INCORPORATION OF RISK VARIABLES IN ECONOMETRIC SUPPLY RESPONSE ANALYSIS

by

Bruce Traill

August 1976 No. 76-27

INCORPORATION OF RISK VARIABLES IN ECONOMETRIC SUPPLY RESPONSE ANALYSIS

Bruce Traill*

Many agricultural economists (e.g., D. Gale Johnson) think that risk has an important and adverse influence on farmers' production decisions. Others (e.g., Robinson) believe that the possible adverse effects of risk have been exaggerated. Until recently little attention has been paid to quantifying the effect of risk beyond the individual farm level. In an econometric framework the inclusion of risk variables in supply response analysis may be desirable in order to predict production more accurately, to improve the estimates of parameters in the model, and as an aid to quantifying the impact of government programs that affect uncertainty as well as prices.

In trying to measure risk at an aggregate level, one is confronted with the problem that much of the uncertainty facing individual farmers may not be captured. Thus there is a dilemma in that the more highly aggregated the level of study, the more useful the results, yet the more likely that important risk elements are missed. The two previous studies incorporating risk variables in acreage response analysis (Behrman's and Just's) took a disaggregated approach (areas smaller than states). In this paper, onion acreage response is examined at two highly aggregated levels, namely the total United States crop and the late summer crop which comprises almost two thirds of the U.S. acreage.

The past econometric approaches to risk analysis may be classified under two headings: the simple approach of Behrman and of Freebairn and Rausser who use a moving standard deviation of past actual prices (yields) over a specified period as an explanatory variable, and the complex distributed lag approach of Just (1974) who defines lags on lags and estimates the equations using a two dimensional maximum likelihood search over the unknown lag parameters. In the following sections, an alternative method of incorporating and estimating equations with risk variables is outlined

---

* Bruce Traill is a graduate research assistant, Department of Agricultural Economics, Cornell University. Paper presented at the annual meeting of the American Agricultural Economics Association, August 15-18, 1976, Pennsylvania State University. The author would like to thank Professors Tomek and Mount for their valuable assistance.

and some empirical results from onion acreage response are presented. These are compared with results using standard deviations of past prices. For the sake of brevity, only price risk is considered, though risk associated with yields can also be incorporated into the model quite easily. To keep the model and results manageable, substitutes which may compete for the same resources are not incorporated into the model. This should not cause great concern. Supply equations are often estimated in the absence of substitute crops' prices when no single substitute is very important, and the response of acreage to a price change is interpreted as the result of substitution among a large number of other crops and perhaps as a change in total acreage planted to all crops. A change in risk can be interpreted in the same manner. If a particular crop warrants inclusion as a substitute, the model may readily be extended, though the number of explanatory variables increases rapidly as the number of substitutes increases.

Methodology

The proposed model specifies acreage planted to be a function of the (subjectively determined) expected price of the crop and of its (subjectively determined) expected riskiness.

$$A_t = \beta_0 + \beta_1 P_t^* + \beta_2 W_t^* + \ldots + \epsilon_t$$

where A_t = acreage planted to the crop in period t,

P_t^* = expected price in period t,

W_t^* = expected risk in period t.

In defining W_t^* it is important to make a distinction between uncertainty and instability. Unstable prices, if known in advance, are not uncertain and so would not induce any risk averse reaction. That is, uncertainty results only from a divergence between actual and expected outcomes, and a correct definition of the risk variable should recognize this fact. The standard deviation-type variables used by Behrman and by Freebairn and Rausser define risk merely in terms of the instability of price. It is implicitly assumed that all variation is unexpected.

In this paper, an observation on risk is defined as being a function of the difference between expected and observed prices ($P_t^* - P_t$). There is however, no theoretical basis on which to choose the exact functional form of the relationship. As Dillon says "Like beauty, risk lies in the eyes of the beholder!" (p. 28). But most commonly investigators look at either the squared or absolute deviations of the differences.

Assuming expected price is a function of past observations on price ($P_t^* = \sum_{i=1}^{n} \gamma_i P_{t-i}$) and assuming expected risk is a function of past observations on risk (say $W_t^* = \sum_{j=1}^{m} \delta_j |P_{t-j}^* - P_{t-j}|$), then the model becomes $A_t = \beta_o + \beta_1 \sum_{i=1}^{n} \gamma_i P_{t-i} + \beta_2 \sum_{j=1}^{m} \delta_j |P_{t-j}^* - P_{t-j}| + \ldots$ The estimation problem involves defining P_{t-j}^* and forming the risk variable. Clearly the parameters of the lagged prices must be known.

A possible way of estimating the equation utilizes an iterative procedure as contrasted to the maximum likelihood method employed by Just. The procedure is relatively simple to apply and allows considerable flexibility in defining the risk variable and in selecting the form of the lag structures.

In the first step, an equation is run omitting the risk variable but including the price variable in some distributed lag form as well as other supply shifters. The estimated coefficients of the lagged prices are then used to form a first-round estimate of P_t^* which is employed to define the risk variable. Using an Almon-type distributed lag, for which results are presented, the model

$$A_t = \beta_o + \beta_1 P_t^* + \ldots$$

yields estimates

$$A_t = b_o + c_1 P_{t-1} + c_2 P_{t-2} + c_3 P_{t-3} + \ldots$$

(A three period lag is illustrated for simplicity.)

Constraining the weights to sum to one,

$$\hat{P}_t^* = \frac{c_1}{c_1+c_2+c_3} P_{t-1} + \frac{c_2}{c_1+c_2+c_3} P_{t-2} + \frac{c_3}{c_1+c_2+c_3} P_{t-3} .$$

In the next step, observations on risk are calculated (in this paper, the absolute deviations definition is used as it provides the best results) and the full equation estimated. Of course, inclusion of the risk variable results in a change in the estimated coefficients of the lagged prices, meaning that the values of $\hat{P}_t^*$ used in the initial definition of the risk variable are not those implied by the reestimated equation. Therefore, expected price is recomputed using the new price weights, the risk variable redefined, and the equation estimated once more. The process is repeated until the change in $\hat{P}_t^*$ from one round to another is small.

Pros and Cons of the Estimating Procedure and Risk Definition

The principal advantage of the method described is that it allows risk to be specified in terms of deviations between actual and expected prices, yet is relatively simple to apply and affords considerable flexibility in specifying the lag structures and the form of the risk variable. However, since the definition of the risk variable depends on expected price, the definition of expected price is critically important. The assumption that the lag structure on prices is generated solely by an expectations mechanism may be invalid, and if this is the case (e.g., if an adjustment process is also an important cause of a lagged response of acreage to price), then the definitions of expected price and of risk are incorrect. In this case, a simple risk definition not relying on an expected price generated within the model may be more reliable.

A further problem in specifying the risk variable is the likelihood of an interaction between price and risk. The model assumes the same response to a given expected price risk at all levels of expected price. More reasonably, one might argue that a large expected risk causes a greater response when expected price is low than when it is high. To examine whether this is important, an alternative risk variable is defined as the probability of price falling below some arbitrarily low level. This variable is calculated by fitting a moving (lognormal) probability distribution to prices and thus depends on both estimated parameters of the distribution[1/], but not on expected price as implied by the price lag structure.

Empirical Results

Results using 1950-74 data for the late summer and the total United States onion crops are presented in tables 1 and 2. The decision to study onion acreage response is based largely on the high level of variability of onion prices; coefficients of variation of prices for the late summer and U.S. crops are around 40 percent over the period 1946-74. For yields, on the other hand, they are only around 6 percent (trend corrected), which provides some justification for examining only price risk. Also, there appears to be no single major substitute for onions which would warrant inclusion in the model.

The conventional supply response equations, the starting point for the inclusion of risk, are based on some experimentation. They include a curvilinear trend and a dummy variable for 1951. The price-acreage relationship was specified in terms of both Almon and geometric-form distributed lags, and an Almon specification (second degree polynomial and four period lag) was selected as providing the most reasonable results. (However, the hump-shaped lag structure on prices for the late summer crop was not expected *a priori*. One would anticipate the most recent prices as being most important in the formation of expectations; hence the possibility that an adjustment process is also operative might be considered.)

In fitting the risk variables by the iterative procedure, an Almon specification with a second degree polynomial and a four period lag was first selected (the same as for the price lags), followed by experimentation to select the "best" lag length. A two period lag seemed appropriate for risk, and the lag structure on prices was not altered sufficiently to warrant a change in its specification. In fact, all changes in coefficients after

---

[1/] To be more precise, the procedure involves first, testing for randomness of prices using a runs test, and then fitting a lognormal distribution to the whole price series, testing for goodness of fit using a χ^2 test. Given that the lognormal distribution provides an adequate description of the data, a moving (two-period) mean and standard deviation are computed and the area in the lower tail is calculated (in this case representing the probability of price falling below 2). This estimated probability is entered as an observation for the succeeding year.

Table 1. Onion Acreage Response, the United States, 1950-74

	Type of Model			
	Conventional	Final Round Iterative	Moving Standard Deviation	Moving Probability Distribution
Constant	117.089 (10.620)	121.681 (11.037)	121.544 (11.236)	121.099 (10.554)
Ln T	-19.498 (-10.433)	-20.767 (-10.313)	-20.036 (-11.085)	-20.041 (-10.481)
D_{1951}	-30.864 (-3.837)	-29.209 (-3.353)	-28.811 (-3.715)	-30.821 (-3.864)
P_{t-1}	5.356 (3.739)	5.969 (3.957)	5.528 (4.041)	4.553 (2.878)
P_{t-2}	3.390 (2.912)	3.982 (3.269)	3.526 (3.172)	3.281 (2.832)
P_{t-3}	2.110 (1.667)	2.276 (1.706)	2.035 (1.686)	2.318 (1.828)
P_{t-4}	1.515 (.905)	.848 (.506)	1.053 (.652)	1.665 (1.000)
Sum	12.371 (3.471)	13.075 (3.714)	12.142 (3.574)	11.817 (3.313)
Abs_{t-1}		-3.105 (-1.428)		
Abs_{t-2}		-2.391 (-1.058)		
Sum		-5.496 (-1.741)		
$\hat{\sigma}_{p,t}$			-4.010 (-1.715)	
$Pr(P \leq 2)_{,t}$				-13.650 (-1.150)
R^2	.856	.879	.876	.866
d	1.078	1.089	1.336	.933
s	6.150	5.976	5.858	6.098
U_2	.410	.379	.424	.368

Table 1. (Continued)

Dependent variable is 1000 acres planted (mean = 113.4). T is a trend variable, equal to 1 in 1950. D_{1951} is a dummy variable for 1951. P_{t-i} (i=1,2,3,4) are prices in \$/cwt. deflated by an index of prices paid by farmers for all production items (mean ≈ 3.50). $Abs_{t-i} = |P^*_{t-i} - P_{t-i}|$ (mean ≈ 0.8). $\hat{\sigma}_{p,t}$ is a moving standard deviation of deflated prices for the two years preceding t (mean = .632), $Pr(P \leq 2)_{,t}$ is an estimated probability of deflated price falling below \$2/cwt., calculated by fitting a lognormal probability density function to the previous two years' prices (mean = .058), d is the Durbin-Watson Statistic. s is the standard error of the estimate, U_2 is Theil's inequality coefficient. Figures in parentheses are t ratios. The Almon lag on prices used a second degree polynomial, with no constraints on the end-points. For the iterative procedure, convergence was obtained after three rounds.

Table 2. Onion Acreage Response, Late Summer Crop, 1950-74

	Type of Model			
	Conventional	Final Round Iterative	Moving Standard Deviation	Moving Probability Distribution
Constant	54.132 (9.057)	51.397 (8.187)	52.597 (8.527)	53.739 (9.092)
T	-1.391 (-4.844)	-1.430 (-4.850)	-1.405 (-4.833)	-1.443 (-5.055)
T^2	.053 (4.464)	.052 (4.223)	.052 (4.331)	.053 (4.475)
P_{t-1}	.974 (2.466)	1.571 (2.490)	1.393 (2.393)	.979 (1.776)
P_{t-2}	1.387 (2.566)	1.934 (2.919)	1.624 (2.706)	1.509 (2.767)
P_{t-3}	1.322 (2.590)	1.700 (2.695)	1.428 (2.470)	1.500 (2.654)
P_{t-4}	.750 (1.534)	.867 (1.678)	.805 (1.621)	.951 (1.901)
Sum	4.408 (2.590)	6.072 (3.049)	5.249 (2.829)	4.939 (2.886)
Abs_{t-1}		-.993 (-1.251)		
Abs_{t-2}		-1.070 (-1.216)		
Sum		-2.063 (-1.567)		
$\hat{\sigma}_{p,t}$			-.974 (-1.100)	
$Pr(P \leq 2)_{,t}$				-5.982 (-1.481)
R^2	.806	.834	.822	.831
d	1.338	1.454	1.396	1.463
s	2.052	2.027	2.040	1.990
U_2	.681	.658	.668	.667

Table 2. (Continued)

Variables are defined as for the U.S. crop (Table 1). Means: Acreage Planted, 61.90; Price, deflated, 3.2; Abs, 0.86; $\hat{\sigma}_p$, 0.822; Pr($P \leq 2$), 0.085. For the iterative procedure, convergence was obtained after two rounds.

the first round were relatively minor in nature.

The coefficients on these risk variables have the anticipated signs, but do not dramatically improve the fit of the estimated equations. Graphs 1 and 2 show the contributions of each of the variables in the final round equation to explaining the variation in the dependent variable (all variables are expressed as deviations from means and prices and risk are condensed to P_t^* and W_t^*).[2/] Evidently, prices are substantially more important than risk in explaining changes in onion acreage. For the total U.S. crop, there was relatively little change in the magnitudes of the price coefficients in moving from the conventional to the risk model, but for the late summer crop the sum of the price coefficients increased 38 percent.

In calculating values for the standard deviation and probability distribution risk definitions, a moving two year period is used, this being the same as the length calculated on risk by the iterative procedure. As tables 1 and 2 indicate, in terms of the signs and magnitudes of the models' coefficients and in terms of goodness of fit statistics (R^2, S, U_2), there is little to choose between the alternatives. Although superficially this might lead one to conclude that there is no point in proceeding with the more cumbersome iterative procedure, it should be remembered that without having first run the iterative model there would have been no basis for choosing a two year moving period for these variables. In fact, the results are very sensitive to the length of period chosen. For example, a four year moving period (the same as the period over which price expectations are formulated) yields perverse signs on coefficients of both the standard deviation and probability distribution risk variables. Thus even employing these seemingly simpler models, one is faced with a cumbersome search for the optimal length of moving period.

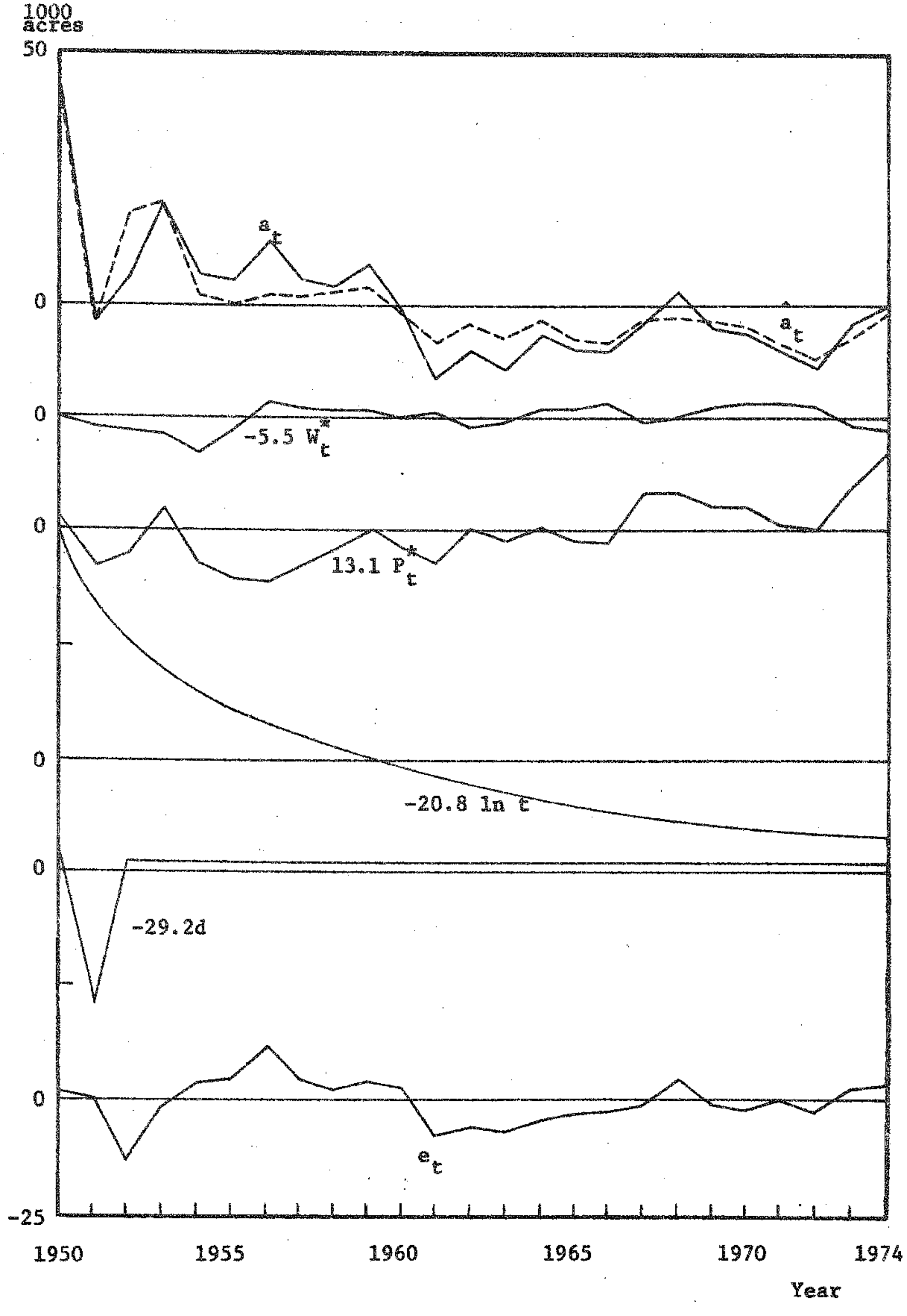
Conclusion

While general conclusions can't be drawn from the results of a single empirical study, it appears possible to incorporate risk variables by the method described and obtain seemingly reasonable results even at a highly

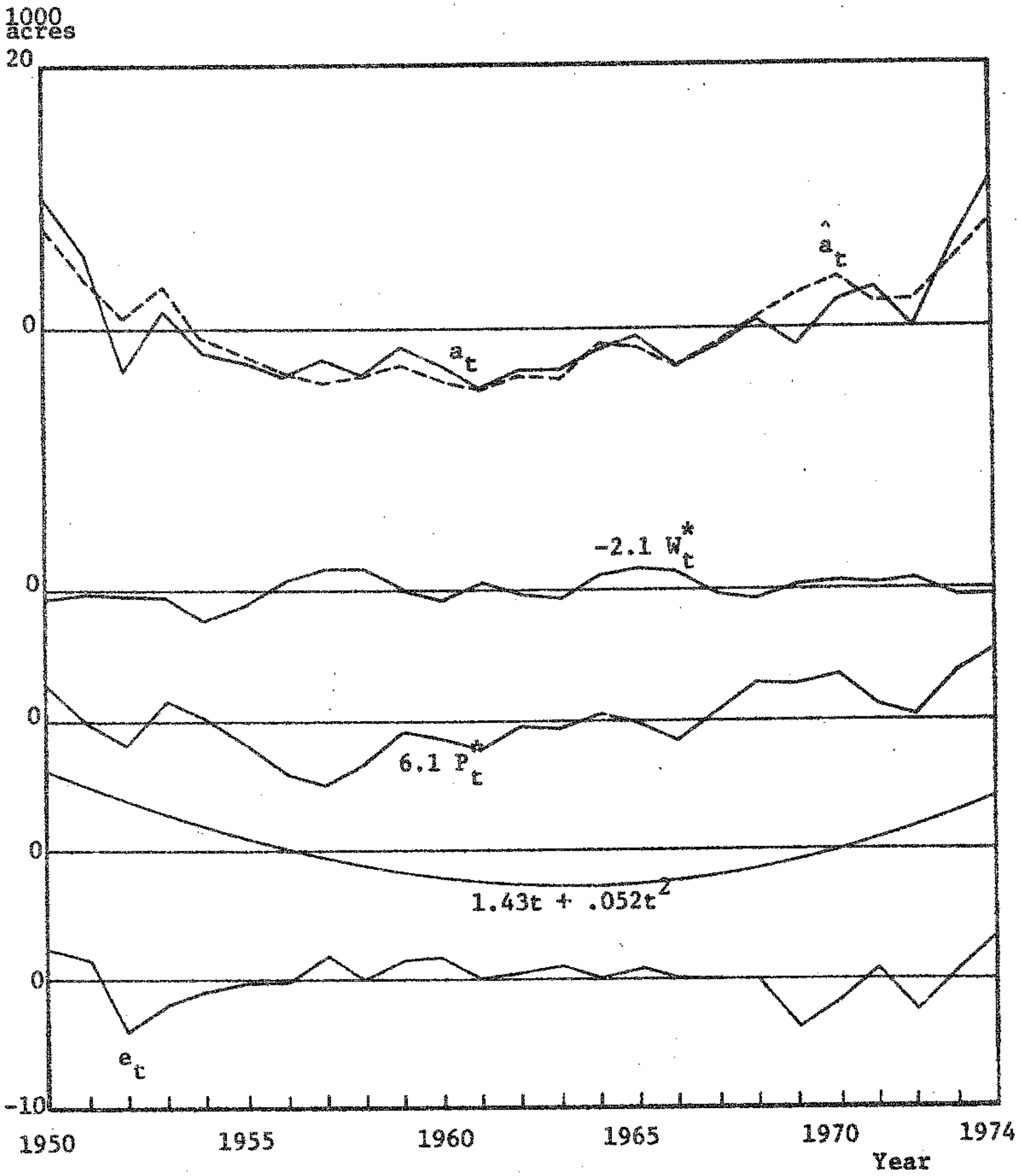
---

[2/] These diagrams are proposed and discussed in Theil (pp. 182-185).

Graph 1. Contribution of Independent Variables to Explaining Variation in the Dependent Variable, The United States



Graph 2. Contribution of the Independent Variables to Explaining Variation of the Dependent Variable, Late Summer Crop



aggregated level. Also, price coefficients for the late summer crop change substantially with the inclusion of risk variables, which may indicate that risk should be considered if accurate estimates of supply elasticities are desired. On the other hand, predictive power of the models does not appear to be greatly affected (a trace of the predicted acreage from the conventional model and from each of the three risk models is virtually indistinguishable). Depending on the purpose of the model and the anticipated importance of risk in the crop under consideration, one could argue, given its relative ease, for at least including risk variables through one or two iterations.

References

Behrman, J. R. Supply Response in Underdeveloped Agriculture. A Case Study of Four Major Annual Crops in Thailand 1937-1963. North Holland, Amsterdam, 1968.

Dillon, J. L. "An Expository Review of Bernoullian Decision Theory in Agriculture: Is Utility Futility?" Review of Marketing and Agricultural Economics, 39(1971):3-80.

Freebairn, J. W. and G. C. Rausser. "Effects of Changes in the Level of U.S. Beef Imports." Amer. J. Agr. Econ. 57(1975):676-688.

Johnson, D. Gale. Forward Prices for Agriculture. Univ. of Chicago Press, 1947.

Just, R. E. "An Investigation of the Importance of Risk in Farmers' Decisions." Amer. J. Agr. Econ. 56(1974):14-25.

__________. "Risk Response Models and Their Use in Agricultural Policy Evaluation." Amer. J. Agr. Econ. 57(1975):836-843.

Robinson, K. L. "Unstable Farm Prices: Economic Consequences and Policy Options." Amer. J. Agr. Econ. 57(1975):769-777.

Theil, Henri. Principles of Econometrics, John Wiley and Sons, 1971.