AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Staff Papers Series

P81-12 June 1981

THE USE OF GENETICS PRINCIPLES IN RESEARCH EVALUATION:

AN EXAMPLE WITH SOYBEANS

George W. Norton and W. B. Sundquist

Department of Agricultural and Applied Economics

University of Minnesota
Institute of Agriculture, Forestry and Home Economics
St. Paul, Minnesota 55108

The Use of Genetics Principles in Research Evaluation -- An Example with Soybeans

by

George W. Norton and W. B. Sundquist*

*The authors are assistant professor, Dept. of Agricultural Economics, Virginia Polytechnic Institute and State University, and professor, Dept. of Agricultural and Applied Economics, University of Minnesota.

The authors would like to thank Yoav Kislev and Jean Lambert for their helpful comments on an earlier draft of this paper without implicating them in any remaining errors. This research was supported in part by funds from the North Central Region Agricultural Experiment Stations and by Cooperative Research, SEA, USDA.

Abstract

The Use of Genetics Principles in Research Evaluation --
An Example with Soybeans

This paper explores the potential use of quantitative genetics principles in evaluating economic returns to plant breeding research. Basic factors affecting genetic progress are described along with possibilities for quantifying them in relation to research expenditures. An example with soybeans illustrates how this information can be incorporated into ex ante research evaluation.

Staff papers are published without formal review within the Department of Agricultural and Applied Economics.

The Use of Genetics Principles in Research Evaluation -- An Example with Soybeans

Much of the work on agricultural research evaluation has focused on estimation of rates of return to agricultural research in the aggregate, to research on groups of commodities such as cash grains or poultry, and, more recently, to research on individual commodities such as wheat, corn, and soybeans. For example, Sundquist, Cheng, and Norton using a production function approach estimated the marginal internal rate of return to agricultural experiment station soybean research in the United States to be in the neighborhood of 100% for research expenditures incurred during the early 1970's. This "ex-post" estimate provides an "ex ante" guide to the potential payoff of future soybean research expenditures provided the physical productivity of research and its economic value remains fairly stable over time.[1]

Research on commodities such as soybeans does not take place, however, at the aggregate level but in programs and projects in plant breeding, plant pathology, soil fertility, moisture control, economics, etc. Benefit cost analysis has been used to evaluate certain of these programs such as plant breeding and crop protection (Araji, Sim, and Gardner; Easter and Norton). Others such as economics have received very little attention primarily due to problems of measuring their output. Even plant breeding research progress, however, has been treated very much like a black box by those economists evaluating research returns. Research dollars flow in and increased yields flow out with very little understanding of how this genetic progress occurs. Projections of future yield increases are sometimes based on scientists'

estimates, but economic analysts have no good way of judging whether or not the scientists were unrealistic in their assumptions and/or their expectations.

Kislev and Rabiner in a study of animal breeding research suggested making use of quantitative genetics principles to help economists better understand how genetic progress occurs. In this paper we explore the potential for using a few selected genetics principles to aid in economic evaluation of plant breeding research using soybeans as an example. Basic factors affecting genetic progress are described and possibilities for quantifying their level in relation to research dollars expended are discussed. Finally, we show how this information can be incorporated into ex ante research evaluation.

Soybean Selection

Soybean improvement programs place major emphasis on development of disease and nematode resistant high-yielding varieties that are also high in oil and protein (Brim). Yield as well as oil and protein content are "quantitative genetic traits" while disease resistance is a "qualitative trait". The basic distinction between quantitative and qualitative traits is that the former varies by degree in the plant or animal and is controlled by many genes while the latter tends to either be present or not and is controlled by one or very few genes. Plant breeding programs for self fertilized crops such as soybeans normally involve three levels of evaluation (Hanson and Brim). First, individual plants or rows of progeny (offspring) of single plants are evaluated in one environment and selected for certain qualitatively inherited characteristics such as disease resistance and for certain readily

inherited quantitative characteristics such as maturity, lodging, and plant height. Second, a detailed evaluation of lines for key measurable characteristics, particularly yield, is made to reduce a set of genotypes to a few select lines. Finally, the selected genotypes are submitted to regional testing for evaluation under a more diverse set of environmental conditions. The term genotype is used to refer to the inherent or true genetic attributes of plants while the term phenotype refers to those attributes which are physically observable or measurable. While the genotype is unobservable it, none-the-less, determines the properties the individual transmits to its offspring. The phenotype, on the other hand, can be affected significantly by the environment.[2]

Plant and animal breeders refer to heritability as the fraction of phenotypic variance that arises from genetic affects. Conceptually, phenotypic variance (σ_p^2) can be thought of as the sum of the genotypic variance (σ_g^2) and environmental variance (σ_e^2). Heritability (H) in its broadest or simplest form is then:

$$H = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$$

Because of the nonadditivity of some contributors to genetic variance, a more useful formula is:

$$H = \frac{\sigma_a^2}{\sigma_g^2 + \sigma_e^2}$$

where σ_a^2 includes only the additive genetic component. Components of

(σ_g^2) which are not additive and are retained only in part during selection include what geneticists refer to as dominance and epistatic effects.[3] Evidence indicates, however, that dominance effects are small in soybeans and the distinction may be of little importance in practice because of large standard errors associated with the heritability estimates (Brim and Cockerham, Sprague). Of more significance is the fact that heritability estimates calculated for soybeans have varied in their unit of evaluation. A plant, a full plot, and replicated full plots in one or more environments have been used as the evaluation unit and these differences affect heritability estimates. Heritability is therefore not a stable parameter but varies with the precision with which the environmental variance is estimated (Sprague, p. 332). Nonetheless, enough heritability estimates have been made to give plant breeders a rough idea of their expected values for different quantitative characteristics in soybeans. Table 1 provides estimates presented by Johnson and Bernard which are based on observations of several experiments by soybean breeders and geneticists.

Table 1. Expected Heritabilities

Character	Selection Unit	
	F_2 Plant	Mean of F_4 or later generation two replications within two environments
Yield	5	38
Height	45	75
Maturity	55	78
Resistance to Lodging	10	54
Percent Oil	30	67
Percent Protein	25	63

These estimates of expected heritabilities provide a guide to the relative ease of selecting for various characteristics. For example, yield is less heritable than oil content. More importantly these heritabilities can be combined with the selection differential to provide an estimate of the rate of genetic progress expected with selection. The selection differential (D) is the difference between the mean of the selected individuals or families and the mean of the population. The expected genetic progress E(G) is then $E(G) = (\frac{HD}{Y})$ where Y is the number of years per cycle of selection (Fehr).[4] If the difference between the mean of the selected group and the population mean was say 2 bushels per acre and the heritability was .4 with four years per cycle, then the breeder could expect to gain .2 bushels per acre per year by his selection.

The value of D depends on the proportion of the population group included among the selected group and the phenotypic standard deviation of the character (σ_p). The intensity of selection (i) which equals D/σ_p depends on the proportion of the population included in the selection group and if the distribution of phenotypic values is normal, it can be determined from tables of the properties of the normal distribution (Falconer, p. 194). For example, in the case of five percent selection in a large sample from a normal distribution, $i = 2.06$ and $D = 2.06\sigma_p$.

Selection intensity affects the cost of a selection program. The more intense the selection, the lower the proportion selected and the higher the selection differential which in turn increases expected genetic progress E(G). One does not want to reduce excessively the number of plants selected because of a problem referred to as genetic

drift.[5] Therefore to decrease the proportion selected it is necessary to increase the total number of lines measured out of which the selection is made. This involves a cost in time and labor for measurement and increases the complexity of the selection program from a management standpoint.

Another major determinant of the variable cost of a selection program is the number of locations in each test. Varying the number of locations affects the heritability estimate through an effect on the phenotypic variance. Additional locations will reduce the phenotypic variance if the interaction between genetic and environmental effects is large, thereby increasing heritability and expected genetic progress.

The number of years per cycle of selection, i.e., the time interval from when parents are crossed until selection from this progeny are available for crossing again, also influences G_y. Investing in winter nursery programs in greenhouses or tropical nurseries are ways of reducing Y.

There are also possibilities for increasing G by increasing the genetic variance through changes in breeding procedures such as increasing the number of two way crosses evaluated, using intermating populations, or exotic germplasm (Fehr). Genetic engineering may soon become economically feasible. But it's direct impact on plant breeding programs appears to be some years away. Thus, assuming that traditional soybean breeding procedures are not altered materially, additional research dollars can affect genetic progress by allowing for an increase in selection intensity, increase in locations for each test, and reductions in years per cycle.[6]

To the extent that scientists have made heritability estimates and have knowledge of the cost of increasing the number of lines selected, the number of locations, or nursery facilities, this information can be used to give a rough prediction of genetic progress expected from additional research expenditures. This can then be combined with information on lags, adoption rates, prices, acreages, etc., in ex ante benefit-cost analysis.

An additional factor to consider when evaluating soybean breeding research is the sensitivity of soybeans to length of photoperiod (day length). As a method of describing soybean responsiveness to daylength, ten maturity groups have geen established for identifying the region of adaptation for soybean varieties in the U.S. and Canada. Groups 00, 0, and I are adapted to the more northern latitudes of the continental U.S. and Canada while succeeding groups through group VIII are adaptable farther south. A germplasm collection is maintained at Urbana, Illinois for material of group IV or earlier. Material for groups V through VIII has been maintained at Stoneville, Mississippi.

The implications of this photoperiod sensitivity are three fold. First, varieties released by an experiment station in one state are often adopted by farmers in all states raising soybeans at the same general latitude. For example, Chippewa 64 of group I maturity was raised in parts of several states stretching from South Dakota to northern Pennsylvania in the 1960's and early 1970's. Therefore, in ex ante plant breeding research evaluation, the area assumed to be affected by one state's research should encompass portions of several states. Second, the probability that the work conducted in one state will

eventually lead to a variety which is released is reduced because of work going on in other states. Unnecessary duplication has been minimized, however, due to the high degree of communication and exchange of plant material among plant breeders from states at similar latitudes. Third, because plant breeders are working in several states over which a variety would be adopted, the plant population from which lines are selected is the total of the populations being tested in these individual states. While one might argue that it would be more efficient to concentrate this work in fewer locations, it is also likely that trials in several locations reduces the phenotypic variation due to genotype-environmental interaction.

The current allocative efficiency for research dollars spent reducing years per cycle at one location versus increasing the total population from which lines are selected at several locations may be an important question. Research dollars should be allocated such that the expected progress for the last dollar spent on increasing selection intensity at either one location or at several locations equals the expected progress from the last dollar spent on methods to reduce time per cycle of selection.

A numerical example - We turn now to a possible use of the concepts presented above in evaluating soybean breeding research in a particular state. Assume there is a proposal to add an additional $100 thousand annually for the purpose of increasing plant breeding research on soybeans in Minnesota. Furthermore, assume that this research will be directed primarily at selection for yield improvement.[7] An evaluation is requested of the projected economic return to this proposed public investment.

A two step procedure can be followed. First, estimate the expected increase in genetic progress and second, incorporate the increase into a benefit cost analysis. Using the formula for expected rate of genetic progress per year $E(G) = \frac{HD}{Y}$ described earlier, the current expected rate of genetic progress is calculated. Then comparisons are made of the relative expected increases in yield achieved by spending the additional resources on alternative research strategies that affect G, e.g., shortening the time per cycle, increasing locations, increasing selection intensity, etc.

Assume that G is measured in Kg/ha, that heritability (H) is .4, and that 10 percent is selected in a large sample from a normal distribution so that the selection intensity (i) equals 1.75 and the selection differential (D) equals $1.75\sigma_p$. Assume that in recent improvement experiments σ_p was 70 Kg[8/] and that years per cycle of selection currently equal 4. Therefore E(G) for current research is $E(G) = \frac{(.4)(1.75)(70)}{4} = 12.25$ Kg.

Assume that it is estimated that with an additional $100 thousand the number of years per cycle of selection could be reduced from 4 to 3 causing an increase in the expected rate of genetic progress of 4.08 Kg per year. Assume it is estimated that with $100 thousand the population from which selections are made could be increased so that a 5 percent selection is made instead of 10 percent increasing (i) to 2.06 and E(G) by 2.17 Kg per year. Assume that after these and other uses for the additional funds (such as adding locations to reduce σ_p^2 and increase H) are compared that the greatest expected gain is from decreasing the time per cycle of selection. This rate of 4.08 Kg per year can then be

incorporated with other assumptions of prices, lags, area affected, etc. in a benefit cost analysis.

Assume: (1) a lag of eight years between the first year of additional research expenditures and the first release of a new variety, (2) an initial soybean price of $6.00 per bushel and a 1% reduction in price for every 1% increase in production[9/] for the United States, (3) an adoption pattern of 40% in the first year after variety release, 70% in the second year, 90% in the third year, and 95% in succeeding years (Easter and Norton), (4) the additional funding starts in year 1 and occurs for 10 years and all benefits drop to zero in year 18, (5) a probability that 50% of the work will be successful. The other 50% will be superceded by work carried on in other states.[10] (6) Harvested acreage affected by the additional breeding research is the soybean acreage in group I in the United States which equals approximately 12 million acres. This acreage is assumed to remain constant. (7) The higher yielding soybeans will not entail additional production costs on the part of the farmers or social costs on the part of society.

These assumptions and estimates are used to calculate an internal rate of return to the new soybean research investment using the following formula:

$$\sum_{t=1}^{18} \frac{\beta_t - C_t}{(1+r)^t} = 0$$

where $\beta_t = A_t \Delta G V_t P_t$ = expected benefits in year t, A_t = area planted to soybeans in year t affected by additional genetic improvement, ΔG = increase in yield per acre due to new research, V_t = expected price of soybeans in year t, P_t = probability of success, r = internal rate of return, and C_t = cost of the additional research in year t.

An internal rate of return of 66% was calculated for the additional research investment in this numerical example. This example is only illustrative, however, as many of the parameter values assumed need to be developed more carefully by plant breeders before conducting an actual analysis of returns to soybean breeding research.

Sensitivity analysis - When conducting this type of analysis it is useful to analyze the sensitivity of the results to changes in assumptions, lags, spillover effects, prices, etc.

In Table 3, effects are shown of assuming that research benefits (1) affect only Minnesota, (2) affect Minnesota plus bordering states in Group I maturity, and (3) are realized with a 10 instead of an 8 year lag.

Table 3. Sensitivity Analysis of Returns to Soybean Breeding Research

Area Affected by Research	Research Lag	Internal Rate of Return (%)
(1) All Group I	8	66
(2) Minnesota	8	50
(3) Minnesota, Iowa, Wisconsin, South Dakota	8	60
(4) All Group I	10	50

Note that the effect of the extra 2 year lag is the same as the effect of assuming that no benefits are realized outside of Minnesota. While acreage affected in the latter case is reduced from 12 million to 4.9 million acres, the 2 year longer lag has an equally great effect because of the discounting of future returns. Sensitivity analysis in which prices are varied is not shown here, but the effect of varying price 10% would be identical to varying acreage impacted by 10%.

Further Research

This paper has suggested a conceptual framework for economic evaluation of plant breeding research using genetics principles. The next step is to test how well it fits the historical yield development using actual data. Since not all breeding is aimed at increasing yields, this will require information from geneticists or plant breeders on the interaction among traits. Part of breeding research is aimed at "maintenance" of yields. Therefore, if data can be obtained from breeders to estimate actual phenotypic standard deviations, years per cycle, proportions selected, etc. from their experiments, then the potential rate of yield increase can be calculated. Much of the gap at the experimental level between potential and actual rates of yield increase can then be attributed to needs of maintenance.[11, 12]

Conclusions -In this paper we have explored the potential for using a few quantitative genetics principles as an aid in evaluating the economic returns to plant breeding research. Though our example dealt with soybeans, the conceptual model should be applicable to other crops as well. Economists are increasingly called upon to evaluate the merits of proposed agricultural research. Unless a limited understanding is developed by economists (or other evaluators) of how progress occurs in an area such as plant breeding, they have very little basis for judging (1) what bounds to place on expected yield increases, (2) what the constraining factors are to increased genetic progress and how large a yield increase could be expected by targeting resources at relieving alternative constraints, and (3) the degree of spillover between states or other geographical units or the relevant distance over which the benefits can be expected to spread.

Answers to these questions add significant information to the research evaluation process. Not only do they facilitate calculation of rates of return but they can potentially provide guidance for such questions as (1) the appropriate degree of research concentration within a region and (2) the relative efficiency of such research strategies as expanding winter greenhouse facilities versus expanding the population from which lines are selected. Moreover, data availability for such evaluations will be enhanced if plant breeders attempt to systematically acquire and record the data required to estimate expected genetic gain for a range of feasible research strategies.

The high rates of return obtained in nearly all research evaluation studies (Evenson, Waggoner, and Ruttan) indicates a historical underinvestment in agricultural research. The type of analysis illustrated in this paper can be used as a basis to provide further evaluation of the economic benefits to prososed research on individual commodities and on specific research programs.

Footnotes

1. If the rate of genetic improvement for yields does not deteriorate over time, this will lend stability to the research coefficient. Other factors affecting the coefficient are susceptibility of varieties to insects and diseases or changes in the quality of scientists or support facilities. Even if the physical productivity of research does remain stable over time, its economic value may not due to shifts in product demand which cause price effects.

2. A comprehensive theoretical depiction of the contributors to yield variance in an experimental context is provided by Comstock and Moll.

3. "The word additive applies to relationships of genes at the same locus... with additive gene action the phenotype faithfully reflects the genotype, assuming no environmental effects ... Applied to two alleles A and a, affecting height, the increase of AA over aa is twice that of Aa.... Dominance applies to the relationship of alleles at the same locus. Complete dominance assumes, in our example of height, that Aa and AA are the same height ... If overdominance were present in the example of height, the genotype AA might be 26" in height, aa would be 10 inches and Aa would be more than AA, perhaps 30 inches. In other words there is some interaction between A and a to increase height... Originally and usually with qualitative factors, epistasis is the

term used for the situation where a gene covers up the effects of another gene at another locus. In quantitative inheritance, the term has been extended to include all situations where genes interact at two or more loci." (Briggs and Knowles, pp. 86-90).

4. A cycle of selection is the time interval from when parents are crossed until selection from their progeny are available for crossing again.

5. Genetic drift refers to the accidental spreading of a gene. "Gene frequencies are subject to fluctuation about their mean from generation to generation. If a population is large, the numerical fluctuations are small and have little or no effect. On the other hand, if the population is small, random fluctuations could lead to complete fixation (or alternatively, loss) of one allele or another" (Gardner).

6. Even when genetic engineering procedures can be incorporated operationally into plant breeding programs it may be possible to partition their genetic effects into one or more of these three categories, or to add others.

7. This example abstracts from the fact that a great deal of plant breeding research is directed at maintaining current yield through breeding in disease resistance and is conducted concurrently with efforts to select for higher yield.

8. 70 Kg was the σ_p assumed in a similar example by Fehr.

9. This price flexibility of 1 approximates that found in the National Academy of Science Report.

10. This assumption may bias the return downward if the plant breeders at the same latitude are drawing from the same plant population and are sharing plant materials.

11. This application of the conceptual model was suggested by Yoav Kislev.

12. A second gap, the nature of which also needs exploring, occurs between experimental and farm level yields.

References

Araji, A., R Sim, and R. Gardner, "Returns to Agricultural Research and Extension Programs. An Ex-Ante Approach," Amer. J. of Agr. Economics, Vol. 60, No. 5. December 1978, pp. 964-968.

Briggs, F. and P. Knowles, Introduction to Plant Breeding, Reinhold Publishing Corp., New York, 1967.

Brim, C., "Quantitative Genetics and Breeding," Chapter 5 in Soybeans: Improvement, Production, and Use, Editor B. E. Caldwell, Am. Society of Agronomy, Madison Wisconsin, 1973.

Brim, C. and C. Cockerham, "Inheritance of Quantitative Characters in Soybeans," Crop Science, No. 1, 1961, pp. 187-190.

Comstock R. E., and R. Moll, "Genotype-Environment Interactions," Statistical Genetics and Plant Breeding. NAS, NCR Publication 982:164-196. 1963.

Easter, K. and G. Norton, "Potential Returns from Increased Research Budget for the Land Grant Universities," Agricultural Economic Research, Vol. 29, No. 4, October 1977, pp. 127-133.

Evenson, R., P. Waggoner, and V. Ruttan, "Economic Benefits from Research. An Example from Agriculture," Science, 14 Sept. 1979, Vol. 205, pp. 1101-1107.

Falconer, O., Introduction to Quantitative Genetics, The Ronald Press, New York, 1961.

Fehr, W. "Description and Evaluation of Possible New Breeding Methods for Soybeans," in World Soybean Research, Proceedings of the World Soybean Research Conference, ed. by L. D. Hill, Interstate Printers and Publishers, Danville, Ill. 1976.

Gardner, E. J., Principles of Genetics. John Wiley Press, New York, 1960.

Hanson, W., "Heritability" in Statistical Genetics and Plant Breeding, ed. by W. Hanson and H. Robinson, Pub. 982 NAS-NRC, Washington, D.C. 1963.

Hanson, W. and C. Brim, "Optimum Allocation of Test Material for Two-Stage Testing with an Application to Evaluation of Soybean Lines", Crop Science No. 3, 1963, pp. 43-49.

Johnson, H. and R. Bernard, "Soybean Genetics and Breeding," in Soybean ed. by A. G. Norman, Academic Press, New York, 1963.

Kislev, Y. and U. Rabiner, "Economic Aspects of Selection in the Dairy Herd in Israel", Australian Journal of Agricultural Economics, Vol. 23, No. 2, August, 1979, pp. 128-146.

National Academy of Science, Pest Control: An Assessment of Present and Alternative Technologies, Vol. II Corn/Soybeans Pest Control, Washington, D.C. 1975.

Sprague, G., "Quantitative Genetics and Plant Improvement," Chapter 8 in Plant Breeding, ed. by K. J. Frey, 1967, pp. 315-354.

Sundquist, W., C. Cheng, and G. Norton "Measuring the Returns to Research for Corn, Wheat, and Soybeans," Staff Paper P80 - Department of Agricultural and Applied Economics, University of Minnesota, August 1980.