MEASURING RETURNS TO RESEARCH EXPENDITURES FOR CORN, WHEAT AND SOYBEANS

W. B. Sundquist, Cheng-Ge Cheng and George W. Norton*

Introduction

A wealth of information exists on _ex post_ returns to research investments for agriculture in the aggregate and for several agricultural commodity groups of commercial importance in the United States. Many of the investments made in agricultural research and, consequently, many of the allocative decisions on research funding, are, however, commodity-specific. Probably the main reason for the limited empirical literature on returns to commodity-specific research is that of the limited production input data available for individual commodities. Data from the _U.S. Census of Agriculture_, for example, do not permit commodity disaggregation beyond that employed by Bredahl and Peterson. Their analysis which utilizes commodity groups including cash crops, dairy, livestock, and poultry, has been updated by Norton using 1974 Census data. The research reported in this paper represents "early stage" progress in an effort to extend the cross-sectional production function-type analyses of Bredahl, Peterson, and Norton to three specific crops--corn, wheat, and soybeans. The production functions estimated are for the 1977 crop year.

Model, Variables and Data

The functional form used for each commodity is the familiar Cobb-Douglas production function. Specification of individual variables in the functions is shown in Appendix A. Since the _U.S. Census of Agriculture_ does not report production input categories for specific agricultural commodities, we undertook to find alternative data sources from which we could develop commodity-specific production function formulations using individual states as observations. One source of such data is that provided in the nationwide set of farm enterprise budgets developed by Krenz, _et. al._, in the National Economics Division, ESCS, USDA. These so-called FEDS budget data have been developed annually since 1974 for all production areas in the United States for which the major farm commodities are produced commercially. The FEDS enterprise budget data are developed, drawing heavily on survey data, for each major substate production area though in some cases, such areas are specified as an entire state. We have weighted and aggregated these enterprise data for 1977 in a manner so as to develop category totals for each state.

While the FEDS budgets are readily available as a data source, they are not without some serious shortcomings for production function analyses. For example, the machine and labor inputs for a specific enterprise budget include, as is desirable for our purposes, only the machinery and labor that are used for that enterprise. But, these input categories have a high degree of multicollinearity within the total set of enterprise budgets for each crop because, though based on farm survey data, they depict a fairly uniform complement of machinery and a fairly standardized set of production practices. Moreover, the machinery and labor input categories are highly correlated with land because each acre of land used for production of a specific crop has a rather standard package of machinery and labor inputs applied to it. Thus, high intercorrelations between these input categories results. And, the budget data, being weighted estimates of per-acre means for each state, do not depict the full variance which is actually present among input categories on individual farms.

Though the above mentioned high intercorrelations between production input categories cause problems in estimating production functions using these categories of land, labor, and machinery as independent variables, this problem increasingly exists even independent of the FEDS set of data. Much of the per-acre variance in input categories between individual farms, state subregions, states, and multi-state regions has disappeared over time as commercial farmers have developed farming operations which are highly mechanized and fairly standardized. And, very little

*W.B. Sundquist is Professor, Cheng-Ge Cheng is Graduate Research Assistant and George W. Norton is formerly Research Associate, Department of Agricultural and Applied Economics, University of Minnesota.

unemployed or underemployed labor now remains on U.S. farms. This suggests that much of the earlier day variance in the use of at least some farm inputs has vanished as the shift to mechanized and otherwise modernized production methods on commercial farms in the United States has become virtually complete. And, this phenomenon will be reflected in the data on input categories whatever their source.1/ If, in fact, labor, machinery, and land inputs are approaching the relationship of technical complements, the only feasible solution to the statistical estimation problem caused by multicollinearity may be that of using a single input category as a proxy for the several or, alternatively, this can be accomplished by converting each of the several categories into single composite measure (dollars) and using the aggregate value of this new input variable.

The research variable included in the production functions for 1977 is an average of 1970-72 research expenditures for each crop from the CRIS data. Centered on 1971, the research expenditures are thus lagged six years from the 1977 production year. Research expenditures for corn average $330,000 with a range of more than $1 million for the 23 states included in the analysis. Iowa, Indiana, and Illinois are the top three states in corn research expenditures. These states are also the top soybean research states, along with Arkansas. Soybean research expenditures average $211,000 for the 26 states included in the analysis with a range of $500,000. Kansas and North Dakota are the top wheat research states in terms of expenditures. The mean wheat research expenditure is $185,000 for the 34 states included in the analysis with a range of $868,000. Data sources for the production functions are shown in Appendix Table B.

Regression Results

Initial production functions were specified for each crop for 1977 using state observations in order to estimate total crop value as a function of land, rainfall, fertilizer, chemicals, pesticides, labor, machinery expenses, and research expenditures (see Table 1). July rainfall was highly significant in the soybean and corn equations but not in the wheat function. It is difficult to determine a single critical weather period for wheat due to the presence of both spring and winter wheat. Moreover, growing seasons and crop maturity dates vary significantly between individual spring wheat production areas and between individual winter wheat areas. And, temperature is also likely to be quite important. Fertilizer is significant in both the corn and wheat equations but not in the soybean equation. This may be due, in part at least, to the ability of soybeans to fix nitrogen.

The chemicals variable is not significant in any of the equations. This is not surprising because pesticides are normally applied in greater quantities in areas where pest problems are more serious. Since no data are available on the severity of the pest problem, there is, in effect, a specification error which biases downward the effect of the chemicals variable and reduces its significance. Also, as previously mentioned, there is a high degree of multicollinarity among chemicals, fertilizer, labor, land, and machinery in the FEDS data set. The coefficients for the latter three variables are not significant in these three equations with the exception of land in the soybean equation. Research is not significant in any of the equations. It too, however, is highly correlated with the land, labor, machinery, fertilizer, and chemicals variables.

Before attempting to deal with the problem of multicollinearity, an effort was made to account for major differences in land quality among regions through a set of crop-specific regional slope dummies on the land variable. These dummies are defined in Appendix C. No significant differences were found for corn, but the Northern and Southern Plains regions showed significantly lower quality land than the Corn Belt region for wheat (see Table 2). And, the Southern and Delta regions showed lower quality land than the Corn Belt for soybeans.

Table 1. Initial Research Production Functions*

	Land	Labor	Machinery	Fertilizer	Chemicals	Rain	Research	Constant	$\bar{R}^2$
Corn	.24	.18	.30	.34	-.03	.19	-.03	1.84	
	(1.68)	(.85)	(1.69)	(5.17)	(-.37)	(8.24)	(-.60)	(4.09)	.994
Wheat	.288	.18	.16	.37	.019		.049	1.48	
	(1.01)	(.91)	(.56)	(3.76)	(.64)		(.73)	(1.35)	.955
Soybeans	1.77	-.97	.23	-.046	.022	.035	-.01	-.43	
	(4.13)	(-1.65)	(.286)	(-.68)	(.23)	(.63)	(-.11)	(-.26)	.977

*Numbers in parentheses are t-values.

Table 2. Research Production Functions with Land Dummies*

	Corn	Wheat	Soybeans
Land	.21	.082	.480
	(.95)	(.24)	(1.44)
Fertilizer	.29	.14	-.014
	(3.56)	(.94)	(-.313)
Chemical	-.03	.025	-.076
	(-.38)	(.75)	(-1.35)
Labor	.06	.48	-.475
	(.20)	(2.16)	(-1.16)
Machinery	.55	.34	1.17
	(2.17)	(1.06)	(2.35)
Rain	.19	--	.068
	(7.9)	--	(2.02)
Research	-.025	.07	-.047
	(-.49)	(.92)	(-.83)
Slope 1	.007	-.02	-.032
	(1.12)	(-1.68)	(-3.40)
Slope 2	-.007	-.013	-.01
	(-.93)	(-.82)	(-1.01)
Slope 3	--	-.05	-.041
	--	(-2.45)	(-5.62)
Slope 4	--	-.039	--
	--	(-1.95)	--
Slope 5	--	-.026	--
	--	(-1.40)	--
Constant	1.1	.357	-5.127
	(1.6)	(.256)	(-4.37)
$\bar{R}^2$	.994	.960	.993

*Numbers in parentheses are t-values.

In an effort to reduce the multicollinearity problem mentioned above and with the likelihood of a high degree of technical complementarity existing among them, the land, labor, and machinery variables were value weighted and added together. Land quality dummy variables were no longer included because the use of land price weights was assumed to pick up at least some of the land quality differences. Functions were reestimated for corn, wheat, and soybeans and the results are shown in Table 3. While the land-labor-machinery aggregate was highly significant in the wheat and soybean functions, it was not significant in the corn equation. The chemicals variable increased in significance in all cases while rainfall and research both increased in significance in the wheat equation. A large amount of multicollinearity still exists, however, among chemicals, fertilizer, research, and the land-labor-machinery aggregate variable.

Another set of regressions was run in which all "traditional variables," i.e., land, labor, machinery, chemicals, and fertilizer, were value aggregated. This input aggregate was included in a production function with rainfall and research as the other independent variables (see Table 4). In this case, the traditional variable aggregate is highly significant for each crop. Rainfall was significant at the 95% level for corn and wheat. Research had a positive coefficient in all cases and was significant at the 95% level for wheat, and the 90% level for soybeans. The research coefficients were roughly of the same magnitude for each of the three crops.

Additional regressions were run in which a "spillover" variable was added in an attempt to pick up the spillover effects of research across state boundaries. While spillover of research occurs for all three crops to some extent, it is thought to be most pronounced, or at least of a different form, for soybeans because varieties are very latitude specific. Varieties raised in Iowa, for example, are also raised in Pennsylvania. Any simply constructed spillover variable is, of course, somewhat arbitrary and open to criticism. But, inclusion of even a crude variable seems preferable to ignoring the existence of spillover. The specification used in this study is described in Appendix A. The results of adding a spillover variable to the production function are shown in Table 5.

These results indicate that inclusion of the research spillover variable improved the soybean and wheat functions substantially but the crude specification of the spillover variable for corn was not a particularly useful addition to the function for that crop.

The specified research spillover variable for soybeans is highly significant. It has a t-test of 4.51 and the adjusted R^2 for that equation increases from 0.886 to 0.940. The research coefficient itself decreased from 0.28 to 0.23. Other specifications for the soybeans equation were tried with the spillover variable included and the coefficient of the research variable for soybeans remained highly significant and stable.

The addition of a research spillover variable for wheat increases the R^2 for the wheat function slightly from that of Table 4 and it strengthens the significance of the rain and research variables in the equation while leaving the coefficient on the "traditional inputs" variable highly significant.

Table 3. Research Production Functions with Land, Labor, and Machinery Aggregated*

	Land, Labor, Machinery	Fertilizer	Chemicals	Rain	Research	Constant	$\bar{R}^2$
Corn	-.14	.78	.37	.28	-.08	3.54	.962
	(-.95)	(3.98)	(2.35)	(5.29)	(-.79)	(3.40)	
Wheat	.47	.31	.12	.03	.12	.75	.948
	(4.91)	(2.63)	(4.78)	(1.86)	(1.82)	(.85)	
Soybeans	.58	-.10	.31	.20	.06	-3.83	.889
	(3.56)	(-.76)	(1.63)	(2.46)	(.28)	(-2.61)	

*Numbers in parentheses are t-values.

Table 4. Research Production Functions with all Traditional Variables Aggregated*

	Land, Labor, Machinery, Chemicals, Fertilizer	Rain	Research	Constant	$\bar{R}^2$
Corn	.72	.33	.20	1.98	.883
	(5.16)	(3.80)	(1.20)	(1.09)	
Wheat	.75	.02	.27	-.17	.899
	(8.69)	(.94)	(3.37)	(-.14)	
Soybeans	.66	.26	.28	-4.63	.886
	(5.25)	(3.50)	(1.74)	(-3.25)	

*Numbers in parentheses are t-values.

Table 5. Research Production Functions with Research Spillover Variable*

	Land, Labor, Machinery, Chemicals, Fertilizer	Rain	Own Research	Research Spillover	Constant	$\bar{R}^2$
Corn	.68	.34	.20	.06	1.89	.878
	(4.26)	(3.76)	(1.22)	(.53)	(1.02)	
Wheat	.67	.02	.27	.14	-.249	.914
	(7.75)	(1.07)	(3.57)	(2.49)	(-.22)	
Soybeans	.64	.15	.24	.33	-7.95	.940
	(6.93)	(2.63)	(2.01)	(4.51)	(-6.25)	

*Numbers in parentheses are t-values.

Marginal Products and Rates of Return

Notwithstanding the danger of placing too much confidence in the exact size of the research coefficients, the coefficients from Table 5 were utilized to compute the marginal products of experiment station research. Estimates of national marginal products of research for corn, wheat, and soybeans are obtained by multiplying the research coefficient for each commodity by its respective average product of research.2/ These estimates were then prorated or discounted by dividing them by three to take account of the contributions of extension and private research. Arguments supporting this procedure are presented in Bredahl and Peterson but alternative proratings of benefits are easy to make and it seems unlikely that the proportional contribution of public research is less than that of private research or extension. The resulting long-run marginal product approximations are shown in Table 6.

Table 6. Marginal Products and Internal Rates of Return*

	Marginal Products	Assumed Lag	IRR(%)
Corn	97	6	115
Wheat	59	6	97
Soybeans	103	6	118

*Calculated from the equations in Table 5 which include spillover variables for all three crops. IRRs calculated from equations in Table 4 (without spillover variables) differ only slightly in magnitude from these.

The calculation of internal rates of return3/ requires that the future returns be discounted. A mean lag of six years is assumed for research on each of these crops. This is consistent with empirical studies such as Evenson's on the length of the lag. Breeding research probably has a somewhat longer lag but other types of crop research probably have a shorter lag. Two facts stand out with regard to the IRRs shown in Table 6. First, they are extremely high and, even if discounted severely for possible error, suggest underfunding of public research for these crops. Second, the IRR is of the same general order of magnitude for each crop. The latter fact suggests that research dollars are probably being allocated reasonably efficiently among the three crops. Moreover, the interstate allocation of research dollars for the three crops appears consistent with the relative economic importance of the three crops at least for those states where these specific crops are of major economic significance.

Summary and Conclusions

The regression results presented in this paper illustrate the data problems involved in trying to use the production function approach in individual commodity research evaluations. Yet, decisions relative to the allocation of research funds are often commodity-specific and even specific to such research functions as plant breeding, analyses of the effects of soil fertility, mechanization of production and/or harvesting, disease and/or insect control, improved marketing systems, etc. And, where feasible, efforts to evaluate the results of these and other lines of agricultural research provide additional insights into the potential payoff for alternative research expenditures.

One should not place excessive confidence in the exact size of the research coefficients reported in Tables 4 and 5 or in the internal rates of return calculated from them (Table 6) due to the possible specification error resulting from the aggregation of several input variables. Also, we are measuring returns to an annual flow of research expenditures when a portion of these returns might reasonably be attributed to prior period investments in the research system and/or to investments in more basic or general purpose research. Moreover, our measure of output is for a single year only and it will vary some between years. With these cautions in mind, we believe that the evidence indicates that returns to agricultural research continue to be high and well in excess of their investment cost. Moreover, the returns to research on corn, wheat, and soybeans are high and funds appear to have been allocated reasonably efficiently across these crops in the early 1970s. The exact magnitude of those returns, however, is open to some question because of data problems for inputs and uncertainty as to the contributions of private research and to the education and information dissemination functions.

Finally, we conclude at this stage of our analysis that the most useful estimates of research returns for individual crops are probably obtained via a production function formulation which aggregates all or most traditional production inputs4/ but which provides for separate specification of major weather effects and research expenditures, the latter including some operational measure of spillover between geographical areas. In our judgment, improved specification of weather and research spillover variables should be the subjects of additional research. If feasible, so should the separation of genetic-related research from that of improved cultural and husbandry practices. And, the current controversy relative to displacement of labor via mechanization suggests a strong case for separating out mechanization research. We believe, however, that the latter is probably a more relevant issue in the labor-intensive specialty crops than for corn, wheat, and soybeans.

Footnotes

1/For example, the following simple correlations exist between land, labor, and machinery on a per farm basis for aggregate cash grain in the 1969 and 1974 Agricultural Census data:

	1969			1974	
	Machinery	Land		Machinery	Land
Land	0.77		Land	0.90	
Labor	0.46	0.80	Labor	0.73	0.61

2/Geometric mean levels of output and research are used in calculating the average products.

3/Previous authors have used varying formulas for computing the IRRs to research (Davis). Differences in these formulas stem from the assumptions made about the distribution of benefits over time. In this study the assumption was made that all benefits occur in the sixth year after the research expenditures which should provide underestimates of the IRRs.

4/We are also exploring other alternatives, including the use of ridge regression techniques, to deal with the problem of high intercorrelations among independent (input) variables.

References

{1} Bredahl, M., and W. Peterson. "The productivity and allocation of research: U.S. agricultural experiment stations." American Journal of Agricultural Economics 58:684-692 1976.

{2} Davis, J. Stability of the Research Production Coefficient for U.S. Agriculture. Unpublished Ph.D. Dissertation, University of Minnesota, 1979.

{3} Evenson, R. "The contribution of agricultural research to production." Journal of Farm Economics 49:1415-1425 December 1967.

{4} Griliches, Z. "Research costs and social returns: Hybrid corn and related innovations." Journal of Political Economy 66:419-431 October 1958.

{5} Norton, G. and J. Davis. "Review of Methods Used to Evaluate Returns to Agricultural Research." Staff paper P79-16, Department of Agricultural and Applied Economics, University of Minnesota, May 1979.

{6} Norton, G. W. "The Productivity and Allocation of Research: U.S. Agricultural Experiment Stations, Revisited." Paper prepared for the Symposium on Methodology for Evaluation of Agricultural Research, May 12-13, 1980, St. Paul, Minnesota.

{7} Krenz, R. "Firm Enterprise Data System Budgets." USDA, ERS. Prepared at Oklahoma State University, 1977.

APPENDIX A--VARIABLES FOR INDIVIDUAL CROP PRODUCTION FUNCTIONS

Variable

1. Output - Total value of crop sold
Multiplied each state's production by national average price. USDA publication Crop Production, Annual Summary gives production data and USDA publication Crop Values gives price data.

2. Land - Area planted to crops
Found in USDA publication Crop Production, Annual Summary.

3. Labor - Value of labor
Multiplied hours of machinery labor from "FEDS" budgets by farm wage rate for the United States found in USDA publication Farm Labor.

4. Fertilizer - Average U.S. prices for N, P, and K used to sum up N P & K into one variable. These prices were found in USDA publication Costs of Producing Food Grains, Feed Grains, Oilseeds, and Cotton.

5. Chemicals - Value of herbicides and insecticides from FEDS crop budgets deflated by the price of the appropriate herbicide and insecticide for each crop. For example, the corn herbicide value was deflated by the ratio of the national to the state price of atrazine.

6. Machinery - Sum of (1) fuel and lube, (2) service flow of machinery stock and (3) custom hire of machinery
(1) Fuel and lube - Value of fuel and lube from crop budgets deflated by the weighted national average price of gasoline and diesel fuel divided by the weighted state price of gasoline and diesel fuel in each state.
(2) Ownership costs from FEDS budgets.
(3) Custom hire from FEDS budgets.

7. Weather - July rainfall (deviations from normal).

8. Soils - Slope dummies on land variables based roughly on 1957 Yearbook of Agriculture land groups, for wheat and more aggregated groups for soybeans and corn.

9. Research - Total expenditure on research for particular commodity from Inventory of Agricultural Research FY 1970-1972 average.

10. Research Spillover - Soybeans: Research expenditures on soybeans for other states at the

same latitude which fall within the same recognized "soybean group." If only a portion of a state is included in the same "soybean group," a production-weighted proportion of that state's research is included in the spillover variable.

Corn and wheat: Research expenditures on corn and wheat for bordering states which fall within the same geoclimatic region. The wheat regions were based on those delineated by Davis and the corn regions were based on corn maturity zones. If only a portion of a bordering state is included in the same geoclimatic region, a production-weighted porportion of the state's research is included in the research variable.

APPENDIX TABLE B
DATA SOURCES

1. United States Department of Agriculture, Agricultural Statistics, 1971, Washington, D.C.

2.__________, 1977 Annual Prices Summary, Crop Reporting Board, ESCS, USDA, June 1978.

3.__________, Cost of Producing Selected Crops in the United States - 1976-1977 and Projections for 1978, ESCS, USDA, March 1978.

4.__________, 1978 Crop Production, Annual Summary, Crop Reporting Board, ESCS, USDA, January 1978.

5.__________, Crop Value 1976-1977-1978, Crop Reporting Board, ESCS, USDA, February 1978.

6.__________, Farm Labor, Crop Reporting Board, ESCS, USDA, February 1978.

7.__________, Farmers' Use of Pesticides, ESCS, Ag Economic Report No. 418, 1978, Washington, D.C.

8.__________, Inventory of Agricultural Research, FY 1970, Vol. II, Science and Education Staff, 1970.

9.__________, Inventory of Agricultural Research, FY 1971, Vol. II, Science and Education Staff, 1971.

10.__________, Inventory of Agricultural Research, FY 1972, Vol. II, Science and Education Staff, 1972.

11.__________, Federal Enterprise Data System, 1977 budgets, ESCS, USDA, (obtained from Ronald Krenz, Oklahoma State University).

12.__________, The Yearbook of Agriculture 1957, Soil, USDA, Washington, D.C.

13.__________, Unpublished data on deviation from normal July and annual rainfall provided by Michael Weiss at USDA.

APPENDIX C:
States Included in Land Dummies

Wheat:	34 states
Slope 1.	KY, MD, N.Y., N.C., PA, TENN, VA
Slope 2.	ALA, ARK, GA, MISS, S.C.
Slope 3.	COLO, KANS, OKLA, TEX
Slope 4.	MONT, NEBR, N. DAK, S. DAK
Slope 5.	ARIZ, CALIF, IDAHO, N. MEX, OREG, WASH
Land	ILL, IND, IOWA, MICH, MINN, MO, OHIO
Corn:	23 states
Slope 1.	DEL, KY, MD, N.J., N.Y., N.C., PA, TENN, VA
Slope 2.	COLO, KANS, NEBR, TEX
Land	ILL, IND, IOWA, MICH, MINN, MO, N. DAK, OHIO, S. DAK, WIS
Soybeans:	26 states
Slope 1.	ALA, GA, OKLA, S.C., TENN, TEX
Slope 2.	DEL, KY, MD, N.J., N.C., VA
Slope 3.	ARK, LA, MISS
Land	ILL, IND, IOWA, KANS, MICH, MINN, MO, NEBR, OHIO, S. DAK, WIS