Journal of Productivity Analysis, 21, 297–315, 2004

Quantifying Structural Change in U.S. Agriculture: The Case of Research and Productivity

JAMES F. OEHMKE
Department of Agricultural Economics, Michigan State University, East Lansing, MI 48824

DAVID E. SCHIMMELPFENNIG des@ers.usda.gov
Economic Research Service, U.S. Department of Agriculture, 1800 M Street, NW, Washington, DC 20036

Abstract

Previous work on structural change in agriculture has failed to distinguish long-run trends from structural breaks leading to new trends. We measure structural changes as statistically significant breaks in either stochastic or deterministic time trends, and apply these measures to agricultural productivity and research. Productivity has a break in 1925 accompanying agriculture's early experience with the Great Depression. Research trends shifted in 1930 as the Depression and new technology began to strongly influence efficient farm size and capitalization. After modeling lags between research and productivity impacts in a vector autoregression (VAR), we compare our results to earlier work by developing a procedure to estimate the rate of return to research from the impulse response function of the VAR.

JEL Classification: Q16

Keywords: agricultural research, multi-factor productivity, stationarity, structural breaks

Structural change has had a profound impact on U.S. agriculture and agricultural productivity. Evenson (1978) and Huffman and Evenson (1993) anecdotally identify structural shifts in productivity in 1925. Evenson (1980) describes a "significant new pattern of investment" in the agricultural research system that occurred between 1925 and 1935, that Huffman and Evenson (1993) refer to as an institutional transformation. Both of these shifts can be associated with falling output and land prices during the 1920s when thousands of farmers went bankrupt (Cochrane, 1979). A technological revolution began in the 1930s, starting with the redesign of the tractor, that increased the commercialization and capitalization of farming (Rasmussen and Stanton, 1993).[1]

Structural change is often associated with changes in farm size, farm income, and industry structure. Despite the substantial evidence of structural change in U.S. agriculture, analyses rarely incorporate structural change (Antle, 1984; Huffman and

Evenson, 2001 are the exceptions). Pardey and Craig (1989), Huffman and Evenson (1993), Maaki et al. (1999), among others, attempt to quantify the contribution of research to productivity change without modeling structural shifts in either the productivity or the research data.

This paper makes two contributions to the agricultural productivity literature. We present the first estimates of the timing and scope of aggregate agricultural structural change in a time series sense. Second, we apply the results to that hornet's nest of controversy, the effects of research on agricultural productivity. Allowing for the possibility of structural change provides improved econometric estimates of this relationship.

We measure structural change as a statistically significant break in either stochastic or deterministic trends. This approach is counter to earlier work identifying "continuous" (McCorriston and Sheldon, 1991) or "gradual structural change" (Goodwin and Brester, 1995). Continuous or gradual change could be considered a trend. Without reference to structural change, technological advance is often modeled as a Hicks neutral deterministic trend, possibly with a trend-squared term to capture non-linearities. Researchers include these trend terms in their models well aware that they might be picking up, possibly several, trended unobserved effects. To help remedy this problem we identify underlying deterministic and stochastic trends using appropriate time series tests, and then characterize the timing and extent of structural breaks. Our approach attempts to eliminate spurious break points identified through the Alston and Chalfant critique.

The remainder of the paper is organized as follows. The next section discusses conceptual and modeling issues related to the general application of time-series techniques to discerning and modeling structural breaks. The following section applies the tests to research and multi-factor productivity (MFP). The next section discusses the implications of the results for modeling the relationship between research and productivity. The last two sections use an impulse response function to estimate the rate of return to research, and make some concluding comments.

1. Time-Series Modeling of Unit Roots and Structural Shifts

The existing literature on time-series modeling of agricultural productivity suffers from two problems: inadequate univariate testing for unit roots (even were no structural shifts present) and inadequate attention to structural shifts.

The possibility of unit roots in time-series models has raised issues about parameter interpretation and inference (Granger and Newbold, 1974; Sims et al., 1990) and spurious regression, for example when two increasing series are correlated even though they are increasing independently with uncorrelated growth segments (see Hamilton, 1993 or Kennedy, 1992 for a general discussion). Prior investigations of agricultural research and productivity provide tests only of the null hypothesis that the series have unit roots, usually relying on Dickey and Fuller (1981) (DF), augmented Dickey–Fuller (ADF) tests, or Phillips–Perron (1988) tests. The power of these tests is notoriously low, (Phillips and Ouliaris, 1990; Schmidt and Phillips, 1992), leading Maddala and Kim (1998) to suggest that "they should be discarded" (p. 99). Consequently, it may be expected that these tests will fail to reject the null hypothesis of non-stationarity for a number of stationary series. For example, Maaki et al. (1999) conclude that U.S. public research and extension expenditures are non-stationary, even though their estimated autocorrelation is only 0.59.[2] Alston et al. (1998) apply the DF and ADF tests and find that research is non-stationary in levels and stationary in logarithms, which is troubling since time series results should be the same whether a variable is in levels or logs. These prior analyses are also incomplete because none provide a test of the null hypothesis of stationarity (see Maddala and Kim's, (1998), discussion of confirmatory analysis).

A second shortcoming of the agricultural productivity literature is the failure to allow for structural change. The statistical importance of structural change emerged following Nelson and Plosser's (1982) benchmark paper, in which they found that most U.S. macroeconomic time series had unit roots, implying that shocks to any of these series would have permanent effects. However, they did not allow for the possibility of structural shifts. In the spirit of the Lucas (1976) critique, Perron (1989) developed unit-root tests that include the possibility of structural breaks and re-examined the Nelson and Plosser (1982) data. Perron (1989) concluded that "most macroeconomic time series are not characterized by the presence of a unit root and that fluctuations are indeed transitory. Only two events (shocks) have had a permanent effect on the various macroeconomic variables: the Great Crash of 1929 and the oil price shock of 1973" (p. 1362). The implication of this finding is that without unit roots, macroeconomic data can be modeled as consisting of transitory fluctuations around a trend, albeit a trend with breaks or shifts in 1929 and 1973. This raises the distinct possibility that existing findings of unit roots in agricultural research and productivity data could be reversed if one allowed for the possibility of structural breaks or shifts.

2. Univariate Analysis of Unit Roots and Structural Change

2.1. Data

Huffman and Evenson (1993) present U.S. MFP and real research expenditure data from 1889 to 1990 (their Tables A7.1 and 4.1, respectively). The research variable is the sum of U.S. Department of Agriculture federal appropriations for research and State Agricultural Experiment Station expenditures converted to constant dollars. Productivity is an index of real U.S. farm output, divided by an index of inputs under farmer control (i.e., it does not include rainfall). The output and input quantity indexes are Tornqvist–Theil indexes, constructed from 34 different output and 13 different input categories. The time-series of productivity estimates is the Tornqvist approximation to the Divisia index.[3] These series are constructed consistently over a long enough period that the probability of multiple breaks could reasonably be explored, providing clear examples of the tests we propose. Data are converted to natural logarithms and used throughout. We address the question of stationarity and the presence of shifts in this data in two parts. First, we identify any structural breaks, modeling them in ways that are compatible with the later tests. Second, we apply stationarity and non-stationarity tests that allow for a known break.

2.2. *Identifying and Modeling the Structural Shifts*

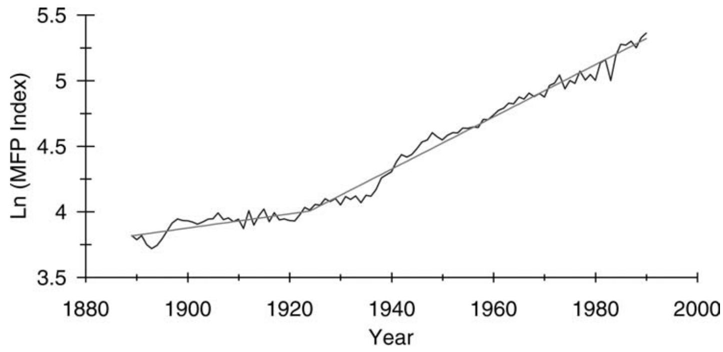
Visual inspection of the data suggests that the trends in each of the series shift, in the mid-1920s for MFP (Figure 1a) and about five years later for research (Figure 1b). To validate or reject this visual perception, we follow Perron's (1989) model B by specifying the regression equation,

$$y_t = \mu + \beta t + \gamma \mathrm{DT}_t^* + \varepsilon_t, \tag{1}$$

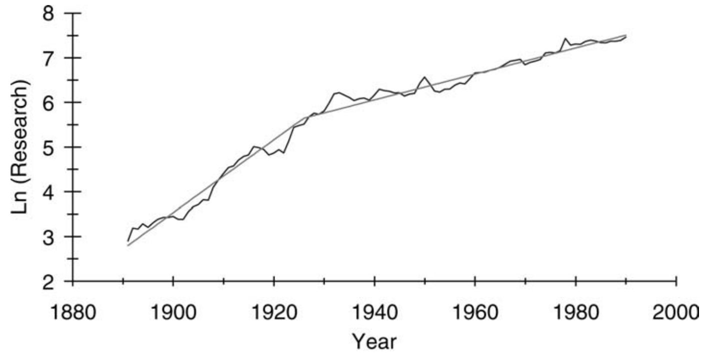
where y_t is either R&D or MFP, μ is the intercept, β is the coefficient on the time trend, t represents the time period, DT_t^* is the second trend defined as in Perron (1989) to take on the value $t - T_B$ if $t > T_B$ and 0 otherwise where T_B is the time at which the trend shift occurs, γ is the coefficient on the second trend, and ε_t is an error term. We estimate equation (1) using OLS regression for values of T_B in the neighborhood of historically suspected or visually inspected break dates. The values of T_B which minimize the sum of squared errors are the maximum likelihood estimates of the time at which the shift occurs, given the OLS assumptions.[4] The resulting T_B values are 1930 for research and 1925 for MFP.[5] Having obtained the maximum likelihood estimate for T_B, T_B^*, we then test for statistical significance of the parameter γ in regression (1) with DT_t^* defined in terms of the break time T_B^*.

For the research shift in 1930 and the MFP shift in 1925, the OLS results are presented in Table 1. Each regression has a high adjusted R^2, and the coefficients on the intercept, trend and trend shift variables are each statistically significant at the 1% level. This indicates that the inclusion of the trend shift is statistically important. Moreover, the trend shifts are important in magnitude. The trend shift coefficient in the research regression is about one-half the size of the trend (in absolute value), leading to a noticeably slower rate of expenditure growth after the shift. The MFP trend shift coefficient is two or three times as large as the trend coefficient. This is consistent with the noticeably faster growth in MFP after the shift.

The magnitude of the model coefficients is plausible. Research expenditures grew on average 7.6% per year prior to the break. This growth rate is consistent with the needs of a young and growing set of research institutions; following the break, in a period of relative "maturity" (Huffman and Evenson, 1993), the growth rate fell to about 2.7%. MFP growth was relatively slow prior to the break, at 0.5% per year, rising to 2.0% following the break. The latter is consistent with descriptions of agriculture as characterized by rapid technical change (Huffman and Evenson, 1993).



(a)



(b)

Figure 1. (a) Logarithm of multi-factor productivity index, actual values and modeled structural shift, 1890–1990. (b) Logarithm of research expenditures, actual values and modeled structural shift, 1890–1990. Source: Huffman and Evenson (1993).

For completeness, we also tested for the possibility of intercept shifts (in 1930 and 1925) in addition to the trend shifts (model 2 in Table 1). As can be seen by the *t*-statistics, neither intercept shift variable is statistically significant at any reasonable significance level. This indicates that the structural breaks were not discontinuous and the intercept-shift variables are excluded from further analysis. We conclude that we should include trend shifts in our modeling of the relationship between research expenditures and MFP below.

The regression models were also used to generate (linear) predicted values of the dependent variables. These predicted values are plotted against actual values in Figures 1a and 1b. These figures provide visual confirmation of what the regression results tell us: that the data series do indeed contain trend shifts during the identified decade.

Table 1. OLS regression results on structural shifts in research expenditures and MFP.

Independent Variable (t-stat)	Dependent Variable			
	ln(Research)		ln(MFP)	
	Model 1	Model 2	Model 1	Model 2
Intercept	2.861***	2.868***	3.818***	3.814***
	(73.06)	(67.83)	(226.62)	(208.55)
Intercept Shift[a]		0.025		− 0.011
		(0.44)		(− 0.489)
Trend	0.076***	0.075***	0.005***	0.006***
	(55.50)	(41.36)	(8.28)	(6.35)
Trend Shift[a]	− 0.049***	− 0.049***	0.015***	0.014***
	(− 24.98)	(− 23.56)	(16.72)	(14.76)
n	102	102	102	102
adjusted R^2	0.98	0.98	0.99	0.99

[a] Shifts at 1930 in the research series, and 1925 in the MFP series.
*** Denotes significance at the 1% level.

2.3. *Testing for Non-Stationarity and Stationarity*

To test for non-stationarity in the presence of a trend shift, Perron (1989) estimates the equation

$$y_t = \mu + \alpha y_{t-1} + \beta t + \gamma \mathrm{DT}_t^* + \sum_{i=1}^{k} c_i \Delta y_{t-i} + \varepsilon_t, \tag{2}$$

where α is the autocorrelation between y_t and y_{t-1}, and the summation includes augmented difference terms as in Said and Dickey (1984). This augmentation is used primarily to generate cleaner estimates of the autocorrelation. The null hypothesis of non-stationarity implies that $\alpha = 1$, which can be checked by a simple t-test with critical values determined from Perron's (1989) Monte Carlo simulations (specifying a particular non-stationary data generating process may imply additional restrictions, see Perron (1989) for details). Thus, for structural shifts in the trend, the Perron (1989) testing procedure is the same as an ADF procedure with a trend shift term DT_t^*.

When no trend shift is included, the ADF test results fail to reject the null hypothesis of non-stationarity for both the research and MFP series (upper portion of Table 2). Including the trend-shift term in the regression equation, increases the absolute value of the test statistics for the MFP series, but only the regression with six lags ($k = 6$ in equation (2)) rejects non-stationarity at any conventional significance level. For the research series, non-stationarity is rejected at the 10% or better significance level in each regression.

Interpreting the ADF test results is somewhat problematic, given the aforementioned low power of this type of test. For the research data, in which rejection of

Table 2. Augmented Dickey–Fuller, Perron, and Hwang–Schmidt tests of the null hypothesis of non-stationarity, with and without a trend shift.

ADF/Perron Tests Number of Lags	ln(MFP) Assuming No Structural Break (ADF)	ln(MFP) With a Structural Break (Perron)	ln(Research) Assuming no Structural Break (ADF)	ln(Research) With a Structural Break (Perron)
4 lags	– 1.17	– 3.51	– 3.07	– 4.72***
6 lags	– 1.29	– 3.77*	– 2.44	– 4.13**
8 lags	– 1.9	– 2.72	– 2.06	– 3.77*

Hwang–Schmidt Test Statistic $\tilde{\tau}_S$ Autocorrelation (ρ*) for H_1	ln(MFP) Assuming No Structural Break	ln(MFP) With a Structural Break	ln(Research) Assuming No Structural Break	ln(Research) With a Structural Break
0.8	– 2.67	– 5.36***	– 9.01***	– 10.62***
0.85	– 2.63	– 5.35***	– 8.99***	– 10.61***
0.9	– 2.55	– 5.34***	– 8.94***	– 10.60***
0.95	– 2.31	– 5.29***	– 8.75***	– 10.56***

*, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.
[a] Represents number of lagged values included in the regression model.
[b] For the model including the structural break, critical values are based on Perron.

non-stationarity is consistent across different lag lengths, it is reasonable to conclude that the data are stationary around a trend shift. For the MFP data, only one of the tests is significant at the 10% level (for $k = 6$). This is insufficient to reject non-stationarity with any comfort. However, coupled with the low power of ADF tests, it is sufficient to raise questions about failing to reject the null of non-stationarity. Consequently, we search for a more powerful test.

Hwang and Schmidt (HS) (1996) develop a unit root test which is well suited to the current situation. In particular, they provide a test based on alternative methods of detrending the data series, which typically improves the power of the test. "Monte Carlo experiments show a clear gain in power, relative to other unit root tests such as the Dickey–Fuller tests, over a large and empirically relevant range of the parameter space" (HS, p. 227). Part of the HS test procedure is to specify an empirically plausible, stationary, value for the autoregressive parameter, and to use this value in constructing the alternative hypothesis. We use a range of values, from 0.80 to 0.95, which are consistent with most previously estimated autocorrelation coefficients, as well as the range over which the HS test exhibits improved power.

HS assume a data generating process $y_t = \psi + \xi t + u_t$, $u_t = \rho u_{t-1} + \varepsilon_t$, $t = 1, \ldots, T$, where $\varepsilon_t \sim \mathrm{NID}(0, \sigma^2)$, that is the same as the one used by Schmidt and Phillips (SP) (1992). This is fortunate because Amsler and Lee (1995) studied the introduction of a known break into SP and were able to provide us with critical values for the HS test with a known break.[6] For the MFP series, the HS test fails to reject the null hypothesis of non-stationarity without a trend shift (lower portion of Table 2). When the trend shift is introduced, the HS tests reject non-stationarity for

Table 3. Univariate stationarity tests of U.S. public agricultural research expenditures and agricultural MFP, 1889–1990, null hypothesis of stationarity.

	KPSS Test Statistics					
	ln(MFP) Assuming No Structural Break		With Structural Break	ln(Research) Assuming No Structural Break		With Structural Break
Lags[a]	Level	Trend	Trend	Level	Trend	Trend
0	9.9242***	1.7300***	0.2279***	9.4686***	2.0687***	0.1804**
1	5.0461***	0.9342***	0.147**	4.8135***	1.0669***	0.1022
2	3.4505***	0.6460***	0.1101	3.2491***	0.7284***	0.0768
3	2.5882***	0.4993***	0.0912	2.4655***	0.5582***	0.0649
4	2.0882***	0.4101***	0.0799	1.9950***	0.4559***	0.0586
5	1.7594***	0.3502***	0.0727	1.6816***	0.3878***	0.055
6	1.5250***	0.3072***	0.0683	1.4580***	0.3393***	0.0529
7	1.3496***	0.2751***	0.0657	1.2950***	0.3031***	0.0517
8	1.2310***	0.2499***	0.064	1.1606***	0.2750***	0.0513

[a] Number of lags used in construction of the long-run variance estimate.
** and *** denote significance at 5% and 1% levels, respectively.

all specified values of the autoregressive parameter. This leads to the conclusion that the MFP series is stationary around a trend shift.

For the research expenditure series, the HS tests reject the hypothesis of non-stationarity for all versions of the test. This result is surprising for the case when only a simple linear trend is modeled, and suggests that the HS tests were unable to detect the emergence of the trend shift when it was not explicitly modeled. When the trend shift is explicitly modeled, the HS tests continue to reject non-stationarity. The conclusion is that the research expenditure series is stationary around a trend shift.[7]

Kwiatkowski et al. (KPSS) (1992) argue for the usefulness of testing not just the null hypothesis of non-stationarity, but also the null hypothesis of stationarity, particularly when using non-stationarity tests with low power. Low power raises the probability of making a Type II error, but by switching the null and alternate hypotheses, the joint probability of getting it wrong in both formulations is greatly reduced. If the HS and KPSS tests give different answers, the tests are inconclusive. The KPSS test can also be used in situations where the data series contain structural shifts (Lee et al., 1997) and critical values are obtained from GAUSS code developed by those authors.[8] We calculate the KPSS test statistic for both research and MFP, in three trend models: no trend, a simple linear trend, and the trend-shift model. For the models with no trend or with a simple linear trend, the KPSS test unambiguously rejects the null hypothesis of stationarity (Table 3). However, the KPSS test fails to reject stationarity for the trend-shift models. These results confirm the findings of the HS tests, namely that the research expenditure and MFP data are stationary around a trend-shift.

Heuristically, the findings of stationarity are much more palatable than findings of non-stationarity. The previous finding that research expenditures are non-stationary means that shocks to research expenditures, such as the U.S. Department of

Agriculture (USDA) "rightsizings" in the mid-1990s, are expected to lead to permanent changes in the level of research expenditures. However, in the late 1990s we saw a USDA re-investment in agricultural research, with increasing real agricultural research funds for the first time in over a decade. Thus, the "rightsizing" shock seems not to be permanent.

Nor is the finding that MFP contains a unit root intuitive. This finding means that any increase in MFP is permanent. Such a result is contrary to descriptions of maintenance research impacts on productivity. For example, rust-resistant wheat varieties have been characterized as losing effectiveness after 3–5 years, as rusts mutate into forms to which the variety is not resistant. It is also contrary to empirical findings that the proportion of research expenditures directed toward maintaining previous productivity gains ranges from 24.0% in the East South Central region to 44.1% in the Mid Atlantic region (Andusei and Norton, 1990). Townsend and Thirtle (2001) establish similar maintenance research requirements for livestock productivity in South Africa.

The significance of these results is that correct stationarity properties are critical to proper estimation of structural relationships. Time-series models which do not explicitly identify structural shifts may classify all innovations as permanent when in fact the only permanent event, as we have uncovered, is a structural shift: the remaining innovations are transitory. A number of authors, including Alavalapati et al. (1997), Fernandez–Cornejo and Shumway (1997), Machado (1995), and Schimmelpfennig and Thirtle (1994) find cointegrating relationships between research, MFP and other variables. These results could be influenced by a number of factors (including short series) and as we have shown the failure correctly to account for structural change and spurious long-run relationships (Phillips, 1986).

3. The Impact of Research on Productivity

3.1. VAR Analysis of Research and Productivity

Since research and productivity are trend-shift stationary they cannot be cointegrated (Campos et al., 1996), but this does not rule out a possible long-run equilibrium relationship involving research and productivity and other non-stationary variables. There is some theoretical guidance on how to proceed. The present accepted theoretical structure is that research impacts should be determined from accumulated stocks of research. Two problems with empirically implementing the theory as Griliches (1998, p. 33) points out, are that distributed lags suffer from multicollinearity because annual R&D is highly correlated from one year to the next, and simultaneity because future output probably depends on past R&D, and future R&D probably depends on past output. Vector-autoregression (VAR) methods rely on impulse response functions rather than individual coefficients to test the impacts of shocks to one variable on the system, thus ameliorating multicollinearity problems. VARs also allow for simultaneity by being constructed with both variables appearing as the dependent variable and neither having to be exogenous.

VARs have been used for the analysis of research impacts by Khatri et al. (1996) in South Africa and by Akgungor et al. (1996) in Kenya. Alston et al. (1996), Makki et al. (1999), and Schimmelpfennig and Thirtle (1994) use VARs for the analysis of U.S. research impacts. Following the previous literature, the same number of lags of both variables are used in each equation and trend and trend-squared terms enter the tested VARs as exogenous variables.[9] Research and productivity are pre-filtered to remove the previously identified breaks. The residuals from the least-squares regression of each variable on a break dummy and a constant are then treated symmetrically in an unrestricted VAR. The VAR model is then,

$$
\begin{aligned}
(EQ.1)\mathrm{MFP}_{1925t} &= c_1 + \alpha_1 t + \beta_1 t^2 + \sum_{i=1}^{k} \delta_{1i}\mathrm{MFP}_{1925t-i} \\
&\quad + \sum_{i=1}^{k} \gamma_{1i}\mathrm{RES}_{1930t-i} + \varepsilon_{1t}, \\
(EQ.2)\mathrm{RES}_{1930t} &= c_2 + \alpha_2 t + \beta_2 t^2 + \sum_{i=1}^{k} \delta_{2i}\mathrm{MFP}_{1925t-i} \\
&\quad + \sum_{i=1}^{k} \gamma_{2i}\mathrm{RES}_{1930t-i} + \varepsilon_{2t},
\end{aligned}
\tag{3}
$$

where the subscripts 1925 and 1930 indicate that MFP and RES (research) have been break filtered, $c_{1,2}$ are individual equation constants, t is a time trend as before, $\delta_{1,2}$ and $\gamma_{1,2}$ are the estimated lag coefficients, some of which may be set equal to zero in the model selection process, and $\varepsilon_{1,2}$ are the remaining errors in each equation.

The estimated equations in Table 4 were selected on the criteria of goodness of fit (variance dominance), data coherence, and parameter parsimony. The model selection criteria considered included adjusted R-squared, Akaike Information Criterion (AIC), and Schwartz Criterion (SC), with the AIC and SC considered simultaneously for individual equations and for the VAR as a system.[10] Data coherence is evaluated by examining the autocorrelation coefficients of the residuals of competing models, and their Q statistics, to ensure that the residuals are white noise. The selected models indicate that the first, second, and 24th lag of research are highly significant in the MFP equation, and the 24th lag of MFP is significant in the research equation, in addition to the significant own lags. To evaluate the structural stability of the estimated equations before and after the break dates, they were also estimated for the sub-samples 1889–1929 and 1930–1990 and the 24th lag is still significant in both equations for both time periods.[11] To evaluate the dynamic stability of the VAR system (as a set of equilbrium conditions) and to allow the implementation of a new procedure for calculating rates of return, we estimate an impulse response function for this VAR system in the next section.

Table 4. Vector autoregression (VAR) of MFP and research (1913–1990)[a].

Independent Variable (Period Lags) [*t*-stat]	Dependent Variable	
	MFP (EQ. 1)	RES (EQ. 2)
MFP(−1)	0.315***	− 0.122
	[2.83]	[−0.54]
MFP(−2)	0.287***	− 0.082
	[2.68]	[−0.38]
MFP(−24)	0.061	0.596***
	[0.57]	[2.72]
RES(−1)	0.124**	1.056***
	[2.17]	[9.13]
RES(−2)	− 0.132***	− 0.163*
	[−2.40]	[−1.46]
RES(−24)	0.018**	− 0.003
	[1.66]	[−0.15]
C	0.070*	0.275***
	[1.57]	[3.04]
TREND	− 0.001*	− 0.004***
	[−1.54]	[−3.08]
Adj. *R*-squared	0.489931	0.966192
F-statistic	11.56569***	315.3658***
Log Likelihood	135.7914	80.76653
Akaike Information Criterion (AIC)	− 3.276702	− 1.865808
Schwarz Criterion (SC)	− 3.034988	− 1.624095
System Estimation Diagnostic Statistics (EQ. 1 and EQ. 2)		
Log Likelihood	208.1317	
AIC	− 4.926453	
SC	− 4.443026	

*, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.
[a] The sample size is reduced by the number of lags. Data is pre-whitened to remove previously identified structural shifts.

3.2. *The Rate of Return Measure*

A wide range of economic models have been employed in measuring the effects of agricultural research (Alston et al., 1995). A result commonly produced by these models is the rate of return to agricultural research, a huge literature that has recently been surveyed by Alston et al. (2000) and Evenson (2002). The study by Alston et al. (2000) found 43 studies estimating 329 supply shifts based on econometric modeling of the relationship between research and productivity; and an additional 58 studies that econometrically model the relationship between research and production. There are some deep disagreements about the measurement of productivity and research and most of these estimates of returns to research are sensitive to changes in assumptions concerning the lag between research and production or productivity (Fan, 2000). Commonly appearing lag structures are trapezoidal (Huffman and Evenson, 1992) and polynomial (Thirtle and Bottomley, 1989) distributed lags. These techniques are designed to economize of degrees of freedom, an important consideration in studies with shorter series.[12]

The cost of being able to summarize the lagged effects of R&D on productivity in terms of only a few parameters is the arbitrary structure imposed on the relationship. These structures assume a smooth, continuous impact of R&D on productivity over time.[13] The creation of research stock variables using moving averages or perpetual inventory models have similar smoothing effects (Griliches, 1998). We instead, like Chavas and Cox (1992), have followed an unstructured approach; and like Pardey and Craig (1989) used a bivariate VAR in the previous section to specify the lagged relationship between research expenditures and MFP.[14] The benefit of this approach is that we are now in a position to implement a new method for estimating returns to research from the impulse response functions of the VARs that can be compared to a mountainous pre-existing literature.

3.3. *Impulse Responses*

An impulse response function (IRF) captures the change in the dynamic path of the system in response to an unanticipated shock to the value of either variable in a VAR. Since both research and MFP are endogenous in the VAR system, we cannot arbitrarily increase the value of research by one unit to calculate a marginal effect, as we would do in a traditional econometric system either explicitly or implicitly by using the estimated coefficient on the research variable.[15] In a VAR system the appropriate comparative exercise is to simulate the system at rest (in a steady state), and introduce an exogenous shock by introducing a non-zero error term to the equation of interest. In the context of the system of equations (3), we introduce a shock $\varepsilon_{R,0}$, independent of the error term $\varepsilon_{2,0}$, to the research equation at time zero; following general practice, the shock is equal in size to one standard deviation of the research variable. The impulse response functions give the magnitude of the research and MFP deviations from the steady-state at each point in time. The dynamic path traced out by the impulse responses represents the deviation from the steady-state path due to the one-time exogenous shock to research (although both the MFP and research variables change endogenously along this dynamic path). This path is used to assess the effects of the research shock.

The estimated impulse response functions for MFP show an early positive effect of research on MFP at 99% confidence, peaking in year 3 (Figure 2). This is consistent with the idea of applied or problem-solving research having a fairly short time lag before its effects are felt. This effect diminishes quickly to zero (the negative responses are not statistically significant and appear because the function makes a smooth transition to the next significant point). At year 24 a second significant but smaller positive effect is seen, which we believe represents the effects of basic research on productivity. As noted in the previous section, the estimation is not precise enough to know that this occurs exactly after 24 lags, but the model consistently shows a statistically significant effect in the 20–25 year lag range. This result of

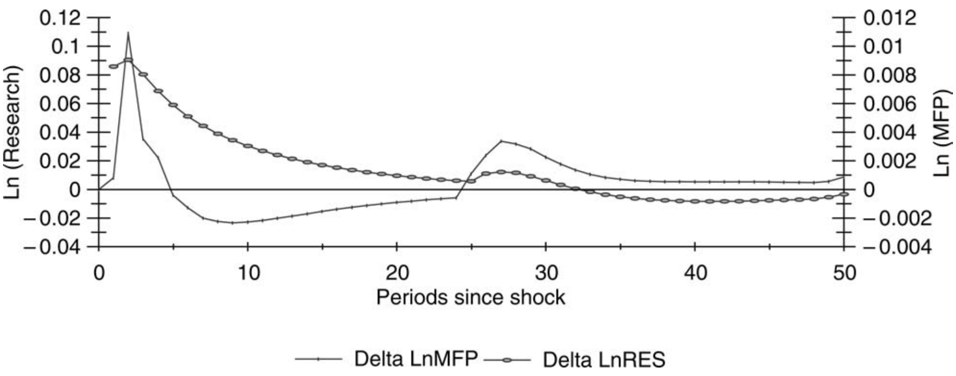


Figure 2. Impulse responses of research and MFP to a shock in research.

effects after long lags is consistent with the findings of Pardey and Craig (1989). Eventually, the estimated impulse responses tend to zero. Thus, the estimated responses show that the representation of MFP in the VAR is stationary, which is consistent with the univariate analysis presented earlier.

The response of research shows that the initial shock to research is followed by a period of increased research expenditures, but the size of the increase diminishes over time again at 99% confidence (Figure 2). There is a small significant upward blip in research spending about 20 years after the initial shock, probably associated with the impact of research on MFP with a long time lag. This might indicate that the fruition of basic research effects on MFP may spur some new research spending. As with MFP, there are periods where the research response is negative, but these responses are not statistically significant, and eventually the responses die off as is consistent with stationarity of the VAR representation of research, and the VAR system as a whole appears to be stable returning to steady state equilibrium.

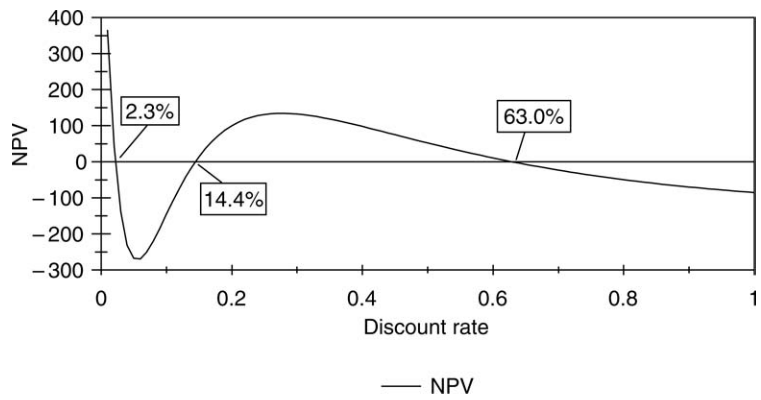


Figure 3. Marginal net present value of research.

3.4. Measuring the Impacts of Research

Since the research and MFP variables are in natural logarithms it is simple to calculate their percentage changes in levels, but the magnitude of the change will depend on the starting point, or "without-change", assumption. We evaluate changes relative to the sample mean values of the logarithms. Thus, the effect of a research shock on the level of research t periods later is $\Delta\mathrm{RES}_t = \exp\{\ln(R) + I_{R,t}\} - \exp\{\ln(R)\}$, where $\Delta\mathrm{RES}_t$ represents the incremental change in research at time t (not the first difference of the data series), $I_{R,t}$ is the impulse response of research at time t, and $\ln(R)$ is the average value of research. Since the research variables are measured in constant dollars, the result is a time series of incremental research expenditures generated from the initial shock.

We follow a similar procedure to generate $\{\Delta\mathrm{MFP}_t\}$, the incremental changes in the MFP series at each time t. However, these are changes in an index-number series, and must be translated into dollar or equivalent terms to be comparable to the research series. We make this conversion using Alston et al.'s (1995) social surplus model and approximation forumulae for research benefits in a closed economy (Table A.5.1.2), where the change in MFP replaces their change in yield. The incremental changes in research expenditures and productivity (converted to social surplus) are compared using standard rate-of-return (ROR) calculations.

3.5. The Rate of Return

The research impacts from the VAR in a previous section shows positive effects on productivity in the first few years and 20–25 years later. Since there are annual costs of doing research, the incremental net benefit stream switches sign several times. This results in multiple ROR candidates that are each solutions equating benefits and costs with net present value (NPV) equal to zero. Examining NPV as a function of the discount rate shows that at very low discount rates the NPV is positive (Figure 3). The NPV crosses the axis (from above) at 2.3%, making this value one ROR solution, and is then negative for discount rates between 2.3 and 14.4%. The NPV crosses the axis from below at 14.4%, and remains positive until the discount rate reaches the third ROR solution of 63.0%.

Discounting the incremental net benefit stream at each of the ROR solutions shows that in every case there is a positive current net value after three or four lags, resulting from the short-term positive impact of research on MFP. The ROR calculation implicitly treats this positive net current value as if it can be reinvested at the ROR (Robison and Barry, 1998). In the case of changes in social surplus, the research agencies do not receive monetized payments based on the increase in social surplus, and so cannot reinvest. The relevant question is whether the recipients of this increase in social surplus—firms and consumers—can earn the ROR on reinvested money. The real ROR solutions of 14.4 and 63.0% are higher than real returns on investment capital over the sample period. The estimated ROR of 2.3% is the only solution which is consistent with the estimated impulse response function, the assumptions implicit in the ROR calculation, and historical real returns to capital investment.

The estimated ROR of 2.3% is substantially lower than those estimated by other authors. Alston et al. (2000) reference 23 studies which calculate real returns to research for all U.S. agriculture or for all crop and livestock agriculture in the United States.[16] Of these studies, only Alston et al. (1996) list an estimated ROR as low as 2.3%: they give a range of possible RORs from −1 to 160% (Alston et al., 2000). We believe that the 1998 study by the same authors updates these estimates; in the later study the estimated range is 3.1–147.0%, so that our 2.3% estimate is perhaps the lowest estimated ROR to U.S. agricultural research.

Although there are a number of differences among these studies, we believe that there are two primary reasons for the magnitude higher estimates in the other studies. The first is that we use long lag lengths (24 years in our case). This topic has been discussed in Makki et al. (1999) who also obtain low long-term rates of return in the United States. The second reason is that we remove a second trend in the data series, particularly the strong positive second trend in the MFP data that represents substantially higher productivity growth since 1925. Neglecting this second trend (or using less accurate detrending that may only partially account for the second trend) means that this trend is likely to be counted as a research impact. To the extent that we remove this second trend, we reduce our estimated impact of research by that amount, and generate a smaller ROR estimate.

The question whether or not research caused the second MFP trend requires a subjective answer. There is no objective statistical procedure for discovering the causal relationship between two trends—or whether a causal relationship even exists. The timing of the structural breaks (research five years after MFP) seems to indicate that the breaks occurred for historical and institutional reasons not related to the lags between research and MFP determined in the VAR. In fact, we have not estimated structural change in the relationship between research and productivity except for two sub-periods. We have estimated how structural change in each of the variables influences the estimated ROR to research. Whether a well-functioning agricultural research system contributed to the second upward trend in MFP—and it may have—is a judgment call. The estimated ROR of 2.3% when the trends in MFP are not counted as research impacts serves to illustrate how important that judgment call is.

4. Summary and Conclusions

The important impacts that structural changes have had on U.S. agriculture can be tested for and modeled using recent advances in time series econometrics. We measured structural change as a statistically significant break in either stochastic or deterministic time trends, and jointly determined the stationarity and structural shift properties of the data. It may be useful to use this approach in the future to improve time series projections of supply responses to changes in price and cost conditions, demand for production inputs over time, or participation rates in government assistance programs, among other things.

Our results show that research and productivity are stationary, with structural shifts in 1930 and 1925, respectively. The shift in the trend of productivity is associated with pre-Depression changes in market conditions, while the research trend-shift represents initial movements forward on technologies that altered capitalization and returns to scale in many agricultural enterprises. Correctly accounting for these structural breaks is shown to have more than historical interest. Previous results for research and MFP that did not model structural shifts found both variables to be non-stationary. The difference between these findings in practical applications is quite large because non-stationarity implies that shocks have permanent effects, whereas our results show that only the previously unaccounted structural breaks are permanent. In addition, when structural breaks are identified, standard tests for cointegration and closely related error-correction representations are unreliable.

We combine the improved time series information with theoretical considerations to construct a VAR of research and MFP. We use the associated impulse response function to estimate a rate of return to research. Our estimate is predictably low compared to earlier estimates. When long-run trends, and breaks in the trends, are controled for statistically, research does not get credit for the trend or trend shift in productivity, in contrast to earlier studies. Since it is impossible statistically to verify the existence of a causal relationship between two trends, future work could seek new and innovative ways to determine whether research expenditures are causing the trend in agricultural productivity.

Acknowledgment

Partial support for this work was provided by the Michigan Agricultural Experiment Station and Cooperative Agreement No. 43-3AEL-9-80097 between USDA and Michigan State. The views expressed are not necessarily those of the U.S. Department of Agriculture.

Notes

1. Additional events that we investigate are the oil price shocks of the 1970s and early 1980s.
2. This is similar to Perron's analysis of wage data, in which he shows that the DF test is unable to distinguish between an estimated autocorrelation coefficient of 0.304 (indicating stationarity, with a fairly rapid decay rate) and the unit-root null with hypothesized autocorrelation value 1.0 (non-stationarity).
3. For more details see Huffman and Evenson (1993).
4. This "grid search" is appropriate to the discrete nature of the time periods.
5. Neither series exhibits significant shifts due to the oil price shocks in the 1970s or early 1980s. This is probably due to the insulating effect of farm programs. We test for a break in the ten-year

neighborhood of a suspected break, but do not do a grid search of all years in the series because this would open up the results to Alston–Chalfant's (1991) spurious breaks critique.

6. Their Gauss program is available from Schimmelpfennig on request.
7. Similar results were obtained by applying the Schmidt–Phillips test, which is a Lagrange Multiplier test. These results are available from the authors on request.
8. Available on request from Schimmelpfennig. The existence of these programs that allow the calculation of critical values in the presence of known breaks, was a large reason for the selection of these specific unit root tests.
9. The absence of stochastic trends in the variables established in the previous section does not rule out the possibility of linear or non-linear deterministic trends.
10. The $\text{AIC}_i = (\text{MLL}_i - h_i)$ and $\text{SC}_i = (\text{MLL}_i - 0.5\text{h}_i(\ln N))$, where MLL_i is the maximized log-likelihood for model i, h_i is the number of parameters and N is the sample size.
11. Trend-squared terms were not significant in any of the VAR specifications.
12. The latter study for example used data for 1965–1980.
13. This is particularly true for popular second degree polynomials with end point constraints.
14. Efforts to model lagged impacts of research on output and factors of production in a constant elasticity of substitution production function, led to an unsatisfactory loss of shorter term research impacts. The long run research effects were surprisingly unchanged.
15. In a co-integrated system, one could use the estimated coefficients of the cointegrating relationship to capture the comparative steady-state effect of a change in research levels.
16. By first author and date, the studies are: Evenson (1968) 112%; Peterson (1977) 34–51%; Lu (1978) 15%; Davis (1979) 28.7–100.00; Knutson (1979) 28.0–49.7%; Evenson (1980) 65–110%; White (1982) 6.98–36.0%; Braha (1986) 47.2; Chavas (1991) 36.0–41.0; Norton (1991) 30.0; Chavas (1992) 17.0–28.0; Yee (1992) 38–58%; Deininger (1993) 27.2–384.4; Huffman (1993) 40.6–73.5; Alston (1994) 19.5–21.4; White (1995) 40.4%; Yee (1995) 40.0–46.0; Alston (1996) −1–260%; Evenson (1996) 71–83%; Evenson (1996) 43–67%; Gopinath (1996) 12.7–52.8; Maaki (1996) 6.0; Alston (1998) 3.1–147.0%. For complete references see Alston et al. (2000).

References

Akgungor, S., D. Makanda, J. Oehmke, R. Myers and Y. Choe. (1996). "A Dynamic Analysis of Kenyan Wheat Research and Rate of Return." Conference Proceedings of Global Agricultural Science Policy for the Twenty-First Century, Contributed Papers. Melbourne, Australia: ICMS Pty. Ltd.

Alavalapati, J. R. R., W. L. Adamowicz and M. K. Luckert. (1997). "A Cointegration Analysis of Canadian Wood Pulp Prices." *American Journal of Agricultural Economics* 79(August), 975–986.

Alston, J. M. and J. A. Chalfant. (1991). "Unstable Models from Incorrect Forms." *American Journal of Agricultural Economics* 73(November), 1171–1181.

Alston, J., B. Craig and P. Pardey. (1996). "Research Lags and Research Returns." Conference Proceedings of Global Agricultural Science Policy for the Twenty-First Century, Invited Papers. Melbourne, Australia: ICMS Pty. Ltd.

Alston, J., B. Craig and P. Pardey. (1998). "Dynamics in the Creation and Depreciation of Knowledge, and the Returns to Research." EPTD Discussion Paper No. 35. Washington DC: IFPRI August (revised 1999).

Alston, J., C. Kang, M. Marra, P. Pardey and J. Wyatt. (2000). "Meta-Analysis of Rates of Return to Agricultural R&D: Ex Pede Herculem." Research Report No. 113, Washington DC: IFPRI, June.

Alston, J. M., G. W. Norton and P. G. Pardey. (1995). *Science Under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*. Ithaca: Cornell University Press.

Amsler, C. and J. Lee. (1995). "An LM Test for a Unit Root in the Presence of a Structural Change." *Econometric Theory* 11(June), 359–368.

Andusei, E. O. and G. W. Norton. (1990). "The Magnitude of Agricultural Maintenance Research in the United States." *Journal of Production Agriculture* 3(January–March), 1–6.

Antle, J. M. (1984). "The Structure of U.S. Agricultural Technology: 1910–78." *American Journal of Agricultural Economics* 66, 414–421.

Campos, J., N. R. Ericsson and D. F. Hendry. (1996). "Cointegration Tests in the Presence of Structural Breaks." *Journal of Econometrics* 70, 187–220.

Chavas, J. P. and T. L. Cox. (1992). "A Nonparametric Analysis of the Influence of Research on Agricultural Productivity." *American Journal of Agricultural Economics* 74, 583–591.

Cochrane, W. W. (1979). *The Development of American Agriculture: A Historical Perspective*. Minneapolis, Minnesota: University of Minnesota Press.

Dickey, D. A. and W. A. Fuller. (1981). "Likelihood Ratio Statistics for Autoregressive Time Series with Unit Roots." *Econometrica* 49(June), 1057–1072.

Evenson, R. E. (1978). "A Century of Productivity Change in U.S. Agriculture: An Analysis of the Role of Invention, Research and Extension." Discussion Paper No. 296. New Haven, CT: Economic Growth Center, Yale University.

Evenson, R. E. (1980). "A Century of Agricultural Research and Productivity Change Research, Invention, Extension, and Productivity Change in U.S. Agriculture: An Historical Decomposition Analysis." A. A. Araji (ed.), *Research and Extension in Agriculture*. pp. 146–228, Moscow ID: Department of Agricultural Economics, University of Idaho.

Evenson, R. E. (2002). "Private and Public Research and Extension." B. L. Gardner and G. C. Rausser (eds.), *Handbook of Agricultural Economics*. Volume 1, Part 1. New York: North-Holland/Elsevier Science Publishers.

Fan, S. (2000). "Research Investment and the Economic Returns to Chinese Agricultural Research." *Journal of Productivity Analysis* 14(2), 163–182, September.

Fernandez-Cornejo, J. and C. R. Shumway. (1997). "Research and Productivity in Mexican Agriculture." *American Journal of Agricultural Economics* 79(August), 738–753.

Goodwin, B. K. and G. W. Brester. (1995). "Structural Change in Factor Demand Relationships in the U.S. Food and Kindred Products Industry." *American Journal of Agricultural Economics* 77(February), 69–79.

Granger, C. W. J. and P. Newbold. (1974). "Spurious Regressions in Economics." *Journal of Econometrics* 2, 111–120.

Griliches, Z. (1998). "Issues in Assessing the Contribution of Research and Development to Productivity Growth." *Bell Journal of Economics* 10(Spring 1979), 92–116, reprinted in Z. Griliches *R&D and Productivity: The Econometric Evidence*, Chicago: University of Chicago Press.

Hamilton, J. (1993). *Time Series Analysis*. Princeton: Princeton University Press.

Huffman, W. E. and R. E. Evenson. (2001). "Structural and Productivity Change in U.S. Agriculture, 1950–1982." *Agricultural Economics* 24, 127–147.

Huffman, W. E. and R. E. Evenson. (1993). *Science for Agriculture*. Ames: Iowa State University Press.

Huffman, W. E. and R. E. Evenson. (1992). "Contributions of Public and Private Science and Technology to U.S. Agricultural Productivity." *American Journal of Agricultural Economics* 74(August), 752–756.

Hwang, J. and P. Schmidt. (1996). "Alternative Methods of Detrending and the Power of Unit Root Tests." *Journal of Econometrics* 71, 227–248.

Kennedy, P. (1992). *A Guide to Econometrics*, 3rd edition. Cambridge: MIT Press.

Khatri, Y. J., C. Thirtle and J. van Zyl. (1996). "Public Research and Development as a Source of Productivity Change in South African Agriculture." *South African Journal of Science* 92, 143–150.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin. (1992). "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We that Economic Time Series Have a Unit Root?" *Journal of Econometrics* 54(Oct.–Dec.), 159–178.

Lee, J., C. J. Huang and Y. Shin. (1997). "On Stationarity Tests in the Presence of Structural Breaks." *Economics Letters* 55(August), 165–172.

Lucas, R. E. Jr. (1976). "Econometric Policy Evaluation: A Critique." K. Brunner and A. Meltzer, (eds.), *The Phillips Curve and Labor Markets*. Volume 1, Carnegie-Rochester Conferences on Public Policy, *Journal of Monetary Economics*, supplement, pp. 19–46.

Makki, S., L. Tweeten and C. Thraen. (1999). "Investing in Research and Education versus Commodity Programs: Implications for Agricultural Productivity." *Journal of Productivity Analysis* 12(1), 77–94, August.

Machado, F. S. (1995). "Testing the Induced Innovation Hypothesis Using Cointegration Analysis." *Journal of Agricultural Economics* 46(September), 349–360.

Maddala, G. S. and I. M. Kim. (1998). *Unit Roots, Cointegration, and Structural Change*. Cambridge, U.K.: Cambridge University Press.

McCorriston, S. and I. M. Sheldon. (1991). "Government Intervention in Imperfectly Competitive Agricultural Input Markets." *American Journal of Agricultural Economics* 73(August), 621–632.

Nelson, C. R. and C. I. Plosser. (1982). "Trends and Random Walks in Macroeconomic Time Series." *Journal of Monetary Economics* 10, 139–162.

Pardey, P. G. and B. Craig. (1989). "Causal Relationships Between Public Sector Agricultural Research Expenditures and Output." *American Journal of Agricultural Economics* 71(February), 9–19.

Perron, P. (1989). "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Econometrica* 57(November), 1361–1401.

Phillips, P. C. B. (1986). "Understanding Spurious Regressions in Econometrics." *Journal of Econometrics* 33, 311–340.

Phillips, P. C. B. and S. Ouliaris. (1990). "Asymptotic Properties of Residual Based Tests for Cointegration." *Econometrica* 58, 165–193.

Phillips, P. C. B. and P. Perron. (1988). "Testing for a Unit Root in Time Series Regression." *Biometrika* 75, 335–346.

Rasmussen, W. D. and B. F. Stanton. (1993). "The Structure of Agriculture in an Historical Context." In A. Hallam (ed.), *Size, Structure, and the Changing Face of American Agriculture*. Chapter 3, Boulder, CO: Westview Press.

Robison, L. J. and P. J. Barry. (1998). *Present Value Models and Investment Analysis*. East Lansing, MI: Michigan State University Press.

Said, S. E. and D. A. Dickey. (1984). "Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order." *Biometrika* 71, 599–607.

Schimmelpfennig, D. and C. Thirtle. (1994). "Cointegration and Causality: Exploring the Relationship between Agricultural R&D and Productivity." *Journal of Agricultural Economics* 45(May), 220–231.

Schmidt, P. and P. Phillips. (1992). "LM Tests for a Unit Root in the Presence of Deterministic Trends." *Oxford Bulletin of Economics and Statistics* 54, 257–287.

Sims, C. A., J. H. Stock and M. W. Watson. (1990). "Inference in Linear Time Series Models with Some Unit Roots." *Econometrica* 58(January), 113–144.

Townsend, R. and C. Thirtle. (2001). "Is Livestock Research Unproductive? Separating Health Maintenance from Improvement Research." *Agricultural Economics* 25(September), 177–189.

Thirtle, C. and P. Bottomley. (1989). "The Rate of Return to Public Sector Agricultural R&D in the UK, 1965–80." *Applied Economics* 21(August), 1063–1086.