DYNAMIC ANALYSIS OF RETURNS TO RESEARCH I: A TIME SERIES METHODOLOGY

by

Young Chan Choe
James F. Oehmke

Staff Paper No. 92-54

July 6, 1992

DYNAMIC ANALYSIS OF RETURNS TO RESEARCH I: A TIME SERIES METHODOLOGY

Young Chan Choe[1]
James F. Oehmke

INTRODUCTION

This paper introduces a time-series econometric methodology for estimating rates of return to research. While previous analyses of research often use time-series data, time-series econometric models such as vector autoregressions (VARs) or error correction models (ECMs) have not been applied. The advantages of VAR models are that

> VAR models are relatively simple to specify and [to] estimate. Only a minimal set of justifying restrictions is employed and no restrictions are placed on the parameters of the reduced form ... given the uncertainty surrounding the underlying economic structure of the market, the unrestricted reduced form VAR provides flexibility which allows the model to be consistent with a wide range of alternative economic structures (Myers et al.).

The advantages of the ECM are similar.

Three advantages of time-series econometric models are important for the estimation of rates of return to research. First, these methods extend the index number method to allow statistical measurement of the relationship between research and the index number. For example, the Akino-Hayami formula requires the researcher to know how much yield increase is attributable to the research program; other techniques attribute increases in an index of total factor productivity to research. While this information may be available for a single, high-yielding variety in a homogeneous agroclimatic zone, it is less readily available for programs

[1] Ph.D. candidate and Associate Professor, respectively, Department of Agricultural Economics, Michigan State University.

with multiple varietal releases; agronomic marketing, management and other research programs whose impacts are not embodied into a particular input such as an improved seed; or for programs which address production in diverse or volatile agroecological environments. Once time-series econometric methods have identified the impact of research or yields and area planted (or other index numbers), the standard Akino-Hayami formulae can be used to compute benefits.

The second advantage of time-series econometrics is that it does not impose market efficiency restrictions on the estimated equations, in the manner that production-function methods sometimes do. Production function methods typically include a research variable as an exogeneous variable in the reduced form specification of a production, cost or profit function (Norton and Davis). Standard econometric estimation of these functions rely on first-order conditions derived from profit maximization (Capalbo and Antle). If there is some sort of market failure or government distortion, then the first-order conditions may not be valid. Consequently, the regression results may be inappropriate. Because time-series models do not impose assumptions about market efficiency they do not suffer from this problem.[2]

The third advantage of the time-series econometric methodology is that it has a parsimonious parameterization, but still allows for long and variable lags between the research activity and the impact. That such lags exist has been known at least since Griliches' (1958) investigation of hybrid corn. Recently, Pardey and Craig have found that lag lengths of thirty years or more are appropriate. Traditional methods require use of thirty or more lags of an annual research variable (e.g., expenditures) to make this determination, placing onerous requirements on the data and reducing degrees of freedom. Time-series methods allow for

[2] In traditional models and time-series models, the econometrician should take care to account for structural change or changes in the policy regime.

persistence of exogenous influences with relatively few explicit lags. Hence, these methods can model research dynamics with much shorter data sets than are required by traditional methods.

SPECIFYING THE ECONOMETRIC MODEL

The Vector Autoregression Model

The first step in building a minimally restricted time series model to study the impacts of research activity on agriculture is to choose the set of variables to be included. Research expenditures, and production or productivity measures such as yields or total factor productivity are natural choices for pursuing the study objectives. Other variables may be included based on researcher's judgement.

A general linear econometric model for the vector y_t takes the form

$$By_t = \sum_{i=1}^{p} B_i y_{t-i} + Au_t \tag{1}$$

where u_t is a vector of serially uncorrelated orthogonal error terms with identity covariance matrix $E\left(u_t u_t'\right) = I$; and A, B, and B_i $(i=1, 2, \ldots p)$ are unknown parameter matrices. It is assumed that the vector y contains g to individual variables. The A and B matrices represent contemporaneous interactions between variables in the system, while the B_i matrices represent the system dynamics. The goal is to use identification restrictions to estimate the contemporaneous parameters A and B while leaving the system dynamics relatively unrestricted. Then the estimated model can be used to trace out the effects of shocks (exogenous changes) to research expenditures on the variables in the system.

The reduced form of (1) is

$$y_t = \sum_{i=1}^{p} C_i y_{t-i} + v_t \tag{2}$$

where $C_i = B^{-1}B_i$; $v_t = B^{-1}Au_t$ and $E\left(v_t v_t'\right) = B^{-1}AA'B^{-1'}$. Thus, model (1) is underidentified in general because there is no way to solve for A, B, and the B_i when only C_i parameters and the covariance matrix of v_t are estimated from the reduced form (Cooley and Leroy). The usual solution in analysis of VARs has been to assume that A is diagonal and B is lower triangular with a unit diagonal (a normalization). The model then takes on a recursive structure with each element of u_t contemporaneously influencing variables situated below it in the vector, but not those situated above it. Identification takes the form of choosing an appropriate recursive order (Sims, 1980; Orden).

More recently, attention has shifted to alternative sets of restrictions on the A and B matrices which allow for simultaneous interactions between variables in the system (Sims, 1986; Bernanke). Orden and Fackler use this approach to investigate the effects of monetary policy on agriculture. The idea is that a recursive ordering is just one of many possible sets of identification restrictions, and that prior information on the nature of the economic system being modeled can be used to develop alternative identification schemes. Like the recursive ordering, these alternative identification schemes focus on contemporaneous relationships between variables (i.e. on restrictions on the A and B matrices) and are generally just identifying so there is no objective basis for choosing among them. We can investigate the effects of research policies under a number of alternative identification schemes to check the robustness of results.

Cointegration and the Error Corrections Model

It is now well known that problems arise in estimation and inference on the reduced form model (2) when some elements of y_t are nonstationary. Subtracting y_{t-1} from both sides of (2) and rearranging gives an alternative representation of the reduced form model

$$\Delta y_t = \sum_{i=1}^{p-1} \pi_i \, \Delta y_{t-i} - \pi \, y_{t-p} + v_t \tag{3}$$

where $\pi = I - \sum_{i=1}^{p} C_i$ and $\pi_i = \sum_{j=1}^{i} C_j - I$. If the matrix π has full rank, then the system is stationary (no unit roots), and OLS applied to each equation in (2) is an appropriate estimator. Similarly, if π has rank zero, then π is the null matrix and we are left with a VAR in first differences. This is the case where there are as many unit roots in the system as there are variables in y_t, and again OLS applied to each (differenced) equation is an appropriate estimator.

Another possibility is that π has positive rank but is not full rank. In this case, π can be expressed

$$\pi = \alpha\beta' \tag{4}$$

where α and β are matrices, such that although some elements of y_t are nonstationary, the linear combinations of y_t given by $\beta' y_t$ are stationary. In this case the elements of y_t are said to be cointegrated with cointegrating vectors β (Engle and Granger). Intuitively, the number of unit roots driving the system y_t is greater than zero but less than the number of variables in the system, which shall be denoted by g. The cointegrated nonstationary variables are driven by the

same unit root and therefore move together in a long run equilibrium relationship. It is important to impose relevant cointegration restrictions during estimation to ensure an efficient estimation procedure and allow the usual inference using standard asymptotic distribution theory (Phillips, 1991).

Substituting (4) into (3) gives the error correction model

$$\Delta y_t = \sum_{i=1}^{p-1} \pi_i \Delta y_{t-i} - \alpha z_{t-p} + v_t \tag{5}$$

where $z_{t-p} = \beta' y_{t-p}$ represents deviations from the long run equilibrium relationship given by $\beta' y_{t-p}$. In this paper, we treat (5) as a reduced form for the structural econometric model

$$\Delta y_t = \sum_{i=1}^{p-1} \pi_i \Delta y_{t-i} - \alpha z_{t-p} + B^{-1} A u_t \tag{6}$$

where the A and B matrices and the u_t vector are as defined previously. Notice that cointegration implies restrictions on the long run dynamics of y_t but does not help us to identify the A and B matrices representing contemporaneous interactions among the y_t variables. This does not pose a problem, however, because the A and B matrices can be identified using exactly the same methods discussed earlier in the context of stationary VAR models. The only difference is that the resulting model features explicit restrictions on the long run effects of shocks to u_t as well as on their contemporaneous impact.

Model Selection

Univariate Tests of Stationarity

Selection of the appropriate model depends on the stationarity properties of the data. Several tests are available. Each of these tests performs well (has high power) on particular

types of data but has lower power on other types of data. Consequently, since the econometrician rarely knows the true population, it is advisable to use all available methods and to compare the results.

Unit Root Tests

The likelihood ratio tests developed by Dickey and Fuller (Fuller, 1976; Dickey and Fuller, 1979 and 1981) and Stock and Watson (1986, 1988) are the most commonly used unit root (nonstationarity) tests. Both tests are based on the significance of the parameter a_o in the OLS regressions

$$\Delta x_t = \mu + bt + a_0 x_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta x_{t-i} + e_t \tag{7}$$

where $\Delta x_t = x_t - x_{t-1}$ and x is a particular series in y. Critical values are reported in Fuller (1976) for the Dickey-Fuller test and Stock and Watson (1988) for their test. The same methods are used for testing larger numbers of unit roots in economic time series by replacing Δx_t with appropriately differenced series in the test equation.

There are no criteria for determination of the lag length, p, that are based on distributional theory. The standard procedures call for comparison of log-likelihood values, or modifications of the values each as Akaike's information criterion (AIC) or the Schwarz Bayesian criterion (SBC). Many econometric texts (e.g. Judge et al.) have discussions of these measures. Schmidt provides a critique of using information criteria which are not based on distributional theory, but does not provide alternative ways of model specifications.

Sen (1986) found the Dickey-Fuller and Stock-Watson techniques lack power when testing for more than one unit root. Dickey and Pantula (1987) provide a test for multiple unit roots, which starts with a large number of unit root and tests downwards. Dickey and Pantula

reparameterized a k-th order autoregressive model to get

$$\Delta^k x_t = \sum_{i=1}^{k} \psi_i \, \Delta^{i-1} x_{t-i} + e_t \tag{8}$$

The appropriate null hypothesis testing for the *k-th* unit root will be H_k, where $H_k : \psi_1 = \psi_2 = \ldots = \psi_k = 0, \ \psi_{k+1} < 0$, and the alternative hypothesis will be H_{k-1}. Test statistics used in the Dickey-Pantula test will be same as the Dickey-Fuller test statistics and the equations also can be augmented by introducing lagged dependent variables, as in the Dickey-Fuller test when auto-correlation appears in the error terms.

Though the Dickey-Fuller test is widely used for economic time series, weaknesses of the test are also well recognized. First, the test ignores the possibility that the true data generating process may have moving average terms as well as autoregressive terms. Schwert (1987, 1988) and Lee and Schmidt (1990) found low power for the Dickey-Fuller test in the presence of moving average errors in a process, causing too many rejections in case of negative moving average errors and too few rejections in case of positive moving average terms are experienced. Second, critical values of the Dickey-Fuller test are not valid when the nuisance parameters μ and b appear in a process (Schmidt, 1988; Nankervis and Savin, 1985; Evans and Savin, 1984). The meaning of the nuisance parameters depends on whether or not the unit root is present. Under the null hypothesis, μ represents a deterministic time trend and b represents a quadratic time trend. Under the alternative hypothesis, μ represents a deterministic level and b represents a deterministic time trend.

Phillips and Perron developed a test to solve the first problem by allowing a wide variety of weakly dependent and heterogeneously distributed time series (Perron, 1986; Phillips, 1987; Phillips and Perron, 1988). The test statistics are corrected by using Newey and West (1987)

type error covariance corrections and are valid under autocorrelation or conditional heteroscadasticity of the errors. The Phillips-Perron unit root test considers the following three OLS regression equations.

$$x_t = \bar{\mu} + \bar{b}(t-\frac{n}{2}) + \bar{a}\, x_{t-1} + \bar{e}_t \tag{9}$$

$$x_t = \mu^* + a^* x_{t-1} + e_t^* \tag{10}$$

$$x_t = a' x_{t-1} + e_t' \tag{11}$$

The test procedure starts with the null hypothesis

$$H_0^1: \bar{\mu} = \bar{b} = 0 \text{ and } \bar{a} = 1,$$

with the test statistic $Z(\Phi_2)$ (see Appendix 1). If the first null hypothesis is rejected, then test the second null hypothesis

$$H_0^2: \bar{b} = 0 \text{ and } \bar{a} = 1$$

with the test statistic $Z(\Phi_3)$. If H_0^2 is also rejected, then test the third null hypothesis

$$H_0^3 : \bar{a} = 1$$

with test statistic $Z(t_{\bar{a}})$. Rejection of H_0^3 implies no unit root and acceptance of H_0^3 implies a unit root in the series. The trend and drift terms are also important. If H_0^1 and H_0^2 cannot be rejected, then the procedure develops into a more powerful test with equation (10). The new null hypothesis will be

$$H_0^4 : \mu^* = 0,\ a^* = 1$$

and the test statistic will be $Z(\Phi_1)$. Rejection of H_0^4 leads to a test of the individual unit root hypothesis

$$H_0^5 : a^* = 1$$

with the test statistic $Z(t_{a^*})$. Acceptance of H_0^4 leads to estimation of equation (11) and the individual unit root hypothesis is

$$H_0^6 : a' = 1$$

The test statistic is $Z(t_{a'})$. No trend and no drift terms will be included in the series.

Schmidt and Phillips (1989) developed a test valid with the nuisance parameters (trend and drift terms in the Dickey-Fuller setup). They modified a unit root test based on Bhargava's (1986) parameterization. The Schmidt-Phillips unit root test starts with the OLS regression equation

$$\Delta x_t = \eta + \phi \overline{S_{t-i}} + \varepsilon_t \tag{12}$$

where

$$\overline{S_{t-1}} = x_{t-1} - \mu + b(t-1)$$
$$\mu = mean\ \Delta x = \frac{(x_n - 1)}{n-1}$$
$$b = x_1 - \mu$$

The nuisance parameters μ and b can be estimated by

$$\Delta x_t = \mu + bt + e_t \tag{13}$$

The test statistic is the usual t statistic for the OLS estimator $\hat{\phi}$ of ϕ in equation (12), and its critical values appear in appendix II. The test equation can be augmented with lagged dependent variables in case autocorrelation is present. In that case, the test statistic will be the ratio of the t statistic to the discount factor, ρ, where $\rho^2 = \sigma_\epsilon^2 / \sigma 2$, $\sigma_\epsilon^2 = \lim n^{-1} E\left(\Sigma e_t^2\right)$, $\sigma^2 = \lim n^{-1} E\left(S_n^2\right)$ and $S_n^2 = \left(\sum e_j\right)^2$. The meaning of the nuisance parameters μ and b in the Schmidt-Phillips test does not depend on the truth of the unit root hypothesis.

Tests and Estimation of Cointegrating Relationships

If unit roots are found, tests of cointegration among the nonstationary series are the next step in specifying a time series model. Several measures are available to test cointegration. A two step method proposed by Engle and Granger (1987) focuses exclusively on a bivariate system and is based on the autocorrelation function of residuals. Engle and Yoo (1987) found that the critical values of the augmented Dickey-Fuller test used by Engle and Granger increase with the number of variables in the system. Although Engle and Granger argued for the efficiency of the two step estimation method in the two variable cointegration case, the properties of the least square estimators are dubious when there are more than two variables. The regression estimates also don't consider the error structure of the underlying processes. Stock (1987) showed that the least square estimators are biased and not asymptotically normally distributed. Stock and Watson (1986 and 1988) developed a test based on the principal component estimators. Though the test is applicable to a system with more than two variables,

Engle (1987) discredited the test since it is not based on likelihood theory. Hoffman and Rasche (1989a) are skeptical to the method. They argued that inference from the Stock-Watson test is highly sensitive to arbitrary lag length specifications in the structure used to estimate the principal components. Also, available critical values generated by numerical simulation may not be applicable to a wide range of variables.

A consistent maximum likelihood estimation procedure is established by Johansen (1988, 1989) by using moments and cross moments matrices of the OLS residuals from the auxiliary regressions

$$\begin{aligned} \Delta y_t &= \sum_{i=1}^{p-1} \Gamma_i \, \Delta y_{t-i} + \tilde{\omega}_{0t} \\ y_{t-p} &= \sum_{i=1}^{p-1} \Phi_i \, \Delta y_{t-i} + \tilde{\omega}_{mt} \end{aligned} \tag{14}$$

The residuals are used to form the moment matrices where n is the sample size. The canonical correlations of $\hat{\omega}_{0t}$ and $\hat{\omega}_{mt}$ can be found by solving the eigenvalues of $S_{m0} \, S_{00}^{-1} \, S_{0m}$ with respect to S_{mm}, as given by the solution to

$$|\lambda S_{mm} - S_{m0} \, S_{00}^{-1} \, S_{0m}| = 0, \tag{15}$$

where $\hat{S}_{ij} = \frac{1}{n} \sum_{t=1} n \, \hat{\tilde{\omega}}_{it} \, \hat{\tilde{\omega}}_{jt}' \quad ; \qquad i, j = 0, m.$

The eigenvalues are arranged from largest to smallest ($\hat{\lambda}_1 > \hat{\lambda}_2 > \ldots > \hat{\lambda}_g$), and the corresponding eigenvectors are denoted by $\hat{D}$, normalized so that $\hat{D}' S_{mm} \hat{D} = I$.

The likelihood ratio test statistics for the hypothesis that there are at most $h < g$ cointegrating vectors is

$$\hat{\Lambda} = -n \sum_{j=h+1}^{g} \ln(1 - \lambda_i) \tag{16}$$

where λ_i are the g-h smallest eigenvalues calculated above. This statistic is distributed approximately as $\delta\chi^2(f)$ where $f = 2(g-h)^2$ and $\delta = 0.85 - 0.58f$. The critical values for Johansen-Juselius test statistics appear in appendix III.

An estimate of the space spanned by the cointegrating vectors is obtained from

$$\hat{\beta} = (\hat{D}_1, \hat{D}_2, \ldots, \hat{D}_h)$$
$$\hat{\pi} = -S_{0m}\hat{\beta}\hat{\beta}'$$

where $\hat{D}_i$ $(i = 1, 2, \ldots, h)$ are the eigenvectors corresponding to the h largest eigenvalues. Johansen shows how these procedures can be interpreted within the context of full information maximum likelihood estimation, which leads to increased efficiency compared to the two-step procedures of Engle and Granger.

ESTIMATING THE MODEL

Estimation Procedure

After the stationarity tests, the reduced form time series model will be estimated by one of three forms: a VAR in levels, a VAR in first differences, or an ECM. A VAR in first differences and the ECM can also be reparameterized to have the form of the VAR in levels as appears in (2). Once the model has been reparameterized, Fackler's two step estimation methods are applied to identify the structural form (1). Standard VAR techniques will then be applied to find impacts of research activity on other variables, and vice versa.

MEASURING RETURNS TO RESEARCH

The point of departure is the model of research benefits developed by Akino and Hayami (AH). In this model the benefits are the increase in social surplus in the agricultural output market, caused by research-induced shifts in the supply curve. Akino and Hayami provide formulae for estimating these benefits from data on quantities, elasticities, and the impacts of research on yield and area planted to new varieties. The time series model presented in the current paper provides information on these impacts.

The AH approach requires calculation of the supply shifter k by the formula

$$k = \left(\frac{Yield_{new} - Yield_{old}}{Yield_{new}} \right) \left(\frac{Area_{new}}{Total\ Area} \right),$$

where $Yield_{new}$ represents improved variety yields, $Yield_{old}$ represents traditional variety yields (using all inputs applied to the improved variety, such as fertilizer), $Area_{new}$ represents the area planted to the improved variety and *Total Area* represents the total land area planted to the crop. Thus, k represents the average percentage yield increase (measured with improved variety yields as the base), for the total area planted to the crop (which can be attributed to the introduction of the new variety). The advantage of this calculation is that the data requirements are quite modest, particularly when only one new variety has been introduced.

Three disadvantages of this calculation are important to the current paper. First, if research leads to the introduction of multiple new varieties, the calculation of k becomes more complex. In this case the analogous expression for k is

$$k^{*} = \sum_{All\ new\ varieties} \left(\frac{Yield_{new\ variety} - Yield_{old}}{Yield_{new\ variety}} \right) \left(\frac{Area_{new\ variety}}{Total\ Area} \right)$$

Data on yield increases are often available. However, data on area planted to each variety, particularly for open pollinated varieties, is often difficult or impossible to obtain.

A second disadvantage is that it is very difficult to obtain either yield increase or area data for non-varietal research, such as improved management or agronomic recommendations. For example, AH attempt to isolate the effect of varietal introductions from the effect of fertilizer use, and use only the varietal effects to calculate the returns to research. While this is important in valuing the impacts of a breeding program, it is of little use in assessing research which results in improved non-seed inputs (e.g. chemicals) and recommendations for their use.

The third disadvantage is the implicit assumption that research does not change total area.[3] Successful research may increase area planted through development of new varieties or agronomic practices that make it profitable to grow the crop on marginal lands that previously were uncultivated. Successful research may also increase area by increasing the relative profitability of a particular commodity, so that area planted increases through the replacement of less profitable crops.

When the time period in which the research inputs take place is relatively short, the disadvantages of the AH method may be minor: In a short time period, probably only a few new varieties were introduced, and area planted to a particular crop may not change dramatically. The disadvantages become more important as the relevant time period is lengthened. The time-series, econometric methodology described above can eliminate these disadvantages when reasonably long time-series of data on average yields (for the commodity in the geographic region at which the research is targeted), area and research expenditures are available.

---

[3] To see this, compare the case in which a new technique realizes 100% adoption but total area does not change with the case in which adoption is 100% but total use doubles. The *AH* supply shiftover remains the same in each case, since the second term of the right hand side in the expressiion for k does not change.

These three variables constitute the vector y described above. Statistical analysis as described above is used to determine the appropriate form of the reparameterization (2). The estimation analysis of (2) consists of regressing each of the current values of the variables on past values of all three variables, using appropriate estimation procedures. This estimation is used to calculate the impacts of research on yields and area.

In a VAR model, all variables in the system are endogenous. Therefore, the impact of research cannot be identified by traditional methods such as treating research as exogenous and arbitrarily increasing these expenditures. The simplest nontraditional method is the use of the dynamic multiplier. In this method the moving average error term, u_t, in equation (1) is considered to be the unexpected shock. These terms can be used to introduce unexpected shocks or exogenous changes to a variable. The resulting path of the system is distinguished from the endogenous path. For example, to see what happens when research expenditures increase by 1 percent at time t, arbitrarily set u_t to equal to 0.01 times actual expenditures. Then, the estimated structural form is transformed into the moving average representation

$$y_t = \sum_{s=0}^{\infty} D_s \, u_{t-s} \tag{17}$$

where D_s is a $(g \times g)$ matrix of moving average parameters. This representation implicitly ignores research shocks occurring prior to time t (which is appropriate of this time is the start of the program under investigation).

The moving average representation equation can be derived by either inversion or successive substitution of equation (2). The moving average parameters represent the net effect of particular shock to economy (ie. research shock to productivity), assuming no additional shocks occur. For example, the i-th row and the j-th column of D_s (that is, $D_{ij,s}$) captures the

response after s time periods of y_i to a unit shock in y_j. The moving average parameters, D_s, are often called the impulse response weights or dynamic response weights. The dynamic multiplier is defined to be the sum of the moving average parameters of the variables of interest (see Judge et. al. p.675). If research is totally ineffective, then the impulse response weights and the dynamic multiplier will equal zero. Henceforth, it will be assumed that research has some nonzero effect.

The variables of interest are research expenditures, yields, and area planted. In the VAR the shock to research expenditures at time t typically is allowed to affect all endogenous variables, although possibly after some time lag. This means that the exogenous increase in research expenditures will affect the area and yields obtained in the future, perhaps through increases in funding for successful projects, and that the shock will affect those future research expenditures determined endogenously in the model. Intuitively, the increase in research expenditures may lead to more examples of success, which in turn will lead the budget administrator to allocate more funds to research. These increased expenditures will lead to additional yield impacts. The moving average parameters, D_s, represent the combined direct and indirect effects of the exogenous increase in research expenditures at time t on area, yields and expenditures at time $t + s$.

The increase in research costs will include the initial 1 percent increase at time t, plus any increases at times $t + s$, for $s = 1, ...,$ as determined by the multiplier D_s. The increase in benefits at time $t + s$ can be calculated from the increase in yields (or other productivity measure) at time $t + s$ as determined by the dynamic multiplier D_s.[4]

[4] For a discussion of methods for translating increase in yields into increases in social benefits see Oehmke et al.

Sources of Fluctuations

A historical decomposition of the endogenous variables into components due to particular shocks provides information on sources of fluctuations in research expenditures and impact variables (yields) for the given period. The historical decomposition of $y_{t+\tau}$ centered at time t is

$$y_{t+\tau} = \sum_{s=0}^{\tau-1} D_s u_{t+\tau-s} + \sum_{s=\tau}^{\infty} D_s u_{t+\tau-s} \tag{18}$$

The matrices D_s are again interpreted to be dynamic multipliers. The first term on the right-hand-side contains the shocks to be realized between time t+1 and t+τ. These shocks are not observed at time t. The second term contains the influence of all shocks that occurred at or before time t. These shocks are known to the econmetrician (the infinite series can be replaced by a finite series and an initial condition). The second term represents the forecast of $y_{t+\tau}$ based on information available at time t, and is called the base projection.

The decomposition of variance for the forecast of future expected values of variables into components due to a particular shock provides information on the sources of unpredictable fluctuations. To show this, take the expectation formed at time t of equation (18). The first right hand side term becomes zero in this case since the unexpected future policy shocks $(u_{t+1}, u_{t+2}, \ldots, u_{t+\tau})$ are not known at time period t. Thus, this expectation is

$$E_t Y_{t+\tau} = \sum_{s=t}^{\infty} D_s u_{t+\tau-s} \tag{19}$$

The forecast error is the difference between the known value and the expected value of a variable (i.e. (18)-(19) = $\sum_{s=0}^{\tau-1} D_s u_{t+\tau-s}$). The forecast variance is

$$\sum_{s=0}^{\tau-1} E\left[u'_{t+\tau-s}\, D'_s\, D_s\, u_{t+\tau-s}\right] \tag{20}$$

Each summand in this expression represents the contribution of the shock $u_{t+\tau-s}$ to the variance of $y_{t+\tau}$. For example, for $s=s_o$, the $(i, j)^{th}$ element of the forecast variance represents the contribution of the shock at time $t+\tau-s_o$ to variable yi to the variance at time $t+\tau$ of variable yj.

Explaining the past

The decomposition of the forecast error and variance can be used retrospectively to determine the effect of previous shocks on the time path of the economy. For example, suppose that at time t, the econometrician wishes to look retrospectively at departures from the expected path between time $t_o < t$, and time t_1. This might be the case if an unexpected decrease in research funding occurred at time t_o, and the econometrician was to evaluate the effects of the decrease on yield, area and total research funds through time t_1. In this case the expected time path of y could be projected using information available at time t. The difference between this projection and the actual path at time t_o+1 provides estimates of the u_{t_o+1}. Continuing forward provides estimates of all the errors between time t_o and t_1, which can then be used in equation (20).

Appendix I. Definition and Critical Values for Phillips-Perron Test Statistics

The test statistics are defined by

$$Z(t_{\bar{\alpha}}) = (S_u/S_{nl})t_{\bar{\alpha}} - (1/2S_{nl})(S_{nl}^2 - S_u^2)[n^{-2}\Sigma(x_{t-1} - \bar{X}_{.1})^2]^{-1/2}$$

$$Z(t\alpha^*) = (S_u/S_{nl})t\alpha^* - (1/2S_{nl})(S_{nl}^2 - S_u^2)[n^{-2}\Sigma(x_{t-1} - \bar{X}_{.1})^2]^{-1/2}$$

$$Z(t\alpha') = (S_u/S_{nl})t\alpha' - (n^3/4(3D_x)^{1/2}S_{nl})(S_{nl}^2 - S_u^2)$$

$$Z(\Phi_1) = (S_u^2/S_{nl}^2)\Phi_1 - (1/2S_{nl}^2)(S_{nl}^2 - S_u^2)\{n(a^* - 1)$$

$$- 1/4(S_{nl}^2 - S_u^2)[n^{-2}\Sigma(x_{t-1} - \bar{X}_{.1})^2]^{-1}\}$$

$$Z(\Phi_2) = (S_u^2/S_{nl}^2)\Phi_2 - (1/3S_{nl}^2)(S_{nl}^2 - S_u^2)[n(a' - 1)$$

$$- (n^6/48D_y)(S_{nl}^2 - S_u^2)]$$

$$Z(\Phi_3) = (S_u^2/S_{nl}^2)\Phi_3 - (1/2S_{nl}^2)(S_{nl}^2 - S_u^2)[n(a' - 1)$$

$$- (n^6/48D_y)(S_{nl}^2 - S_u^2)]$$

where

$$t_{\bar{\alpha}} = (\bar{\alpha} - \alpha)\{\Sigma(x_{t-1} - \bar{X}_{.1})^2\}^{1/2}/\bar{S}$$

$$t_{\alpha^*} = (\alpha^* - \alpha)\{\Sigma(x_{t-1} - \bar{X}_{.1})^2\}^{1/2}/S^*$$

$$t_{\alpha'} = (\alpha' - \alpha) / (S'^2C_3)^{1/2}$$

$$\Phi_1 = (2S^{*2})^{-1}\{nS_0^2 - nS^{*2}\}$$

$$\Phi_2 = (3S'^2)^{-1}(nS_0^2 - nS'^2)$$

$$\Phi_3 = (2S'^2)^{-1}(nS_0^2 - \bar{n}(x_0 - \bar{X}_{.1})^2 - nS'^2)$$

$$\alpha^* = \Sigma(x_t - \bar{X}_0)x_{t-1}/\Sigma(x_{t-1} - \bar{X}_{.1})^2$$

$$a' = (\Sigma x_{t-1}^2)^{-1}\Sigma x_t x_{t-1}$$

$$\mu^* = \bar{X}_0 - a^{*x}\bar{X}_{.1}$$

$$\bar{X}_{.i} = n^{-1}\sum_t x_{t-i} \quad (i=0,1)$$

and $\bar{S}$, S^* and S' are the standard errors of regression (9), (10) and (11), respectively.

S_0 is S' when a' = 1. S_u^2 is a consistent estimator of $\sigma_e^2 = \lim n^{-1} \Sigma E(e_t^2)$ and S_{nl}^2 is a consistent estimator of $\sigma^2 = \lim n^{-1}E(S_n^2)$ under the appropriate null hypothesis, where $s_n = \Sigma e_t$. n represent the number of observations. The consistent estimation of σ^2 concerns the appropriate choice of truncation lag parameter. Though the choice will be an empirical matter, Perron (1986) recommended to inspect the sample autocorrelation of first differenced data. In this paper, the LR in first differences is used together with the recommendation. C_i is the (i,i) element of the matrix $(X'X)^{-1}$ and D_x denotes the determinant of the (X'X) which is presented as

$$D_x = (n^2(n^2-1)/12)\ \Sigma x_{t-1}^2 - n(\Sigma t x_{t-1})^2 + n(n+1)\ \Sigma t x_{t-1}\ \Sigma x_{t-1}$$
$$-(n(n+1)(2n+1)/6)\ (\Sigma x_{t-1})^2.$$

The critical values for the test statistics are presented as

Test Statistics	Percentiles			
	10%	5%	2.5%	1%
$Z(\Phi_2)$	4.03	4.68	5.31	6.09
$Z(\Phi_3)$	5.34	6.25	7.16	8.27
$Z(t_{\tilde{\alpha}})$	-3.12	-3.41	-3.66	-3.96
$Z(\Phi_1)$	3.78	4.59	5.38	6.43
$Z(\alpha*)$	-2.57	-2.86	-3.12	-3.43
$Z(\alpha')$	-1.62	-1.95	-2.23	-2.58

Appendix II. Critical Values for Schmidt-Phillips Test Statistics

n	Percentiles			
	10%	5%	2.5%	1%
25	-2.85	-3.18	-3.50	-3.90
50	-2.80	-3.11	-3.39	-3.73
100	-2.77	-3.06	-3.32	-3.63
200	-2.76	-3.04	-3.30	-3.61
500	-2.76	-3.04	-3.29	-3.59
1000	-2.75	-3.02	-3.28	-3.58
2000	-2.75	-3.02	-3.27	-3.56

Appendix III. Critical Values for Johansen-Juselius Cointegration Test Statistics

h	m	percentiles		
		5%	2.5%	1%
g-8	18	131.4	136.3	142.2
g-7	7	103.1	120.0	112.7
g-6	6	78.1	82.0	86.6
g-5	5	57.2	60.3	63.9
g-4	4	38.6	41.2	44.5
g-3	3	23.8	26.1	28.5
g-2	2	12.0	13.9	15.6
g-1	1	4.2	5.3	5.3

References

Akino, M. and Hayami, Y., 1975. Efficiency and equity in public research: Rice breeding in Japan's economic development. Am. j. Agric. Econ., 57:1-10.

Bernanke, Ben S. "Alternative Explanations of the Money-Income Correlation." Carnegie-Rochester Conference Series on Public Policy. 25(1986):49-100.

Bessler, David A. "Relative Prices and Money: A Vector Autoregression on Brazilian Data."Amer. Journ. Agr. Econ. 66(1984):25-30.

Bhargava, A. :On the Theory of Testing for Unit Roots in Observed Time Series." Rev. of Econ. Studies. 52(1986):369-384.

Blanchard, Oliver J. and Mark Watson. "Are All Business Cycles Alike?" Paper Presented at Conference on Business Cycle. Nat. Bur. Econ. Res., Puerto Rico. Mar. 22-25, 1984.

Bonnen, J. T. "Historical Sources of U.S. Agricultural Productivity: Implications for R & D Policy and Social Science." Amer. Journ. Agr. Econ. 65(1983):958-966.

Bredahl, M. and W. Peterson. "The Productivity and Allocation of Research: U.S. Agricultural Experiment Stations." Amer. Journ. Agr. Econ. 58(1976):684-692.

Capalbo, S.M. and J.M. Antle (Eds.), 1988. Agricultural productivity: measurement and explanation. Washington, D.C.: Resources for the Future (Baltimore: Johns Hopkins University Press).

Cooley, Thomas F. and Stephen F. LeRoy. "Atheoretical Macroeconometrics: A Critique." J. Monetary Econ. 16(1985):283-308.

Dickey, D. A. and W. A. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." Journ. Amer. Stat. Assoc. 74(1979):427-431.

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

Dickey, D. A. and S. G. Pantula. "Determining the Order of Differencing in Autoregressive Processes." Journ. Bus. and Econ. Stat. 5(1987):455-461.

Engle, Robert F. "On the Theory of Cointegrated Econometric Time Series." Mimeo. 1987.

Engle, Robert F. and C. W. J. Granger. "Cointegration and Error Correction: Representation, Estimation, and Testing." Econometrica. 55(1987):251-276.

Engle, and B. S. Yoo. "Forecasting and Testing in Co-integrated Systems." Journ. Econometrics. 35(1987):143-59.

Evans, G. B. A. and N. E. Savin. "Testing for Unit Roots 2." Econometrica. 52(1984):1241-1270.

Evenson, R. E., P. E. Waggoner and V. W. Ruttan. "Economic Benefits from Research: An Example from Agriculture." Science. 205(1979):1101-1107.

Fuller, W. A. Introduction to Statistical Time Series. Wiley. N.Y. 1976.

Griliches, Z. 1958. Research costs and social returns: Hybrid corn and related innovations. J. Polit. Econ., 66:419-431.

Hoffman, Dennis and Robert H. Rasche. "Long-Run Income and Interest Elasticities of Money Demand in the United States." Working Paper No. 2949. Nat. Bur. of Econ. Res., Inc. 1989a.

Hoffman, Dennis and Robert H. Rasche. "The Demand for Money in the U.S. During the Great Depression: Estimates and Comparisons with the Post War Experience." Unpublished Working Paper. Nat. Bur. Econ. Res., Dec. 1989b.

Johansen, Sören. "Statistical Analysis of Cointegration Vectors." Journ. of Econ. Dynamics and Control. 12(1988):231-254.

Johansen, Sören. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." Preprint 3. Inst. of Math. Stat., Univ. of Copenhagen. Jan. 1989.

Johansen, Sören and Katarina Juselius. "Hypothesis Testing for Cointegration Vectors with an Application to the Demand for Money in Denmark and Finland." Preprint 2. Inst. of Math. Stat., Univ. of Copenhagen, Feb. 1988.

Johansen, Sören and Katarina Juselius. "The Full Information Maximum Likelihood Procedures for Inference on Cointegration with Applications." Preprint 4. Inst. of Math. Stat., Univ. of Copenhagen, Jan. 1989.

Judge, G. , W. Griffith, R. Hill, H. Lutkepohl, and T. Lee. The Theory and Practice of Econometrics. 2ed. Wiley, N.Y., 1985.

Lee, J. and P. Schmidt. "Unit Root Tests with ARMA Errors." Unpublished Working Pap., Dept. of Econ., M.S.U., E. Lansing, MI. 1990.

Myers, R. J., R. R. Piggott and W. G. Tomek. "Estimating Sources of Fluctuations in Australian Wool Market: An Application of VAR Methods." Amer. Journ. Agr. Econ. Forthcoming. 1991.

Nankervis, J. C. and N. E. Savin. "Testing the Autoregressive Parameters with t Statistics." Journ. Econometrics. 27(1985):143-161.

Newey, W. K. and K. D. West. "A Simple Positive Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." Econometrica. 55(1987):703-708.

Norton, G. W. "The Productivity and Allocation of Research: U.S. Agricultural Experiment

Stations, Revisited." N. Cent. Journ. Agr. Econ. 3(1981):1-12.

Norton, G.W. and Davis, J.S., 1981. Evaluating returns to agriculture research: A review. Am. J. Agric. Econ., 63:685-699.

Oehmke, J. F. "Persistent Underinvestment in Public Agricultural Research." Agri. Econ. 1(1986):53-65.

Oehmke, J. F. and X. Yao. "A Policy Preference Function for Government Intervention in the U.S. Wheat Market." Amer. Journ. Agr. Econ. 72(1990):632-640.

Orden, David. "Agriculture, Trade, and Macroeconomics: the U.S. Case." Journ. Policy Modeling, 8(1986):27-51.

Orden, D. and P. F. Fackler. "Identifying Monetary and Other Macroeconomic Impacts on Agricultural Prices in VAR Models." Staff Paper 88-36. Dept. of Agr. Econ., Virginia Polytechnique Institute and State Univ., Blacksburg, Virginia. 1988.

Perron, P. "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach." Unpublished. Dept. of Econ. Universite' de Montre'al. 1986.

Phillips, P. C. B. "Time Series Regression with A Unit Root." Econometrica. 55(1987):277-301.

Phillips, P. C. B. "Optimal Inference in Cointegrated Systems." Econometrica. 59(1991):283-306.

Phillips, P. C. B. and P. Perron. "Testing for a Unit Root in Time Series Regression." Biometrica. 75-2(1988):334-346.

Robertson, John and David Orden. "A Vector Error-Correction Model of Money and Price Dynamics in New Zealand." Working Paper, Dept. Agr. Econ., Virginia Tech., Blacksburg, Virginia. Sept. 1989.

Rose-Ackerman, S. and R. E. Evenson. "The Political Economy of Agricultural Research and Extension: Grants, Votes and Reapportionment." Amer. Journ. Agr. Econ, 67(1985):1-14.

Schmidt, Peter. "Dickey-Fuller Tests with Drift." Econometrics and Economic Theory," Pap. No. 8717. Dept. of Econ. M.S.U., E. Lansing, MI. 1988.

Schmidt, Peter and Peter C. B. Phillips. "Testing for a Unit Root in the Presence of Deterministic Trends." Unpublished. Dept. of Econ. M.S.U., E. Lansing, MI. Oct. 1989.

Schwert, G. W. "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data." Journ. Monetary Econ. 20(1987):73-103.

Schwert, G. W. "Tests for Unit Roots; A Monte Carlo Investigation." Journ. Bus. and Econ. Stat. 7(1988):147-159.

Sen, D. L. "Robustness of Single Unit Root Test Statistics in the Presence of Multiple Unit

Roots." Unpublished Ph.D Dissertation, Dept. of Stat., N. Carolina State Univ., 1986.

Sims, C. A. "Macroeconomic and Reality." Econometrica. 48(1980):1-48.

Sims, C. A. "Are Forecasting Models Usable for Policy Analysis?" Fed. Res. Bank of Minn. Quart. Rev. Winter. 1986. pp. 1-16.

Sims, C. A., J. Stock, and M. Watson. "Inference in Linear Time Series Models with Some Unit Roots." Working Paper. Dept. of Econ., Univ. of Minn., Minneapolis, Minn. Nov. 1986.

Stock, J. H. "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors." Econometrica. 55(1987):1035-1056.

Stock, J. H. and M. W. Watson. "Does Real GNP have a Unit Root?" Economic Letters. 22(1986):147-151.

Stock, J. H. and M. W. Watson. "Testing for Common Trends." Journ. Amer. Stat. Assoc. 83(1988):997-1107.

Stock, J. H. and M. W. Watson. "Interpreting the Evidence on Money-Income Causality." Journ. Econometrics. 40(1989):161-181.

Zetner, R. P. and W. L. Peterson. "An Economic Evaluation of Public Wheat Research and Extension Expenditure in Canada." Canadian Journ. Agr. Econ. 32(1984):327-353.

[File: c:\docs\oehmke\pap\dynamica.pap]