The Effects of Research on Wheat Yields and Yield Variances

by

Senhui He

Fred C. White

Stanley Fletcher

(Senhui He is a post-doctoral associate, Fred C. White and Stanley Fletcher are professors, of the Department of Agricultural and Applied Economics, University of Georgia. The contact author is Senhui He, phone 770-228-7231 ext:109, fax 770-228-7208, email shenhe@gaes.griffin.peachnet.edu.)

Selected Paper Prepared for Presentation at the Southern Agricultural Economics Association Annual Meeting, Mobile, Alabama, February 1-5, 2003.

The Effects of Research on Wheat Yields and Yield Variances

by

Senhui He, Fred C. White, and Stanley Fletcher

Abstract:

This paper investigates the impacts of research investment on wheat production. A two-way error component random effect model was estimated using panel data to examine the effects of research investment on mean yield and yield variance. Results show that, for the time period between 1974 and 1995, research investment enhanced yield stability but had no significant impact on mean yield.

Key words: yield, yield variance, research investment.

The Effects of Research on Wheat Yields and Yield Variances

Since 1950, agricultural output of the United States increased more than 80% while the production inputs under the control of farmers increased only about 10%. On the other hand, real expenditures on public agricultural research increased by about 110% over the same period. The search for the sources of agricultural output growth prompted researchers to examine the impacts of investment in agricultural research.

As far as crop production is concerned, research investment may affect production in many aspects. For instance, wheat varietal research may result in a new variety of higher yield, of which the best known example is the identification and application of the semidwarfing characteristic. However, other improvements, such as systematic breeding for resistance to various rust fungi, the development of broad-habitat varieties and breeding for specific quality characteristics, have also been important. Nevertheless, apart from the determinants of quality, many genotypic characteristics have their main expression through yield. As a matter of fact, the issue of yield has been in the core of research in crop production all the time. As for the agricultural extension services, its main purpose is to promote what gained from research. Therefore, yield performance is an extremely useful summary statistics, reflecting a diversity of objectives of the respective research.

Many economic studies have investigated the effects of research on mean yield and yield variance, the two main characteristics of crop production. Some researchers (Mehra; Hazell) found that research tends to increase both mean yield and yield variance while others (Singh and Byerlee; Anderson, Findlay and Wan) found evidence that research tends to reduce yield variance. Traxler et. al think that yield performance maybe related to crop varieties released in different time. Some researchers suggest that research during the green revolution tends to increase both mean yield and yield variance.

This study investigates the effects of research investment on mean yield and yield variance of wheat. Five major classes of wheat are considered in the study, including hard red winter, hard red spring, soft red winter, white, and durum. Effects of research investment during the 1974-1995 period are examined with an intention to find out how research investment after the Green Revolution affects yield performance. The goal is accomplished through the estimation of a production function for panel data based on the stochastic specification by Just and Pope.

The Framework

It is assumed in this study that an input can have a positive, a negative, or an insignificant effect on mean yield or yield variance or both. Further, it is assumed that the effect of an input on yield variance is not tied to the effect of the input on mean yield *a priori*. To capture the stochastic characteristics of the effects of inputs on yield performance, and taking the panel nature of the data into consideration, the stochastic production function developed by Just and Pope (1979) is modified to express the yield as a function of a set of production factors including research investment

$$\begin{aligned}
& y_{it} = f(x_{it}) + u_{it}, \\
& where \\
& u_{it} = h^{\frac{1}{2}}(x_{it})(m_i + l_t + e_{it}), \\
& f(x_{it}) = \prod_{k=1}^{K} x_{kit}^{a_k} \\
& h(x_{it}) = \prod_{k=1}^{K} x_{kit}^{b_k}
\end{aligned} \tag{1}$$

where y_{it} denotes wheat yield of class i at time t; x_{it} is a 1×K input vector used by class i at time t, including both conventional production inputs and the research investment. The f in the deterministic part of the production function, denoted by f(x_{it}, á), is a function transforming inputs into yields and á is a vector of parameters indicating the effects of the inputs on yield. The h function in the stochastic part u_{it} can be of any form that should remain the same for all stages of estimation and â is a vector of parameters indicating the effects of inputs on yield variance. The error component μ_i denotes the time-invariant class-specific effect which has a distribution iidN(0,ó$_{\mu}^{2}$), ë$_t$ represents the class-invariant time-specific effect which has a distribution iidN(0,ó$_{ë}^{2}$), and ε_{it} is the disturbance term which is distributed iidN(0,ó$_{\varepsilon}^{2}$). The error components are mutually uncorrelated for all i and t. All the error components are heteroskedastic because their covariance matrixes are affected by the explanatory variables. This implies that the class-specific effects such as land quality, the time-specific effects such as weather, and the remainder disturbance are all affected by the application levels of the inputs.

With the assumption that μ_i is time invariant and ë$_t$ is class invariant, the stochastic part is characterized by

$$
\begin{aligned}
&E(u_{it}) = 0,\\
&E(u_{it}^2) = (s_m^2 + s_l^2 + s_e^2)h(x_{it}) = s^2 h(x_{it}),\\
&E(u_{it}u_{js}) = s_m^2 h^{\frac{1}{2}}(x_{it})h^{\frac{1}{2}}(x_{js}), \forall\ i = j,\ t \neq s,\\
&E(u_{it}u_{js}) = s_l^2 h^{\frac{1}{2}}(x_{it})h^{\frac{1}{2}}(x_{js}), \forall\ i \neq j,\ t = s,\\
&E(u_{it}u_{js}) = 0, \forall\ i \neq j,\ t \neq s.\\
&where\\
&s_m^2 = \frac{1}{NT(T-1)/2}\sum_{i=1}^{N}\sum_{s=1}^{T-1}\sum_{t=s+1}^{T}\frac{m_{it}m_{is}}{h_{it}h_{is}},\\
&s_l^2 = \frac{1}{NT(N-1)/2}\sum_{t=1}^{T}\sum_{j=1}^{N-1}\sum_{i=j+1}^{N}\frac{m_{it}m_{jt}}{h_{it}h_{jt}}
\end{aligned}
\qquad \textbf{(2)}
$$

where N is the total number of classes and T is the total number of years considered in the study. To obtain asymptotically efficient estimate of á, the non-spheric nature of the random component part must be accounted for. This can be done in the following steps. First, run a regression on the pooled data to obtain a consistent estimate of á by minimizing the sum of the squares of the residuals with respect to á. The residuals from the regression are consistent estimates of U because á̂ is a vector of consistent estimates (Just and Pope, 1978).

The consistent estimates of u are used to obtain the estimates of â. Taking the log of the squared residuals results in

$$\ln(u_{it}^{2}) = \sum_{k=1}^{K} b_k \ln(x_{kit}) + \ln(m_i + l_t + e_{it})^2 . \tag{3}$$

Following Griffiths and Anderson's computational procedures, equation (3) can be expressed as

$$\ln(u_{it}^{2}) = b_o + \sum_{k=1}^{K} b_k \ln(x_{kit}) + v_{it} . \tag{4}$$

where $â_o = \ln [E(μ_i + ë_t + ε_{it})^2]$, $v_{it} = \ln(μ_i + ë_t + ε_{it})^2 - â_o$. Let s represent for $μ_i + ë_t + ε_{it}$, then s is distributed N(0 , ó). Denote Z = s / ó. Z is a normalized variable with zero mean and unit variance. Z^2 is distributed as a $÷^2$ random variable with a degree of freedom of one. Harvey showed that the log of z^2 has a mean of -1.2704 and a variance of 4.9349. Thus, in our study, $E(\ln(s^2) - \ln(ó^2)) = E[\ln(μ_i + ë_t + ε_{it})^2 - \ln ó^2] = E(â_o + v_{it} - \ln (ó^2)) = -1.2704$ and $\mathrm{var}(â_o + v_{it} - \ln(ó^2)) = 4.9349$. Since $E(v_{it}) = 0$, we have $E(â_o + v_{it} - \ln(ó^2)) = â_o - \ln ó^2 = -1.2704$. Hence, $ó^2 = \exp(â_o + 1.2704)$. On the other hand, since $â_o$ and $\ln(ó^2)$ are constant values, $\mathrm{var}(â_o + v_{it} - \ln(ó^2)) = \mathrm{var}(v_{it}) = 4.9349$. The variance and covariance of z look like

Since ln(z) are monotonic transformation of z, ln(z_{it}) and ln(z_{js}) are correlated in the same pattern as z_{it} and z_{js}. Further, since ln(z_{it}^2) is the sum of v_{it} and some constants, the variance and covariance matrix of v and ln(z) are exactly the same. The variance and covariance matrix of v looks like

$$\begin{aligned}
&E(v_{it}v_{js}) = t^2 = 4.9349 \ \forall\ i = j, t = s,\\
&E(v_{it}v_{js}) = t_m^2 \ \forall\ i = j, t \neq s,\\
&E(v_{it}v_{js}) = t_l^2 \ \forall\ i \neq j, t = s,\\
&E(v_{it}v_{js}) = 0 \ \forall\ i \neq j, t \neq s,\\
&where\\
&t_m^2 = \sum_{g=1}^{\infty}\left(\frac{s_m^2}{s^2}\right)^{2g}\frac{g!\Gamma(0.5)}{g^2\Gamma(g+0.5)},\\
&t_l^2 = \sum_{g=1}^{\infty}\left(\frac{s_l^2}{s^2}\right)^{2g}\frac{g!\Gamma(0.5)}{g^2\Gamma(g+0.5)}
\end{aligned} \qquad (6)$$

$$\begin{aligned}
&E(z_{it}z_{js}) = 1 \ \forall\ i = j, t = s,\\
&E(z_{it}z_{js}) = \frac{s_m^2}{s^2} \ \forall\ i = j, t \neq s,\\
&E(z_{it}z_{js}) = \frac{s_l^2}{s^2} \ \forall\ i \neq j, t = s,\\
&E(z_{it}z_{js}) = 0 \ \forall\ i \neq j, t \neq s.
\end{aligned} \qquad (5)$$

With the information about the variance and covariance matrix v, we can obtain asymptotically efficient estimate for â. Further, with consistent and asymptotically efficient estimate for â, asymptotically efficient estimate for á can be obtained through a feasible generalized least squares (FGLS) regression of equation (1).

Data and Model

The data are obtained from various sources. Data on the yields of hard red winter wheat, hard red spring wheat, soft red winter wheat, white wheat and durum wheat were obtained from various issues of *Wheat Situation and Outlook Yearbook*. Data on annual, national-level expenditures on wheat research from 1970 to 1995 were obtained from the Current Research

Information System, USDA. Data on the costs of wheat seed, fertilizers-lime-gypsum, chemicals, custom operations, fuel-lube-electricity, repairs, hired labor, unpaid labor and capital replacement were collected from various issues of *Costs of Production of Major Field Crops* and *US Agricultural Statistics*.

An empirical model is constructed where yield is specified as a function of research investment, machinery, material, labor, and weather. Yield is the number of bushels of wheat produced per harvested acre. Over the years covered by this research, hard red winter wheat yields fluctuated sharply without obvious trend of either increase or decrease. Yields of all the other classes of wheat show a trend of increase.

The research variable is formed of annual national-level research expenditures for hard red winter wheat, hard red spring wheat, soft red winter wheat, white wheat and durum wheat. White wheat includes club white, western white and soft white. The research expenditures consist of USDA approved federal research expenditures, other federal research expenditures and non-federal research expenditures. The research expenditures are deflated by price deflator for research and measured in million dollars.

Following the practices of previous studies, the machinery, the material, and the labor variable are constructed using the factor costs. The machinery variable is formed of costs on custom operation, repair, and capital replacement. Capital replacement represents a charge sufficient to maintain a machinery investment and production capacity through time. It is the major contributor in the machinery variable.

The materials variable consists of wheat seed, fertilizer-lime-gypsum, chemicals, fuel-lube-electricity. Data on seed quantities used per acre and seed prices of both hybrid varieties and home produced seed were used to form per acre seed cost. Chemical fertilizer, especially nitrogen, plays such an important role in crop production that it is considered to be a major technology component of the green revolution (Taxler et al.) Therefore, the quantity of chemical fertilizers used by major producing states and the corresponding prices were used to form the regional per planted acre fertilizer cost. Fuel costs for tractor are computed on basis of power takeoff, horsepower size and fuel consumption. For other kinds of machines, fuel consumption is gauged at the hourly rate. The total amount of fuel used for production is the sum of all the fuel used for each machine based on the required hours of machine use. The fuel cost is the product of the total amount used per acre and the corresponding fuel prices at state level.

The labor variable includes hired labor and self-employed labor. Agriculture in the United States is mechanized and labor in crop production is mostly involved in machine handling. Machinery labor requirement is directly related to machine time requirement. Labor requirement for other activities, such as for irrigation and hand operation, are also included in the total per acre labor cost.

The weather variable is formed of monthly precipitation of wheat producing states weighted by the production of the states. According to reports in wheat situation and outlook over the years, precipitation often affects wheat production significantly while extreme temperature was occasionally reported to affect wheat production to some extent. Thus, precipitation is used to represent weather. The weather variable for winter wheat is formed of the weighted sum of precipitation in August, September, October, November, December of previous years and January, February and March of the current year. For spring wheat, the weather variable is formed of the weighted sum of precipitation in March, April, May, June and July.

Results

We follow the computational procedures of Griffiths and Anderson in our empirical estimation. First, we run a regression based on equation (1). Second, the residuals from the regression of equation (1) are used to obtain asymptotically efficient estimate of â through a regression of equation (4). Finally, the estimate of â are then used to specify an FGLS regression based on equation (1) to obtain asymptotically efficient estimate of á. The estimates of á and â indicate the effects of the production factors on mean yield and yield variance respectively.

Research is known to have a lag effect. In the estimation, research is treated as a polynomial distributed lagged variable. The lag effect is composed of two stages, the lag between research investment and research discoveries and the lag between research discoveries and their application in production. No effort is made to distinguish these lags. The lag effect is assumed to take quadratic form. Investment in current year is assumed to have no effect. Then, the effect increases until it reaches the peak. After that, with the new discoveries of later research are applied in production, the effect decreases until it totally disappears.

The estimation results are presented in Table 1. The results show that all the production factors have a significant impact on yield performance except weather. Machinery and labor are found to increase yield. A plausible explanation is that with more use of machinery and labor, the crop is better cultivated during the growing season. Yield is generally higher when the crop is better cultivated. The results indicate that labor and machinery tend to reduce yield variance. With more labor and machinery, the producers may plant, cultivate and harvest the crop in a timely fashion, and thus reduces the risk of unexpected loss and stabilizes yield. For example, with more labor and machinery, a producer may be able to get his crop in before the coming of unusual adverse weather and thus avoid the unexpected loss due to bad weather.

The results indicate that research investment does not have a statistically significant impact on mean yield. Research activities may have their main expression through development of new varieties. Development of new varieties in different time may focus on different characteristics of the crop in accordance with demand. Traxler et. al, using experimental data to examine the effects of genetic improvement on wheat yields and yield variance, found that from 1950 to 1979, release of new varieties has positive effect on yield. They found that the rate of yield improvement slowed down in the post-green-revolution years. Further, the yield improvement reached a plateau in the 80s. Our research covers the period from 1978 to 1995 when the yield improvement slows down and reaches a plateau. One possible reason is that greater effort is made to achieve secondary characteristics such as improving quality, increasing insect and disease resistance and shorting crop duration (Traxler et. al.) rather than to increase yield alone. While their study focused on research in genetic improvement alone, the results from this study also suggest that the focus of general research activities may change over time.

Many researches addressed the issue whether research, especially research in genetic improvement, increases or decreases yield variance. Contradictory findings are reported in previous studies. Some researchers (e.g. Mehra; Hazell) reported that research increased yield variance while others (e.g. Singh and Byerlee; Anderson, Findlay and Wan) found evidence that research reduced yield variance. Our results suggest that research investment during 1978-1995 helped to stabilize yield.

Traxler et. al found that yield variance peaked around 1970's release of new varieties. New varieties released in the 1970s are products of the green revolution technology which culminated an era in which wheat breeders achieved rapid yield increases accompanied by higher yield variances. Their results showed that in the years in which the most rapid yield increases were achieved, yield variance also increased, while in the period when progress was made in reducing yield variance, mean yield improvement was slower. It seems there exists a trade off between the effects of research on yield and yield stability. They reported that yield variance decreased after 1970, a finding consistent with our result. It could be that the emphasis of research activities shifted from yield increase to yield stability.

Concluding Remarks

Research activities related to crop production generally have their main expression through their impacts on yield performance and crop quality. The focus of research activities may change over time. Many factors can affect the focus of research activities, including market demand. For example, when demand exceeds supply, yield increase may be a priority. When there is a strong demand for quality improvement, more efforts may be made to improve quality. When risk related to yield variance poses a serious problem to crop cultivators, yield stability may be emphasized in research activities.

There seems to be a change in the effects of wheat research on yield performance at the end of the green revolution. While earlier studies found that the green revolution technology is characterized by increasing both yield and yield variance, we find that investment in wheat research after the green revolution increase yield stability, but does not have a significant impact on yield. It could be that research activities during the green revolution focused mainly on increasing yield while research activities after the green revolution takes more objectives into consideration, especially the intent to reduce the risk related to yield variability.

References:

Anderson J.R., C. J. Findlay, and G.H. Wan. "Are Modern Cultivars More Risky? A question of Stochastic Efficiency." *Variability in Grain yields.* Anderson and P.B.R. Hazell, eds., pp.301-308. Baltimore: Johns Hopkins University Press, 1989.

Griffiths, William E., and Jock R. Anderson " Using Time-Series and Cross-Section Data to Estimate a Production Function With Positive and Negative Marginal Risks" *Journal* of *the American Statistics Association.* Vol. 77, No 379 (Sept. 1982) 529-536.

Harvey, A. C. "Estimating Regression Models with Multiplicative Heteroscedasticity" *Econometrica.* 44 (1976) 461-465.

Hazell, P.B.R. "Source of Increased Instability in India and U.S. Cereal Production." *American Journal of Agricultural Economics* 66(May 1984) 302-311.

Just, Richard E., and Rulon D. Pope "Stochastic Specification of Production functions and Economic Implications" *Journal of econometrics 7 (1978) 67-86.*

Just, Richard E., and Rulon D. Pope "Production Function Estimation and Related Risk Consideration" *American Journal of Agricultural Economics* (May 1979):276-284.

Mehra, S. "Instability in India Agriculture in the Context of the New Technology." International Food Policy Research Institute, Research Report No. 25, Washington DC, July 1981.

Singh, A. J., and D. Byerlee. "Relative Variability in wheat yields Across Countries and Over Time." Journal of Agricultural Economics. 41(January 1990):21-32.

Traxler, Greg, Jose Falck-Zepeda, J. I. Ortiz-Monasterio R., and Ken Sayre "Production risk and the Evaluation of Varietal Technology" *American Journal of Agricultural Economics* 77(February 1995) 1-7.

Table 1. Estimates of the effects of production factors on Mean Yield and yield variance.

	Mean Yield		Yield Variance	
Variable	Coefficient	T-Value	Coefficient	T-Value
Constant	0.889	2.856***	3.024	0.806
Research	0.034	1.168	-0.603	-1.919*
Material	0.453	6.124***	1.165	1.281
machinery	0.18	2.317***	-2.145	-1.978**
Labor	0.167	2.709***	-1.565	-1.805*
Weather	0.04	0.699	-0.341	-0.469
R^2	0.51		0.661	

Note: * denotes significant at 0.1 level, ** denotes significant at 0.05 level, * ** denotes significant at 0.01 level.