Effects of Soybean Checkoff Research Expenditures on U.S. Soybean Yields and Net Revenue: A Time Series Analysis

David A. Bessler*

TAMRC Commodity Market Research
Report No. CM-02-09

April 2009

* David A. Bessler is a TAMRC Associate and a professor in the Department of Agricultural Economics, Texas A&M University, College Station, TX 77843-2124.

Effects of Soybean Checkoff Research Expenditures on U.S. Soybean Yields and Net Revenue: A Time Series Analysis

Texas Agribusiness Market Research Center (TAMRC) Commodity Market Research Report No. CM-02-09, April 2009 by Dr. David A. Bessler.

ABSTRACT:
Statistical methods are used to study relationships between research expenditures (adjusted for inflation) made from producer checkoff programs and soybean yields and net revenues in the United States for the years 1978 – 2007. Results presented are for yield and net revenue data and research expenditures for the entire United States. We find that research expenditures over the years 1994 to 2007 are responsible for a 0.95 bushel per acre per year increase in soybean yields. We calculate net producer revenues to be about $17 per acre higher than would have been the case without the soybean research checkoff expenditures.

ACKNOWLEDGEMENTS:

The Texas Agribusiness Market Research Center (TAMRC) has been providing timely, unique, and professional research on a wide range of issues relating to agricultural and agribusiness markets and products of importance to Texas and the nation for forty years. ***TAMRC*** *is a market research service of AgriLife Research and AgriLife Extension. The mission of* ***TAMRC*** *is to provide high quality, objective, and timely market research to support strategic agribusiness decision-making at all levels along the supply chain from producers to processors, wholesalers, retailers, and consumers. Major* ***TAMRC*** *research divisions include International Market Research, Consumer and Product Market Research, Commodity Market Research, and Contemporary Market Issues Research.*

Effects of Soybean Checkoff Research Expenditures on U.S. Soybean Yields and Net Revenue: A Time Series Analysis

EXECUTIVE SUMMARY

Statistical methods are used in this study to analyze the relationships between research expenditures (adjusted for inflation) made by the soybean checkoff program and U.S. soybean yields and net revenues in the United States for the years 1978 – 2007. Data on yield, acreage, production, price and cost of production are taken from open record sources as published in TAMRCb (2009). The checkoff research expenditures data are those collected for the United Soybean Board by Keith Smith and Associates and published in TAMRCa (2009). Observations on the consumer price index (CPI), to convert nominal expenditures to real expenditures, are obtained from the Bureau of Labor Statistics.

There was a gap in the research expenditure data for the years 1996 – 1999. Data points for those years were forecasted and used by following a random walk model. Results presented here are for yield and net revenue data and research expenditures for the entire United States. We do not break-out responses of individual states or regions. Following recent literature on assessment of aggregate research benefits, we construct a "research stock" variable, which is a weighted average of research expenditures for research expenditures made in the previous four to ten years. Models studying the lagged relationship between this research stock variable and U.S. aggregate yields and U.S. aggregate net revenues are specified using information criteria of Hannan and Quinn.

We find that research expenditures over the years 1994 to 2007 are responsible for a 0.95 bushel per acre per year increase in soybean yields. Further, we consider conditional forecasts of net revenues under the condition that soybean checkoff dollars are set to zero. Under this "counterfactual" scenario, we calculate that net revenues are actually about $17 per acre higher than would have been the case in the absence of the expenditures.

Table of Contents

Effects of Soybean Checkoff Research Expenditures on U.S. Soybean Yields and Net Revenue: A Time Series Analysis

This report presents the results of our investigation of the effects of U.S. checkoff research expenditures on U.S. soybean production yields and associated net revenues. The data used in our study are observed in time sequence and thus time series econometric methods are applied. The checkoff research data are those obtained from Keith Smith and Associates and published in TAMRCa. These data are measured yearly over the period 1978 through 2007. Soybean yield, acreage, production, price and cost of production data are obtained from USDA and other government and other open source records (TAMRCb). Observations on the consumer price index (CPI), to convert nominal expenditures to real expenditures, are obtained from the Bureau of Labor Statistics. A gap in the data for the years 1996 – 2000 existed, data points for these years were interpolated following a random walk model.

The analysis is presented in six sections. The first section offers a graphical and statistical summary of each of the major data series used in the analysis. The second section offers a summary of the methods used for model selection and the model selected for further analysis. The third section presents results summarizing the effects of checkoff expenditures on soybean yields and net revenues. The fourth section offers results on a simulation of net returns under the conditional that research stock (research expenditures) are held at zero. A short discussion ends the main body of the report. Two appendixes complete the report.

Graphical Presentation of and Summary Statistics on U.S. Soybean Yields, Net Revenues and Research Expenditures

We begin by offering summary statistics on each data series and their time series plots in this section. Then in the following section, these initial results are followed by formal tests of their non-stationary behavior over time. Results on identifying the optimal lagged relationship between research expenditure and yield response in the third section.

Table 1 gives summary statistics on U.S. soybean per acre yields, planted acreage, production, gross revenue, net revenue, checkoff research expenditures and a computed series of research stock from data measured annually over the years 1978 – 2007. Mean yields are 35.22 bushels per planted acre, with a low yield of 25.7 bushels per acre observed in 1983 and a high of 43.00 bushels per acre observed in 2005. Average plantings of soybeans over this period is 66,726,483 acres; with a low of just under 58 million acres in 1990 and a high of 75.5 million acres in 2006. Production averaged 2.3 billion bushels over our observation period, with a low of just over 1.5 billion bushels in 1998 and a high of about 3.2 billion bushels in 2006. These production levels translated, at nominal prices, to average gross revenues of just over 14 billion dollars. The low of 9.2 billion dollars occurred in 1986 and the high of 26.8 billion dollars in 2007. Net revenues (gross revenues – cost of production) averaged approximately 9.2 billion dollars, with the associated low of 6.3 billion dollars in 1986 and the high of 20 billion dollars in 2007. Checkoff research expenditure data averaged 11.5 million dollars over the 1978 – 2007 period, with a low

Table 1. Descriptive Statistics on 1979- 2007 U.S. Soybean Data.

Series	Mean	Standard Deviation	Minimum (Date)	Maximum (Date)
Yield (Bushels/Acre)	35.22	4.91	25.7 (1983)	43.00 (2005)
Acreage	66,726,483.	6,135,523.	57,795,000. (1990)	75,522,000 (2006)
Production (Bushels)	2,314,733,414.	458,790,844.	1,548,841,000. (1998)	3,188247,000 (2006)
Gross Revenue (Dollars)	14,002,480,356.	3,696,803,836.	9,273,683,187. (1986)	26,886,152,800. (2007)
Net Revenue (Dollars)	9,198,920,184.	2,926,108,926.	6,309,987,387. (1986)	20,076,999,490. (2007)
Research Expenditures (Dollars)	11,520,342.	9,791,589.	2,291,268. (1979)	30,257,513 (2007)
Research Stock (Dollars	5,937,522.	5,217393.	1,407,870 (1979)	20,033519. (2007)

Statistics calculations presented here begin in the year 1979 because we need previous data as start-up to begin calculation research stock. Actual research expenditures provided by Keith Smith and Associates begin in 1978. Data on other series go back much earlier, but are not used here, as we wish to focus attention on the relationship between research expenditures and soybean production and revenues.

of about 2.3 million in 1979 and a high of 30.26 million dollars in 2007. Our computed research stock variable (see footnote to table 1 for its formula) averaged 5.9 million dollars with a low of 1.4 million dollars in 1979 and a high of 20 million dollars in 2007.

The justification for this last variable rests on two main factors. First is the recognition that research benefits are not immediate, a lag exists from the time the expenditures are made and possible real time adoption of results in the field. Second, research results from many years ago may still be yielding benefits for several years into the future. To accommodate both factors we compute a "research stock" variable as a weighted average of research expenditures over the previous ten years. More specifically, Stock of Research measured at year t (S_t) is a convex combination of research expenditures for the years t-4, t-5, t-6, t-7, t-8, t-9, t-10, with respective weights of .10, .20, .20, .20, .20, .05, and .05. Other weight patterns were explored and gave generally the same results (see Alston, Norton and Pardey 1998, chapter 3 for a general discussion of this issue).

Plots of each series used in our analysis are given in Figure 1. Generally, we see a saw-tooth upward growth in bushels per acre from 1978 to 2007; which coheres well with the upward movement in soybean production. Interesting is the bi-modal plot of U.S. soybean acreage, reaching local maxima in 1979, then generally declining until 1990, thereafter growing (generally), reaching its maximum in 2006, with a sharp drop in 1007. Net returns generally trended downward from 1979 until 1990, thereafter increasing until 1996, and then dropping sharply through 2001, increasing sharply thereafter. The two research variable, actual expenditures and research stock follow the same general modest growth path up to 1999, with rather sharp increases thereafter. The research stock variable is a "smoothed" version of the actual research expenditure variable.

The general upward trend in several of the variables suggests that we may be working with non-stationary variables. That is to say, the series do not revert to their historical mean values. If such is the case we will either model the series jointly as an error correction model or model them jointly as a levels VAR. However, in the latter case we need to be sure the resulting model has stationary residuals. Table 2 offers a summary of tests of the null hypothesis of non-stationarity for each series (Dickey and Fuller 1981). We fail to reject the null for each, suggesting that either we consider an error correction model on how these series interact or, if we use a vector autoregression, we offer evidence on the stationarity of resulting innovations. For more on this latter modeling strategy see Nerlove, Grether and Carvalho(NGC) 1979. There they suggest under a levels vector autoregressions we follow a "plausible principle of model formulation is that nonstationarities in the "explanatory" variable ought, if possible, to explain nonstationarities in the dependent variable." (NGC, p. 232). For additional motivation for study with the vector autoregression, see the idea of "balancing regression" in Granger (1981).

Loss Metrics on Yield and Net Revenue Responses to Previous Levels of Research Expenditures

Modeling the relationship between research expenditures and yields and net returns requires an appreciation for possible causal relations generating the observed data. In the question under consideration here the research dollars are generated by checkoff money, which in-turn is based

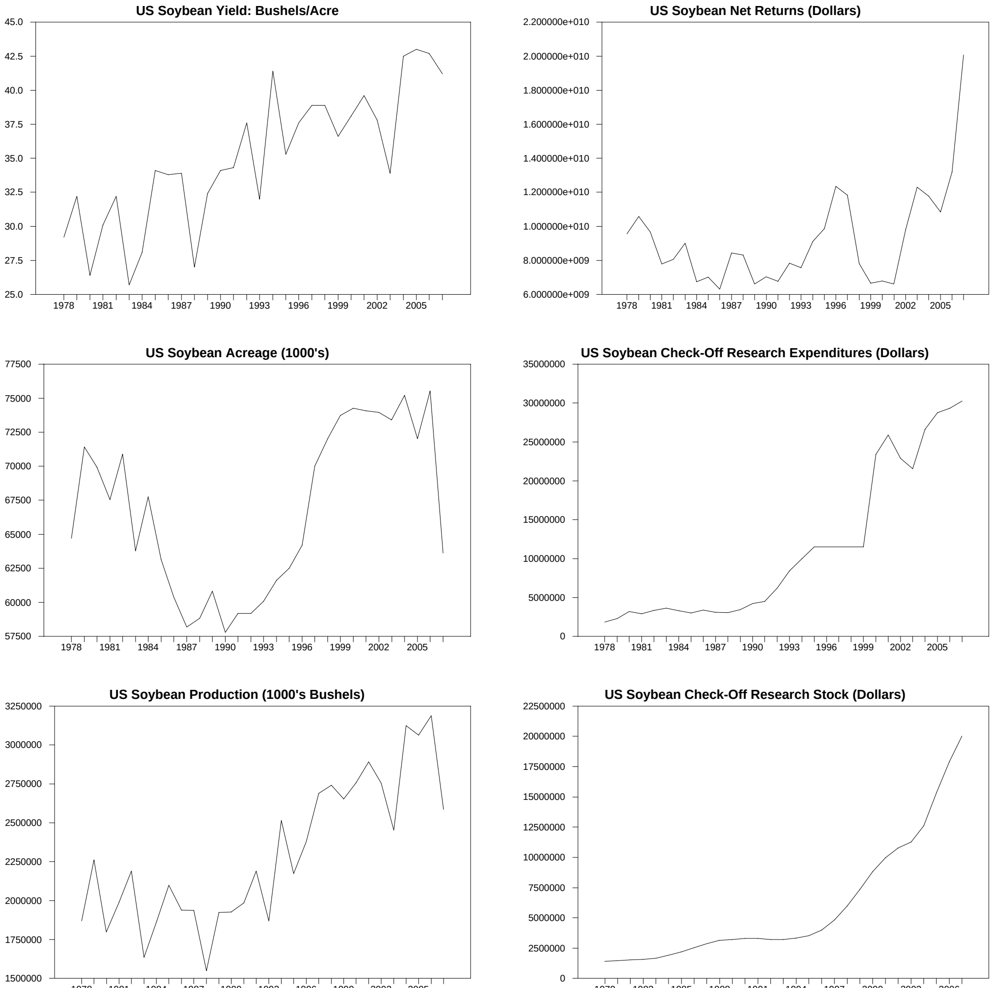


Figure 1. Time Series Plots of Soybean Yields, Acreage, Production, Net Returns, Checkoff Research Expenditures and Checkoff Research Stock, 1979 – 2007.

Table 2. Tests on Non-Stationarity of U.S. Soybean Data, 1978- 2007.

Series	t-stat
Yield (Bushels/Acre)	-2.32*
Acreage	-1.59*
Production (Bushels)	-1.83*
Gross Revenue (Dollars)	+1.21*
Net Revenue (Dollars)	+.13*
Research Expenditures (Dollars)	+.39*
Research Stock (Dollars	+8.65*

Note on Table 2: The test is on the null hypothesis that the data on the series listed in the left-hand column are non-stationary in levels (non-differenced data). The test for each series is based on an ordinary least squares regression of the first differences of the levels of each state's yield on a constant and one lag of the levels of that state's yield. The t-test is associated with the estimated coefficient on the lagged levels variable from this regression. Under the null hypothesis (non-stationary yields) the t-statistic is distributed as a non-standard student-t. Critical values are given in Fuller [9]. The 5% critical value is –2.89. We reject the null for observed t values less than this critical value. An asterisk (*) signifies failure to reject the null at the 5% level of significance.

on actual yields (production) of soybean producers. A simple model summarizing the relationships among variables A,B,C,D, and E is as follows:

A: (Soybean Production (yields)$_{t-k}$) →
B: (Checkoff Dollars$_{t-k+1}$) →
C: (Research Expenditures$_{t-k+1}$) →
D: (Soybean Production Practices and Input Levels$_{t}$) →
E: (Soybean Production (yields and net revenues)$_{t}$)

The time sequence is important. There will be a non-trivial lag between research expenditures and changes (if any) in production. Further, research findings manifest themselves through changes in either input levels or production practices, so one needs to respect the chain: C → D →E. If we include D as an explanatory variable (along with C) in an equation summarizing the causal relationship between C and E, variable D will block (distort) any estimated relationship between C and E (this is known as blocking the front door path between C and E in Pearl's terminology (Pearl (2000)).

Finally, we should note from the above schematic that research expenditures are themselves not exogeneous (independent). They depend on previous production levels according to the checkoff program (checkoff dollars available for research depend on production in previous year). Production or yield levels made in the year before the research expenditures must be included in any equation estimated to summarize the relation between research expenditures and yields (this is known as 'blocking the back-door path in Pearl's terminology (Pearl (2000)).

The general model we explore here is the levels vector autoregression as given in equation (1):

$$\begin{bmatrix} Y_t \\ R_t \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} a_{111} & a_{121} \\ a_{211} & a_{221} \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ R_{t-1} \end{bmatrix} + ... + \begin{bmatrix} a_{11k} & a_{12k} \\ a_{21k} & a_{22k} \end{bmatrix} \begin{bmatrix} Y_{t-k} \\ R_{t-k} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \qquad [1]$$

Here current values of Y (Yields) and R (research expenditures) are expressed by past values of of yields and research, with lags 1 to some upper value k, and two white noise innovation terms ε_{1t} and ε_{2t}. The problem in specification is how may lags are behind the generation of Y and R. And what delay pattern is behind the affect of R on Y?

A similar model was explored for net returns (NR), as given in equation (2):

$$\begin{bmatrix} NR_t \\ R_t \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} a_{111} & a_{121} \\ a_{211} & a_{221} \end{bmatrix} \begin{bmatrix} NR_{t-1} \\ R_{t-1} \end{bmatrix} + ... + \begin{bmatrix} a_{11k} & a_{12k} \\ a_{21k} & a_{22k} \end{bmatrix} \begin{bmatrix} NR_{t-k} \\ R_{t-k} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \qquad [2]$$

Here u_{1t} and u_{2t} are current period innovations (information shocks) in net revenues (NR) and research expenditures (R), respectively.

Our engine for lag specification in models 1 and 2 is taken from the specification search literature, balancing fit and forecasts (or parsimony) in possible models. The metrics used is that of Hannan and Quinn, given as:

$$\Phi_k = \ln(\sigma_k^2) + [2k(\ln(\ln(T)))]/T \qquad [3]$$

The metric has two components: the first term, $\ln(\sigma_k^2)$ falls as we add more terms to the right hand side of equation (2), where ln is the natural logarithm transformation and σ_k^2 is the residual error variance from fit versions of the model with k lags. The second terms on the right hand side of Φ_k is a penalty functions, which increase with more complex models. We select that model specification that minimizes the metric (sum of both terms). For more discussion on this metric see Geweke and Meese (1981).

To construct a Research Stock variable, it is crucial to have a good idea of the delay between actual expenditures of checkoff dollars and their adoption and use in the field. Equation (3) can be used for such a purpose. In figure 2, we offer loss metric on the period of delay between U.S. checkoff research expenditures and U.S. soybean yields. We see a delay of four periods (not three and not five) results in the lowest Hannan and Quinn measure. Accordingly we adopt a four period delay in constructing our research stock variable. We (somewhat arbitrarily) select an upper bound of ten periods (we cold not search over higher lags due to degrees of freedom limitations). Our stock variable is constructed as follows:

$$S_t = .1R_{t-4} + .2R_{t-5} + .2R_{t-6} + .2R_{t-7} + .2R_{t-8} + .05R_{t-9} + .05R_{t-10} \qquad [4]$$

Other weighting schemes to map past research into a current stock variable were studied, all gave similar results as we report here. [We use the same stock measure (equation (4)) for both yields models and net revenue models below.]

Results given below are based on models relating values of this research stock variable (S_t) to current values of soybean yields and soybean net revenues. Both revenues and research stock variables are expressed in "real" U.S. dollars, where the Consumer Price Index (CPI) is used to convert nominal U.S. dollars to real dollars.

Table 3 gives Hannan and Quinn Loss metrics on alternative lags of U.S. soybean yields on past values of soybean yields, both without and with the current value of research stock in the equation. Table 3 explores the model generating yields, as represented by equation (5):

$$Y_t = c_1 + a_{111}Y_{t-1} + a_{112}Y_{t-2} + ... + a_{115}Y_{t-5} + a_{12}S_t \qquad [5]$$

Notice that Hannan and Quinn Loss is minimized at one lag of yields and current value of research stock variable (Φ is minimized with a value of 2.63 at one lag of yields and current research stock).

A similar search was conducted on Net Returns; results are reported in Table 4. Here we find two lags of net returns and current value of research stock generate the data (Φ is minimized with a value of 43.08 at two lags of net returns and the current value of research stock). Similar search over models generating research stock were conducted, were we find four lags of the research stock variable and 5 to 9 lags of yields and returns. We do not allow earlier lags of yields to affect research stock because research stock is itself (by rules governing the generation

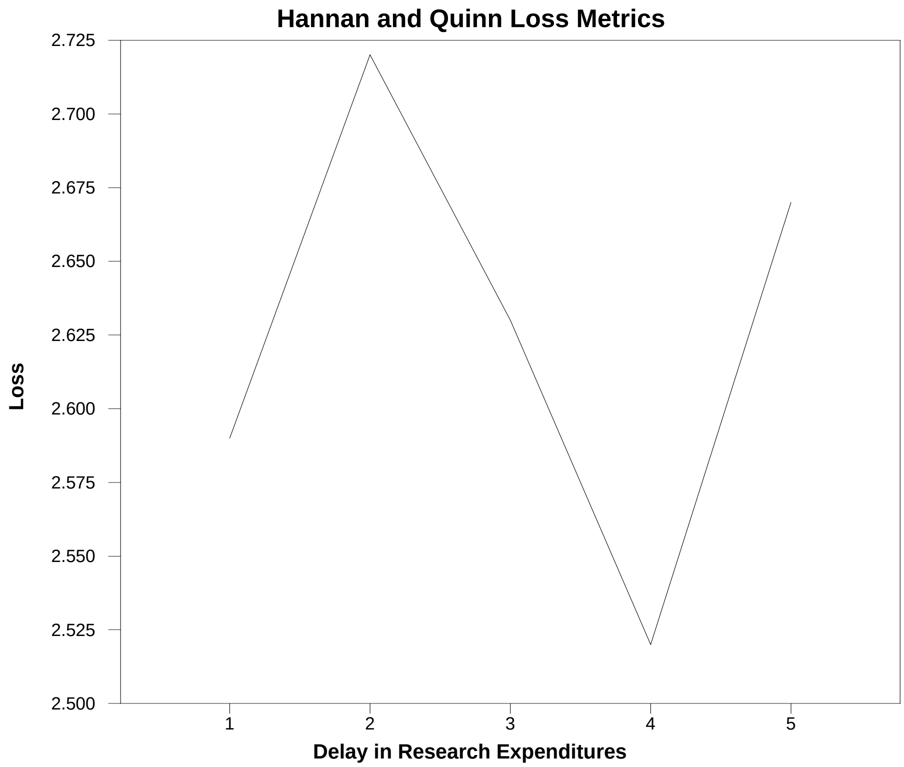


Figure 2. Hannan and Quinn Loss in Delay of Affect of Checkoff Research Expenditures on U.S. Soybean Yields.

Table 3. Statistical Loss Metrics on Oder of Lags in Generating U.S. Soybean Yields, without and with Research Stock, 1978- 2007.

Lags = k	Hannan and Quinn Loss (Φ)
[With No Research Stock]	
Constant Only	3.00
-1	2.75
-2	2.77
-3	2.88
-4	2.89
-5	2.86
[With Research Stock]	
Constant	2.70
-1	2.63 *
-2	2.74
-3	2.89
-4	2.94
-5	2.96

Hannan and Quinn's measure on lag length (k) of a levels vector autoregression

$$\Phi = \log(\sigma^2_k) + (2.00)(k+1) \times (\log(\log T))/T$$

where σ^2_k is the error variance estimated with k+1 (k+2 in the lower panel) regressors in each equation, T is the total number of observations on each series log is the natural logarithm. We select that order of lag that minimizes the loss metric. The asterisk ("* ") indicates minimum of the loss column.

Table 4. Statistical Loss Metrics on Oder of Lags of Net Returns in Generating U.S. Soybean Net Returns, without and with Current Stock of Research, 1978- 2007.

Lags = k	Hannan and Quinn Loss (Φ)
[With No Research Stock]	
Constant Only	43.99
-1	43.32
-2	43.40
-3	43.53
-4	43.67
-5	43.86
[With Research Stock]	
Constant	43.49
-1	43.11
-2	43.08 *
-3	43.23
-4	43.37
-5	43.52

Hannan and Quinn's measure on lag length (k) of a levels vector autoregression:

$$\Phi = \log(\sigma^2_k) + (2.00)(k+1) \times (\log(\log T))/T$$

where σ^2_k is the error variance estimated with k+1 (k+2 in the lower panel) regressors in each equation, T is the total number of observations on each series log is the natural logarithm. We select that order of lag that minimizes the loss metric. The asterisk ("* ") indicates minimum of the loss column.

of checkoff funds) generated by yields at earlier periods. The estimated models for both equation (1) and (2) are given in appendix II. Reported there as well are corresponding Dickey-Fuller statistics on innovations from each equation: -4.22 for the yield equation and -3.29 for the net returns equation (see as well appendix I Tables 1 and 2). Recall from above that since we are modeling series that are individually non-stationary, we want to be confident that the innovations series are stationary. These calculated statistics are less than the 5% critical value on tests based on estimated residuals of approximately -3.17 (Granger and Newbold 1976 Table 8.8). Accordingly, we see no evidence that the estimated models do not generate stationary innovations (residuals).

Innovation Accounting on Yield Response and Net Returns and Research Expenditures

Given the fit relationships between research stock and yield and net revenues we can calculate the long run relationships between both yields and net returns and research stock using standard innovation accounting techniques (Sims 1980). More specifically estimated versions (estimated parameters a_{ijk}) of equations 1 or 2 can be inverted to express current values of yields (Y_t) and/or net revenues (NR_t) in terms of current and previous (back to infinity) information.

$$\begin{bmatrix} Y_t - c_1 \\ R_t - c_2 \end{bmatrix} = \begin{bmatrix} b_{110} & 0 \\ b_{210} & b_{220} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \begin{bmatrix} b_{111} & b_{121} \\ b_{211} & b_{221} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{bmatrix} + ... + \begin{bmatrix} b_{11k} & b_{12k} \\ b_{21k} & b_{22k} \end{bmatrix} \begin{bmatrix} \varepsilon_{t-k} \\ \varepsilon_{t-k} \end{bmatrix} + ... \quad [6]$$

Where here we have allowed current shocks in research (or research stocks) affect current period yields. Two forms of accounting on this "moving average" form of the vector autoregression are informative on our soybean research assessment question. We can calculate the forecast error variance for both series (from equation [6]) and then decompose it into that part composed of information shocks in yields (ε_{1t+k}, for k going out to any distant horizon) and that part composed from information shocks in research stock (ε_{2t+k}). This will give us a sense of how important research stock is in the yield generating process (and of equation (2) how important research stock is for the net revenue process). For details on the decomposition of forecast error variance see Sims (1980).

At a ten year horizon, research stock accounts for about 25% of the uncertainty (variation) in soybean yields. This percentage is less than 5% at three years or less. At the ten year horizon variation in research Stock accounts for about 14% of the variation in net revenues (a considerably smaller portion than that found for yields). The numbers are summarized in Table 5. One might expect such numbers, as we would expect research stock to manifest itself as a relatively greater influence on yields than on net revenues, as the former is a component of the latter. Many other consideration, besides yields go into the net revenue calculations (price volatility is particularly noteworthy). The 25% portion that research stock contributes to yield variation at the long horizon is interesting and suggests that research is a non-trivial mover of future yields (the 14% contribution of research stock to net revenues is not trivial either, but only indicates that many other sources of uncertainty contribute to uncertainty in future net revenues). In terms of yield variability the other 75% (in addition to the 25% accounted for by research) is presumably dominated by weather variability.

A further look at the numbers reported in Table 5 allows one to assess the stationarity of the innovations from the estimated VAR models. Notice for both panels (yields and net revenues) that the standard error of the forecast errors increases (as is always the case as we go further out in the time horizon) at a decreasing rate – the standard errors at horizons 10 are not explosive extrapolations of earlier standard errors.

Equation (6) can also be used to explore how a particular series (yields and net revenues in our case) evolves through time. That is to say at every date we can add-up (or decompose) the series into a base component based on known information at a particular date, plus new information (expressed as our information shocks ($\varepsilon_{1,t-k}$ and $\varepsilon_{2,t-k}$)). In our case we consider how yields and net revenues have evolved from 1994 to 2007. So the value of the actual series, say yields, can be written as that part projected on information known before 1994 and the new information (acquired at or subsequent to 1994) arising from shocks in research stock and shocks in all other information (weather). [An application of this form of innovation accounting is given in Yang and Bessler (2008).]

Figures 3 and 4 give historical decomposition of 1994 -2007 U.S. soybean yields and U.S. soybean net revenues, respectively. Each figure offers three panels. The upper panel in each plots as a solid line the series of interest (solid line yields for figure 3 and net revenues for figure 4). Each upper panel also gives a dashed or broken line. This second line represents the base forecast calculated from information known at 1993. The vertical distance between the solid line (say, from figure 3, actual yields in any year) and the broken line can be decomposed into that part due to new information arising from research stock and that arising from all other sources. The research stock component is plotted in the middle panel of each figure. The component arising from all other sources is plotted in the bottom panel of each figure.

Taking figure 3 first, we see projected yields based on information known at 1993 are generally below actual yields (the solid line is above the broken line in almost all years in the upper panel of figure 3). The one exception to this last sentence is the year 2003, were actual yields are considerably below their projections. The year 2003 was generally a wet year in the corn-belt, but other management issues may as well have been responsible for low yields (Nafziger 2004 offers a discussion of soybean yields for 2003). The results given in figure 3 show that the low yield in 2003 was associated with these other factors (see the lower panel in figure 3) and not related to research expenditures. The high yield year of 1994 also stands out as unusual in figure 3. Again this result was not associated with research expenditures, as illustrated by the high number (greater than 7.0 bushels per acre higher than the base projection) in the year 1994 in the lower (other factors) panel in figure 3.

Interesting other work shows the trend growth rate in soybean yields at about .4 bushels per acre per year based on 1970 – 2005 data (Streit 2005). Our growth rate of .95 bushels per acre per year is more than twice this earlier estimate. At least two reasons exist for this difference. We are studying more recent data, where a strong increase in research expenditures has been recorded (see figure 1). More importantly our number (.95) is a growth rate associated only with research. We do not consider other factors (in the middle panel of figure 2). Several of these "other factors" will have a negative contribution to the overall growth rates (pests, weather, etc). From figure 3 it is the lower panel which captures the influence of these other factors. These

Table 5. Forecast Error Variance Decomposition of U.S. Yields and U.S. Net Revenues.

Horizon	Std. Error	Due to Own Series	Due to Information in Research Stock
		Yields	
0	3.12	99.65	0.35
1	3.15	98.52	1.48
5	3.23	93.94	3.46
10	3.63	74.59	25.41
		Net Returns	
0	9.4	99.92	0.08
1	13.4	99.55	0.45
3	14.1	96.29	3.71
10	15.5	85.57	14.43

Numbers in each grouping (Yields and Net Revenues) reflect the partition of forecast error variance of each series (grouping: yields or net revenues) into that portion that is accounted for from variation (new information) arising in either its own history or arising from new information from research stock. The numbers sum to 100 in any particular row. These are derived from the moving average representation of the VAR models given in Appendix II for yields and net revenues.

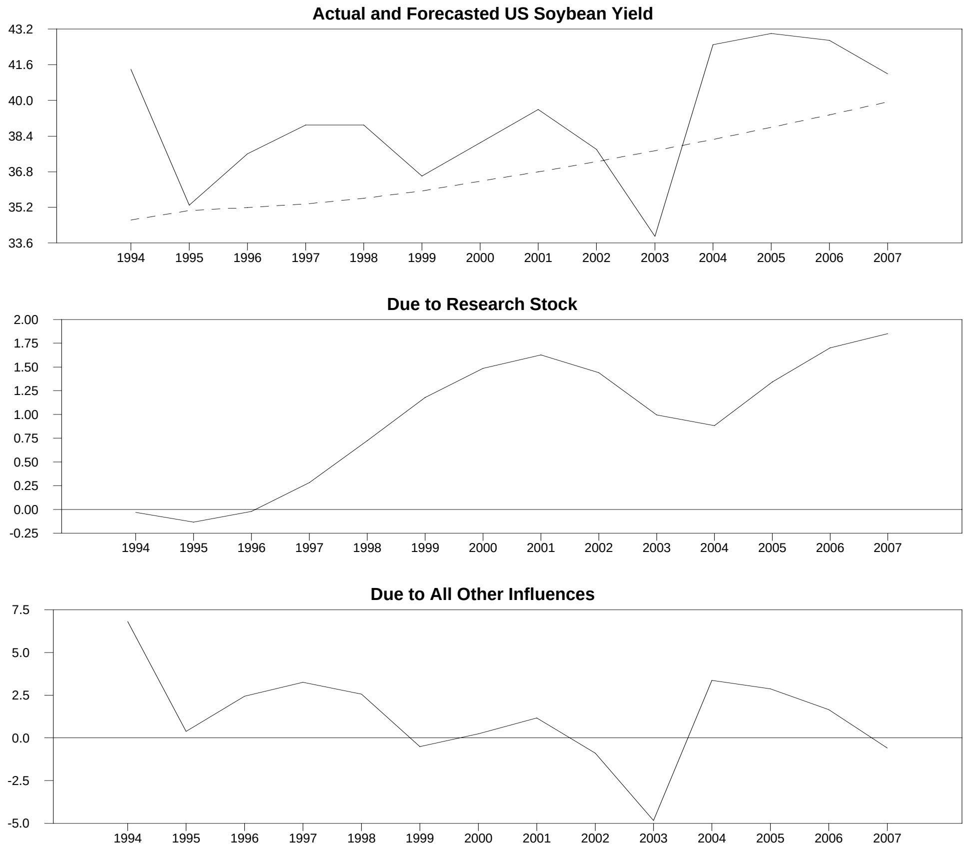


Figure 3. Decomposition of Actual Soybean Yield into Two Component Parts: that due to new information emanating from real research expenditures and that due to all other influences.

Note: in the upper graph actual soybean yield per acre given in the solid line; while forecasted yields, with research expenditures and any other new information acquired between 1994 and 2007 set equal to zero, are given by the broken line (----). The difference in any year between the actual yield and forecasted yield is decomposed into two parts: that due to research expenditures over the period 1994 – 2007 and that due to all other new information (e.g. weather).

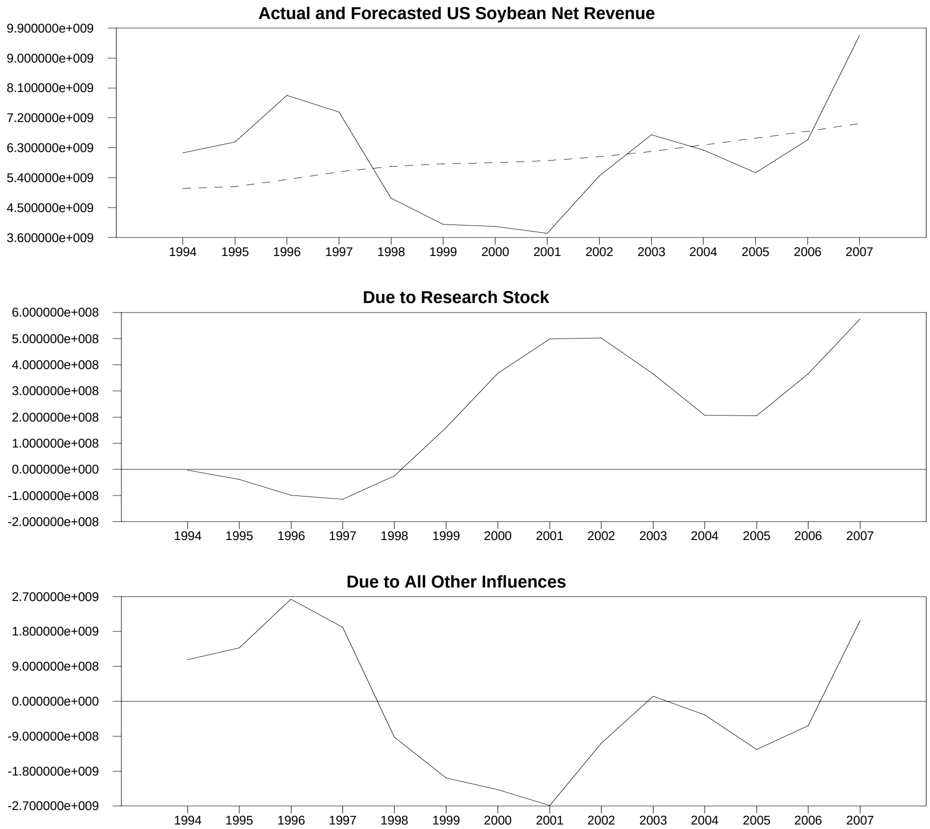


Figure 4. Decomposition of Real U.S. Soybean Net Returns (total real revenue minus total real variable cost) into two component parts: that due to new information emanating from real research expenditures and that due to all other influences.

Note: in the upper graph actual real net returns are given in the solid line; while forecasted net revenues, new information acquired between 1994 and 2007 on research and or other sources set equal to zero, are given by the broken line (----). The difference in any year between the actual net revenues and forecasted net revenues is decomposed into two parts: that due to new research expenditure information over the period 1994 – 2007 and that due to all other new information. CPI was used to reduce nominal returns to real returns.

factors generally contribute to lower yields over time. That is, the trend around the contribution of these other factors is actually negative (the aggregate affect of other factors not related to research expenditures moves yields down through time). Clearly the message from figure 3 is research has a positive affect on soybean yields and all other influences, those not associated with research (such as adoption of new management practices based on research recommendations) have, in the aggregate, a negative impact on yields.

Figure 4 offers a similar historical decomposition on net returns to soybean producers. The upper panel of figure 4 shows a modestly upward growth in net revenues. The base projection with information known at 1993 (dashed line) begins below the actual net revenue plot for years 1994 to 1997. The dashed line actually moves above the actual net revenue line until 2007, where is returns to its position below the actual net revenue line. The middle and bottom panels of figure 4 account for the vertical differences in the two lines in the upper panel. The middle panel shows that research stock contributed positively to net income over the years 1994 - 2007. Generally, however, the contribution of research stock to net return is rather modest during much of the 1990's, but shows strong contribution after 1998. The other factors affect on net revenues (as shown in the lower panel of figure 4) were generally negative over the year 1998 – 2006. These other factors contributed positively to net revenues in 2007. Much of the influence of the other factors affect was undoubtedly related to price variation. Any explanation for the large positive in the other factors panel in the year 2007 must seriously consider the affect of unprecedented prices in that year (average price was $10.40 per bushel in 2007, increasing from 2006 average price of $6.43/bushel). Similarly the low net returns of the late 1990's and into 2000 and 2001 can be explained well in the other factors panel. The negative values in this other factors panel in 1998, 1999 and 2000 are most likely due to market price. Prices in the mid 1990's were at or above $6.50 per bushel, but fell to under the loan at $4.93/bushel in 1998, $4.63/bushel in 1999 and $4.54/bushel in 2000.

Conditional Forecasts with Research Stock Set to Zero

Figures 3 and 4 are associated with an experiment where we keep research expenditures on their pre-1994 path. These figures and the associated calculations do not set research expenditures at zero or at any other fixed level. Rather they allow the research path over the years 1994 – 2007 to behave just as it did up through 1993. The innovations associated with research and other factors represent new information coming forth in 1994 or thereafter. An alternative experiment is considered here, where we actually set research expenditures (research stock) at zero over the 1994- 2007 period. We must caution that such "simulations" of counterfactuals are not well understood. We have less confidence with respect to this last simulation than those represented by figures 3 and 4. Clearly, in the research set equal to zero case we are forecasting out of the range of our data, as research stock has never (over our data set) been at zero. Yet, if we don't oversell accompanying results they may be helpful in stimulating questions of what might actually happen under such a scenario. In figure 5 we plot the result of such an exercise. Here we calculate the forecast of net revenues under the condition that research stock over the years 1994 – 2007 is zero. The upper panel gives both the actual net revenue (solid line) and the conditional forecast (broken line (---)). The picture is what we would expect. Actual net revenues are almost everywhere above the conditional forecast line. Put on a per acre basis,

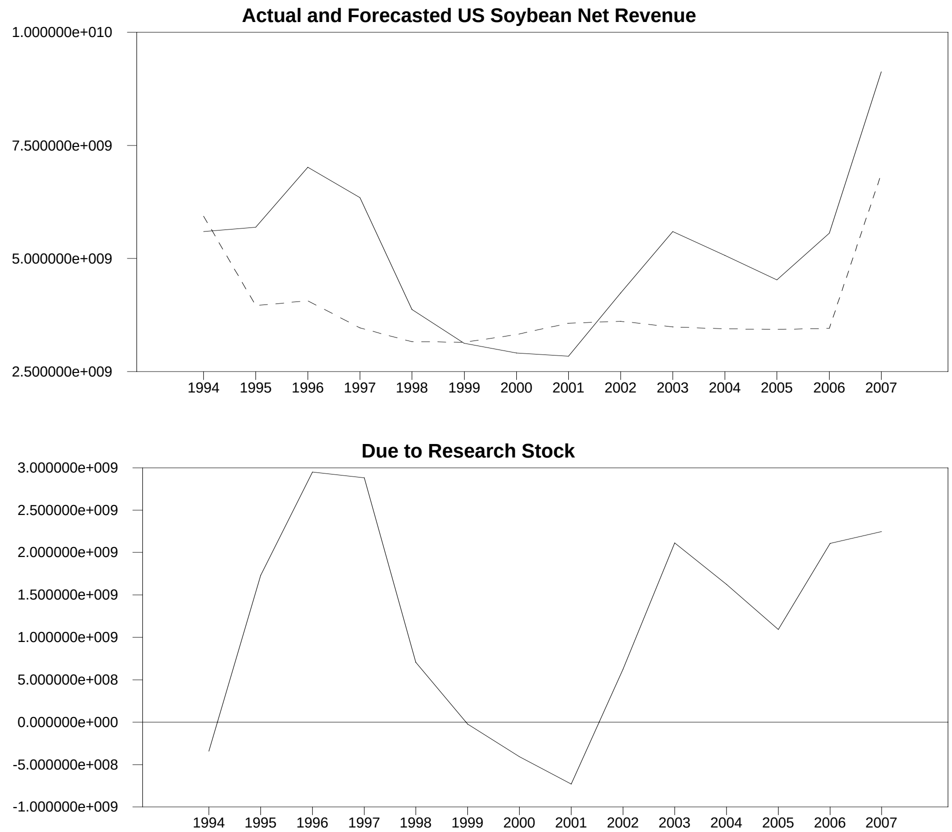


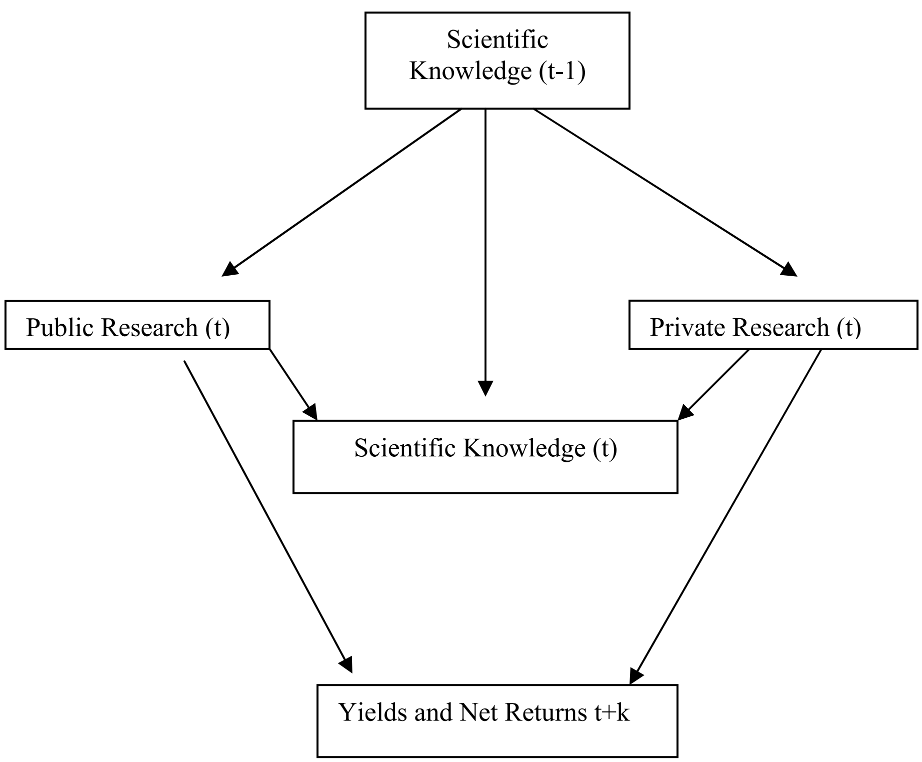
Figure 5. Actual and Conditional Forecast of Real U.S. Soybean Net Returns (total real revenue minus total real variable cost) Under the Condition that Research Stock over the years 1994 – 2007 is held at Zero.

Note: in the upper graph actual real net returns are given in the solid line; while forecasted net revenues, under the condition that research stock between 1994 and 2007 on is zero, are given by the broken line (----). The difference in any year between the actual net revenues and forecasted net revenues is plotted in the lower panel as the vertical distance between the two lines plotted in the upper panel. CPI was used to reduce nominal returns to real returns.

average difference between actual net revenues and the counterfactual net revenues (net revenues when research stock is set to zero) is about $17. per acre (in real dollars).

Discussion

Herein we have studied the relationships between soybean checkoff research expenditures (and a derived research stock variable) and soybean yields and net revenues for the U.S. We used data from 1978 – 2007 to inform us on potential impacts of the research funding on its two of its ultimate targets: yields and revenues. We found the research dollars do indeed have a positive effect on both. However, we have ignored an important player in the soybean research world, private sector research. Clearly this aspect of the returns to research needs consideration. Frey (1996) suggests that scientist years devoted to soybean breeding are divided amongst three main groups: 45 for state experiment stations, 10 for the USDA and 101 for private industry. We do not necessarily endorse Frey's numbers (but we have no reason to doubt them). The numbers do suggest that any study of returns to research needs to consider the possibility of an omitted variable that, if not considered, could bias any associated calculations. A very general model of returns to research may well look as follows:



The numbers we have focused on in the work presented here ignore the "common cause" of both private sector research and public sector research. Both sectors employ scientists trained at U.S. Land Grant Institutions. They take the same botany, genetics, chemistry, mathematics, experimental design courses. Indeed they may even be the same people. Advances in breeding germplasm enhancement, and cultivar development made with checkoff program monies have a common cause as those made at private companies. To ignore this backdoor path (Pearl 2000) in our attempt to measure the effectiveness of research dollars spent in the public sector is to over-state the contribution of the checkoff dollars. Our numbers are most surely picking up contributions from the larger private sector effort.

The ratio of increase in discounted net revenues associated with checkoff dollars to discounted soybean dollars over the 1994 – 2007 period is 118.52 – a huge number. If we recognize that the denominator in such calculations does not nearly represent the magnitude of the "causal" research dollars that have generated observed yield increases over time, we should adjust the above discounted return ratio by a factor of 1/3 or 1/4 or even more. A ratio of returns to expenditures in the neighborhood of 30 or 40 to one is probably closer to an accurate representation of the effectiveness of the soybean checkoff dollars.

Table 1. Descriptive Statistics on 1979- 2007 U.S. Soybean Data.

Series	Mean	Standard Deviation	Minimum (Date)	Maximum (Date)
Yield (Bushels/Acre)	35.22	4.91	25.7 (1983)	43.00 (2005)
Acreage	66,726,483.	6,135,523.	57,795,000. (1990)	75,522,000 (2006)
Production (Bushels)	2,314,733,414.	458,790,844.	1,548,841,000. (1998)	3,188247,000 (2006)
Gross Revenue (Dollars)	14,002,480,356.	3,696,803,836.	9,273,683,187. (1986)	26,886,152,800. (2007)
Net Revenue (Dollars)	9,198,920,184.	2,926,108,926.	6,309,987,387. (1986)	20,076,999,490. (2007)
Research Expenditures (Dollars)	11,520,342.	9,791,589.	2,291,268. (1979)	30,257,513 (2007)
Research Stock (Dollars	5,937,522.	5,217393.	1,407,870 (1979)	20,033519. (2007)

Statistics calculations presented here begin in the year 1979 because we need previous data as start-up to begin calculation research stock. Actual research expenditures provided by Keith Smith and Associates begin in 1978. Data on other series go back much earlier, but are not used here, as we wish to focus attention on the relationship between research expenditures and soybean production and revenues.

Table 2. Tests on Non-Stationarity of U.S. Soybean Data, 1978- 2007.

Series	t-stat
Yield (Bushels/Acre)	-2.32*
Acreage	-1.59*
Production (Bushels)	-1.83*
Gross Revenue (Dollars)	+1.21*
Net Revenue (Dollars)	+.13*
Research Expenditures (Dollars)	+.39*
Research Stock (Dollars	+8.65*

Note on Table 2: The test is on the null hypothesis that the data on the series listed in the left-hand column are non-stationary in levels (non-differenced data). The test for each series is based on an ordinary least squares regression of the first differences of the levels of each state's yield on a constant and one lag of the levels of that state's yield. The t-test is associated with the estimated coefficient on the lagged levels variable from this regression. Under the null hypothesis (non-stationary yields) the t-statistic is distributed as a non-standard student-t. Critical values are given in Fuller [9]. The 5% critical value is –2.89. We reject the null for observed t values less than this critical value. An asterisk (*) signifies failure to reject the null at the 5% level of significance.

Table 3. Statistical Loss Metrics on Oder of Lags in Generating U.S. Soybean Yields, without and with Research Stock, 1978- 2007.

Lags = k	Hannan and Quinn Loss (Φ)
[With No Research Stock]	
Constant Only	3.00
-1	2.75
-2	2.77
-3	2.88
-4	2.89
-5	2.86
[With Research Stock]	
Constant	2.70
-1	2.63 *
-2	2.74
-3	2.89
-4	2.94
-5	2.96

Hannan and Quinn's measure on lag length (k) of a levels vector autoregression

$$\Phi = \log(\sigma^2_k) + (2.00)(k+1) \times (\log(\log T))/T$$

where σ^2_k is the error variance estimated with k+1 (k+2 in the lower panel) regressors in each equation, T is the total number of observations on each series log is the natural logarithm. We select that order of lag that minimizes the loss metric. The asterisk ("* ") indicates minimum of the loss column.

Table 4. Statistical Loss Metrics on Oder of Lags of Net Returns in Generating U.S. Soybean Net Returns, without and with Current Stock of Research, 1978- 2007.

Lags = k	Hannan and Quinn Loss (Φ)
[With No Research Stock]	
Constant Only	43.99
-1	43.32
-2	43.40
-3	43.53
-4	43.67
-5	43.86
[With Research Stock]	
Constant	43.49
-1	43.11
-2	43.08 *
-3	43.23
-4	43.37
-5	43.52

Hannan and Quinn's measure on lag length (k) of a levels vector autoregression:

$$\Phi = \log(\sigma^2_k) + (2.00)(k+1) \times (\log(\log T))/T$$

where σ^2_k is the error variance estimated with k+1 (k+2 in the lower panel) regressors in each equation, T is the total number of observations on each series log is the natural logarithm. We select that order of lag that minimizes the loss metric. The asterisk ("* ") indicates minimum of the loss column.

Table 5. Forecast Error Variance Decomposition of U.S. Yields and U.S. Net Revenues.

Horizon	Std. Error	Due to Own Series	Due to Information in Research Stock
		Yields	
0	3.12	99.65	0.35
1	3.15	98.52	1.48
5	3.23	93.94	3.46
10	3.63	74.59	25.41
		Net Returns	
0	9.4	99.92	0.08
1	13.4	99.55	0.45
3	14.1	96.29	3.71
10	15.5	85.57	14.43

Numbers in each grouping (Yields and Net Revenues) reflect the partition of forecast error variance of each series (grouping: yields or net revenues) into that portion that is accounted for from variation (new information) arising in either its own history or arising from new information from research stock. The numbers sum to 100 in any particular row. These are derived from the moving average representation of the VAR models given in Appendix II for yields and net revenues.

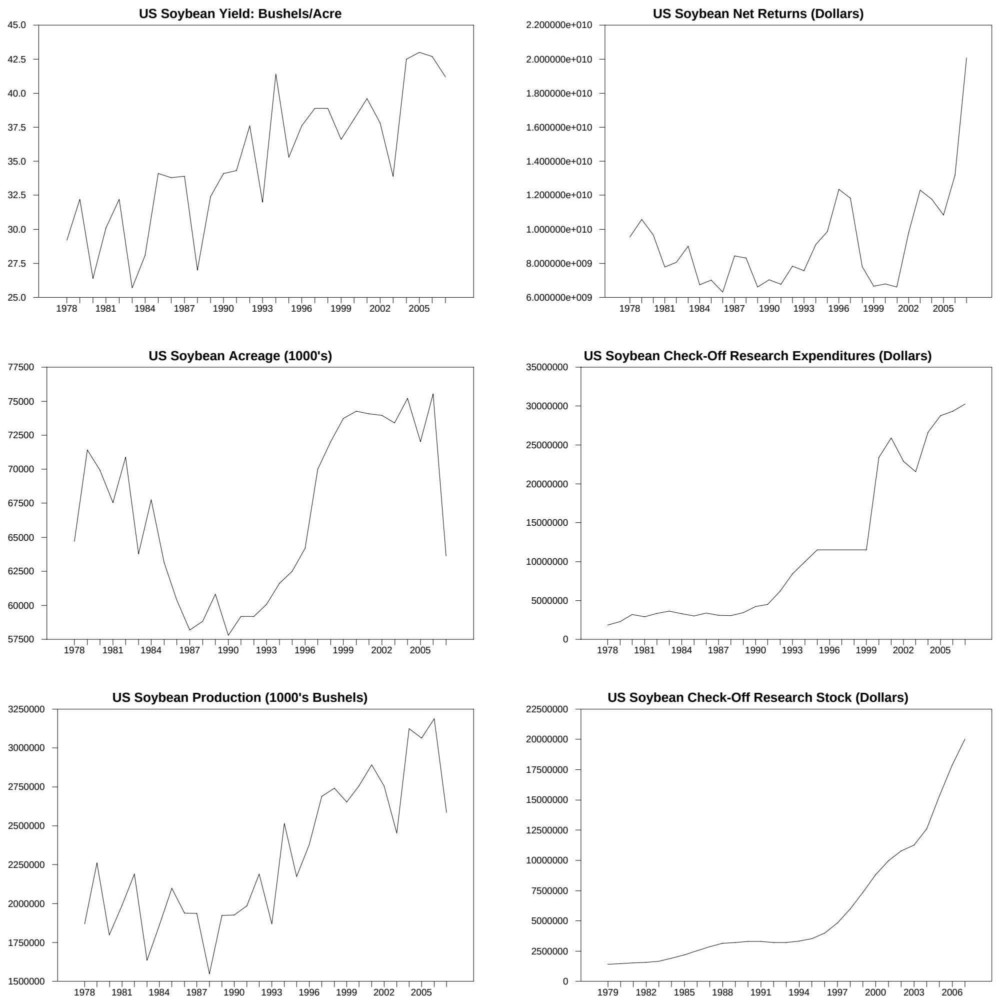


Figure 1. Time Series Plots of Soybean Yields, Acreage, Production, Net Returns, Checkoff Research Expenditures and Checkoff Research Stock, 1979 – 2007.

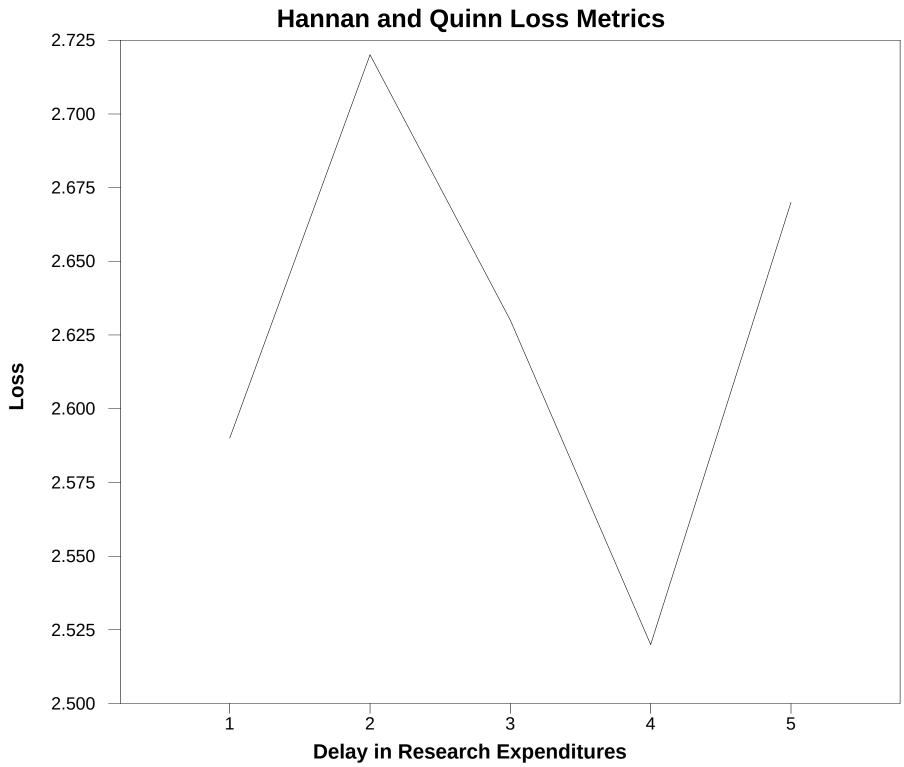


Figure 2. Hannan and Quinn Loss in Delay of Affect of Checkoff Research Expenditures on U.S. Soybean Yields.

Figure 3. Decomposition of Actual Soybean Yield into Two Component Parts: that due to new information emanating from real research expenditures and that due to all other influences.

Note: in the upper graph actual soybean yield per acre given in the solid line; while forecasted yields, with research expenditures and any other new information acquired between 1994 and 2007 set equal to zero, are given by the broken line (----). The difference in any year between the actual yield and forecasted yield is decomposed into two parts: that due to research expenditures over the period 1994 – 2007 and that due to all other new information (e.g. weather).

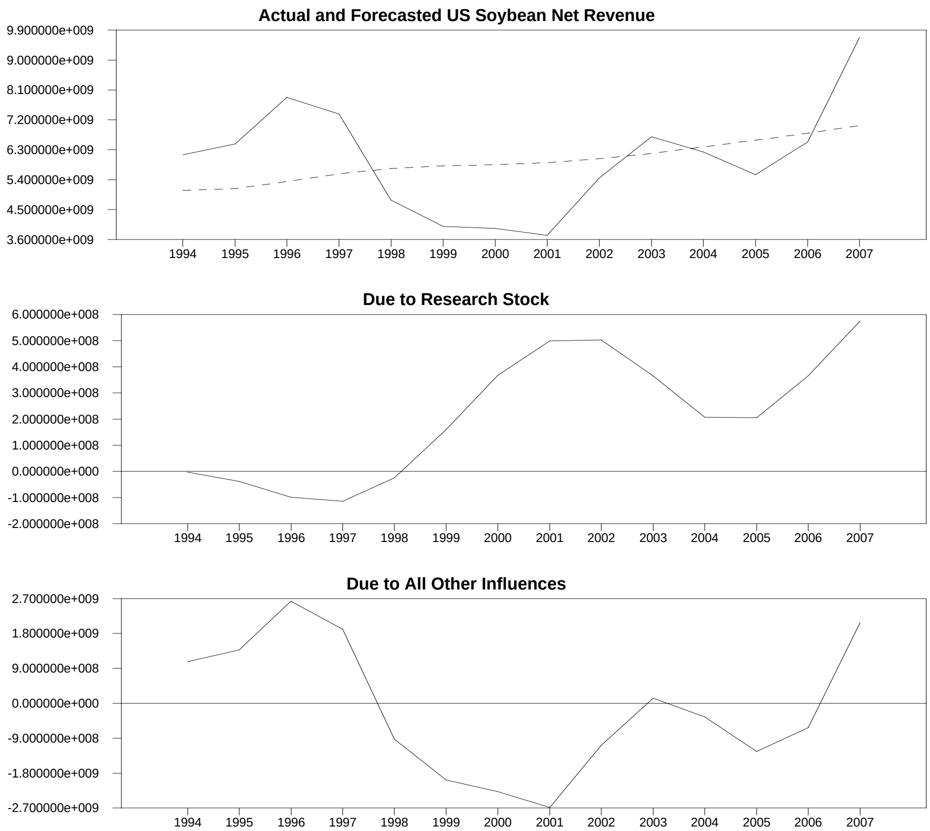


Figure 4. Decomposition of Real U.S. Soybean Net Returns (total real revenue minus total real variable cost) into two component parts: that due to new information emanating from real research expenditures and that due to all other influences.

Note: in the upper graph actual real net returns are given in the solid line; while forecasted net revenues, new information acquired between 1994 and 2007 on research and or other sources set equal to zero, are given by the broken line (----). The difference in any year between the actual net revenues and forecasted net revenues is decomposed into two parts: that due to new research expenditure information over the period 1994 – 2007 and that due to all other new information. CPI was used to reduce nominal returns to real returns.

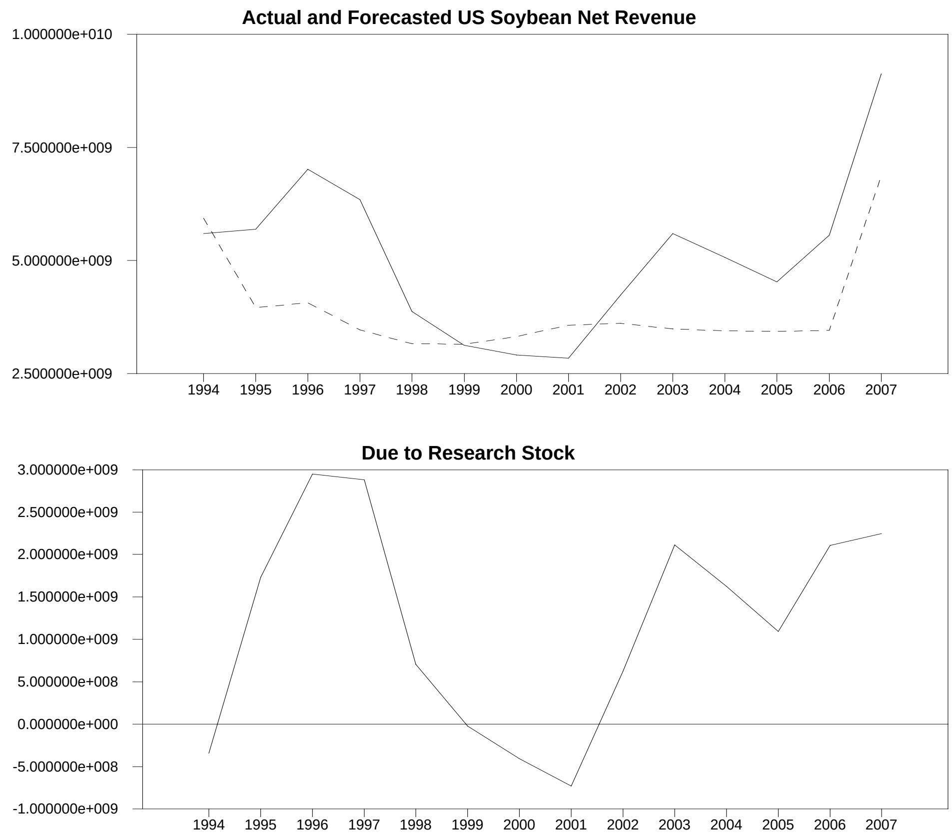


Figure 5. Actual and Conditional Forecast of Real U.S. Soybean Net Returns (total real revenue minus total real variable cost) Under the Condition that Research Stock over the years 1994 – 2007 is held at Zero.

Note: in the upper graph actual real net returns are given in the solid line; while forecasted net revenues, under the condition that research stock between 1994 and 2007 on is zero, are given by the broken line (----). The difference in any year between the actual net revenues and forecasted net revenues is plotted in the lower panel as the vertical distance between the two lines plotted in the upper panel. CPI was used to reduce nominal returns to real returns.

References

Alston, J., G. Norton and P. Pardey. 1998. *Science under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*, CAB International in association with International Service for National Agricultural Research, Wallingford, United Kingdom.

Dickey, D.A. and W.A. Fuller. 1981. "The Likelihood Ratio Statistics for Autoregressive Time Series with A Unit Root." *Econometrica*, 49, 1057-1072.

Frey, K.J. (1996) National plant breeding study – I. Human and financial resources devoted to plant breeding research and development in the United States in 1994, Iowa State University, Agricultural and Home Economics Report 98: Ames.

Fuller, W. A. 1976. *Introduction to Statistical Time Series*, New York Wiley.

Geweke J. and R. Meese. 1981. "Estimating Regression Models of Finite but Unknown Order," *International Economic Review*, 55-70.

Granger, C. 1981. "Some Properties of Time Series Data and Their Use in Econometric Model Specification," *Journal of Econometrics*, 121-130.

Granger, C. and P. Newbold. Forecasting Economic Time Series, New York: Academic Press, 19786.

Nazfziger, E. "Chasing High Corn and Soybean Yields," College of Agriculture, Consumer and Environmental Science, University of Illinois, Urbana, 2004.

Nerlove, M., D.M Grether and J.L. Carvalho. 1979. Analysis of Economic Time series: A Synthesis, New York: Academic Press.

Pearl, J. Causality, Cambridge, U.K. Cambridge University Press, 2000.

Sims, C. (1980) "Macroeconomic economics and reality," Econometrica 48:1-48.

Streit, L. "Climbing above the Yield Plateau," Presentation at the 2005 North Central American Society of Agronomy, march 16, 2005.

Yang, J.* and D.A. Bessler. 2008. "Contagion Around the 1987 Stock Market Crash," European Journal of Operational Research 184:291-310.

Appendix A Tables on Stationarity of Innovations from fit VARs.

Table A1. Tests on Non-Stationarity of Innovations from Each Equation of the Levels VAR on Yields and Research Stock, 1988- 2007.

Equation	t-stat
Yields	-4.22
Research Stock	-2.86

Note on Appendix Table 1: The test is on the null hypothesis that the data on the series listed in the left-hand column are non-stationary in levels (non-differenced data). The test for each series is based on an ordinary least squares regression of the first differences of the levels of the innovations from the equation listed in the left-hand column on a constant and one lag of the levels of that variable. The t-test is associated with the estimated coefficient on the lagged levels variable from this regression. Under the null hypothesis (non-stationary yields) the t-statistic is distributed as a non-standard student-t. Critical values are given in Fuller [9]. The 5% critical value is –2.89. We reject the null for observed t values less than this critical value. An asterisk (*) signifies rejection of the null at the 5% level of significance.

Table A2. Tests on Non-Stationarity of Innovations from Each Equation of the Levels VAR on Net Returns and Research Stock, 1988- 2007.

Equation	t-stat
Net Returns	-3.29
Research Stock	-4.22

Note on Appendix Table 2: The test is on the null hypothesis that the data on the series listed in the left-hand column are non-stationary in levels (non-differenced data). The test for each series is based on an ordinary least squares regression of the first differences of the levels of the innovations from the equation listed in the left-hand column on a constant and one lag of the levels of that variable. The t-test is associated with the estimated coefficient on the lagged levels variable from this regression. Under the null hypothesis (non-stationary yields) the t-statistic is distributed as a non-standard student-t. Critical values are given in Fuller [9]. The 5% critical value is –2.89. We reject the null for observed t values less than this critical value. An asterisk (*) signifies rejection of the null at the 5% level of significance.

Appendix B: Computer Programs in RATS language.

I. Yields

II. Net Revenues

RATS Output on U.S. Yields Response to Research Stock

Dependent Variable YIELDS - Estimation by Least Squares
Annual Data From 1987:01 To 2007:01
Usable Observations 21 Degrees of Freedom 18
Centered R**2 0.393373 R Bar **2 0.325971
Uncentered R**2 0.993019 T x R**2 20.853
Mean of Dependent Variable 37.085714286
Std Error of Dependent Variable 4.100400677
Standard Error of Estimate 3.366400168
Sum of Squared Residuals 203.98770159
Durbin-Watson Statistic 1.977264

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	29.255021074	7.987663400	3.66253	0.00178127
2. YIELDS{1}	0.099299351	0.245491467	0.40449	0.69061691
3. NOMSTOCK	0.000000962	0.000000417	2.30654	0.03318568

Dependent Variable NOMSTOCK - Estimation by Least Squares
Annual Data From 1987:01 To 2007:01
Usable Observations 21 Degrees of Freedom 17
Centered R**2 0.993213 R Bar **2 0.992015
Uncentered R**2 0.998492 T x R**2 20.968
Mean of Dependent Variable 4349000.0225
Std Error of Dependent Variable 2382305.2630
Standard Error of Estimate 212876.6316
Sum of Squared Residuals 7.70380e+011
Durbin-Watson Statistic 1.258660

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	-614984.3490	445070.3979	-1.38177	0.18493276
2. NOMSTOCK{1}	1.7284	0.1818	9.50853	0.00000003
3. NOMSTOCK{2}	-0.7286	0.2073	-3.51512	0.00265554
4. YIELDS{8}	21779.4816	15363.1116	1.41765	0.17436250

Covariance\Correlation Matrix

	R_YIELDS	R_NOMSTOCK
R_YIELDS	9.71	0.4604090159
R_NOMSTOCK	274839.38	36684753565.24

DICKEY FULLER test on Yields Innovations

Dependent Variable No Label (103) - Estimation by Least Squares
Annual Data From 1988:01 To 2007:01
Usable Observations 20 Degrees of Freedom 18
Centered R**2 0.497619 R Bar **2 0.469709
Uncentered R**2 0.497632 T x R**2 9.953
Mean of Dependent Variable -0.022459246
Std Error of Dependent Variable 4.607359712
Standard Error of Estimate 3.355129193
Sum of Squared Residuals 202.62405426

Regression F(1,18)	17.8294
Significance Level of F	0.00051200
Durbin-Watson Statistic	1.749616
Q(5-0)	0.454961
Significance Level of Q	0.99367978

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	0.056844808	0.750464746	0.07575	0.94045660
2. R_YIELDS{1}	-0.998423139	0.236453890	-4.22249	0.00051200

DICKEY FULLER test on Research Stock Innovations

Dependent Variable No Label (104) - Estimation by Least Squares
Annual Data From 1988:01 To 2007:01

Usable Observations	20	Degrees of Freedom	18
Centered R**2	0.312386	R Bar **2	0.274185
Uncentered R**2	0.312601	T x R**2	6.252
Mean of Dependent Variable	-3888.2516		
Std Error of Dependent Variable	225871.8666		
Standard Error of Estimate	192430.9810		
Sum of Squared Residuals	6.66534e+011		
Regression F(1,18)	8.1775		
Significance Level of F	0.01041365		
Durbin-Watson Statistic	1.608909		
Q(5-0)	6.764820		
Significance Level of Q	0.23872669		

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	747.63196	43059.40367	0.01736	0.98633819
2. R_NOMSTOCK{1}	-0.63622	0.22249	-2.85963	0.01041365

Decomposition of Variance for Series YIELDS

Step	Std Error	INNOV_1	INNOV_2
1	3.122120971	99.65182	0.34818
2	3.155439237	98.52045	1.47955
3	3.187453319	96.56065	3.43935
4	3.231530681	93.94457	6.05543
5	3.285275203	90.89599	9.10401
6	3.346148923	87.61888	12.38112
7	3.411973392	84.27077	15.72923
8	3.481020306	80.96086	19.03914
9	3.553352108	77.73210	22.26790
10	3.630249269	74.59400	25.40600

Decomposition of Variance for Series NOMSTOCK

Step	Std Error	INNOV_1	INNOV_2
1	191532.644	0.00000	100.00000
2	382464.112	0.00000	100.00000
3	577459.184	0.00000	100.00000
4	768172.929	0.00000	100.00000
5	950845.023	0.00000	100.00000
6	1124036.125	0.00000	100.00000

7	1287488.963	0.00000	100.00000
8	1441539.310	0.00000	100.00000
9	1589931.948	0.18227	99.81773
10	1736932.656	0.66292	99.33708

Normal Completion

RATS Output on Net Returns Response to Research Stock

```
Dependent Variable NETRETURN - Estimation by Least Squares
Annual Data From 1988:01 To 2007:01
Usable Observations     20      Degrees of Freedom    16
Centered R**2     0.537347      R Bar **2   0.450599
Uncentered R**2   0.975508      T x R**2      19.510
Mean of Dependent Variable      5900531606.5
Std Error of Dependent Variable 1431287146.2
Standard Error of Estimate      1060892446.3
Sum of Squared Residuals        1.80079e+019
Durbin-Watson Statistic             1.638511

   Variable         Coeff       Std Error      T-Stat     Signif
*************************************************************************
1.  Constant        2.3829e+009  1.5353e+009 1.55206  0.14020166
2.  NETRETURN{1}    1.0072       0.2485      4.05385  0.00092161
3.  NETRETURN{2}   -0.5309       0.2531     -2.09727  0.05221238
4.  STOCK         166.5915     102.8573      1.61964  0.12484979


Dependent Variable STOCK - Estimation by Least Squares
Annual Data From 1988:01 To 2007:01
Usable Observations     20      Degrees of Freedom    13
Centered R**2     0.995573      R Bar **2   0.993529
Uncentered R**2   0.999034      T x R**2      19.981
Mean of Dependent Variable      4440232.0373
Std Error of Dependent Variable 2406261.4538
Standard Error of Estimate       193562.5862
Sum of Squared Residuals        4.87064e+011
Durbin-Watson Statistic             2.023521

   Variable            Coeff       Std Error      T-Stat     Signif
*************************************************************************
1.  Constant        -30662.3873  271514.0235     -0.11293  0.91181044
2.  STOCK{1}          2.3419       0.2724         8.59798  0.00000101
3.  STOCK{2}         -2.1761       0.6044        -3.60057  0.00322835
4.  STOCK{3}          1.0315       0.6821         1.51223  0.15440012
5.  STOCK{4}         -0.1505       0.3538        -0.42549  0.67743850
6.  NETRETURN{8}      4.2888e-005  5.7381e-005    0.74742  0.46811189
7.  NETRETURN{9}     -4.1029e-005  3.6611e-005   -1.12067  0.28270417


Covariance\Correlation Matrix
               R_NETRETURN      R_STOCK
R_NETRETURN  9.003942e+017  -0.1386291460
R_STOCK     -2.052811e+013  2.435321e+010
```

Dickey-Fuller on Net Returns

```
Dependent Variable No Label(103) - Estimation by Least Squares
Annual Data From 1989:01 To 2007:01
Usable Observations     19      Degrees of Freedom    17
Centered R**2     0.389246      R Bar **2   0.353319
Uncentered R**2   0.394912      T x R**2       7.503
Mean of Dependent Variable       120027821.8
```

```
Std Error of Dependent Variable 1274371144.6
Standard Error of Estimate      1024804056.5
Sum of Squared Residuals        1.78538e+019
Regression F(1,17)                   10.8345
Significance Level of F           0.00430775
Durbin-Watson Statistic             1.768327
Q(4-0)                              0.216712
Significance Level of Q           0.99453679

   Variable            Coeff       Std Error      T-Stat     Signif
************************************************************************
1.  Constant        21318757.9884 237011009.2741  0.08995  0.92937950
2.  R_NETRETURN{1}  -0.9151       0.2780         -3.29157  0.00430775
```

Dickey-Fuller on Research Stock

```
Dependent Variable No Label(104) - Estimation by Least Squares
Annual Data From 1989:01 To 2007:01
Usable Observations     19      Degrees of Freedom    17
Centered R**2     0.512103      R Bar **2   0.483403
Uncentered R**2   0.512149      T x R**2       9.731
Mean of Dependent Variable        -2214.8567
Std Error of Dependent Variable  233986.1430
Standard Error of Estimate       168176.8124
Sum of Squared Residuals        4.80818e+011
Regression F(1,17)                   17.8434
Significance Level of F           0.00057109
Durbin-Watson Statistic             1.969931
Q(4-0)                              0.811684
Significance Level of Q           0.93687490

   Variable          Coeff       Std Error      T-Stat     Signif
************************************************************************
1.  Constant      -4029.87421  38584.79934     -0.10444  0.91804092
2.  R_STOCK{1}      -1.01918      0.24127      -4.22415  0.00057109


Estimation by Simplex

   Variable                     Coeff
***************************************
1.  A12                      842.93000000

Iteration limit hit. Conv. Criterion =       NA
Coefficient Estimates for Structural Decomposition
 Row Col  Value    Std.Error
LR Test of Overidentification
Chi-Square(   1 ) =       0.36870 Signif. Level = 0.5437154
```

Decomposition of Variance for Series NETRETURN

```
Step  Std Error   INNOV_1   INNOV_2
   1    949247120  99.92499   0.07501
   2   1349848127  99.54685   0.45315
   3   1434662972  98.35649   1.64351
   4   1450740935  96.28531   3.71469
```

5	1496009730	94.28048	5.71952
6	1533986685	92.70435	7.29565
7	1550483043	91.29876	8.70124
8	1563281617	89.83624	10.16376
9	1582430494	88.02589	11.97411
10	1608454313	85.56537	14.43463

Decomposition of Variance for Series STOCK

Step	Std Error	INNOV_1	INNOV_2
1	156055.146	0.00000	100.00000
2	397392.171	0.00000	100.00000
3	651532.010	0.00000	100.00000
4	868859.950	0.00000	100.00000
5	1042525.005	0.00000	100.00000
6	1191042.110	0.00000	100.00000
7	1338298.189	0.00000	100.00000
8	1502818.905	0.00000	100.00000
9	1694381.301	0.05769	99.94231
10	1915461.849	0.30351	99.69649

Ratio of Discounted Net Returns to Expenses

ENTRY	DISCRATIO
1994:01	-1.46771821034
1995:01	-12.74467212721
1996:01	-30.05465168188
1997:01	-39.74830454502
1998:01	-32.34351460833
1999:01	-11.94307537119
2000:01	11.48898562093
2001:01	30.59544647866
2002:01	42.73086033102
2003:01	47.49380196723
2004:01	46.96457999108
2005:01	45.42906834399
2006:01	46.17433706998
2007:01	49.20227451257

Normal Completion