Project Number: 88-AM2

THE EFFECTS OF UNCERTAINTY AND ADJUSTMENT COSTS ON INVESTMENT IN THE ALMOND INDUSTRY

by

Jeffrey Dorfman

and

Dale Heien

June 1988

THE EFFECTS OF UNCERTAINTY AND ADJUSTMENT COSTS ON INVESTMENT IN THE ALMOND INDUSTRY

Research on investment analysis peaked in the late sixties concurrent with professional interest in aggregate econometric models. The past few years have witnessed a renewed interest in investment analysis with emphasis on uncertainty, adjustment costs and informational imperfections. Recent analyses dealing with asymmetric information include Bernstein and Nadiri (1986), Fazzari and Athey (1987), Greenwald, Stiglitz and Weiss (1984), Myers and Majluf (1984), Sinai and Eckstein (1983), and Stiglitz and Weiss (1986). Abel (1983), Hartman (1972, 1973, and 1976), Lucas and Prescott (1971), and Pindyck (1982) analyze investment under uncertainty with adjustment costs. Recent empirical studies in this vein are Abel and Blanchard (1986), Craine (1975), Kokkelenberg and Bischoff (1986), Meese (1980), and Pindyck and Rotemberg (1983). Evaluative studies of empirical investment models include Bernanke, Bohn and Reiss (1988), Chirinko and Eisner (1983), Clark (1979) and Uri (1982).

This paper presents a model of investment behavior which incorporates uncertainty and adjustment costs. This formulation is based on maximizing the expected present value of profits. Under the assumptions of quadratic adjustment costs and fixed coefficient technology, the time path and the determinants of investment under uncertainty are derived. The model is then tested on a specific industry, almond production, which has undergone considerable investment over the past fifteen years and for which there is substantial uncertainty. In Section I the theory is developed and related to the existing literature. Section II presents a brief discussion of the almond industry, specifies the model and its error structure, and details the estimation procedure. Section III gives the econometric estimates of the model and a prediction interval test which compares the performance of the model to similar specifications without uncertainty and adjustment costs. Also presented are the short and long run impacts of changes in various investment determinants.

I. A TECHNOLOGY SPECIFIC INVESTMENT MODEL

The neoclassical investment model assumes that firms choose investment levels to maximize the expected present value of the firm. This neoclassical paradigm as implemented by Jorgenson[1], uses the shadow price of capital services, referred to as the user cost of capital, to define the optimal level of capital stock. The use of this shadow price implies a high degree of perfection in the capital markets. In this formulation, the present price and output serve as expectations for future price and output levels. While there is nothing inherently incorrect with this formulation, it does involve specific assumptions about capital markets, expectations, and the production process, i.e. Cobb-Douglas. Maximization of the present value of the firm leads to optimal demand paths for variable factors as well as capital stock. Subsequent generations of this neoclassical formulation have specified more flexible functional forms, added adjustment costs, and made alternative assumptions about the formation of expectations.

Although still neoclassical in nature, the model developed here departs from previous studies in several ways. Previous studies using the neoclassical model have taken output as exogenous. Output is a choice variable along with input levels and is inappropriate on the right hand side of an investment equation. Also, we do not rely on a user cost of capital concept. By assuming a fixed coefficient technology for almonds, the maximization problem can be solved while treating output as an endogenous variable.[2] The exogenous variables in this model are the prices of factor inputs and the output price. Because the technology in almond production is approximately Leontief, the main consideration in this industry is not factor substitution, but rather the decision to change the level of output. In

---

[1] Literature on the neoclassical model is quite large. An excellent survey of the neoclassical, and other investment models, can be found in Jorgenson (1971). A survey of more recent developments in the neoclassical model can be found in Berndt, Morrison and Watkins (1981).

[2] Output can also be determined endogenously through the specification of a profit function, cf. LeBlanc and Hrubovcak (1986). For an alternative consideration of the treatment of output see Pindyck and Rotemberg (1983), page 1068.

order to develop a theoretical model utilizing Leontief technology, we build on Wickens and Greenfield's (1973) study of the coffee market. The model derived here departs from Wickens and Greenfield by introducing uncertainty.

To derive the investment equation we begin by noting that current output can be written as a weighted sum of past investment levels, or

(1) $Q_t = \sum_{i=0}^{n} y_{it} I_{t-i}$

where Q_t is output in t, I_{t-i} is the new investment in almond acreage in year t-i, y_{it} is the yield (output per acre) of i year old trees in year t, and n is the lifespan of the trees. The above formulation implies that the fixed amounts of the other factors of production are also supplied. By assuming that weather affects yields each year and that there was no technical change during the period studied, the y_{it} can be considered random variables with time-constant means, or

(2) $E(y_{it}) = y_i$ for all t.

Firms are assumed to maximize the expected present value of net revenue. Each firm faces two constraints: a production function and a time trajectory for output. The equation for output simply brings to bear the physical realities of almond production. Almond trees do not produce a harvestable crop until the third or fourth year. Yields then increase rapidly until full production is reached, hold steady for about twenty years, and then begin to decline. Because the age profile (vintage) of the trees evolves in a known pattern and because there is a lag between investment and production there is a restriction on changes in expected output. Weather, of course, makes actual output move in a more discontinuous and random manner. Thus, the problem of maximizing the expected present value can be written as

(3a) $\max \Pi = E\{ \sum_{t=0}^{\infty} (1+r)^{-t} [(p_t - c_t) Q_t - C_t - g(I_t)] \}$

(3b) s.t. $Q_t = \sum_{i=0}^{n} y_i I_{t-i}$

(3c) $$\dot{Q}_t = \sum_{i=0}^{n} (y_i - y_{i-1})\, I_{t-i}$$

where p is the price of output, c is cost per acre (described in more detail in Section II), C is fixed costs, r is the discount rate, and g(I) represents the adjustment costs of additions to the capital stock. Prices and costs are not known as of t and are considered to be random variables.

Substituting (3b) into (3a) allows a simplification of the Hamiltonian (H), which now can be written as

(4) $$H_t = E\{\, \pi_t \sum_{i=0}^{n} y_i I_{t-i} - C_t - g(I_t) + \lambda_t \sum_{i=0}^{n} (y_i - y_{i-1})\, I_{t-i} \,\}$$

where $\pi_t = (p_t - c_t)$ and the λ_t's are the costate variables. The necessary conditions for an optimum are

(5) $$\partial H/\partial I_t = E\{\sum_{i=0}^{n} (1+r)^{-i}\, \pi_{t+i}\, y_i - g'(I_t) + \sum_{i=0}^{n} (1+r)^{-i} (y_i - y_{i-1})\, \lambda_{t+i} \} = 0$$

(6) $$\dot{\lambda}_t - r\lambda_t = -\partial H_t/\partial Q_t = -E(\pi_t)$$

(7) $$\dot{Q}_t = \sum_{i=0}^{n} (y_i - y_{i-1})\, I_{t-i}$$

where the dot represents a time rate of change. Investment will change output for a number of future periods, so that (5) is the sum of a set of $\partial H_t/\partial I_t = 0$ conditions.

Assuming that $g'(I_t)$ is known, (5) can be solved for $g'(I_t)$. To further simplify, note that the first sum on the right hand side of (5) is simply the expected present value of one acre of almonds planted in year t. By defining this expected present value of investment as

(8) $$EPVI_t = E\{\sum_{i=0}^{n}(1+r)^{-i}\, \pi_{t+i}\, y_i \},$$

the relation (5), is now

(9) $$g'(I_t) = EPVI_t + E\{\sum_{i=0}^{n} (1+r)^{-i} (y_i - y_{i-1})\, \lambda_{t+i} \}.$$

The λ_t's are the shadow price of a marginal relaxation of the associated constraint. This makes each λ_t the value of an extra change in the change in output. From (6), it is clear that in a steady-state equilibrium λ would equal zero. Therefore, values of λ that differ from zero indicate a system that is not in equilibrium. The system will be away from equilibrium when firms realize that actual values differ from their expectations or when their expectations about the future change. Therefore, we employ a measure of uncertainty as a proxy for the weighted sum of the (unobservable) λ_{t+i}.

The uncertainty measure chosen as a proxy for the second term on the right hand side of (9) is the variance of the EPVI's. This choice is motivated by two considerations. First, as the λ's change, so do the EPVI's, through the relations in (5) and (6). Thus if the λ's are large, indicating that the system is away from equilibrium, the EPVI's should exhibit a large variance. Second, Abel (1983), Hartman (1972, 1973, 1976), and Pindyck (1982) showed that changes in uncertainty cause changes in investment levels. In general, the uncertainty considered was a mean preserving spread in prices. The equivalent here would be a mean preserving spread in the π_t. Variation in π will be closely related to the variance of the EPVI (σ^2).

Given the above discussion, the relation for investment can be written

(10) $\quad g'(I_t) = EPVI_t + \gamma\sigma^2$

where the parameter γ is added since σ^2 will at least differ by a scale factor from the var(π_t) and because it is a proxy for the weighted sum of the λ's. Assuming $g(I_t)$ to be quadratic (i.e., $g(I_t) = bI_t + cI_t^2$), the final equation becomes

(11) $\quad I_t = \beta_0 + \beta_1 EPVI_t + \beta_2\,\sigma^2_t.$

While the variance of the expected present value of investment offers a reasonable measure of uncertainty, it should be cautioned that the coefficient for this variable cannot be interpreted as a risk aversion coefficient in the normal Pratt-Arrow sense. This is due to the multi-period nature of the investment decision. With serially connected payoffs over a finite asset life, conventional results about attitudes toward risk do not necessarily apply {Newberry and Stiglitz (1981)}. We have avoided assumptions regarding maximization of the utility of present value for this reason. While no claims are made here that the investment model deals with risk, it does have an explicit treatment of uncertainty.

Several earlier studies have shown the complexities involved in the response of investment to changes in the level of uncertainty. Hartman (1972, 1973) demonstrated that a mean preserving spread in agents' subjective prior joint distributions of prices can increase investment in risk neutral firms. Because each firm's short-run profit function is convex in prices, the increased uncertainty raises the expected marginal revenue of a given capital stock. Abel (1983) provided explicit solutions by specifying Cobb-Douglas technology and convex adjustment costs. Pindyck (1982) used more general production and adjustment cost functions and showed that an increase (decrease) in future uncertainty leads to a higher optimal path for the capital stock if the marginal cost of adjustment is convex (concave) in investment. Theoretical developments, then, provide a clear basis for including a measure of uncertainty in the investment equation. Theory, though, cannot provide an *a priori* sign for this variable. Nor can the agents' attitude toward risk be inferred from the sign {Pindyck (1982)}. Thus, the empirical results will show agents' investment response to changes in uncertainty, but not the agents' attitudes toward risk.

II. ECONOMETRIC SPECIFICATION AND ESTIMATION

The almond industry was selected for a number of reasons. First, it is possible to obtain detailed microdata on costs, plantings, yields and prices for this industry. Second, there has been a high level of investment in almond orchards over the past twenty years. Third, the technology of growing and harvesting almonds has changed little over this period, obviating the need to deal with technological change. Fourth, production of almonds is characterized by near Leontief technology thus reducing the scope of factor substitution and facilitating the computation of present value. Fifth, the industry is reasonably competitive and is not dominated on the supply side by any single large producer. Sixth, returns in the almond industry are characterized by considerable variation.

This variation occurs mainly because of substantial weather induced fluctuations in yields. Hence, substantial uncertainty exists regarding future levels of profits. Lastly, output is homogeneous which further minimizes aggregation problems.

After planting, four years are required before almond trees reach sufficient maturity to yield a harvestable crop. These acres are then included in bearing acreage. Hence, it is possible to separate the delivery lag from the expectational lag, a problem which Abel and Blanchard (1986) noted for the neoclassical model. The main costs incurred during this period, in addition to acquiring the land, are the cost of the seedling trees, installation of an irrigation system, planting, cultivation, pruning, fertilization and management. In order to bring one acre of new almond trees into production, approximately $7000 must be spent on these depreciable capital items. The model here deals only with new acreage and does not consider the sale of existing orchards as investment. There are slightly over 400,000 acres in almonds in California for a capital stock value of almost three billion dollars.

Large scale almond production is a relatively new phenomenon in the U.S. As a result there are very few orchards approaching twenty years of age, the typical productive life of an almond tree. Hence, removals over the period under consideration, 1970-1985, were minimal.[3] Due to this four year lag, equation (11) can now be presented in its final form

$$(12) \quad NI_{it} = \beta_o + \beta_1 PVI_{it-4} + \beta_2 PVI_{it-5} + \ldots + \beta_5 PVI_{it-8} + \beta_6 VAR_{it} + w_{it}$$

where

$NI_{it} = (A_{it} - A_{it-1}) / A_{it-1}$ = percentage additions to bearing acreage in region i,

PVI_{it-j} - present value of an acre of almonds trees in region i in period t-j,

A_{it} - bearing acreage of almonds in the ith region in time period t,

$VAR_{it} = var(PVI)_{it} = \sigma^2$, and

w_{it} = stochastic error term for region i in time period t.

---

[3] According to estimates made by Bushnell and King (1986), removals are currently about one percent of bearing acreage. We consider this to be minimal in view of the fact that there are more removals presently than earlier in the study period.

Thus, the agents' expectations for the EPVI are represented by a weighted sum of five years' calculated PVI_{it}'s. In effect, the expected present value (EPVI) is a distributed lag of current and past present values (PVI's). This formulation can be considered as a two stage expectations calculation. The calculated PVI's take past prices and costs into account to form expectations for future prices and costs, but are not formed as a result of any moving average or ARMA of past PVI's. However it is also reasonable to assume that agents base expectations of the present values on more than one year's calculation. Hence, we specify the EPVI as a distributed lag in the PVI's. In a manner similar to that used by Clark (1979), investment is divided by capital stock lagged so that both left and right hand side variables are measured on a per acre basis (since the right hand side is measured on a present value per acre basis). Dividing only the left hand side of (12) also avoids problems of spurious correlation first pointed out by Kuh and Meyer (1955). The PVI are defined by

(13) $PVI_{it} = (PV_{it} + D_{it} - L_{it}) / f_t$.

PV is the present value of cash revenues minus costs on one acre of almonds over a twenty year horizon viewed from time period t. Costs include non-adjustment cost investment expenditures as well as production costs. These costs include the cost of trees, irrigation systems, equipment, and other depreciable assets. D is the present value of tax savings due to accelerated depreciation and the investment tax credit. L is the current price of irrigated farmland in region i and serves as a measure of the opportunity cost of the land. At any point in time over the sample period the level of investment will be a function of the level of PVI in nominal terms. However, in order to use the PVI over time in a regression equation, they must be converted to constant dollars. The deflator used for this purpose is f, prices paid by farmers. This index is composed of prices of various farm inputs. Since profits are often re-invested by farmers, this is an appropriate index. As a practical matter, it is highly correlated with the CPI and the GNP deflator. A more precise explanation of how PV and D were computed and the data sources is given below. VAR_{it} is computed by the standard sampling variance formula using the eight most recent PVI's.

For purposes of estimation, California was classified into nine growing regions. Seven of these regions correspond to counties while the remaining two are groups of counties, one in the north and one in the south, whose involvement in almonds is small. The following assumptions were made about the error term, w_{it},

(14a) $E(w_{it}) = 0$

(14b) $E(w_{it}w_{ks}) = 0, \qquad \forall t \neq s$

(14c) $E(w_{it}^2) = \sigma_{ii}, \qquad i = 1,\ldots,9$

(14d) $E(w_{it}w_{kt}) = \sigma_{ik}, \qquad i,k = 1,\ldots,9,\ i \neq k.$

Further spatial and temporal error covariance assumptions were made about the σ_{ik}. First, the assumption was made that there is no autocorrelation between the error terms, even within a region. While this assumption may seem simplistic in a model with lagged variables, it should be noted that all lagged variables are exogenous. Also, a sample autocorrelation coefficient was calculated (r=.232) and was not statistically significant. As a further test, the model was estimated using the appropriate Prais-Winsten transformation which provided no improvement over the model without the correction. Hence, there was no autocorrelation correction in the final version of the model. The errors however, are assumed to have particular non-zero spatial covariances. The nine regions mentioned above are grouped into three super-regions (1, 8; 2, 3, 4, 9; 5, 6, 7; see Table 1 for the numbering key) corresponding to the Sacramento Valley, Northern San Joaquin Valley, and Southern San Joaquin Valley, respectively. On this basis the spatial error assumptions are: each region has its own distinct error variance given by (14c), any two regions in the same super-region have the same error covariance, and any two regions in two different super-regions have the same error covariance. For example, this means that $\sigma_{23} = \sigma_{24}$ and $\sigma_{25} = \sigma_{37}$. These covariance assumptions are based on the spatial nature of the data set. If almonds are profitable in one county, some of the investment thereby encouraged may take place in another county. Therefore, non-zero covariances between regions should be expected and will likely depend on the proximity of the two regions. Such reasoning leads to covariance assumptions of the type made in (14d).

Applying the proposed error structure to the model, estimates of the parameters in the error covariance matrix, $\Omega = E(ww')$, were obtained from an OLS regression of the model. The OLS residuals were used to estimate each σ_{ik} parameter, being careful to correct for the fact that the regional subsets of the residuals do not have zero means. The final estimation of the model was then performed using GLS. The estimate of Ω obtained by this method is consistent due to the consistency of the OLS residuals. Lastly, it should be noted that $\Omega = \Sigma \otimes I$, where Σ is a 9x9 matrix of the σ_{ik} defined above, I is an 8x8 identity matrix and $\otimes$ is the Kronecker delta product. Such an error structure, based here on the regional nature of the problem, is equivalent to the error structure of Zellner's Seemingly Unrelated Regression (SUR). Zellner (1962) showed that such a procedure is not only consistent, but also efficient for estimating systems of equations such as these. Here the "system" is simply the nine equations, each representing one county or county group. Although the variables in each of the equations are identical, the actual values of those variables differ across equations because costs, technology, and prices vary by county. Hence, there is a gain in efficiency captured by exploiting the familiar SUR error structure. The estimates yielded by Zellner's SUR are also unbiased. This fact can be important in the small to medium sized samples which are often used in applied work. Kakwani (1967) showed that if the disturbance term has a continuous symmetric probability density function the Zellner estimators are unbiased. This result is based on the fact that Σ is then an even function of the disturbance term, **w**. Therefore, $\Omega = \Sigma \otimes \mathbf{I}$ is also an even function of **w**. This result extends to the model used here even though the β's do not vary across the nine regions.

Estimation of the Production Function. One of the most important components of the PVI is the yield--the output of one acre of almond trees. To estimate the y_i requires the estimation of a production function for almonds. As indicated earlier, there is substantial variation in almond yields, due mainly to weather, but also due to the fact that almonds are an alternate bearing crop. Almonds have a physiological tendency to alternate yearly between relatively heavy and relatively light crops. Because this pattern is weather related, all the trees are on the same cycle. This phenomenon is difficult to observe due to the much larger, weather induced variations. However, any production function must take these variations into account to provide an accurate forecast of expected yield for the present value computation. Since yield is total output divided by the number of acres, estimation of the yield function is equivalent to estimating the production function.

The production function for almonds was estimated using bearing acreage, rainfall, and dummy variables for location and alternate years as regressors. Bearing acreage is the number of acres of almond trees four years of age and older. Once in place, the input requirements for cultivation, harvesting, pruning, etc. are virtually fixed. Hence, we have assumed Leontief technology for our production function. As a result, the usual economic inputs, e.g., labor and materials, do not appear in the production function. The costs of these inputs do, however, appear in the present value computations discussed below. Rainfall is included because the bloom period for almonds falls during California's rainy season (February and March). If it rains too much during the bloom period the bees cannot pollinate the flowers well and a small crop is the result. A dummy variable was included to account for the alternate bearing pattern.

Using a pooled time-series cross-section data base, production was estimated as a function of region (due to climate, soil type, etc.), of alternate years, and of rainfall in that region. Dummy variables were employed for eight of the nine regions, with Butte County being the base region. A dummy variable that took the values of 0 in even years and 1 in odd years was used to model the alternate bearing phenomenon. The variable for rainfall was inches of rainfall in February squared.[4] The production model can be written

(15) $Q_{it} = \alpha_0 + \alpha_1 A_{it} + A_{it}\Sigma\alpha_k R_{kt} + \delta_1 A_{it}FR_{it} + A_{it}FR_{it}\Sigma\delta_k R_{kt} + \mu_1 A_{it}D_t + v_{it}$

where

Q_{it} = tons of almond kernels produced in region i, in year t

A_{it} = bearing acreage, in thousands of acres for region i, in year t

R_{kt} = dummy variable equal to 1 if i = k, 0 otherwise, k=2, 9

D_t = dummy for alternate bearing, equal to 1 in odd years, 0 in even years

FR_{it} = inches of rainfall in February, squared, in region i, in year t

v_{it} = stochastic error term,

and the sums are over k from 2 to 9.

The production relation was also estimated by GLS. The assumptions about the v_{it} were

(16a) $E(v_{it}) = 0$

(16b) $E(v_{it}v_{ks}) = 0, \quad \forall t \neq s$

(16c) $E(v_{it}v_{kt}) = \phi_{ik}, \quad \forall\ i,k = 1,...,9$

This gives an error structure where $E(vv') = \Phi \otimes I$ where I is a 16x16 identity matrix, Φ a matrix of the ϕ_{ik}. The ϕ_{ik} were estimated from the residuals of the OLS regression of the above production model. No spatial restrictions were imposed.

The data used to perform the estimation were county level data on almond acreage and production in California from 1970 to 1985, taken from the relevant County Agricultural Commissioner's Reports. The data on rainfall were collected from the National Oceanographic and Atmospheric Administration published reports. Within each county a weather station nearest the center of the almond growing area was chosen. The data were organized into nine regions: seven counties (Butte, Fresno, Kern, Madera, Merced, San

[4] Both rainfall and rainfall squared were tried with the latter producing better results. The squared effect was also found superior in other weather related studies of almond production by the authors. The month of February was selected because it is the bloom period for almonds throughout the state. Experimentation with later bloom periods for the northern counties did not improve the results.

Joaquin, and Stanislaus) and two groups of counties that grow fewer almonds, North (Colusa, Contra Costa, Glenn, Solano, Sutter, Tehama, Yolo, and Yuba Counties) and South (Kings, San Luis Obispo, and Tulare Counties). Hence, a total of 144 observations were used to estimate the production function. The estimates of the production function are given in Table 1.

Table 1. Generalized Least Squares Estimates of Almond Production Function

Symbol	Variable	Coefficient	t-ratio
α_0	Intercept	$-.402 \times 10^4$	5.84
α_1	Bearing acreage (A)	.780	17.90
δ_1	February rainfall x A	-.003	3.64
μ_1	Alternate bearing dummy x A	-.076	4.13
	Acreage x Regional Dummies		
α_2	A x Fresno	.076	1.71
α_3	A x Kern	.036	.89
α_4	A x Madera	.075	1.65
α_5	A x Merced	.023	.34
α_6	A x San Joaquin	.090	1.58
α_7	A x Stanislaus	.067	2.56
α_8	A x North Region	-.257	8.37
α_9	A x South Region	-.083	1.72
	Acreage x February Rainfall x Regional Dummies		
δ_2	A x FR x Fresno	-.005	1.29
δ_3	A x FR x Kern	-.011	3.94
δ_4	A x FR x Madera	-.011	2.60
δ_5	A x FR x Merced	-.006	1.21
δ_6	A x FR x San Joaquin	-.015	2.82
δ_7	A x FR x Stanislaus	-.010	4.05
δ_8	A x FR x North Region	.0004	.51
δ_9	A x FR x South Region	.0002	.13

$R^2 = .927$ $\bar{R}^2 = .916$ $s = .984$

Next, we present the details of the estimation of the various investment models used in this paper. These models include: a) the complete EPVI model with uncertainty and adjustment costs, b) the EPVI model in the absence of uncertainty, and c) the EPVI model with no uncertainty and no adjustment costs, or, as it is commonly called, the accelerator. All three specifications are estimated with the same data by the GLS regression technique using the error structure outlined above for investment in the period 1978-85, yielding a sample of 72 observations for each model.

The EPVI Model. The EPVI model with uncertainty and adjustment costs is the one developed in Section I and given by equation (12). All present value computations in the model were done over a twenty year time horizon[5], which is the life of a typical almond tree. Hence,

(17) $$PVI_t = \{\sum_{j=0}^{20}[(p_{jt}y_{jt} - c_{jt}l_{jt})(1-m_{jt}) + d_{jt}](1 + r_{jt})^{-j}\}/f_t$$

where

PVI_t - present value of an acre of almond trees in year t, t=1970-1985

p_{jt} - the expectation in year t for almond price in years t + j,

y_{jt} - the expectation in year t for almond yield in years t + j,

c_{jt} - the expectation in year t of a vector of input prices in years t + j,

l_{jt} - the expectation in year t of a vector of input coefficients in year t + j,

d_{jt} - the expected tax savings due to depreciation and investment tax credit in year t + j,

m_{jt} - the expectation in year t of the marginal tax rate in period t + j,

r_{jt} - the expectation in year t of the discount rate in year t + j,

f_t - prices paid by farmers in year t.

---

[5] While we assume a twenty year planning horizon, almond trees can, of course, be left in the ground longer. However, yields begin to decline substantially after twenty years due mainly to the impact of shaking from mechanical harvesters. Moreover, the effect on present value of an additional year past twenty is quite small.

The PVI's were computed for each year, 1970-85, for each of the nine regions. The expectations assumption used was the following

(18) $p_{jt} = (1+g_{pt})^j \, p_{ot} \qquad j=0,\ldots,20$

where g_{pt} is the rate of growth of expected prices in period t. The estimate of g_{pt} is based on the unweighted average of the past three years rates of growth of prices. Hence, prices are assumed to grow at the same rate as they have averaged over the past three years. Prices, and their associated g_{pt}'s, are different for each of the nine regions. We have deleted regional subscripts here in order to avoid confusion. The same procedure was used to estimate the vector of expected input prices, the c_{jt}'s. They also vary by region and are composed of prices for herbicides, pesticides, skilled and unskilled labor, trees, planting labor, water, fertilizer, bees for pollination, a tractor, harvesting labor and equipment, management, and miscellaneous. As a result of the assumptions of Leontief technology and no technical change, the vector of l_{jt}'s which pertain to the variable inputs, do not change over time, but do vary by region and age of orchard. The values of the expected yield, also by region, are obtained using the production function estimated above. In order to estimate these yields, it was assumed that rainfall in the future would be at its historical mean. The data used for the discount rate is the Federal Land Bank long term rate and is the same for all regions and is assumed to remain constant for any one PVI_t calculation. The tax savings from depreciation, d, was estimated by computing the accelerated depreciation (sum-of-the-years' digits) times the marginal tax rate. This category also includes the value of the investment tax credit on the qualifying expenditures. For example, in orchards the trees qualify for the investment tax credit. The marginal tax rate was computed by taking the average of the marginal tax rates on the current dollar equivalents to incomes of twenty and eighty thousand 1965 dollars. Depreciation varies by region since input costs vary by region. Following Sims (1972) the model is estimated with no constraints on the shape of the lag distribution.

The EPVI model presented above has elements of rational expectations theory. The use of the production model for yield expectations is consistent with the rational expectations hypothesis. In order to fully implement that approach it would be necessary to estimate ARIMA type models for each of the input prices and the output price. While this would not be an insurmountable task, the short annual time series available for such an estimation makes that approach unappealing. Hence, we have chosen to use the moving average growth rate assumption although the errors from this method may be autocorrelated. We also note that the focus of this paper is not to test the rational expectations hypothesis.[6] The model given here also has similarities to Tobin's q-theory of investment. If prices were available over time on almond orchards, it would be possible to compare the replacement costs, which are given above, with these prices to determine if investment should take place.[7]

The EPVI Model Without Uncertainty. If the agents' expectations of the future are correct, then the system should always be in equilibrium, given the adjustment costs that are still present. Thus, while the PVI can still differ from zero due to adjustment costs, the λ_t should all be zero. Therefore, the variance of the PVI drops out of the equation, resulting in:

(19) $NI_{it} = \beta_o + \beta_1 PVI_{it-4} + \beta_2 PVI_{it-5} + \ldots + \beta_5 PVI_{it-8} + w_{it}$

This is essentially the model derived in Wickens and Greenfield (1973).

The Accelerator Investment Model. If agents face no uncertainty and there are no adjustment costs, the EPVI model becomes degenerate. The model yields a 'bang-bang' solution, with optimal investment for a given firm either an infinite positive or negative amount. Any limit on investment would be physical, not economic. For the industry as a whole, the optimal behavior in this situation, given the Leontief technology assumed, is

[6] For a test of the rational expectations hypothesis applied to agriculture see Goodwin and Sheffrin (1982).

[7] For examples of the q-theory approach in applied studies see Engel and Foley (1976), Summers (1981), and von Furstenberg (1977).

represented by the accelerator model. The accelerator model can be considered as a special case of the neoclassical investment model. With Leontief technology instead of Cobb-Douglas, the input demand for the capital services will be

(20) $K^{*} = a_1 Q$

where Q is total output. Since K^{*} represents the desired capital stock (A_{it}), given a similar expectational argument, the model can be written as:

(21) $NI_{it} = \gamma_0 + \gamma_1 \Delta Q_{it-4} + \gamma_2 \Delta Q_{it-5} + \ldots + \gamma_5 \Delta Q_{it-8} + w_{it}.$

The accelerator model can, of course, be derived from alternative considerations.[8]

III. THE EMPIRICAL RESULTS

This section presents the GLS estimates for the three models presented above. It also gives the results of the prediction interval test and the elasticities for the exogenous variables. The estimates of the models are given in Table 2. The EPVI model with uncertainty produces a stable lag structure, all coefficients are significant with the exception of PVI_{t-4}. The coefficient on VAR is significant and positive. Thus, when faced with more uncertainty, agents increase their capital stocks.

The first R^2 in Table 2 is for the GLS residuals and is given by

(22) $R^2{:}GLS = \mathbf{w}'\Omega^{-1}\mathbf{w}/\mathbf{y}'\Delta^{-1}\mathbf{y}$

where Ω is as before, $\Delta = \Sigma_n \otimes A_T$, Σ_n is as before, $A_T = I_T - (1/T)\mathbf{ii}'$, $\mathbf{i}$ is the units vector, n is the number of regions, and T is the number of time periods. The R^2 for the untransformed (UT) data is therefore,

(23) $R^2{:}UT = \mathbf{w}'A\mathbf{w}/\mathbf{y}'A\mathbf{y}$

where $A = I_{nT} - (1/nT)\mathbf{ii}''$. Both of the R^2 in the table are adjusted for degrees of freedom by the standard method. The EPVI model with uncertainty (hereafter, simply EPVI) has the best fit by either R^2 criterion. The lag structure for the no uncertainty model (NUM) is

---

8 Both the accelator and the EPVI model without uncertainty were estimated in the same manner as the EPVI model, i.e. with $(A_{it} - A_{it-1}) / A_{it-1}$ as the dependent variable.

consistent with prior expectations and a little less delayed than the one for the EPVI model. The lag structure for the accelerator model (AIM) is sawtoothed and stands in strong contrast to the reasonably smooth lags found for both the NUM and the EPVI models.

Table 2. Generalized Least-Squares Estimates of the EPVI, NUM, and AIM Models

Variable	EPVI	NUM	AIM
Intercept	.681E-01 (11.22)*	.753E-01 (8.95)	.038E-01 (4.54)
** X_{t-4}	.497E-06 (0.68)	-.545E-08 (0.53)	.149E-05 (1.58)
X_{t-5}	.321E-05 (4.86)	.188E-07 (1.77)	-.208E-06 (0.19)
X_{t-6}	.567E-05 (7.40)	.367E-06 (3.32)	.174E-05 (1.42)
X_{t-7}	.499E-05 (6.81)	.262E-06 (2.47)	-.502E-07 (0.03)
X_{t-8}	.452E-05 (5.44)	.152E-08 (1.48)	.552E-05 (2.67)
Var	.117E-08 (3.82)	-----	-----
R^2:GLS	.664	.306	.081
R^2:UT	.386	.225	.025

EPVI - Expected Present Value Model
NUM - No Uncertainty Model
AIM - Accelerator Investment Model

R^2:GLS - See text (22)
R^2:UT - See text (23)

* The numbers in parenthesis are the t-ratios.
** The X's represent the independent variables for each particular model. Hence, for the EPVI and NUM the X's are given by (17), and for AIM by (21)

Out of sample forecasts were made to test the forecasting ability and the empirical consistency of models. If most of the forecasts fall within one standard deviation of the actual values, then a given model can be deemed empirically consistent, even if its forecasts are not particularly accurate. If out of sample forecasts errors routinely exceed one estimated standard deviation , then the validity of the model must be questioned. Also, the calculated confidence levels that are placed on the model and its parameters must be re-examined.[9]

Predictions for net investment in 1986 were made for each region. The results of these predictions are presented in Table 3. Below each prediction is the t-value for the forecast error (i.e. how many standard deviations the forecast fell from the actual value). For a prediction interval test done in this manner, as low a t-value as possible is desired. The covariance matrix of the forecasts for a given model is calculated as:

(24) $$V = \Sigma + R(X'\Omega^{-1}X)^{-1}R'.$$

Here, Ω is the error covariance matrix for the given model, and remembering the error structure employed, $\Omega = \Sigma \otimes I$, X is the in sample data matrix of regressors, and R is the out of sample data matrix {Johnston (1986)}. The standard deviation of a particular forecast is simply the square root of the appropriate diagonal element of V. For example, the t-value for the forecast error of the prediction for region 3 is

(25) $$t_3 = (a - R\hat{\beta})_3/(V_{33})^{1/2},$$

where a is the vector of actual values for net investment, $R\hat{\beta}$ is the vector of predictions, and subscripts refer to rows and columns of vectors and matrices.

As can be seen from Table 3, there is a definite difference forecasts accuracy, both absolute and in terms of standard deviations. For the EPVI and NUM every forecast is within one standard deviation of the actual value, while the AIM shows five forecasts that fall farther than one standard deviation from the actual value. The reader has surely noticed

[9] For an earlier application of this prediction test to investment models see Clark (1979).

that the standard deviation is model and region specific. Thus, a model with poor explanatory power might show itself to be empirically consistent by such a test simply by having a large forecast variance. This appears to be the case with the NUM, but not with the EPVI. Note that the EPVI model has the highest R^2 of the three, both in sample and out. Further, it should be noted that the covariance matrix of the coefficients appears in the formula for the covariance matrix of the forecasts. Table 2 shows that the EPVI model also had the best overall set of t-statistics, which implies a small covariance matrix for the forecasts.

Table 3. Results of Prediction Interval Test

Region	Actual	EPVI	NUM	AIM
1	867	1788	2088	1841
		(.44)*	(.47)	(.60)
2	-471	948	1360	1937
		(.57)	(.36)	(1.31)
3	1980	3781	4174	9244
		(.23)	(.11)	(1.77)
4	2722	1455	1554	1888
		(.57)	(.17)	(.75)
5	1241	2906	3432	2492
		(.35)	(.37)	(.62)
6	-489	1668	2002	1849
		(.56)	(.70)	(2.77)
7	2206	3751	3616	3278
		(.33)	(.20)	(.32)
8	-1847	996	2492	1620
		(.67)	(.47)	(.88)
9	-577	757	1071	1091
		(.94)	(.40)	(1.13)
R^2 - predict vs. actual		.471	.136	.210
χ^2		2.77	1.44	15.98

*t-values for the prediction errors are in parenthesis.

Finally, a Chi-squared test on the set of nine forecasts from each model was performed. The null hypothesis for this test is that all nine forecasts are equal to the actual values. The critical value for a Chi-squared with 9 degrees of freedom and $\alpha = .05$ is 3.32. Clearly, the model without adjustment costs, the AIM, fails this test. Thus, ignoring adjustment costs leads to a model whose validity must be questioned, while ignoring uncertainty leads to a more valid, but less accurate model.

The remaining task is to examine the impact and long run effects of changes in expectations regarding the various exogenous variables entering the computation of the present value of profits. Because of the dimension specific nature of the problem these impacts are best measured as elasticities. Table 4 presents the elasticities of almond investment with respect to changes in the expected rates of growth of output price, the wage rate, the marginal tax rate, the investment tax credit, the discount rate, and the cost of trees--the main capital input.[10] Remembering that trees planted four years ago show up as bearing acreage in the present year, all lags start in t-4. The lags go on to t-10 because our expectations are based on a three year moving average of the rates of growth in output and factor prices. In the EPVI model the elasticity of investment with respect to Var is .088.

Table 4. Cumulative Effects of Changes in Expectations on Almond Investment: Elasticities Evaluated at 1985

Time Period	Output Price	Wage Rate	Marginal Tax Rate	Investment Tax Credit	Discount Rate	Cost of Trees
t-4	.003	-.002	-.002	.000	-.022	-.000
t-5	.030	-.014	-.021	.001	-.188	-.001
t-6	.096	-.046	-.068	.003	-.606	-.002
t-7	.195	-.094	-.137	.008	-1.224	-.004
t-8	.303	-.146	-.213	.015	-1.901	-.007
t-9	.370	-.179	-.261	.021	-2.325	-.008
t-10	.402	-.194	-.284	.024	-2.527	-.009

[10] All calculations are based on the EPVI model whose coefficients are presented above in Table 2.

VI. CONCLUDING REMARKS

The model presented here implements a specification of the neoclassical notion that firms maximize the expected present value of profits utilizing a more specific assumption regarding technology than is customarily done. The model is applied to investment in almond orchards, an industry which is characterized by substantial price and output uncertainty. Under the assumptions of Leontief technology in almond production and prices that grow at the same rate they grew over the past three years, the expected present value of profit for an acre of almonds trees was computed. The model was then estimated using cross-section data from nine almond growing regions over the sixteen year period from 1970-85. The specification of the error structure recognizes the pooled time series cross-section nature of the data as well as the regional relations in almond production. The estimated model was compared to a model without uncertainty and to one without uncertainty and adjustment costs, both of which were estimated on the same data using the same technique and error structure assumptions. The model with uncertainty and adjustment costs showed a better fit over the historical period, had a lag structure more in line with *a priori* expectations and outperformed the other two models in a prediction interval test. The predictions from the EPVI model were all within one standard error of the actual values and collectively passed a Chi-squared test. Comparison of the two alternative models shows that both adjustment costs and uncertainty are important in explaining investment in this industry.

REFERENCES

Abel, Andrew B., "Optimal Investment Under Uncertainty," *American Economic Review* 73 (March 1983) 228-233.

Abel, Andrew B. and Olivier J. Blanchard, "The Present Value of Profits and Cyclical Movements in Investment," *Econometrica* 54 (March 1986) 249-273.

Arrow, Kenneth J., *Risk Perception in Psychology and Economics*, Tech. Report No. 351 (Stanford: Institute for Mathematical Studies in the Social Sciences, Oct. 1981).

Bernanke, Ben, Henning Bohn and Peter C. Reiss, "Alternative Non-Nested Specification Tests of Time Series Investment Models," *Journal of Econometrics* 37 (1988) 293-326.

Berndt, Ernst R., Catherine J. Morrison and G. Campbell Watkins, "Dynamic Models of Energy Demand: An Assessment and Comparison," in E. Berndt and B. Field eds. *Modeling and Measuring Natural Resource Substitution* (MIT Press:1981) 259-289.

Bernstein, Jeffrey I. and M. Ishaq Nadiri, "Financing and Investment in Plant and Equipment and Research and Development," in M. Peston and R. Quandt eds. *Prices, Competition and Equilibrium,* (Philip Allan/Barnes & Noble: 1986) 213-248.

Bushnell, Peter G. and Gordon A. King, *The Domestic and Export Markets for California Almonds*. Giannini Research Report No. 334 (University of California, Sept. 1986).

Chirinko, Robert S. and Robert Eisner, "Tax Policy and Investment in Major U.S. Macroeconomic Econometric Models," *Journal of Public Economics* 20 (1983) 139-166.

Clark, Peter K., "Investment in the 1970's: Theory, Performance, and Prediction," *Brookings Papers on Economic Activity* (1979) 73-113.

Craine, Roger, "Investment, Adjustment Costs, and Uncertainty," *International Economic Review* 16 (October 1975) 648-661.

Engle, Robert F. and Duncan K. Foley, "An Asset Price Model of Aggregate Investment," *International Economic Review* 16 (October 1975) 625-647.

Fazzari, Steven M. and Michael J. Athey, "Asymmetric Information, Financing Constraints, and Investment," *Review of Economics and Statistics* 64 (August 1987) 481-487.

Goodwin, Thomas H. and Steven M. Sheffrin, "Testing the Rational Expectations Hypothesis in an Agricultural Market," *Review of Economics and Statistics* 64 (November 1982) 658-667.

Greenwald, Bruce, Joseph Stiglitz, and Andrew Weiss, "Informational Imperfections in the Capital Market and Macroeconomic Fluctuations," *American Economic Review* 74 (May 1984) 194-199.

Hartman, Richard, "The Effects of Price and Cost Uncertainty on Investment," *Journal of Economic Theory* 5 (Oct. 1972) 258-266.

Hartman, Richard, "Adjustment Costs, Price and Wage Uncertainty, and Investment," *Review of Economic Studies* 40 (Apr. 1973) 259-267.

Hartman, Richard, "Factor Demand with Output Price Uncertainty," *American Economic Review* 66 (Sept. 1976) 675-681.

Intriligator, Michael D., *Mathematical Optimization and Economic Theory* (Englewood Cliffs: Prentice-Hall, 1971).

Johnston, J., *Econometric Methods*. (New York: McGraw-Hill:1984).

Jorgenson, Dale W., "Econometric Studies of Investment Behavior: A Survey," *Journal of Economic Literature* 9 (Dec. 1971) 1111-47.

Kakwani, N.C., "The Unbiasedness of Zellner's Seemingly Unrelated Regression Equations Estimators," *Journal of the American Statistical Association* 62 (March 1967) 141-142.

Kokkelenberg, Edward C. and Charles C. Bishoff, "Expectations and Factor Demands," *Review of Economics and Statistics* 63 (August 1986) 423-431.

Kuh, Edwin and James Meyer, "Correlation and Regression When the Data Are Ratios," *Econometrica* 23 (1955) 400-416.

LeBlanc, Michael, and James Hrubovcak, "The Effects of Tax Policy on Aggregate Agricultural Investment," *American Journal of Agricultural Economics* 68 (November 1986) 767-777.

Lucas, Robert E. and Edward Prescott, "Investment Under Uncertainty," *Econometrica* 39, (September 1971) 659-681.

Meese, Richard., "Dynamic Factor Demand Schedules for Labor and Capital Under Rational Expectations," *Journal of Econometrics* 14 (1980) 141-158.

Myers, Stewart C. and Nicholas S. Majluf, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics* 13 (1984) 187-221.

Newbery, David M. G. and Joseph E. Stiglitz.,*The Theory of Commodity Price Stabilization: A Study in the Economics of Risk*. (Oxford: Clarendon Press, 1981).

Pindyck, Robert S., "Adjustment Costs, Uncertainty, and the Behavior of the Firm," *American Economic Review* 72 (June 1982) 415-427.

Pindyck, Robert S. and Julio Rotemberg, "Dynamic Factor Demands and the Effects of Energy Price Shocks," *American Economic Review* 73 (December 1983) 1066-1079.

Pratt, J., "Risk Aversion in the Small and in the Large," *Econometrica* 32 (Jan.-Apr. 1964) 122-136.

Sims, Christopher A., "The Role of Approximate Prior Restrictions in Distributed Lag Estimations," *Journal of the American Statistical Association* 67 (March 1972) 169-75.

Sinai, Allen and Otto Eckstein, "Tax Policy and Fixed Business Investment Revisited," *Journal of Economic Behavior and Organization* 4 (1983) 131-162.

Stiglitz, Joseph E. and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review* 71 (June 1971) 393-410.

Summers, Lawrence H., "Taxation and Corporate Investment: A q-Theory Approach," *Brookings Papers on Economic Activity* (1981) 67-127.

University of California, Cooperative Extension. Various reports on almond production costs and orchard establishment costs, unpublished, University of California-Davis, 1981-1986.

Uri, Noel P., "Testing for the Stability of the Investment Function," *Review of Economics and Statistics* 64 (February 1982) 117-125.

von Furstenberg, George M., "Corporate Investment: Does Market Valuation Matter in the Aggregate?" *Brookings Papers on Economic Activity* (1977) 347-408.

Wickens, M.R. and J.N. Greenfield, "The Econometrics of Agricultural Supply: An Application to the World Coffee Market," *The Review of Economics and Statistics* 55 (1973) 433-40.

Zellner, Arnold, "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregate Bias," *Journal of the American Statistical Association* 57 (1962) 348-68.

INVESTMENT UNDER UNCERTAINTY:

AN APPLICATION TO THE ALMOND INDUSTRY

by

Jeffrey Dorfman

and

Dale Heien

January 15, 1987

INVESTMENT UNDER UNCERTAINTY:
AN APPLICATION TO THE ALMOND INDUSTRY

Empirical research on investment analysis peaked in the late sixties concurrent with professional interest in econometric models of aggregate demand and income determination. This research was motivated by the notion that investment plays a key role in the determination of aggregate demand. However, in the intervening years interest in macroeconometric models *per se* has diminished. Attention has focused on the implications for these models of various behavioral postulates, mainly rational expectations. The past few years have witnessed a renewed interest in investment analysis, theoretically and empirically, from the point of view of uncertainty, adjustment costs, and informational imperfections. Recent investment analysis dealing with asymmetric information includes Bernstein and Nadiri (1986), Fazzari and Athey (1987), Greenwald, Stiglitz and Weiss (1984), Myers and Majluf (1984), Sinai and Eckstein (1983), and Stiglitz and Weiss (1986). Works by Abel (1983), Hartman (1976), Lucas and Prescott (1971), Pindyck (1982) analyse theoretical issues of investment under uncertainty with adjustment costs. Recent empirical studies in this vein are Abel and Blanchard (1986), Craine (1975), Kokkelenberg and Bischoff (1986), Meese (1980), and Pindyck and Rotemberg (1983). Evaluative studies of empirical investment models include Chirinko and Eisner (1983), Clark (1979) and Uri (1982).

This paper contributes to this renewed interest by presenting a model of investment behavior which emphasizes risk and uncertainty. This model is tested on a specific industry, almond production, which has experienced considerable investment over the past fifteen years. The plan of the paper is as follows. In part I, the model is developed and related to the existing literature, noting several features which give it added realism. Next a brief discussion of the almond industry is presented along with reasons why this industry is well suited to test the model. Part III presents the specification of the model, its error structure, the estimation procedure, and two traditional models for comparison. Also included in this section are the estimates of the production function used to determine yield expectations. Section IV gives a more detailed specification of the model. The last section presents the econometric estimates of the model, a prediction interval test, and the short and long run impacts of changes in various exogenous instruments.

I. INVESTMENT UNDER UNCERTAINTY

The notion that firms undertake investment with an eye to future returns is certainly not new. The neoclassical model assumes that firms choose investment levels to maximize the present value of the firm. This neoclassical paradigm as implemented by Jorgenson[1], uses the shadow price of capital services, referred to as the user cost of capital, to define the optimal level of capital stock. The use of this shadow price implies a high degree of perfection in the capital markets. Also, in this formulation, the present price and output levels serve as expectations for future price and output levels. While there is nothing incorrect with this formulation, it does involve specific assumptions about capital markets, expectations formation, and the production process-in this case Cobb-Douglas. Other, perhaps equally plausible assumptions, would produce alternative implementations of the neoclassical paradigm. Rather than follow that approach, a model is developed here which emphasizes the uncertain and risky nature of investment decisions. In doing so, we do not explicitly define an optimum level or path of capital, but rely on the notion of maximizing the present value of the firm. Investment decisions for which this model is applicable can be characterized by four attributes.

First, and perhaps most importantly, there must be substantial price and output uncertainty. Uncertainty, of course, characterizes decisions in agriculture as it does in many other sectors. Agents possess expectations with respect to output price and factor prices

---

[1] Literature on the neoclassical model is quite large. An excellent survey of the neoclassical, and other investment models, can be found in Jorgenson (1971).

and have knowledge of the production process. Hence, output is treated as a fully endogenous variable and as such does not appear in the empirical investment relation. Second, the investment process must be considered at a very detailed level. This is necessary so that the present value of future profits can be identified and computed. The notion of highly detailed investment demand is contrary to many previous studies which have as their main purpose the determination of large portions of aggregate demand. The purpose here is to test a specific investment theory and not to provide an integral part of a macroeconometric model. Also, this level of detail reduces aggregation problems for the empirical work. Given these conditions, it is possible to compute the expected value of a marginal unit of investment (e.g., a square foot of commercial office space, a particular new machine, an acre of farmland, etc.).[2] Third, although not a strict requirement, the computational burden is eased if the production process under consideration is characterized by near zero elasticities of substitution between factors. However, if the structure of production is known or is capable of being estimated, then, in principle, non-zero elasticities of substitution should present no problem. Fourth, the investment decision under consideration should be capable of being undertaken in fairly small units by many potentially competing firms. This makes it possible to compute the marginal efficiency of capital.

The model to be tested here assumes that the amount of investment undertaken is a function of the expected utility of profits. The expected utility of profits is represented by expected profitability plus a risk factor. Expected profitability is the expectation of the present value of future revenues minus costs, including investment costs and tax benefits, discounted over the anticipated (finite) life of the asset. The measure to represent risk is the variance of the expected profitability which is based on the historical experience for that particular asset. It should be noted that our model says that the level of investment in a

[2] It is of course possible to compute it other ways. See Abel and Blanchard (1986) for example.

particular asset or capital stock is a function of the expected utility of profit from a unit of investment. This implies that as the expected utility becomes larger, more investors will be drawn into the industry and/or those investing will invest larger amounts. This can lead to some questions as to the existence of a finite solution to the investment problem. We refer to this as the problem of closure and discuss it more fully below in reference to the specific industry we test.

The present value of expected profits is given by

(1) $\pi = (\theta + \delta - \tau)$

where θ is the present value of the expected revenue derived minus the expected variable costs of producing it including the present value of all expected expenditures incurred for the investment project itself, δ is the present value of expected tax savings due to accelerated depreciation and the investment tax credit, and τ is the expected opportunity cost of any asset acquired to enable the investment to proceed minus the discounted expected capital gains from the sale of the asset at the end of the horizon. Ignoring risk for the moment, additions to capital stock occur if $\pi > 0$. Reductions, or disinvestment, occur if $\pi < 0$. It is important to note this distinction since it implies that replacement investment *per se* does not occur. Firms replace capital only if it is profitable to do so. They also disinvest, or scrap capital, if that particular asset is no longer profitable. When the asset finally wears out, it is 'replaced' only if it meets the same criterion as a new investment. Hence, we assume a 'one hoss shay' theory of depreciation which is quite accurate for almond production.

Capital stock at any moment of time is given by

(2) $k_t = k_{t-1} + a_t - s_t$

where k is the stock of capital in physical units, a is the gross additions to stock and s is the level of scrappings. As long as technology is Leontief and unchanging over time, the vintage of the capital does not matter and it is possible to aggregate the various vintages in the manner given by (2) in physical terms. From the above discussion,

(3) $a_t = f\{ \pi, var(\pi) \}$

and

(4) $s_t = g\{ \pi, var(\pi) \}$

where var(π) is the variance of expected profits and it is understood that the computation of π in (3) may be substantially different than the computation of π in (4) mainly because of differing time horizons, but also due to other factors. It should be noted that $\partial a/\partial\pi > 0$, but that $\partial a/\partial var(\pi)$ is ambiguous since we assume that investor can be either risk averse or risk inclined.

Given the above, if it pays to invest in one unit of the asset, what stops the investor from investing in an infinite (or, at least, very large) number of units? Two factors will give the system closure. First, although the investor may view input supply as perfectly elastic, it is not. The prices of inputs will eventually rise, and, just as important, input supply may be limited in crucial cases. The availability of inputs places a limit on the system that restricts potential investment. For example, credit may be rationed and/or cash flow limited. The possibility of credit rationing has been identified theoretically in the context of informational asymmetries and has a long history in empirical work on investment, usually under the name of cash flow. We do not investigate credit rationing in this study, but reserve it for future work. The second factor that restricts potential investment to finite amounts is risk. For risk averse investors, any amount of risk is an inhibiting factor, even though they expect to earn a profit. Hence, as initially recognized by Kalecki (1937), risk will be a factor causing the marginal efficiency of investment to fall. In practice, although we compute π for a small unit of investment, it is an average π in the sense that it is not affected by risk or any of the factors causing it to fall. This occurs mainly for two reasons. First, we do not conduct the analysis at the level of the individual almond grower and therefore do not have knowledge of the limitations, in terms of say increasingly unsatisfactory land, which he faces. These factors would result in a downward sloping schedule of π as a function of the level of investment. Second, and more importantly from an estimation point of view, year to year variations in π are large and as such are the main force determining investment.

The above theoretical development suggests a model of investment driven by the expected marginal present value (EMPV) of the investment. When the decisions of many individual investors are aggregated, constraints on inputs and credit availability result in finite investments that should be proportional to the magnitude of the EMPV. Because the model is for actual investment, not desired capital stock or desired investment, there is no need for any sort of partial adjustment process. There is, though, every reason to assume that investment decisions are based on past EMPV's as well as the current EMPV. For this reason, lagged values of the EMPV of almonds were included in the investment equation. In effect, π is represented by a distributed lag in EMPV. Lastly, because risk is believed to affect investment decisions, a measure of risk is included in the specification. Following a common formulation from the optimization literature {c.f. Pratt (1964) and Paris (1979)}, the variance of the EMPV is used as the measure of risk.

While the variance of expected profits will offer a reasonable measure of risk, it should be cautioned that the coefficient for this variable cannot be interpreted as a risk aversion coefficient in the normal Pratt-Arrow sense. This is due to the multi-period nature of the investment decision. With serially connected payoffs over a finite asset life, conventional results about attitudes toward risk do not necessarily apply.[3] It is less clear what the sign of the coefficient tells about the shape of the individual's utility function or his attitude toward risk than when confronted with a one-period independent uncertain decision. It is safe to conclude that a positive coefficient on the variance of the EMPV implies a greater propensity to invest in almonds when a larger degree of year-to-year variation in returns is expected and that a negative coefficient implies the opposite. Therefore, it may be conjectured that a positive (negative) coefficient does imply some type of risk inclination

[3] See Newberry and Stiglitz (1981), chapters 6 & 7.

(aversion), but this may only be for multi-period decisions where a loss in one year may appear to make large profits the next year more likely.

II. INVESTMENT IN ALMOND ORCHARDS

The almond industry was selected as a test for this approach for several reasons. First, it is possible to obtain detailed microdata on costs, plantings, yields and prices for this industry. Second, there has been a high level of investment in almond orchards over the past twenty years. Third, the technology of growing and harvesting almonds has changed little over this period, obviating the need to deal with technological change. Fourth, production of almonds is characterized by near Leontief technology thus reducing the scope of relative price considerations and facilitating the computation of present value. Fifth, the industry is reasonably competitive and is not dominated on the supply side by any single large producer. Sixth, returns in the almond industry are characterized by considerable variation. Hence, there is substantial uncertainty regarding future levels of profits. This variation occurs because of weather induced fluctuations in yields which are as great as 100 percent from year to year. Also, a large portion of the crop is exported. There is substantial variation in export demand due to fluctuations in competing supplies and changes in exchange rates. Seventh, the model presented here is highly dissaggregated: nine producing regions are identified. Output and the capital stock are homogeneous and undifferentiated. As a result aggregation problems are hopefully minimized.

As is the case for many tree crops, most soils are suitable for almonds, the critical factor being the availability of water. Once the seedlings are planted, four years are required before trees reach sufficient maturity to yield a harvestable crop. Hence, it is possible to separate the delivery lag from the expectational lag, a problem which Abel and Blanchard (1986) noted for the neoclassical model. The main costs incurred during this period, in addition to acquiring the land, are the cost of the seedlings, installation of an irrigation system, planting, cultivation, pruning, fertilization and management. These costs, including the opportunity costs of capital expended, are currently about $7000 per acre on a present value basis. There are slightly over 400,000 acres in almonds in California (the only state which raises them) for a capital stock value of almost three billion dollars.

A Federal marketing order exists for almonds and is used to provide grower funds for research, mainly on disease control in the trees. Large scale almond production is a relatively new phenomena in this country. As a result there are very few orchards approaching twenty years of age, the typical life of an almond tree. Hence, removals over the period under consideration, 1970-1985, were minimal.[4] In view of the above discussion, it is possible to write (2) as

(5) $k_t = k_{t-1} + n_t$

where n_t is net additions (gross additions less removals). Hence, net additions are the investment variable to be explained. As indicated above, the geometric depreciation theory where removals are a function of k_{t-1} is not realistic for the almond industry where most trees live twenty years. As a result k_{t-1} does not appear in the investment demand relation. The Almond Board provides funds for the California Crop and Livestock Reporting Service to do a survey of the bearing acreage. Since net additions are estimated as the change in bearing acreage, it is not possible to test that portion of our theory which says that removals as well as gross additions are determined by the utility of expected profits. However, as we note above, removals over this period have been negligible.

III. ECONOMETRIC SPECIFICATION AND ESTIMATION

With the above development in mind the investment model can now be specified as,

(6) $NI_{it} = ß_o + ß(L)\ EMPV_{it-4} + ß_6\ VAR_{it-4} + w_{it}$

where

---

[4] According to estimates made by Bushnell and King (1986), removals are currently about one percent of bearing acreage. We view this to be mimimal in view of the fact that there are more removals presently than earlier in the study period.

$NI_{it} = (A_{it} - A_{it-1}) / A_{it-1}$

A_{it} - bearing acerage of almonds in the ith region in time period t,

$EMPV_{it}$ = the expected marginal present value of profit in region i (defined below) in year t

L- distributed lag operator

VAR_{it-4} = variance of the EMPV (also defined below)

w_{it} = stochastic error term for region i in time period t.

Recognizing that it takes four full years for almond trees to bear, we specify the model as

(7) $NI_{it} = \beta_0 + \beta_1{*}EMPV_{it-4} + \beta_2{*}EMPV_{it-5} + ... + \beta_5{*}EMPV_{it-8} + \beta_6{*}VAR_{it-4} + w_{it}$

where

(8) $EMPV_{it} = (PV_{it} + D_{it} - L_{it}) / f_t$.

Following Sims (1972) the model is estimated with no constraints on the shape of the lag distribution. In a manner similar to that used by Clark (1979), investment is divided by capital stock lagged so that both left and right hand side variables are measured on a per acre basis, i.e. investment per acre and profit per acre. Only the left hand side of (7) is divided by capital stock, thus avoiding problems of spurious correlation first pointed out by Kuh and Meyer (1955).

Assumptions regarding the expectations of future prices and costs are discussed below. Technology is assumed constant once the trees are in place. Expected yield (output per acre) is based on an estimated production function also discussed below. PV is the expected present value of cash revenues and costs on one acre of almonds over a twenty year horizon viewed from time period t. Costs include investment expenditures as well as production costs. D is the expected present value of tax savings due to accelerated depreciation and the investment tax credit. L is the current price of irrigated farmland in region i and is used as a measure of the opportunity cost of the land. All the present values are discounted by the current Federal Land Bank long term interest rate. At any point in time over the sample period the level of investment will be a function of the level of EMPV which is given in nominal terms. However, in order to use EMPV over time in regression analysis, it must be converted to constant dollars for purposes of comparison. The deflator used for this purpose is f, prices paid by farmers. This index is composed of prices of various farm inputs. Since profits are often re-invested by farmers, this is the appropriate index. As a practical matter, it is highly correlated with the CPI and/or the GNP deflator. A more precise explanation of how PV and D were computed and the data sources is given below in Section IV. VAR_{it-4} is computed via the standard sampling variance formula using the eight most recent EMPV's.

For purposes of estimation, California was classified into nine growing regions. Seven of these regions correspond to counties while the remaining two are groups of counties, one in the north and one in the south, whose involvement in almonds is small. The following assumptions were made about the error term, w_{it},

(9a) $E(w_{it}) = 0$

(9b) $E(w_{it}w_{ks}) = 0,$ for all t not equal to s

(9c) $E(w_{it}^2) = \sigma_{ii},$ $i = 1,...,9$

(9d) $E(w_{it}w_{kt}) = \sigma_{ik},$ $i,k = 1,...,9$, i not equal to k.

Further spatial and temporal error covariance assumptions were made about the σ_{ik}. First, the assumption was made that there is no autocorrelation between the error terms, even within a region. While this assumption may seem simplistic in a model with lagged variables, it should be noted that all lagged variables are exogenous. Also, a sample autocorrelation coefficient was calculated (r=.232) and was not statistically significant. As a further test, the model was estimated using the appropriate GLS autocorrelation transformation and it provided no improvement over the model without the correction. Hence, there was no autocorrelation correction in the final version of the model. The errors are, however, assumed to have particular non-zero spatial covariances. The nine regions which our data pertain to are grouped into three super-regions (1,8; 2,3,4,9; 5,6,7; see Table 1 for the numbering key) corresponding to the Sacramento Valley, Northern San

Joaquin Valley, and Southern San Joaquin Valley, respectively. On this basis the spatial error assumptions are: each region has its own distinct error variance given by (9c), any two regions in the same super-region have the same error covariance and any two regions in two different super-regions have the same error covariance. For example, this means that $\sigma_{23} = \sigma_{24}$ and $\sigma_{25} = \sigma_{37}$. These covariance assumptions are based on the regional nature of the data set. It certainly seems plausible that if almonds are profitable in one county, some of the investment thereby encouraged may take place in another county. Therefore, covariances between regions should be expected *a priori*, and would likely depend on how close two regions are to each other. Such reasoning leads to covariance assumptions of the type made in (9).

Applying the proposed error structure to the model, estimates of the parameters in the error variance-covariance matrix, $\Omega= E(ww')$, were obtained from an OLS regression of the model. The OLS residuals were used to estimate each σ_{ik} parameter, being careful to correct for the fact that the regional subsets of the residuals do not have zero means. The estimate of Ω obtained by this method is consistent due to the consistency of the OLS residuals. Lastly, it should be noted that $\Omega = \Sigma \otimes I$, where Σ is a 9x9 matrix of the σ_{ik} defined above, I is an 8x8 identity matrix and $\otimes$ is the Kronecker delta product. Such an error structure, based here on the regional nature of the problem, is equivalent to the error structure of Zellner's Seemingly Unrelated Regression (SUR). Zellner (1962) showed that such a procedure is not only consistent, but also efficient for estimating systems of equations such as these. Here the "system" is simply the nine equations, each representing one county or county group. Although the variables in each of the equations are identical, the actual values of those variables differ across equations because costs, technology, and prices vary by county. Hence, there is a gain in efficiency captured by exploiting the familiar SUR error structure.

What is less well known is that estimates yielded by Zellner's SUR are also unbiased. This fact can be important in the small to medium sized samples which are often used in applied work. Kakwani (1967) showed that as long as the disturbance term has a continuous symmetric probability density function the Zellner estimators are unbiased. This result is based on the fact that Σ is then an even function of the disturbance term, $\mathbf{w}$. Therefore, $\Omega = \Sigma \otimes \mathbf{I}$ is also an even function of $\mathbf{w}$. This fact can be used to show that the estimates obtained by performing SUR are unbiased both in finite and infinite samples. We can extend this result to the model used here even though the ß's do not vary across the nine regions. The final estimation of the model was then performed using GLS, with the estimate of Ω. In addition to the EMPV model, this estimation procedure was implemented on two alternative models. The same data base was used on all three models. The alternative models tested were the neoclassical investment model (NIM) and the accelerator investment model (AIM).

Estimation of Expected Yield. One of the most important components of the EMPV is the yield. This requires the estimation of a production function for almonds. As indicated earlier, there is substantial variation in almond yields, due mainly to weather, but also due to the fact that almonds are an alternate bearing crop. Alternate bearing means that the crop alternates yearly between relatively heavy and relatively light crops. This phenomena is difficult to observe due to the much larger weather induced variations. However, any expected profit model must take these variations into account and provide an accurate forecast of expected yield for the present value computation. Since yield is total output divided by the number of acres, estimation of the yield function is equivalent to estimating the production function.

The production function for almonds was estimated using three variables: bearing acreage, rainfall, and a dummy variable for alternate years. Bearing acreage is the number of acres of almond trees four years of age and older. Once in place, the input requirements for cultivation, harvesting, pruning, etc. are virtually fixed. Hence, the usual economic inputs e.g. labor, materials, do not appear in the production function. The costs of these inputs do, however, appear in the present value computations discussed below. Rainfall was included because the bloom period for almonds falls during California's rainy season (February and March). If it rains too much during the bloom period the bees cannot pollinate the trees well and a small crop is the result. A dummy was included to account for the alternate bearing phenomena.

The model was estimated as a pooled cross-section time series model. The relation specifies that production varies by region (due to climate, soil type, water, etc.), by alternate years, and with the rainfall in that region. In functional notation,

(10) production = f(bearing acreage, region, alternate year, rainfall).

Dummy variables were employed for eight of the nine regions with Butte County being the base region. A dummy variable that took the values of 0 in even years and 1 in odd years was used to model the alternate bearing phenomena. The variable for rainfall was inches of rainfall in February squared. The production model can be written as,

(11) $P_{it} = \beta_0 + \beta_1 A_{it} + A_{it}\Sigma\beta_k R_{kt} + \pi_1 A_{it} FR_{it} + A_{it} FR_{it}\Sigma\pi_k R_{kt} + \mu_1 A_{it} D_{it} + v_{it}$

where

P_{it} = tons of almond kernels produced in region i, in year t

A_{it} = bearing acreage, in thousands of acres for region i, in year t

R_{kt} = dummy variable equal to 1 if the kth region, 0 otherwise, k=2, 9

D_{it} = dummy for alternate bearing, equal to 1 in odd years, 0 in even years

FR_{it} = inches of rainfall in February, squared[5] , in region i, in year t

v_{it} = stochastic error term.

The production relation was estimated by generalized least squares (GLS). The assumptions about the v_{it} were:

(12a) $E(v_{it}) = 0$

(12b) $E(v_{it}v_{ks}) = 0,$ for all s not equal to t.

---

[5] Both rainfall and rainfall squared were tried with the latter producing better results. The squared effect was also found superior in other weather related studies of almond production by the authors. The month of Februaury was selected because it is the bloom period for almonds throught the state. Experimentaiton with later bloom periods for northern counties did not improve the results.

(12c) $E(v_{it}^2) = \phi_{ii}$, $i = 1, ..., 9$

(12d) $E(v_{it}v_{kt}) = \phi_{ik}$, for all $i,k = 1,...,9$

This gives an error structure where $E(vv') = \Phi \otimes I$ where I is a 16x16 identity matrix, Φ a matrix of the ϕ_{ik}. The ϕ_{ik} were estimated from the residuals of the ordinary least squares (OLS) regression of the above production model. Once again, reference to Kakwani's proof yields estimates which are unbiased.

The data used to perform the estimation were county level data on almond acreage and production in California from 1970 to 1985, taken from the relevant County Agricultural Commissioner's Reports. The data on rainfall was collected from the National Oceanographic and Atmospheric Administration published reports. February rainfall was chosen as most representative of the bloom period. Within each county a weather station nearest the center of the almond growing area was chosen. The data were organized into nine regions: seven counties (Butte, Fresno, Kern, Madera, Merced, San Joaquin, and Stanislaus) and two groups of counties that grow fewer almonds, North (Colusa, Contra Costa, Glenn, Solano, Sutter, Tehama, Yolo, and Yuba Counties) and South (Kings, San Luis Obispo, and Tulare Counties). Hence, there are a total of 144 observations used to estimate the production function. The estimates of the production function are given in Table 1.

IV. ESTIMATION OF THE INVESTMENT MODELS

In this section we present more details of the EMPV investment model as well as a brief development of the alternative models tested: a) neoclassical model with Cobb-Douglas technology: and, b) neoclassical model with Leontief technology, or, as it is commonly called-the accelerator. All three specifications are estimated with the same data by the GLS regression technique using an error structure which is equivalent to SUR as outlined above.

Table 1. Generalized Least Squares Estimates of Almond Production Function

Symbol	Variable	Coefficient	t-ratio
β_0	Intercept	$-.402\times10^4$	5.84
β_1	Bearing acreage (A)	.780	17.90
π_1	February rainfall x A	-.003	3.64
μ_1	Alternate bearing dummy x A	-.076	4.13
	Acreage x Regional Dummies		
β_2	A x Fresno	.076	1.71
β_3	A x Kern	.036	.89
β_4	A x Madera	.075	1.65
β_5	A x Merced	.023	.34
β_6	A x San Joaquin	.090	1.58
β_7	A x Stanislaus	.067	2.56
β_8	A x North Region	-.257	8.37
β_9	A x South Region	-.083	1.72
	Acreage x February Rainfall x Regional Dummies		
π_2	A x FR x Fresno	-.005	1.29
π_3	A x FR x Kern	-.011	3.94
π_4	A x FR x Madera	-.011	2.60
π_5	A x FR x Merced	-.006	1.21
π_6	A x FR x San Joaquin	-.015	2.82
π_7	A x FR x Stanislaus	-.010	4.05
π_8	A x FR x North Region	.0004	.51
π_9	A x FR x South Region	.0002	.13

$R^2 = .927$ $\bar{R}^2 = .916$ $s = .984$

The EMPV Model. All present value computations in the model were done over a twenty year time horizon,[6] which is the life of a typical almond tree. Hence,

$$(13)\quad EMPV_t = \{\sum_{j=0}^{19} (p_{jt}y_{jt} - c_{jt}l_{jt} + d_{jt})(1-m_{jt})(1 + r_{jt})^{-j}\}/f_t$$

where

$EMPV_t$ - expected marginal present value in year t, t=1972-1985

p_{jt} - the expectation in year t for almond price in years t + j,

[6] While we assume a twenty year planning horizon, almond trees can, of course, be left in the ground longer. Yields begin to decline substantially after twenty years due mainly to the impact of shaking from the mechanical harvesters. Also, the effect on present value of an additional year past twenty is quite small.

y_{jt} - the expectation in year t for almond yield in years t + j,

c_{jt} - the expectation in year t of a vector of input prices in years t + j,

l_{jt} - the expectation in year t of a vector of input coefficients in year t + j,

d_{jt} - tax savings due to depreciation and investment tax credit in year t + j,

m_{jt} - the expectation in year t of the marginal tax rate in period t + j,

r_{jt} - the expectation in year t of the discount rate in year t + j,

f_t - prices paid by farmers in year t.

The EMPV's were computed for each year, 1970-85, for each of the nine regions. The expectations assumption used was the following

(14) $p_{jt} = (1+g_{pt})^j p_{0t}$ $\quad$ j=0,...,19

where g_{pt} is the rate of growth of expected prices in period t. The estimate of g_{pt} is based on the unweighted average of the past three years rates of growth of prices. Hence, prices are assumed to grow at the same rate as they have averaged over the past three years. Prices, and their associated g_{pt}'s, are different for each of the nine regions. We have deleted regional subscripts here in order to avoid confusion. The same procedure was used to estimate the vector of expected input prices, the c_{jt}'s. They also vary by region and are composed of prices for herbicides, pesticides, skilled and unskilled labor, trees, planting labor, water, fertilizer, bees for pollination, tractor, harvesting labor and equipment, management, and miscellaneous. As a result of the assumption of Leontief technology, the vector of l_{jt}'s which pertain to the variable inputs, do not change over time, but do vary by region and age of orchard. The values of the expected yield, also by region, are obtained using the production function estimated above. In order to estimate these yields, it was assumed that rainfall in the future would be at the historical mean level. The data used for the discount rate is the Federal Land Bank long term rate and is the same for all regions and is expected to remain the same in all future periods. Depreciation, d, was estimated by computing the accelerated depreciation (sum-of-the-years' digits) times the marginal tax rate. This category also includes the value of the investment tax credit on the qualifying expenditures. The marginal tax rate was computed by taking the average of the marginal tax rates on incomes (in constant 1965 dollars) of 20 and 80 thousand dollars. These two income levels represented average farmer and professional investor income levels. Depreciation varies by region since input costs vary by region.

The EMPV model presented above clearly has elements of rational expectations theory. The use of the production model for yield expectations is consistent with the rational expectations hypothesis. In order to fully implement that approach it would be necessary to estimate ARIMA type models for each of the input prices and the output price. While this would not be an insurmountable task, the short annual time series available for such an estimation makes that approach unappealing. Hence, we have chosen to use the moving average growth rate assumption although the errors from this method may be autocorrelated. Some, limited, experimentation showed that the price expectations computed by the moving average method were quite close to those computed by ARIMA models. We also note that the focus of this paper is not to test the rational expectations hypothesis.[7] The model given here also has similarities to Tobin's q-theory of investment. If prices were available over time on almond orchards, it would be possible to compare the replacement costs, which are given above, with these prices to determine if investment should take place.[8]

The Neoclassical Investment Model. The starting point for the neoclassical investment model used here is a Cobb-Douglas production function relating the output of almonds (X) to the capital stock of almond trees (K) and other inputs. Under the assumption of multiperiod optimization the desired level of capital stock is given by

$$(15) \quad K^{*} = pX/q$$

where p is the price of output and q is the user cost of capital here defined as

---

[7] For a test of the rational expectations hypothesis applied to agriculture see Goodwin and Sheffrin (1982).

[8] For examples of the q-theory approach in applied studies see Engel and Foley (1976), Summers (1981), and von Furstenberg (1977).

(16) $q = (d + r)z - m$

where d is the depreciation rate on the investment good, r is the rate of interest, z is the price of the capital good, and m is the present value of that years' tax savings from depreciation. Following the usual transformation the model to be estimated is

(17) $NI_{it} = \beta_o + \beta(L)\ \Delta K^*_{it-4} + w_{it}.$

The neoclassical investment model usually contains a term in K_{t-1}. However, as noted above there has been virtually no replacement demand over this period. Hence, lagged capital stock is omitted.[9] The user cost of capital was computed by adding together the cost of all of the items used to bring the trees to maturity. This includes the cost of the trees themselves, the irrigation system, cultivation, management and all other costs for the first four years. This total was z, the price of the capital good. The depreciation method used was sum-of-the-years' digits and the Federal Land Bank long term rate was used for r. Again, the lag distribution was estimated in an unconstrained manner.

The Accelerator Investment Model. The accelerator model can be considered as a special case of the neoclassical investment model. If instead of the Cobb-Douglas production function, Leontief technology is assumed, then the input demand for the capital services will be

(18) $K^* = a_1 X.$

Again, with the usual adjustment assumptions,

(19) $N_{it} = \beta_o + \beta(L)\ \Delta X_{it-4} + w_{it}.$

The accelerator model can, of course, be derived from alternative considerations.[10]

V. THE EMPIRICAL RESULTS

---

[9] The model given by (17) was estimated with K_{t-1} included and it was not significant.

[10] Both the accelerator and the neoclassical model were estimated in the same manner as the EMPV model, i.e. $(A - A_{it-1}) / A_{it-1}$ as the dependent variable.

This section presents the Generalized Least Squares estimates for the three models presented above. It also gives the results of the prediction interval test and the elasticities for the exogenous variables. The model estimates are given in Table 2. The EMPV model produces a stable lag structure, all coefficients are significant with the exception of X_{t-4}. The coefficient on Var is significant and positive indicating that almond growers are risk inclined. This situation would seem to be consistent with a willingness to undertake risky investments. Omission of the Var variable did not greatly affect the shape of the lag structure or the significance of the individual coefficients.

The first R^2 in Table 2 is for the Generalized Least Squares residuals and is given by

(20) $$R^2\text{:GLS} = \mathbf{w}'\Omega^{-1}\mathbf{w}/\mathbf{y}'\Delta^{-1}\mathbf{y}$$

where Ω is as before, $\Delta = \Sigma_n \otimes A_T$, Σ_n is as before and $A_T = \mathbf{I}_T - 1/T\ \mathbf{ii}'$, $\mathbf{i}$ is the units vector, n is the number of regions, and T is the number of time periods. The R^2 for the untransformed (UT) data is therefore,

(21) $$R^2\text{:UT} = \mathbf{w}'\mathbf{A}\mathbf{w}/\mathbf{y}'\mathbf{A}\mathbf{y}$$

where $\mathbf{A} = \mathbf{I}_{nT} - 1/nT\ \mathbf{ii}''$. Both of the R^2 in the table are adjusted for degrees of freedom by the standard method. The EMPV model has the best fit by either R^2 criterion. In order to determine the effect of including the variance term on the goodness of fit, the NIM and AIM were estimated using both the Var measure from the EMPV model and a variance computed from the appropriate X's for those models. None of the four models estimated in this manner were an improvement in terms of an increased adjusted R^2. The estimates in terms of the lag structure were worse for both NIM and AIM with the variance terms included. The lag structure for the neoclassical model is consistent with prior expectations and a little more delayed than the one for the EMPV model. The lag structure for the accelerator model is sawtoothed and stands in strong contrast to the reasonably smooth lags found for the EMPV and NIM models.

Table 2. Generalized Least-Squares Estimates of the EMPV, NIM, and AIM Models

Variable	EMPV	NIM	AIM
Intercept	.681E-01 (11.22)*	.330E-01 (5.82)	.038E-01 (4.54)
**X_{t-4}	.497E-06 (0.68)	-.627E-08 (0.17)	.149E-05 (1.58)
X_{t-5}	.321E-05 (4.86)	.587E-07 (1.54)	-.208E-06 (0.19)
X_{t-6}	.567E-05 (7.40)	.203E-06 (5.32)	.174E-05 (1.42)
X_{t-7}	.499E-05 (6.81)	.209E-06 (4.86)	-.502E-07 (0.03)
X_{t-8}	.452E-05 (5.44)	.091E-08 (2.14)	.552E-05 (2.67)
Var	.117E-08 (3.82)	-----	-----
R^2:GLS	.664	.398	.081
R^2:UT	.386	.261	.025

* The numbers in parenthesis are the t-ratios.
** The X's represent the independent variables for each particular model. Hence, for the EMPV model the X's are given by (13), for NIM by (15) and (17), and for AIM by (19).

Out of sample forecasts were made to test the forecasting ability and the empirical consistency of models. If most of the forecasts fall within one standard deviation of the actual values, then a given model can be deemed to be empirically consistent, even if its forecasts are not particularly accurate. If out of sample forecasts routinely exceed one standard deviation from the actual values, then the validity of the model as a forecasting tool must be questioned. Also, the calculated confidence levels that are placed on the model and its parameters must be re-examined.[11]

[11] For an earlier application of this prediction test to investment models see Clark (1979).

The predictions were made for each region for the year 1986, giving nine forecasts per model. The results of these predictions are presented in Table 3. Below each prediction is the t-value for the forecast error (i.e. how many standard deviations the forecast fell from the actual value). For a prediction interval test done in this manner, as low a t-value as possible is desired. The covariance matrix of the forecasts for a given model is calculated as:

(21) $$\mathbf{V} = \Sigma + \mathbf{R}(\mathbf{X}'\Omega^{-1}\mathbf{X})^{-1}\mathbf{R}'.$$

Here, Ω is the error covariance matrix for the given model, and remembering the error structure employed, $\Omega = \Sigma \otimes \mathbf{I}$, $\mathbf{X}$ is the in sample matrix of data on the regressors, and $\mathbf{R}$ is the out of sample matrix of data on the regressors.[12] The standard deviation of a particular forecast is simply the square root of the appropriate diagonal element. For example, the t-value for the forecast error of the prediction for region 3 is

(22) $$t_3 = (\mathbf{a} - \mathbf{R}\beta)_3/(\mathbf{V}_{33})^{1/2},$$

where $\mathbf{a}$ is the vector of actual values, $\mathbf{R}\beta$ is the vector of predictions, and subscripts refer to rows and columns of vectors and matrices.

As can be seen from Table 3, there is a definite difference in how close the forecasts lie to the actual values in terms of standard deviations. For the EMPV model every forecast is within one standard deviation of the actual value, while the two alternative models both show a number of forecasts with errors that fall farther than one standard deviation from the actual value. The reader has surely noticed that the standard deviation varies not only with the model chosen, but also with the actual forecast (i.e., region). While it is clear that a model with poor explanatory power might easily show itself to be empirically consistent by such a test simply by having a large forecast variance, this is not why the EMPV model performs better than the other models here. It should be noted that the EMPV model has the highest R^2 of the four models. Further, it should be noted that the covariance matrix of

---

[12] See Johnston (1986).

the coefficients appears in the formula for the covariance matrix of the forecasts. It will be remembered that the EMPV model also had the best overall set of t-statistics for the regression coefficients, which implies a small covariance matrix in relation to the ß's used to calculate the forecasts. With these two facts in mind, it seems safe to conclude in light of the results of the prediction tests performed that the EMPV model has the greatest empirical consistency in terms of out of sample validation.

Table 3. Results of Prediction Interval Test

Region	Actual	EMPV	AIM	NIM
1	867	1788 (.44)*	1841 (.60)	1502 (.72)
2	-471	948 (.57)	1937 (1.31)	1975 (1.66)
3	1980	3781 (.23)	9244 (1.77)	9055 (2.08)
4	2722	1455 (.57)	1888 (.75)	1575 (.99)
5	1241	2906 (.35)	2492 (.62)	3596 (1.00)
6	-489	1668 (.56)	1849 (2.77)	1416 (1.53)
7	2206	3751 (.33)	3278 (.32)	4314 (.78)
8	-1847	996 (.67)	1620 (.88)	1041 (1.00)
9	-577	757 (.94)	1091 (1.13)	973 (.69)
R^2 - predict vs. actual		.471	.210	.293

*t-values for the prediction errors are in parenthesis.

The remaining task is to examine the impact and long run effects of changes in expectations regarding the various exogenous variables entering the computation of the present value of profits. Because of the dimension specific nature of the problem these impacts are best measured as elasticities. Table 4 presents the elasticities of almond investment with respect to changes in the expected rates of growth of output price, the wage rate, the marginal tax rate, the investment tax credit, the discount rate, and the cost of trees - the main capital input.[13] Remembering that trees planted four years ago show up as bearing acreage in the present year, all lags start in t-4. The lags go on to t-10 because our expectations are based on a three year moving average of the rates of growth in output and factor prices. In the EMPV model the elasticity of investment with respect to Var is .088.

Table 4. CumulativeEffects of Changes in Expectations on Almond Investment: Elasticities Evaluated at 1985

Time Period	Output Price	Wage Rate	Marginal Tax Rate	Investment Tax Credit	Discount Rate	Cost of Trees
t-4	.003	-.002	-.002	.000	-.022	-.000
t-5	.030	-.014	-.021	.001	-.188	-.001
t-6	.096	-.046	-.068	.003	-.606	-.002
t-7	.195	-.094	-.137	.008	-1.224	-.004
t-8	.303	-.146	-.213	.015	-1.901	-.007
t-9	.370	-.179	-.261	.021	-2.325	-.008
t-10	.402	-.194	-.284	.024	-2.527	-.009

VI. CONCLUDING REMARKS

The model presented here implements an alternative specification of the neoclassical notion that firms maximize the expected utility of the present value of profits. The expected utility of profits is represented by the expected present value of profits and the variance of profits. The model is then applied to investment in almond orchards, an industry which is characterized by substantial price and output uncertainty. Under the assumptions of Leontief technology in almond production and of future output price and factor prices that

[13] All calculations are based on the EMPV model whose coefficients are presented above in Table 2.

grow at the same rate they grew over the past three years, it is possible to compute the present value of profit for an acre of almonds trees. The model was then estimated by Generalized Least Squares using cross-section data from nine almond growing regions over the sixteen year period from 1970-85. The specification of the error structure recognizes the pooled time series cross-section nature of the data as well as the regional relations in almond production. The estimated model was compared with traditional neoclassical and accelerator investment models which were estimated on the same data using the same technique and error structure assumptions. The EMPV model showed a better fit over the historical period, had a lag structure more in line with *a priori* expectations and also outperformed the other two models in an out of sample prediction interval test. All predictions from the EMPV model were within one standard error of the sample prediction standard error.

REFERENCES

Abel, Andrew B., "Optimal Investment Under Uncertainty," *American Economic Review* 73 (March 1983) 228-233.

Abel, Andrew B. and Olivier J. Blanchard, "The Present Value of Profits and Cyclical Movements in Investment," *Econometrica* 54 (March 1986) 249-273.

Arrow, Kenneth J., *Risk Perception in Psychology and Economics*, Tech. Report No. 351 (Stanford: Institute for Mathematical Studies in the Social Sciences, Oct. 1981).

Bernstein, Jeffrey I. and M. Ishaq Nadiri, "Financing and Investment in Plant and Equipment and Research and Development," in M. Peston and R. Quandt eds. *Prices, Competition and Equilibrium,* (Philip Allan/Barnes & Noble: 1986) 213-248

Bushnell, Peter G. and Gordon A. King, *The Domestic and Export Markets for California Almonds*. Giannini Research Report No. 334 (University of California, Sept. 1986).

Clark, Peter K., "Investment in the 1970's: Theory, Performance, and Prediction," *Brookings Papers on Economic Activity* (1979) 73-113.

Chirinko, Robert S. and Robert Eisner, "Tax Policy and Investment in Major U.S. Macroeconomic Econometric Models," *Journal of Public Economics* 20 (1983) 139-166.

Craine, Roger, "Investment, Adjustment Costs, and Uncertainty," *International Economic Review* 16 (October 1975) 648-661.

Engle, Robert F. and Duncan K. Foley, "An Asset Price Model of Aggregate Investment," *International Economic Review* 16 (October 1975) 625-647.

Fazzari, Steven M. and Michael J. Athey, "Asymmetric Information, Financing Constraints, and Investment," *Review of Economics and Statistics* 64 (August 1987) 481-487.

Goodwin, Thomas H. and Steven M. Sheffrin, "Testing the Rational Expectations Hypothesis in an Agricultural Market," *Review of Economics and Statistics* 64 (November 1982) 658-667.

Greenwald, Bruce, Joseph Stiglitz, and Andrew Weiss, "Informational Imperfections in the Capital Market and Macroeconomic Fluctuations," *American Economic Review* 74 (May 1984) 194-199.

Hall, Robert E., "The Macroeconomic Impact of Changes in Income Taxes in the Short and Medium Runs," *Journal of Political Economy* 86:2 (April 1978) S71-S85.

Hartman, Richard, "Factor Demand with Output Price Uncertainty," *American Economic Review* 66 (Sept. 1976) 675-681.

Johnston, J. *Econometric Methods*. (McGraw-Hill:1984)

Jorgenson, Dale W. "Econometric Studies of Investment Behavior: A Survey," *Journal of Economic Literature* 9 (Dec. 1971) 1111-47.

Kakwani, N.C., "The Unbiasedness of Zellner's Seemingly Unrelated Regression Equations Estimators," *Journal of the American Statistical Association* 62 (March 1967) 141-142.

Kalecki, M., "The Principle of Increasing Risk," *Economica* (New Series) 4 (November 1937) 440-447.

Kokkelenberg, Edward C. and Charles C. Bishoff, "Expectations and Factor Demands," *Review of Economics and Statistics* 63 (August 1986) 423-431.

Kuh, Edwin and James Meyer, "Correlation and Regression When the Data Are Ratios," *Econometrica* 23 (1955) 400-416.

Lucas, Robert E. and Edward Prescott, "Investment Under Uncertainty," *Econometrica* 39, (September 1971) 659-681.

Meese, Richard., "Dynamic Factor Demand Schedules for Labor and Capital Under Rational Expectations," *Journal of Econometrics* 14 (1980) 141-158.

Myers, Stewart C. and Nicholas S. Majluf, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics* 13 (1984) 187-221.

Newbery, David M. G. and Joseph E. Stiglitz.,*The Theory of Commodity Price Stabilization: A Study in the Economics of Risk.* (Oxford: Clarendon Press, 1981).

Paris, Quirino, "Revenue and Cost Uncertainty, Generalized Mean-Variance, and the Linear Complementarity Problem," *American Journal of Agricultural Economics* 61(May 1979) 268-75.

Pindyck, Robert S., "Adjustment Costs, Uncertainty, and the Behavior of the Firm," *American Economic Review* 72 (June 1982) 415-427.

Pindyck, Robert S. and Julio Rotemberg, "Dynamic Factor Demands and the Effets of Energy Price Shocks," *American Economic Review* 73 (December 1983), 1066-1079.

Pratt, J., "Risk Aversion in the Small and in the Large," *Econometrica* 32 (Jan.-Apr. 1964) 122-136.

Sims, Christopher A., "The Role of Approximate Prior Restrictions in Distributed Lag Estimations," *Journal of the American Statistical Association* 67 (March 1972) 169-75.

Sinai, Allen and Otto Eckstein, "Tax Policy and Fixed Business Investment Revisited," *Journal of Economic Behavior and Organization* 4 (1983) 131-162.

Stiglitz, Joseph E. and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review* 71 (June 1971) 393-410.

Summers, Lawrence H., "Taxation and Corporate Investment: A q-Theory Approach," *Brookings Papers on Economic Activity* (1981) 67-127.

University of California, Cooperative Extension. Various reports on almond production costs and orchard establishment costs, unpublished, University of California-Davis, 1981-1986.

Uri, Noel P., "Testing for the Stability of the Investment Function," *Review of Economics and Statistics* 64 (February 1982) 117-125.

von Furstenberg, George M., "Corporate Investment: Does Market Valuation Matter in the Aggregate?" *Brookings Papers on Economic Activity* (1977) 347-408.

Zellner, Arnold, "An Efficient Method of Estimating Seemingly Unrelated Regresions and Tests for Aggregate Bias," *Journal of the American Statistical Association* 57 (1962) 348-68.

REGIONAL AND WEATHER RELATED VARIATIONS IN ALMOND YIELDS

Jeffrey Dorfman and Dale Heien

Almond production is one of agriculture's best examples of the impact of weather on yields. Year to year changes in yields, due mainly to rainfall, can be as great as one hundred percent. As a result, variations in yield are translated directly into variations in production, which is clearly demonstrated in Figure 1. Because a significant proportion of the almond crop is exported it is important to know as early as possible what the total crop will be. This is useful mainly for pricing considerations and for making an optimal allocation between domestic and foreign sales.

Figure 1. Graph of Almond Yields and Production

In a previous paper (California Agriculture, March-April 1987) we reported on a technique to improve the State's objective almond survey which is made in May of each year. The forecast made as a result of the State's survey is quite accurate, but it does not come out until June. Although this is still two months in advance of the time harvest begins, it is desirable to have an earlier forecast for marketing planning purposes. The state does a small survey in February, which is not as accurate as the later one. In this paper we report on results of trying to use econometric methods to forecast the almond crop using data which is commonly available at the same time the first survey is made. For forecasting purposes, the model turned out to be unsatisfactory. There is simply too much random variation to use statistical techniques to forecast crop size so early in the growing season with the reliability required for economic decisions. However, we did uncover some interesting results with regard to the effects of rainfall and other factors on yield.

As mentioned above, the production of almonds is greatly influenced by rainfall; in particular, rainfall in the month of February. The bloom period for almonds falls during California's rainy season (February and March). Almond trees cannot self-pollinate, but must be pollinated by another variety of almond tree. For this reason almond orchards always contain at least two varieties of trees, planted either in alternating rows or two rows of one and one row of the other. Because cross-pollination is necessary, bees are vital to a good crop. If it rains too much during the bloom period the bees do not pollinate the trees well enough and a small crop (as in 1986) is the result.

In order to measure the effect of rainfall and the alternate bearing phenomena we estimated a statistical production function for almonds. A production function measures the relationship between output and inputs. The inputs are represented here by acreage, rainfall, and a variable for the alternate bearing pattern of almond trees. Alternate bearing means that the crop alternates yearly between relatively light and relatively heavy yields due to physiological factors, and almonds are considered moderately alternate bearing. We began with a relation that indicates that yields vary by region, by alternate years, and with the rainfall in that region. Since identically, production equals yield times bearing acreage, the production relationship can be estimated as yield times bearing acreage. This was then translated into a relation which specified production as a function of bearing acreage times regional effects, bearing acreage times the rainfall effect, and bearing acreage times the alternate yield effect. Bearing acreage is the number of acres of almond trees four years of age and older. Technically, the trees are not mature enough to bear nuts until four although a small crop often is obtained now from three year old trees. A qualitative (zero or one) variable for alternate years was included to measure the size of the alternate bearing effect. Weather (and other) induced variations are always so great relative to the alternate yield phenomena in almonds that it has been difficult for pomologists to determine the magnitude of the alternate bearing in almonds.

The production function was estimated as a pooled cross-section time series model: cross-section refers to the fact that data was from seven counties and two other regions composed of groups of counties and time series refers to the fact that these cross-section data are for the years 1970-1985. Hence, the estimation was based on a total of 144 observations. Dummy variables were employed for eight of the nine regions with Butte County being the base region. The variable for rainfall used was inches of rainfall in February squared. Experimentation with rainfall and rainfall squared indicated that rainfall squared performed better. The relationship was estimated by Generalized Least Squares regression and the results are presented in Table 1.

The data used to perform the estimation was county level data on almond acreage and production in California from 1970 to 1985, collected from the relevant County Agricultural Commissioner's Reports. The data on rainfall was collected from the National Oceanographic and Atmospheric Administration published data. February rainfall was chosen as most representative of the bloom period. Within each county a weather station nearest the center of the almond growing area was chosen. The data was organized into nine regions: seven counties (Butte, Fresno, Kern, Madera, Merced, San Joaquin, and Stanislaus) and two groups of counties that grow fewer almonds, North (Colusa, Contra Costa, Glenn, Solano, Sutter, Tehama, Yolo, and Yuba Counties) and South (Kings, San Luis Obispo, and Tulare Counties).

The results of this estimation allow two interesting effects to be calculated, the effect of rainfall during the pollination period and the magnitude of the alternate bearing effect in almonds. The results of the regression show the alternate bearing phenomenon to result in a variation of 152 pounds per acre in the yield of almonds from year to year. This is certainly a significant amount compared to common yields of about 1400 pounds per acre. Still, when this variation of approximately eleven percent is compared to the variation in Figure 1, it is easy to see how this effect was masked by the remaining variation. The rainfall effects are somewhat more complicated as the coefficients involved are allowed to vary by region.

Table 1 presents some figures on the effects of rainfall and regional location on almond yields. The mean and standard deviation of the February rainfall squared are given for each of the nine regions, along with the mean yield. Then the estimated regional variation in yield is shown. This variation is due to such factors as differences in soil, climate, orchard age, and cultural practices. The final column gives the estimated loss in yield in a year with average rainfall in February from that rain. All yields are given in meat pounds per acre.

Table 1. Effects of Rainfall and Location on Yield

Region	Rainfall ($in.^2$)		Regional (lb/acre)		Rain (lb/acre)
	Mean	Std. Dev.	Mean Yield	Effect	Mean Loss
(1)	(2)	(3)	(4)	(5)	(6)
Butte	20.34	25.70	1070.5	0.0	-122.1
Fresno	5.81	7.63	1088.0	152.0	- 92.0
Kern	2.84	6.33	1429.0	72.0	- 79.5
Madera	5.53	6.42	960.0	150.0	-154.8
Merced	6.36	6.92	1023.6	46.0	-114.5
San Joaquin	4.27	4.81	1132.0	180.0	-153.7
Stanilaus	5.77	7.84	1156.2	134.0	-150.0
North	12.83	17.78	688.0	-514.0	- 66.7
South	7.26	9.58	671.6	-166.0	- 46.0

It should be noted when analyzing the figures in Table 1 that the effect of rainfall varies by region. Because areas of the state receive differing amounts of rainfall on average, the trees in some regions seem better able to tolerate rain than the trees in others. Also, a higher average rainfall in a region can be partially translated into a lower average region specific yield. In this way, some of the effect that the rain has on yield in a high rainfall area can be hidden. For easy reference, these estimated regional variations in yield are also included in the table (with Butte County serving as a base and assigned a value of zero).

As a way to see more clearly how rain affects almond yields and how this varies by region, the estimated loss in yield due to one inch of rainfall in February above normal was calculated for each region. This was done using the estimated production function discussed above and presented in the appendix. These results are found in Table 2. It can be seen that the loss in yield from an additional inch of rainfall can be quite large. Also, the loss from the first inch of rain past the normal amount is the largest in three counties with very low average rainfalls (Madera, San Joaquin, and Stanislaus). In fact, it is interesting to note that the loss is smaller in the county with the lowest average rainfall, Kern. This is in part due to this low average. Remembering that the rainfall is squared, it is not difficult to see that with another inch or two, the loss in yield from an extra inch of rainfall in Kern County would be just as large as for the Counties of Madera, San Joaquin, and Stanislaus.

Table 2. Loss in Yield From One Inch Above Normal Rain

Region	Loss (lbs/acre)
Butte	60.1
Fresno	93.1
Kern	122.4
Madera	159.7
Merced	108.8
San Joaquin	184.8
Stanislaus	150.9

North	42.5
South	35.8

With the figures in Tables 1 and 2 it can be seen that although almonds were found to be alternate bearing by an average of 152 pounds per acre, this variation is often masked by the much larger effect that rainfall has on yield. By using Generalized Least Squares regression techniques it was possible to compute the relative magnitudes of both the rainfall and the alternate bearing effects on almond yields.

Appendix. Generalized Least Squares Estimates
of Almond Production Function

Symbol	Variable	Coefficient	t-ratio
ßo	Intercept	$-.4024 \times 10^4$	5.84
ß1	Bearing acreage (A)	.780	17.90
π2	February rainfall x A	-.003	3.64
d1	Alternate bearing dummy x A	-.076	4.13
	<u>Acreage x Regional Dummies</u>		
ß2	A x Fresno	.076	1.71
ß3	A x Kern	.036	.89
ß4	A x Madera	.075	1.65
ß5	A x Merced	.023	.34
ß6	A x San Joaquin	.090	1.58
ß7	A x Stanislaus	.067	2.56
ß8	A x North Region	-.257	8.37
ß9	A x South Region	-.083	1.72
	<u>Acreage x February Rainfall x Regional Dummies</u>		
π2	A x FR x Fresno	-.005	1.29
π3	A x FR x Kern	-.011	3.94
π4	A x FR x Madera	-.011	2.60
π5	A x FR x Merced	-.006	1.21
π6	A x FR x San Joaquin	-.015	2.82
π7	A x FR x Stanislaus	-.010	4.05
π8	A x FR x North Region	.0004	.51

π9	A x FR x South Region	.0002	.13

$R^2 = .927$ $\bar{R}^2 = .916$ $s = .984$

P_{it} = tons of almond kernels produced in region i, in year t
A_{it} = bearing acreage, in thousands of acres for region i, in year t
R_{kt} = dummy variable equal to 1.0 if the kth region, 0 otherwise, k=2, 9
D_{it} = dummy for alternate bearing, equal to 1.0 in odd years, 0 in even years
FR_{it} = inches of February rain, squared in region i, in year t

Causes of almond yield variations

Jeffrey Dorfman ❑ Melody Dorfman ❑ Dale Heien

Statistical analysis of regional, rainfall-related effects could help early-spring forecasting

California almond production is characterized by wide fluctuations in yield, due mainly to a combination of rainfall and a moderate alternate bearing effect. Because annual additions to acreage have been much smaller than the yield variations, the fluctuations have resulted in large swings in production (fig. 1). In this report, we discuss the variations in almond yield caused by rainfall, regional location, and alternate bearing.

Background

A previous report described a technique [illegible]prove the state's objective almond survey prepared in June of each year (*California Agriculture*, March-April 1987). Although quite accurate at predicting crop size, the state's survey and the improved estimate (the "Dorfman-Heien correction") are not available until July. Because a significant proportion of the almond crop is exported, it is essential to be able to predict total crop production as early as possible. Even though the state forecast appears two months before harvest begins, an earlier forecast would be useful for pricing and for determining the optimal allocation between domestic and foreign sales.

We have developed a production function that allows a preliminary crop forecast in February with later adjustments through the state July forecast and the correction. Although an earlier crop forecast is possible, it should be cautioned that its predictions are not as reliable as either the state July forecast or the Dorfman-Heien forecast. However, the estimates do allow for a study of the effects of rainfall and other factors on yield.

The production of almonds is greatly influenced by rainfall during the bloom period. Almonds typically bloom in February and March, during California's rainy season. Almond trees cannot self-pollinate, but m[illegible]t be pollinated by another almond variet[illegible] [illegible]or this reason, almond orchards always contain at least two varieties of trees, planted either in alternating rows or in two rows of one and one row of the other. Because cross-pollination is necessary, bees are vital to a good crop. If it rains too much during the bloom period, pollination by bees is inadequate and the almond crop is small. This situation occurred in 1986.

We collected rainfall data for the period of February 1 through March 15. Tests revealed February rainfall to be as good an indicator of production as the rainfall from February 15 to March 15, which more accurately reflects the bloom period. Because of earlier availability and ease of collection, we chose February rainfall as the variable to be used in forecasting production.

Methods

To measure the effect of rainfall and of alternate bearing, we statistically estimated a production function for almonds. A production function measures the relationship between inputs and output. The inputs were acreage, rainfall, and a variable for the alternate bearing pattern of almond trees. At a simple level, alternate bearing means that the crop alternates between relatively light and relatively heavy yields due to physiological factors. Almonds are considered moderately alternate bearing. The deviation of the past year's yield from the historical average yield was used to model the alternate bearing pattern.

We began with a relation that indicates yields vary by region, by the amount of rainfall in the region, and by last year's deviation from average yield. Since, by definition, yield equals production divided by bearing acreage, the production relationship can be estimated as yield times bearing acreage. This was then translated into a relation that specified production as a function of bearing acreage times regional effects, bearing acreage times the rainfall effect, and bearing acreage times the alternate yield effect. Bearing acreage represents the number of acres of almond trees four years of age and older. Technically, the trees are not mature enough to bear nuts until the fourth year although a small crop often is obtained now from three-year-old trees. The variable for the deviation in the past year's yield from the average yield is of particular interest, because continual weather-induced variations have made it difficult to determine the magnitude of the alternate bearing effect in almonds.

The production function was estimated as a pooled cross-section time-series model. Cross-section refers to the fact that data came from seven counties and two other regions composed of groups of counties. Time series refers to the fact that these cross-section data are for the years 1971-85. The estimation thus was based on a total of 135 observations. Dummy variables were employed for eight of the nine regions with Butte County as the base region. These variables were used to allow the average yield and rainfall sensitivity to vary by region. The variable for rainfall was inches of rainfall in February squared. Experimentation with rainfall and rainfall squared indicated that rainfall squared performed better. The relationship was estimated by Generalized Least Squares regression. The equation and results are presented in the boxed table.

We used county-level data on almond acreage, production, and rainfall from 1970 to 1985 for the estimation. The acreage and production data came from County Agricultural Commissioner's Reports, and the February rainfall data from the National Oceanographic and Atmospheric Administration. We chose a weather station nearest the center of the almond-growing area in each county. The data were organized into nine regions: the seven counties (Butte, Fresno, Kern, Madera, Merced, San Joaquin, and Stanislaus) and two groups of counties that grow fewer almonds—North (Colusa, Contra Costa, Glenn, Solano, Sutter, Tehama, Yolo, and Yuba)—and South (Kings, San Luis Obispo, and Tulare).

Results

The results of this estimation (shown in the boxed table) allow two interesting effects to be calculated: the effect of rainfall during the pollination period and the magnitude of the alternate bearing effect in almonds. We found the alternate bearing

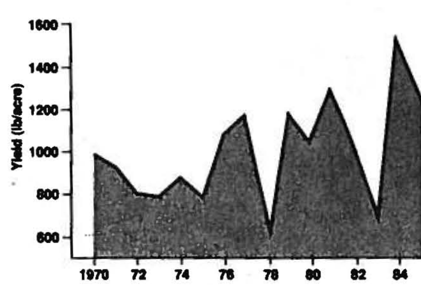

Fig. 1. Large fluctuations in yield are typical of almond production in California.

TABLE 1. Effects of rainfall and location on yield

Region (1)	Rainfall: Average (2)	Rainfall: Std. Dev. (3)	Rainfall: Effect (4)	Location: Effect (5)	Location: Average (6)
	inches	*inches*	*meat lb/acre*	*meat lb/acre*	*meat lb/acre*
Butte	3.54	2.88	−123.9	0.0	1,070.5
Fresno	1.85	1.61	−100.6	152.0	1,088.0
Kern	1.18	1.25	− 63.9	72.0	1,429.0
Madera	1.90	1.44	−129.6	150.0	960.0
Merced	2.10	1.45	−122.6	46.0	1,023.6
San Joaquin	1.73	1.16	−157.6	180.0	1,132.0
Stanislaus	1.87	1.56	−153.5	134.0	1,156.2
North	2.72	2.41	− 68.8	−514.0	688.0
South	2.09	1.76	− 27.8	−166.0	671.6

TABLE 2. Yield loss from rainfall one inch above normal

Region	Loss
	meat lb/acre
Butte	49.2
Fresno	81.4
Kern	75.6
Madera	112.5
Merced	100.2
San Joaquin	164.6
Stanislaus	126.1
North	34.5
South	19.8

phenomenon to be 12.2 percent of the past year's deviation from an average yield. If one year's yield is 10 percent higher than average, the next year's should thus be 1.22 percent below average, holding weather effects constant. Since the average deviation in yield (in absolute value terms) is 249 pounds per acre, the average alternate bearing effect is 30.4 pounds per acre (249 x 0.122). This means that, in an average year, the yield is 30 pounds per acre (2 to 3 percent) larger or smaller than the yield expected, because of the physiological effect on the tree of the past year's crop. This 30-pound alternate bearing effect is in the opposite direction from the past year's deviation in yield from the average. Of course, in a year following a particularly low or high yield, this alternate bearing effect can be considerably larger. In some years, the alternate bearing effect has a magnitude of approximately 100 pounds per acre, or about 8 percent of the yield. These effects vary slightly by region because of differences in each region's average deviation, but all have alternate bearing effects of very similar magnitudes.

Generalized Least Squares Estimates of Almond Production Function

[illegible]

Symbol	Variable	Coefficient	t-ratio
β_0	Intercept	-.3871	4.28
β_1	Bearing acreage (A)	.729	18.25
π_1	February rainfall x A	-.003	3.83
δ_1	Deviation in yield in t-1 x A	-.122	1.84
	Acreage x Regional Dummies		
β_2	A x Fresno	.079	1.59
β_3	A x Kern	.016	.38
β_4	A x Madera	.068	1.19
β_5	A x Merced	.033	.82
β_6	A x San Joaquin	.049	.90
β_7	A x Stanislaus	.073	2.54
β_8	A x North Region	-.281	10.40
β_9	A x South Region	-.094	1.84
	Acreage x February Rainfall x Regional Dummies		
π_2	A x FR x Fresno	-.006	1.38
π_3	A x FR x Kern	.003	2.14
π_4	A x FR x Madera	-.009	2.06
π_5	A x FR x Merced	-.007	2.38
π_6	A x FR x San Joaquin	-.015	3.51
π_7	A x FR x Stanislaus	-.010	3.99
π_8	A x FR x North Region	.0004	.84
π_9	A x FR x South Region	.0011	.53

R^2 = .909 $\bar{R}^2$ = .894 n = 135

When this variation of approximately 2 percent is compared with the fluctuations in figure 1, it is easy to see how this effect was masked by the remaining variation. The rainfall effects are somewhat more complicated, because the coefficients involved are allowed to vary by region.

Table 1 presents some figures on the effects of rainfall and regional location on almond yields. Column 4 gives the estimated loss in yield from average rainfall in February. We calculated this figure by multiplying each region's estimated coefficient for rainfall sensitivity (π_i) by the average February rainfall squared for that region. Column 5 shows the estimated regional variation in yield. These figures are simply the values of the coefficients for the regional dummy variables for acreage (the ß's) converted to meat pounds from meat tons. This variation is due to such factors as differences in soil, climate, orchard age, and cultural practices. The final column presents the average yield for each region.

As indicated in table 1, the effect of rainfall varies by region. Because different areas of the state receive rainfall at different intensities (that is, lots in one day or a slow drizzle for a week), the rainfall in inches does not necessarily represent the same number of days of rain in every region. Since it is primarily the amount of time lost to pollination during rainfall that matters, the differences in the effect of rainfall on the various regions are probably due to differences in the pattern of rainfall. Also, a higher average rainfall in a region can be partially translated into a lower region-specific average yield. In this way, some of the effect that rain has on yield in a high-rainfall area can be hidden. For easy reference, these estimated regional variations in yield are also included in the table (in column 5).

To see more clearly how rain affects almond yields and how this effect varies by region, we calculated each region's estimated yield loss due to 1 inch of rainfall above normal in February. We used the estimated production function previously discussed. To calculate the values presented, the π_i's (remembering to add in the value of π_1, the base sensitivity) from the regression are multiplied by the difference between 1 inch above average rainfall squared and average rainfall squared for each region, then converted to meat pounds from meat tons. For example, for Fresno County, the calculation is:

$$\text{Loss} = (-.003 - .006) \times [(2.845)^2 - (1.845)^2] \times 2000 = -81.4 \text{ meat pounds per acre.}$$

It is evident that the loss in yield from an additional inch of rainfall can be quite large (table 2). Also, the loss from the first inch of rain past the normal amount is the largest in three counties with very low average rainfalls (Madera, San Joaquin, and Stanislaus). It is interesting that the loss is smaller in the county with the lowest average rainfall, Kern. This result is due in part to this low average. Since the rainfall is squared, another inch or two would make the loss in yield from an extra inch of rainfall in Kern County just as large as those for Madera, San Joaquin, and Stanislaus.

Conclusions

By using statistical techniques, we were able to compute the relative magnitudes of the rainfall and alternate-bearing effects in almonds. The results show that, although almonds display an alternate-bearing pattern with an average difference of 30.4 pounds per acre between heavy and light crop years, this variation is often masked by the much larger effect of rainfall on yield. These rainfall effects proved to vary because of the amount and intensity of rainfall in a given region.

Melody Dorfman is Post-graduate Research Assistant, Division of Environmental Studies; Jeffrey Dorfman is a graduate student, and Dale Heien is Professor, Department of Agricultural Economics. All are with the University of California at Davis. *The authors acknowledge the useful comments of Dale Kester, Department of Pomology, UC Davis.*

Improving almond crop forecasts

Jeffrey Dorfman □ Dale Heien

Wide fluctuations in the California almond crop from year to year complicate marketing strategy. Including early-season weather data could improve the accuracy of crop estimates.

The California almond crop fluctuates widely: over the last 10 years it has ranged from 181 to 587 million meat pounds (the weight of the kernels only, excluding the shells) (fig. 1). Such fluctuations make planning a marketing strategy very difficult. In this article, we report on our attempt to use weather information to improve the annual California almond crop forecast.

There are essentially two markets for almonds: one for processing use in cereals, chocolate bars, and the like, and another for direct consumption as raw or smoked nuts. Another dimension is added by export demand, mainly from West Germany, where almonds are used in processing. Because demand by these processors is relatively stable from year to year and the crop varies to such an extent, almond sellers must have a large "swing" market of customers who are flexible in adapting their demand to changing market conditions, especially price. Almond marketers must design their pricing strategies with these factors in mind.

California's almond crop ranges from 181 to 587 million pounds of nut meats per year; demand is relatively stable.

Although almonds can be stored, it is generally desirable to sell all, or almost all, of a given year's crop. In large crop years, some stocks may be held over, but a marketer still must decide what price is needed to sell the desired amount. Almond marketers not only must select a pricing strategy that will clear the market, but also must accurately forecast what that market (harvest) will be. For this they rely on state crop estimates. If these estimates are incorrect, both the marketers and the producers will lose potential profit. Both would benefit from having an early, accurate estimate of the almond crop size.

Model development

Each year the California Crop and Livestock Reporting Service estimates the California almond crop (hereafter referred to as the state estimate). While generally quite good, occasionally the estimate is off by a significant amount from a marketing viewpoint. It takes into account the bearing acreage, a sample of the number of nuts per tree, and various scientific analyses of a large number of sampled nuts. The final state sample is taken in June and the final estimate is released in mid-July. The harvest generally starts in early August, peaks in September, and continues through November or even December. Because of the long harvest season, marketers have to wait months to know the true crop size.

Our goal was to produce an improved early estimate of the California almond crop. Although it seemed unlikely that the state's technical analyses could be improved, the forecasting errors suggested that something might be missing from the model. We began our search with the weather.

A plant needs water, sunlight, and nutrients for growth. Despite the need for water, hard rain during the bloom period can interfere with cross-pollination and severely shrink crop size. The state sample picks up such early-season rain damage, however, so we did not choose rainfall as a weather variable.

Temperature could be considered a fourth class of input. While it is clear that extreme temperatures, both cold and hot, can damage plants, the effects of moderate temperatures are not completely understood. We therefore chose temperature as the weather variable. Temperature also could be used as a proxy for soil moisture because of the proportional relationship between temperature and evapotranspiration.

The next question was how to measure temperature for a crop grown over a large part of the state. The California Almond Growers Exchange (CAGE) provided production records for 10 years (1976-85). Using their data, and treating each locale as typical of the weather throughout the area from which the almonds were gathered, we derived weights for the different almond-growing areas of the state. Next we obtained temperature data for these locales. Finally, computation of a weighted average of the data provided a single value for the entire state. In extending this analysis to other crops, county acreages of bearing orchards could be used to construct the weights.

Two alternative statistics were computed from temperature data. The first was cooling degree days per month. (Cooling degree days are compiled to estimate energy use for air-conditioning in the summer.) One cooling degree day was added for each degree that the daily mear temperature exceeded 65°F. A day with a mean temperature of 70° would be a 5 cooling degree day. Nothing was subtract ed for days with average temperatures below 65°. Each day's value is simply added over the month. The second statis tic computed was the number of days in a month with a high temperature ove 90°F. Both statistics were compiled fo the months of May, June, and July 1976 85 from National Oceanic and Atmo spheric Administration data.

Discussions with Stephen Heinrich and Melody Warfield at CAGE and Un versity of California pomologists Dal Kester and Warren Micke, at Davis, led t the hypothesis that temperatures durin May, June, and July should affect the siz of the crop. There was no definite idea c what mathematical form the relationshi between temperature and crop size woul take. We then specified crop size relatio using alternative mathematical forn and estimated these relations by regre sion techniques. We chose the state est