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 to improve 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 must be pollinated by another almond variety. For 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

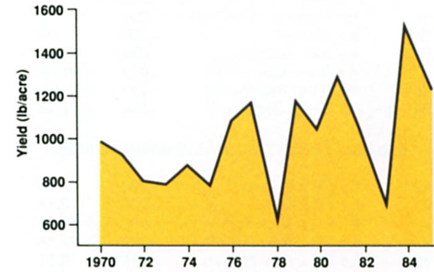

Fig. 1. Large fluctuations in yield are typical of almond production in California.

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.

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.

Generalized Least Squares Estimates of Almond Production Function

$$P_{it} = \beta_0 + \beta_1 A_{it} + \pi_1 A_{it} FR_{it} + d_1 A_{it} D_{it} + \sum_2^9 \beta_i A_{it} R_{kt} + \sum_2^9 \pi_i A_{it} FR_{it} R_{it}$$

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_{it} - dummy variable equal to 1.0 if the kth region, 0 otherwise, i=2, 9
D_{it} - yield in year t-1 in region i minus the region's average yield
FR_{it} - inches of February rain, squared in region i, in year t

Symbol	Variable	Coefficient	t-ratio
ß0	Intercept	- 3671	4.28
ß1	Bearing acreage (A)	. 729	16.23
π1	February rainfall x A	-. 003	3.83
d1	Deviation in yield in t-1 x A	-.122	1.84
	Acreage x Regional Dummies		
ß2	A x Fresno	.079	1.59
ß3	A x Kern	.016	.38
ß4	A x Madera	.056	1.19
ß5	A x Merced	-.033	.92
ß6	A x San Joaquin	.039	.90
ß7	A x Stanislaus	.073	2.54
ß8	A x North Region	-.281	10.40
ß9	A x South Region	-.094	1.84
	Acreage x February Rainfall x Regional Dummies		
π2	A x FR x Fresno	-.006	1.34
π3	A x FR x Kern	-.008	2.14
π4	A x FR x Madera	-.009	2.06
π5	A x FR x Merced	-.007	2.38
π6	A x FR x San Joaquin	-.015	3.51
π7	A x FR x Stanislaus	-.010	3.99
π8	A x FR x North Region	.0004	.64
π9	A x FR x South Region	.0011	.53

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

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.