b_{gj} \quad for \quad j \ne \text{water}, \tag{27}$$

where b_{gi} is total input available by region. Water constraints are incorporated as a restriction on the total water used by region and source,

$$wat_{gw} \le \text{watcons}_{gw} \tag{28}$$

and total water input use,

$$\sum_i x_{g,i,water} \le \sum_w wat_{gw}. \tag{29}$$

SWAP allows for movement along the CES production surface, i.e. substitution between inputs. One intensive margin adjustment commonly observed in agriculture is deficit (stress) irrigation. SWAP endogenously determines potential stress irrigation which is dictated by the shape of the respective CES production function. An upper-bound constraint of 15% stress irrigation (relative to the base condition applied water per hectare) is allowed in the model, to prevent the model from reducing applied water rates below the range normally observed. We define the stress irrigation constraint as

$$\frac{x_{gi,water}}{x_{gi,land}} \ge 0.85\ aw_{gi} \tag{30}$$

Perennial crops are subject to natural retirement or rotation as yields decline in older stands. The average perennial life ($prenlife_i$) is 25 years for almonds and pistachios, other deciduous, and vine crops in SWAP (UCCE, 2011). Subtropical crops have an average life of 30 years. If the time horizon of analysis exceeds 30 years then we expect that farmers have full flexibility to adjust production decisions, including retirement of orchards and vineyards. In the short run we expect farmers devote resources to preserve perennial stands still in prime bearing years. The SWAP model constrains perennial retirement in the short-run (less than the life of the field) to be a proportion of total land use. The proportion is the short-run horizon in years divided by the perennial life. This implicitly assumes that stand age is uniformly distributed and that only older, lower-bearing, fields will be retired. Formally,

$$x_{g,pren,j} \ge \tilde{x}_{g,pren,j}\left(1-\min\left(1,\frac{yr}{prenlife_{pren}}\right)\right), \tag{31}$$

where $pren \subset i$ and yr is the number of years of the analysis. Marques et al. (2005) demonstrate a two-stage formulation to more explicitly address permanent and annual crops for a range of water availability conditions.

We also include a regional silage constraint for dairy herd feed in the model. The silage constraint forces production to meet the regional feed requirements of the California dairy herd. For example, each cow consumes 20.5 kg of silage per day and corn grain yields are 11.01 tonnes per hectare thus each cow requires about 0.11 silage hectares per year. Multiplying the silage hectares per cow per year by the number of cows in each region yields the minimum silage requirement. The default model assumes a constant herd size into the future, though additional information about future of herd sizes could be used. This constraint can be excluded if the policy being assessed causes relatively small changes in water supply relative to existing regional supplies. Formally,

$$x_{g,corn,land} \ge \tilde{x}_{g,corn,land}, \tag{32}$$

where $\tilde{x}_{g,corn,land}$ defines the minimum silage constraint for each region.

Maximizing Equation (26) subject to Equations (27)–(29), where production satisfies Equation (7) by choosing the optimal input allocation for each crop and region yields a unique maximum for the SWAP model. The result of the base model run is used to determine if the model calibrates properly. Constraints defined by Equations (30)–(32) are relaxed in the base model in order to check for proper calibration.

There are three fundamental underlying assumptions which we want to emphasize. First, we assume water is interchangeable among crops in the region. Second, a representative regional farmer acts to maximize annual expected profits, equating the marginal revenue of water to its marginal cost. Third, a region selects the crop mix that maximizes profits within that region. This assumes sufficient levels of water storage and internal water distribution capacity and flexibility.

We use the base program to evaluate the fit of the fully calibrated model. The final test for the fully calibrated model compares the percentage difference in input allocation and production output for the model and the base data. The next stage of testing, test 3 in Fig. 1, compares the value marginal product of inputs and their marginal costs for each regional crop input. This test checks that the calibrated model satisfies the necessary conditions for optimization in the CES model (Howitt, 1995b). Before policy scenarios are run, the elasticities of output supply and input demand should be tested by sensitivity analysis in output prices and input price and quantities.

At this point in the program, the six-step calibration routine is complete and, assuming all conditions are met, the modeler can be confident the model has successfully calibrated. The final stage involves specification of a policy scenario for a second non-linear program. We call this the policy analysis phase. In this phase the modeler specifies the non-linear program (calibrated from the previous six-step procedure) with relevant policy constraints. Policy constraints include adjustments ranging from simple shocks to relative prices to complicated adjustments to production technology and interactions with biophysical models.

3. Policy application of SWAP

3.1. Water markets application background

In theory, fully-flexible agricultural water markets allow water to flow from low to high-value uses such that marginal value product is equalized across regions. Historical water rights holders, such as farmers in the Central Valley, are able to evaluate tradeoffs between production and selling water in the market which leads to more efficient allocation of the resource across the state. In practice, California has a limited history with water markets. The 1991 water bank, managed by the California Department of Water Resources (DWR), was a recent example which was only brought on by catastrophic drought. The water bank was a fairly rigid market institution but managed to buy and sell over 975 million cubic meters of water per year (Mm^3/yr). However, California has yet to adopt more flexible markets. As of 2003, 22 of 58 counties in the state had restrictions to prevent sales of groundwater and less than 3 percent of the water used is sold through markets (Hanak, 2003).

Transfers remain important for managing water in years of shortage. In 2009, the last year of a three-year drought in California, the Water Transfer Database compiled by the UCSB from the Water Strategist reports over 600 Mm^3/yr of water were transferred for agricultural use. Even with limited markets there are strong incentives to transfer water during years with shortage.

An important function for water markets in California is to help agriculture smooth drought losses. This is critically important when planning for a future with climate change and increasing demands for environmental water uses. During drought, water markets reduce land fallowing and stress irrigation. The former increases cultivated area and, in turn, generates additional agricultural jobs for the economy, and the latter increases crop yields. Both effects increase agricultural revenues which create jobs and helps rural communities. Additionally, in response to drought farmers may invest in new wells and increase groundwater pumping which has a cost to both farmers and the environment. Flexible water markets may mitigate this effect by encouraging the sale and transfer of surface water.

The potential for water markets in California is limited by physical and political constraints. Physical constraints include regional connections, conveyance capacity, and existing reservoir operation regulations. Political constraints include legal restrictions on the sale of groundwater and an aversion to water transfers out of the county. Out-of-county water transfers may shift agricultural jobs from the region. For example, farmers may idle land and sell water out of the region which would shift production and jobs out of the county. While this is acceptable from an economic efficiency standpoint it may harm local communities and, as such, may not be politically acceptable.

We evaluate the effect of water markets in the San Joaquin Valley, south of the Sacramento-San Joaquin Delta (the Delta), on groundwater pumping, stress irrigation, and the economy. Policymakers are often interested in estimating the extent of these effects and, in particular, if benefits outweigh the cost of facilitating a water transfer market.

3.2. SWAP Model with water transfers

To simulate a moderate-to-severe drought we reduce water exports from the Delta by 30% which translates into 1350 Mm^3/yr of surface water shortage to regions south of the Delta. Water exports from the Delta include Central Valley Project (CVP1) and State Water Project (SWP) deliveries, which vary by SWAP region as detailed in the data description. We base this scenario on the 2009 drought and environmental pumping restrictions (the Wanger decision) experienced in California. The shortage in 2009 followed dry years in 2007 and 2008 and was estimated to cause the loss of over $350 million in farm revenues, 115,000 ha of land fallowing, and 7500 agricultural jobs (Howitt et al., 2011). It generated significant interest from policymakers and agricultural water markets were frequently discussed.

We link the SWAP model to a hydro-economic network model of California's water infrastructure, incorporate political transfer constraints, and introduce the drought scenario to evaluate the effects of water shortage with and without water markets. We include restrictions to account for political difficulties of water transfers. First, we do not allow the sale of groundwater. This is consistent with regulations adopted by many counties in the Central Valley. Second, we do not allow water transfers out of the county to account for political difficulties from jobs flowing with the water out of the region. The second point is a restrictive assumption and, as such, our estimates represent a lower-bound on the effects of water transfers. We link water supply infrastructure to SWAP by using the hydro-economic network representation of California's water system provided by the CALVIN model (Draper et al., 2003). Fig. 5 illustrates a schematic of the water delivery system in the San Joaquin Valley which includes SWAP regions 10–21C. Fig. 5 includes agricultural demand regions, rivers, dams, other points in the distribution system, and flow volume and direction. We only report flow volumes and label select components of the system to keep the illustration clear.[4] Agricultural demand regions are shown as ovals. Circles and lines indicate various points in the distribution system, including canals, wasteways, dams, and rivers. The lines are both color and style coded to represent ownership by one of four entities including, (i) SWP shown as a red dotted-line, (ii) CVP shown as a purple dashed-line, (iii) intakes shown as a green dashed-dotted-line, and (iv) natural flows shown in solid blue. Arrows denote the direction of flow and relevant maximum flow volumes are reported below labels. For example, the California Aqueduct, managed by the SWP, has a flow capacity of 12,000 Mm^3/yr.

The SWAP water markets model represents transfers that are physically feasible given the existing water network in California. To test the value of an additional infrastructure development, we allow two "wasteways" (Westly and Newman) to be operated in reverse to facilitate transfers from the east side of the Central Valley to the west. The distribution system is also geo-coded so we can estimate distances between regions for potential water transfers. We introduce a transfer cost which is a function of the distance between regions and assumed constant. We assume that import

[4] For further information and schematics, we refer the interested reader to the CALVIN website: http://cee.engr.ucdavis.edu/faculty/lund/CALVIN/.

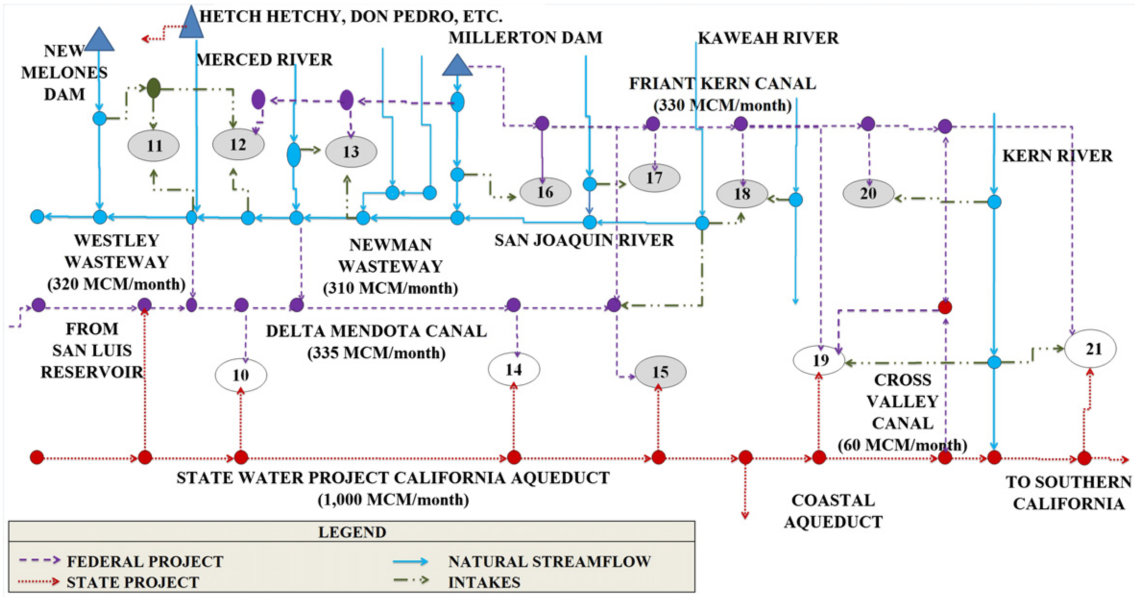


Fig. 5. Water supply and demand network in the San Joaquin Valley of California, flow conveyance capacities in million m^3 per month (MCM/month). Darker ovals represent exporting agricultural regions.

regions pay the transaction cost plus the cost per unit of water transferred. Incorporating these additions to the model, we rewrite the objective function, Equation (26), as

$$\begin{aligned}\max_{x_{gij},wat_{gw}} PS+CS = & \sum_i\left(\xi\alpha_i^1\left(\sum_g y_{gi}\right)+\frac{1}{2}\alpha_i^2\left(\sum_g y_{gi}\right)^2\right) \\ & +\sum_g\sum_i\left(rm_{gi}\left(y_{gi}\right)\right) \\ & -\sum_g\sum_i\left(\delta_{gi}\exp\left(\gamma_{gi}x_{gi,land}\right)\right) \\ & -\sum_g\sum_i\left(\omega_{gi,\text{supply}}x_{gi,\text{supply}}+\omega_{gi,labor}x_{gi,labor}\right) \\ & -\sum_g\sum_w\left(\varpi_{gw}wat_{gw}\right)-\sum_g\sum_w\left(trc\cdot d_{gh}\cdot xwt_{ghw}\right),\end{aligned} \tag{33}$$

where xwt and trc are the amount of water transferred between region g and h and the (constant) transaction cost per million cubic meters per km, respectively. The matrix d_{gh} is the transfer distance between region g and h, as estimated from the geo-coded hydrologic model. Under this specification, the importer pays the cost of the water plus the transaction cost which varies by volume and distance. Additionally, we incorporate water transfer constraints into the model. With the water trade variable xwt_{ghw} the regional water constraint, Equation (5), becomes

$$wat_{gw} \le \text{watcons}_{gw}+\sum_h xwt_{ghw}-\sum_g xwt_{ghw}. \tag{34}$$

We additionally impose the constraint that a region cannot simultaneously import and export a water source to avoid unrealistic arbitrage opportunities,

$$\left(\sum_h xwt_{ghw}\right)\left(\sum_g xwt_{ghw}\right) = 0. \tag{35}$$

Finally, we include physical and political transfer constraints such that only regions that are physically connected and within the same county are able to transfer water. These constraints take the form of a series of equality constraints based on the transfer feasibility matrix shown in Table 1.

In addition to physical and political constraints, the transfer matrix includes identity restrictions such that a region cannot trade with itself. The shaded cells indicate potential for transfers between regions. We define physical conveyance capacity constraints as

$$\sum_w xwt_{ghw} \le cap_{gh}, \tag{36}$$

where cap_{gh} is the maximum water transfers between regions g and h as estimated from the hydro-economic network model. This implicitly assumes water transfers from all sources are through the same facilities. This assumption could be relaxed by imposing additional infrastructure which differentiates between individual water sources.

This analysis holds groundwater pumping constraints fixed at the observed base level. In other words, groundwater pumping may change within regions but it cannot be sold and regions cannot pump in excess of observed capacity (i.e. drill new wells).

We combine the basic policy model constraints, Equations (27) and (29)–(32), and water transfer model constraints, Equations (34)–(36), plus the water transfer restriction matrix and maximize the modified objective function, the sum of consumer and producer surplus, Equation (33). We estimate the benefits of flexible water markets to agriculture in the San Joaquin Valley by comparing model runs with and without water markets under reductions that results in Delta export deliveries being 30% of normal (mean) quantities.

4. Modeling results and discussion

In response to water markets we anticipate changes in production at both intensive and extensive margins. At the extensive margin water markets allow water to flow from lower to higher value uses, thus inducing changes in total irrigated land area and in the crop mix. Regional water use will change as regions buy and sell water. At the intensive margin we expect applied water per unit

Table 1
Water transfer feasibility matrix. Darker boxes indicate institutionally and physically feasible inter-regional water transfers.

	Exports																	
Imports		10	11	12	13	14a	14b	15a	15b	16	17	18	19a	19b	20	21a	21b	21c
	10	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0
	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	12	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	13	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	14a	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
	14b	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
	15a	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
	15b	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
	16	1	0	0	0	1	0	0	0	0		0	0	0	0	0	0	0
	17	1	0	0	0	1	0	0	0		0	0	0	0	0	0	0	0
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	19a	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
	19b	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1
	20	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1
	21a	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1
	21b	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1
	21c	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0

area will change in response to water markets, such as by increased irrigation efficiency. We summarize results in terms of total water transfers, land use change, farm revenue effects, and regional impacts on total employment. We discuss all results by comparison of with and without water markets.

4.1. Water transfers

We model a drought of 30 percent of surface water deliveries through the Delta (1350 Mm^3/yr). Net export regions include 10, 12, 15A, 16, 17, and 21A. These represent regions on the east-side of the Central Valley (see Fig. 2) and regions, such as 10, with priority water rights. Net import regions include 13, 14A, 14B, 15B, 19A, 19B, 20, 21B, and 21C. This is consistent with prior expectations as these regions generally have lower-priority water rights, higher reliance on SWP and CVP deliveries, and limited access to groundwater. West-side regions typically realize higher losses during water shortage. Import regions are concentrated along the west-side of the San Joaquin Valley whereas export regions are on the east-side, where snow runoff and local stream inflows increase available water. Regions 11 and 18 don't trade water due to political (within county) transfer constraints.

Table 2 summarizes total imports and exports by region in response to a 30% reduction in Delta exports. The water supply network in Fig. 5 includes two new conveyance options, not currently used. Since the point of the model was to show the value of expanding the conveyance system, we cannot compare the model results with actual transfers, due to the difference in conveyance options. The results show that a total of 582 Mm^3/yr of water could be transferred between regions during drought, corresponding to just over 40 percent of the total amount of shortage. The largest importer, 303 Mm^3/yr, is region 14A which is located on the west-side of the San Joaquin Valley. This region is Westlands Water District which relies heavily on CVP exports and is consequently one of the most affected regions during drought. The largest exporter is region 10, 271 Mm^3/yr, which is in the northern portion of the west-side of the San Joaquin Valley. This region is largely Settlement and Exchange Contract water users (CVPS) which have higher priority and are rarely shorted during droughts. As such, region 10 typically fares well during drought and this is reflected in SWAP model results.

The level of detail in the SWAP water supply data allows us to estimate individual transfers by water source between regions. For example, region 10 exports 35 Mm^3/yr of CVP1 and 236 Mm^3/yr of CVPS to region 14A during the drought. Thus all of the 271 Mm^3/yr region 10 exports flows into region 14A and the majority of this transfer is from settlement and exchange water. Table 3 shows another example of water transfers between regions in Kern County (19Aa–21C) from local surface water supplies (LOC). We estimate water transfers of 160 Mm^3/yr between regions in Kern County, of which nearly 60 percent (90 Mm^3/yr) comes from local surface water supplies. The largest transfer is from 21A (Central Kern) to 19A (West Kern), 53 Mm^3/yr. From Fig. 5 we can see that this transfer is through the Cross-Valley Canal which has capacity of 60 million cubic meters per month, which exceeds the capacity needed for this trade. In general, Kern County transfers are through the Cross-Valley Canal, Friant Kern Canal, or the Kern River.

The SWAP model linked to the hydrologic network allows us to estimate transfers between regions and surface water sources in the San Joaquin Valley. Next, we evaluate the effect of drought with and without water markets for production, revenue, and employment across regions in the San Joaquin Valley.

Table 2
Estimated water transfers between regions during drought (in Mm^3/yr).

Region	Total imports	Total exports	Net trade
10	72.1	271.8	−199.8
11	0.0	0.0	0.0
12	0.0	0.7	−0.7
13	0.7	0.0	0.7
14A	303.1	0.0	303.1
14B	34.3	0.0	34.3
15A	0.0	46.8	−46.8
15B	12.5	0.0	12.5
16	0.0	86.3	−86.3
17	0.0	17.0	−17.0
18	0.0	0.0	0.0
19a	57.0	0.0	57.0
19B	7.2	0.0	7.2
20	16.4	7.6	8.9
21A	0.0	117.8	−117.8
21B	45.5	16.4	29.1
21C	33.6	17.9	15.8

Table 3
Local surface water transfers in Kern County (in Mm^3/yr).

	Exports		
Imports	Region	20	21A
	19A	3.78	53.19
	21C	3.78	29.87

4.2. Change in land use

In response to drought farmers may shift crop mix and increase land fallowing. The former action results from directing water to lower water use and/or higher value crops and the latter from the decision to devote scarce water to existing crops. For example, farmers may decide allocate scarce water to perennial crops, which would be permanently damaged by shortage, by fallowing fields in annual crops (Marques et al., 2005). We also anticipate more intensive water management on existing fields which our model captures through maximum 15 percent deficit irrigation. The CES production functions capture the corresponding yield effects.

Table 4 summarizes the change in total irrigated hectares by region under the base (no drought) case, drought without water markets, and drought with water markets. In the third column we show the average revenue per unit area by region highlighted to show exporting and importing regions. Interestingly, the average revenues range in the both importing and exporting regions are similar, emphasizing that it is the crop specific marginal revenues that determine exporting and importing regions. In the fourth and fifth columns we summarize the change in total irrigated hectares due to the existence of water markets. Without water markets 87,000 ha are fallowed during the drought, representing just over 4.5 percent of all irrigated land. Allowing for politically feasible (limited) markets prevents 14,000 ha of land fallowing or, in other words, decreases fallowing due to drought by over 16 percent. Finally, some regions increase land fallowing during drought due to the ability to export water. As discussed below, these regions choose to fallow lower-value land and export water out of the region. Table 5 summarizes the total change in irrigated hectares by crop. Fallowing increases for pasture, corn, and rice (region 10) as farmers sell the water to other regions where it is applied to higher value crops.

4.3. Change in farm revenue

The change in farm revenues shows the aggregate effect of a drought water market. Table 6 summarizes the change in farm revenues under base conditions, drought without water markets, and drought with water markets. As with total irrigated land, revenues fall significantly under drought. Without water markets, total farm revenues decrease by over $355 million across the region, or 2.8% of total farm revenues. This includes the effect of increased land fallowing and a shift in crop mix to reflect the increased water scarcity due to drought. If water markets are allowed, farmers can reduce total losses by $104 million in farm revenues across the region. Thus, water markets decrease aggregate farm revenue losses by approximately 30 percent.

Water markets smooth aggregate and regional losses due to drought. We can see these effects in the region-specific revenue changes in Table 6. Farm revenues increase by 18, 15, and 5 percent in regions 14A, 19A, and 15B, respectively. However, revenues fall by 16 percent in region 21A due to water transfers out of the region.

Changes in agricultural revenues will affect other sectors of the economy. These effects are typically modeled with Input-Output (“multiplier”) models, which take SWAP model results and estimate changes in related sectors of the economy. Multiplier models capture a snapshot of a region’s economy and the interrelations that exist among sectors and institutions. These models estimate direct, indirect, and induced effects for relevant sectors of the

Table 5
Irrigated crop area (in hectares) with and without drought and water markets.

Crop	Base land use	Drought without markets	Drought with markets	Additional land due to markets	Percent change (%)
Alfalfa (Lucerne)	210,413	195,961	201,095	5134	2.62
Almonds/pistachios	265,144	263,917	264,252	334	0.13
Corn	205,381	186,221	176,984	−9237	−4.96
Cotton	267,843	245,496	260,690	15,195	6.19
Cucurbits	23,222	22,471	22,834	364	1.62
Dry beans	11,958	9464	11,101	1637	17.30
Fresh tomatoes	10,712	10,666	10,669	3	0.03
Grain	86,228	81,329	83,901	2572	3.16
Onions/garlic	17,407	17,288	17,352	64	0.37
Other deciduous	119,891	119,398	119,455	57	0.05
Other field	150,131	142,744	144,522	1778	1.25
Other truck	72,270	70,957	71,495	538	0.76
Pasture	59,402	53,810	49,196	−4614	−8.57
Potato	9475	9425	9441	16	0.17
Processing tomatoes	70,010	66,625	68,632	2007	3.01
Rice	5153	4752	2555	−2197	−46.24
Safflower	2901	1910	2295	385	20.16
Sugar beet	8482	7913	8337	424	5.36
Subtropical	89,160	88,474	88,486	12	0.01
Vines	184,264	183,406	183,495	90	0.05
Total	1,869,446	1,782,225	1,796,787	14,562	0.82

Table 4
Irrigated crop area with and without drought and water markets (in hectares).

Region	Base land use	Revenue per ha ($/ha)	Drought no markets	Drought with markets	Additional land with markets	Percent change (%)
10	173,557	4367	170,918	162,225	−8692	−5.09
11	97,705	2992	97,691	97,666	−26	−0.03
12	103,714	2853	103,698	103,608	−90	−0.09
13	229,417	4140	226,056	225,957	−98	−0.04
14A	196,046	4321	161,618	190,980	29,362	18.17
14B	15,405	2647	15,390	15,420	30	0.19
15A	254,698	3574	249,822	248,026	−1796	−0.72
15B	7717	3312	7239	7630	392	5.41
16	62,035	5776	62,073	61,630	−444	−0.71
17	106,113	6110	106,283	105,997	−286	−0.27
18	291,202	4776	282,799	282,754	−45	−0.02
19A	34,113	3621	28,419	32,701	4282	15.07
19B	66,417	4353	56,189	56,793	604	1.08
20	84,236	5524	79,262	79,658	397	0.5
21A	78,141	4288	72,151	60,300	−11,851	−16.43
21B	41,478	6534	37,184	39,130	1946	5.23
21C	27,453	6112	25,434	26,312	877	3.45
Total	1,869,446	4370	1,782,225	1,796,787	14,562	0.82

Table 6
Farm revenues with and without drought and water markets (in $1000 2008).

Region	Base revenues	Drought with markets	Drought without markets	Revenue change due to markets	Percent change
10	757,990	657,180	642,610	14,560	2.27
11	292,360	356,050	357,320	−1260	−0.35
12	295,850	394,640	400,380	−5730	−1.43
13	949,900	1,048,370	1,052,320	−3940	−0.37
14A	847,030	781,010	721,260	59,750	8.28
14B	40,770	46,040	48,530	−2490	−5.13
15A	910,340	693,140	707,820	−14,680	−2.07
15B	25,560	55,920	40,200	15,720	39.10
16	358,320	385,120	385,120	0	0.00
17	648,320	547,070	547,070	0	0.00
18	1,390,780	1,223,140	1,223,140	0	0.00
19A	123,540	223,650	168,160	55,490	33.00
19B	289,140	189,540	189,540	0	0.00
20	465,320	565,290	565,290	0	0.00
21A	335,080	168,160	186,660	−18,500	−9.91
21B	271,000	285,370	277,100	8270	2.98
21C	167,800	146,650	144,280	2370	1.64
Total	8,169,190	7,766,420	7,656,870	109,550	1.43

Table 7
Total agricultural jobs change due to water markets.

Region	Additional jobs
10	332
11	−29
12	−131
13	−90
14A	1362
14B	−57
15A	−335
15B	358
16	0
17	0
18	0
19A	1265
19B	0
20	0
21A	−422
21B	189
21C	54
Total	2498

economy. Typical results include changes in sector output, employment, value added, and tax revenues due to changes in crop revenues. We follow the methodology of Howitt et al. (2011) and estimate that water markets would save 2500 total jobs. Table 7 summarizes the change in total jobs by region due to water markets.

4.4. Summary of water market effects

Water markets reduce the localized effects of drought in the San Joaquin Valley, and in particular significantly reduce effects in regions heavily reliant on Delta exports. Of course, the water comes from other regions which must reduce hectares, revenues, and employment. However, regional shifts in employment are within the counties due to political policy constraints added to the model. Water markets allow transfers which preserve county economies and reduce the local and regional effects of water shortage in the San Joaquin Valley. If we allow out-of-county water transfers, revenue losses decrease by an additional $39 million (15 percent), which translates into 890 additional total jobs.

Our analysis with SWAP shows that one way to dampen the effects of drought on California agriculture is the use of water markets for regions south of the Delta. Under unrestricted trading, SWAP results indicate that water markets could reduce total fallowing by 16 percent, total farm revenue losses by 30 percent, and total job losses by 28 percent. In general, if water could be transferred among regions, most regions on the west side of the Central Valley are willing buyers of water while some eastern regions are willing to sell. Even moderate transfers of water between regions within the same county significantly reduce economic drought impacts.

This example highlights the wide range of policy simulations available from calibrated models like SWAP. We linked SWAP with a hydrologic model of the San Joaquin Valley, added a water transfer variable and corresponding constraints, and were able to estimate the effects of drought and water markets.

4.5. Further model development and limitations

Usually, we want to perform sensitivity analysis after checking the results from the policy model. Sensitivity analysis normally focuses on key parameters in the model defined by the analyst. For example, if exogenous yield growth due to technological innovations is incorporated in the model, it may be important to assess the size of the effect. Other important variables for sensitivity analysis include crop prices, groundwater availability, and water costs. Fully-calibrated optimization models like SWAP are well-suited for sensitivity analysis which will be determined by the specific research project.

Extensions and other improvements to SWAP include enriching the dataset of coastal regions and the Colorado River. Future versions of the model will include disaggregate estimation of changes in yields and shifts in future demands that incorporate results from research in-progress. Production cost information is also continuously updated in the SWAP database. Inputs, in addition to fertilizer and other supplies, are being added. Disaggregate inputs to the production function allow for a more accurate representation of the response of farmers to external shocks in policy simulations.

Limitations of the SWAP model and its applications have been discussed elsewhere (Medellin-Azuara et al., 2012). The most important limitations are due to data availability, for example disaggregate input data. One area not explicitly addressed in SWAP is uncertainty in the calibrated model from hydrological conditions. This uncertainty in water supply availability is inherent to the hydrological simulation or hydro-economic optimization models that are used in the calibration stage of SWAP. However, applications of SWAP often quantify the economic effects of water availability as a policy outcome. Uncertainty in SWAP parameters, including crop prices and production costs, is addressed by running sensitivity analyses based on the model application at hand.

Groundwater is an alternative source to augment local surface water supplies and SWP and CVP delivery in many regions. The cost and availability of groundwater therefore has an important effect on how SWAP responds to water shortage. Changes in hydrologic head over time and in response to short run drought are important inputs to the model. However, SWAP is not a groundwater model and does not include any direct way to adjust pumping lifts and unit pumping cost in response to long-run changes in pumping quantities. Economic analysis using SWAP must rely on an accompanying groundwater analysis (such as CVHM) or at least on careful specification of groundwater assumptions.

5. Conclusions

Several conclusions arise from the SWAP calibration and modeling framework and the application presented in this paper. Calibrated programming models such as SWAP provide useful policy insights and a framework to easily accommodate changing market conditions, improved datasets, and increased regional coverage. Such models provide a versatile tool for regional water management and policy as well as a framework for integrating many aspects of regional water and agricultural management. Models like SWAP can be easily linked to agronomic, hydrologic, and other biophysical models which provides the researcher with a rich and flexible modeling framework. We also demonstrated that model output can be linked to multiplier models in order to estimate effects in related sectors of the economy.

From a policy perspective, we used the SWAP model framework to show revenues losses during drought may be significantly reduced through more flexible water allocations and better markets. However in practice, infrastructure and institutional limitations often prevent some economically worthwhile water exchanges. Results from this work can help policymakers by highlighting worthwhile opportunities for water transfers across the state and the associated opportunity costs of these transfers.

The stepwise systematic calibration procedure outlined in the paper has diagnostic check criteria calculated at each stage. This approach enables a sequential and focused approach to diagnosis of problems in model calibration or policy response. The empirical example of the drought water markets shows that the disaggregation scale of SWAP is sufficient to meaningfully interact with detailed water distribution networks. In this sense detailed calibrated economic models such as SWAP can be useful in the management of natural resources, and the economic and environmental tradeoffs that this entails.

Acknowledgments

The authors acknowledge the research support form Ray Hoagland, Farhad Farnan and Tom Hawkins; and appreciate the valuable and thoughtful feedback from Steve Hatchett in developing and improving some of the applications presented in this paper.

References

Beattie, B.R., Taylor, C.R., 1985. The Economics of Production. Krieger Publishing Company.

Blocken, B., Gualtieri, C., 2012. Ten iterative steps for model development and evaluation applied to Computational Fluid Dynamics for Environmental Fluid Mechanics. Environmental Modelling and Software 33 (0), 1–22.

Booker, J.F., Howitt, R.E., Michelsen, A.M., Young, R.A., 2012. Economics and the modeling of water resources and policies. Natural Resource Modeling Journal 25 (1).

Braat, L.C., vanLierop, W.F.J., 1987. Integrated economic-ecological modeling. In: Braat, L.C., vanLierop (Eds.), Integrated Economic Ecological Modeling. Elsevier, Amsterdam.

Buysse, J., Huylenbroeck, G.V., Lauwers, L., 2007. Normative, positive and econometric mathematical programming as tools for incorporation of multifunctionality in agricultural policy modeling. Agriculture Ecosystems and Environment 120, 70–81.

Cai, X., 2008. Implementation of holistic water resources-economic optimization models for river basin management – reflective experiences. Environmental Modeling and Software 23, 2–18.

Draper, A.J., Jenkins, M.W., Kirby, K.W., Lund, J.R., Howitt, R.E., 2003. Economic-engineering optimization for California water management. Journal of Water Resources Planning and Management 129, 155–164.

DWR, California Department of Water Resources, Land and Water Use Data. Available at: http://www.water.ca.gov/landwateruse/ (Last accessed December 2010).

DWR, California Department of Water Resources, 2009. California Water Plan Update, 2009. Bulletin 160-09. Sacramento, California.

Gardner, B., 1983. Water pricing and rent seeking in California Agriculture. In: Anderson, T.L. (Ed.), Water Rights. Ballinger Publishing Company, Cambridge, MA.

Gomann, H., Kreins, P., Kunkel, R., Wedland, F., 2005. Model based impact analysis of policy options aiming at reducing diffuse pollution by agriculture – a case study for the River Ems and a sub-catchment of the Rhine. Environmental Modeling and Software 20, 261–271.

Green, R., Howitt, R.E., Russo, C., 2006. Estimation of Supply and Demand Elasticities of California Commodities. Working Paper. Department of Agricultural and Resource Economics, University of California, Davis, California.

Griffin, R.C., 2006. Water Resource Economics: the Analysis of Scarcity, Policies, and Projects. MIT Press, Cambridge, Mass.; London, England.

Hanak, E., 2003. Who Should be Allowed to Sell Water in California? Third-party Issues and the Water Market. Public Policy Institute of California.

Harou, J.J., Pulido-Velazquez, M., Rosenberg, D.E., Medellin-Azuara, J., Lund, J.R., Howitt, R.E., 2009. Hydro-economic models: concepts, design, applications, and future prospects. Journal of Hydrology 375 (3–4), 627–643.

Hazell, P.B.R., Norton, R.D., 1986. Mathematical Programming for Economic Analysis in Agriculture. MacMillan Publishing, New York.

Heckelei, T., Britz, W., 2005. Models Based on Positive Mathematical Programming: State of the Art and Further Extensions. EAAE Seminar Paper, February 2005, Parma.

Heckelei, T., Wolff, H., 2003. Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximum entropy. European Review of Agricultural Economics 30, 27–50.

Howitt, R.E., 1995a. Positive mathematical-programming. American Journal of Agricultural Economics 77, 329–342.

Howitt, Richard E., 1995b. A calibration method for agricultural economic production models. Journal of Agricultural Economics 46 (2), 147–159.

Howitt, R.E., Medellín-Azuara, J., 2008. Un modelo regional agrícola de equilibrio parcial: el caso de la Cuenca del Río Bravo. In: Guerrero-GarcíaRojas, H.G., Yúnez Naude, A., Medellín-Azuara, J. (Eds.), El agua en México: Implicaciones de las políticas de intervención en el sector. El Fondo de Cultura Económica, México, D.F.

Howitt, R.E., Kaplan, J., Larson, D., MacEwan, D., Medellin-Azuara, J., Horner, G., Lee, N.S., 2009. Central Valley Salinity Report. Central Valley Water Quality Control Board.

Howitt, R.E., MacEwan, D., Medellin-Azuara, J., 2011. Drought, jobs, and controversy: revisiting 2009. University of California Giannini Foundation of Agricultural Economics ARE Update 14 (6), 1–4.

Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. Environmental Modeling and Software 21 (5), 602–614.

Kanellopoulos, A., Berentsen, P., Heckelei, T., van Ittersum, M., Oude Lansink, A., 2010. Assessing the forecasting performance of a generic bio-economic farm model calibrated with two different PMP variants. Journal of Agricultural Economics 61 (2), 274–294.

Lund, J., Hanak, E., Fleenor, W., Howitt, R.E., Mount, J., Moyle, R., 2007. Envisioning Futures for the Sacramento-San Joaquin Delta. Public Policy Institute of California. Available at: http://www.ppic.org (Last accessed June 2011).

Marques, G.F., Lund, J.R., Howitt, R.E., 2005. Modeling irrigated agricultural production and water use decisions under water supply uncertainty. Water Resources Research 41 (8).

Medellin-Azuara, J., Howitt, R.E., Lund, J., Hanak, E., 2008. Economic effects on agriculture of water export salinity south of the Sacramento-San Joaquin Delta. In: Lund, J.R., Hanak, E., Fleenor, W., Bennett, W., Howitt, R.E., Mount, J., Moyle, P. (Eds.), Comparing Futures for the Sacramento-San Joaquin Delta. Public Policy Institute of California, San Francisco, California.

Medellin-Azuara, J., Howitt, R.E., Waller-Barrera, C., Mendoza-Espinosa, L.G., Lund, J.R., Taylor, J.E., 2009. A calibrated agricultural water demand model for three regions in northern Baja California. Agrociencia 43, 83–96.

Medellín-Azuara, J., Harou, J.J., Howitt, R.E., 2010. Estimating economic value of agricultural water under changing conditions and the effects of spatial aggregation. Science of the Total Environment 408, 5639–5648.

Medellin-Azuara, J., Howitt, R.E., MacEwan, D., Lund, J.R., 2012. Estimating impacts of climate related changes to California agriculture. Climatic Change 109 (S1), S387–S405.

Merel, P., Bucaram, S., 2010. Exact calibration of programming models of agricultural supply against exogenous supply elasticities. European Review of Agricultural Economics 37 (3), 395–418.

Merel, P., Simon, L.K., Yi, F., 2011. A fully calibrated generalized constant-elasticity-of-substitution programming model of agricultural supply. American Journal of Agricultural Economics 93 (4), 936–984.

Moore, C.V., Hedges, T.R., 1963. A method for estimating the demand for irrigation water. Agricultural Economics Research 15, 131–135.

Paris, Q., Howitt, R.E., 1998. An analysis of ill-posed production problems using maximum entropy. American Journal of Agricultural Economics 80, 124–138.

Piuleac, C.G., Rodrigo, M.A., Cañizares, P., Curteanu, S., Sáez, C., 2010. Ten steps modeling of electrolysis processes by using neural networks. Environmental Modelling and Software 25 (1), 74–81.

Quinn, N.W.T., Brekke, L.D., Miller, N.L., Heinzer, T., Hidalgo, H., Dracup, J.A., 2004. Model integration for assessing future hydroclimate impacts on water resources, agricultural production and environmental quality in the San Joaquin Basin, California. Environmental Modeling and Software 19, 305–316.

Reclamation, United States Bureau of Reclamation, 1997. Central Valley Production Model, Central Valley Project Improvement Act Draft Programmatic EIS. Technical Appendix Volume 8. Mid-Pacific Region. Sacramento, California.

Reclamation, United States Bureau of Reclamation, 2011. North of Delta Offstream Storage (NODOS) Analysis and SWAP Model Application to National Economic Development (NED). Technical Appendix. Mid-Pacific Region. Sacramento, California.

Scheierling, S.M., Loomis, J.B., Young, R.A., 2006. Irrigation water demand: a meta-analysis of price elasticities. Water Resources Research 42 (1), 9.

Taylor, J.E., Dyer, G.A., Yúnez-Naude, A., 2005. Disaggregated rural economywide models for policy analysis. World Development 33, 1671–1688.

UCCE, University of California Cooperative Extension, Crop Budgets. Available at: http://coststudies.ucdavis.edu (Last accessed February 2011).

UCSB, University of California, Santa Barbara. Water Strategist Transfer Database. Available at: http://www.bren.ucsb.edu/news/water_transfers.htm (Last accessed May 2010).

USDA, United States Department of Agriculture. Available at: http://www.usda.gov (Last accessed June 2011).

vanWalsum, P., Helming, J., Stuyt, L., Schouwenberg, E., Groenendijk, P., 2008. Spatial planning for lowland stream basins using a bioeconomic model. Environmental Modeling and Software 23, 569–578.