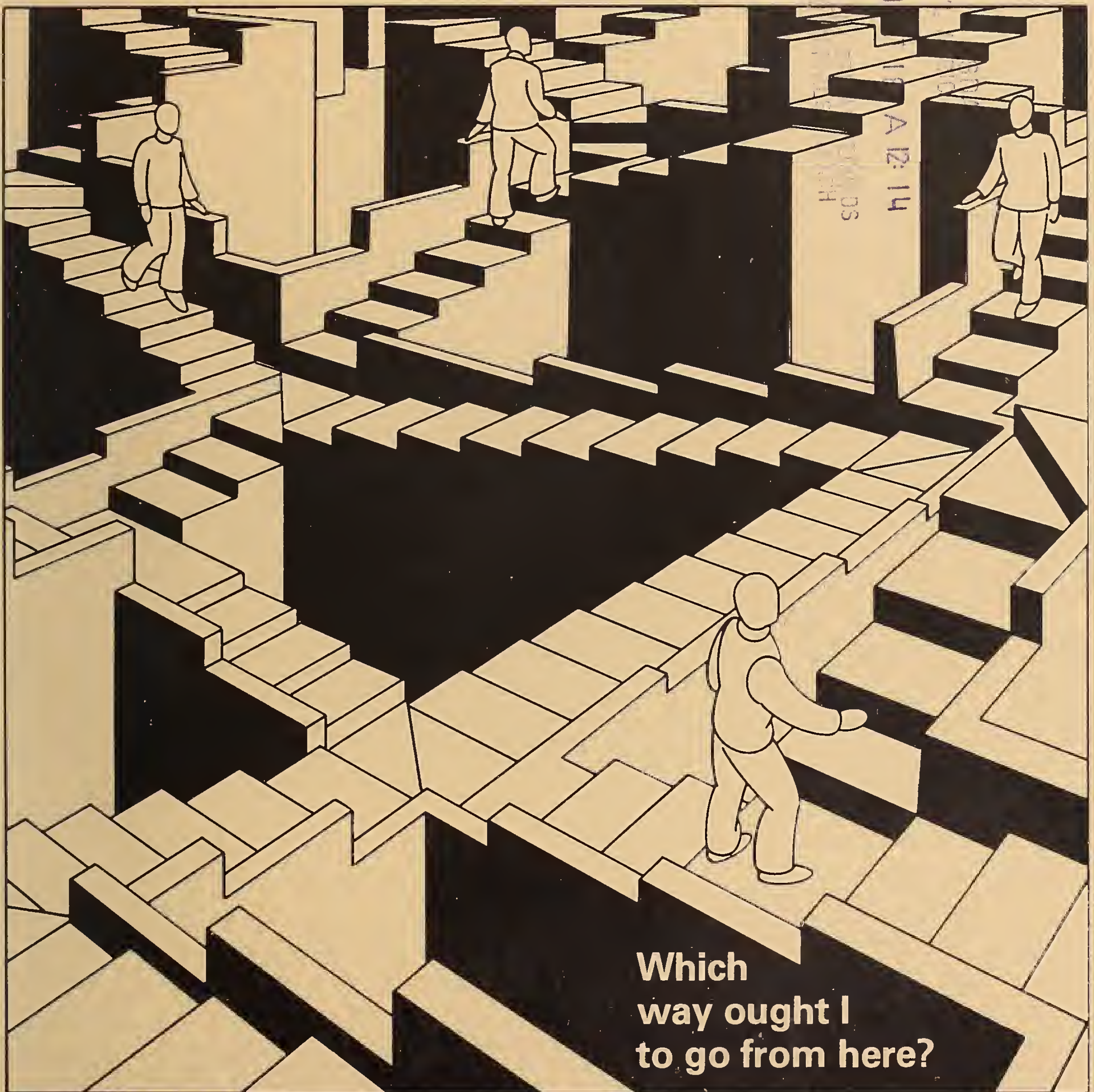
AGRICULTURAL ECONOMICS RESEARCH

APRIL 1980 VOL. 32 No. 2



AGRICULTURAL ECONOMICS RESEARCH

A Journal of the U.S. Department of Agriculture • Economics, Statistics, and Cooperatives Service

Readers interested in a chronological bibliography of materials in Agricultural Economics Research from the first issue in 1949 through the second issue in 1980 should write Ann Bassett, Staff Assistant, ESCS/EDD, U.S. Department of Agriculture, Washington, D.C. 20250.

COVER IDEA: Our appreciation to M.C. Escher for his visual perceptions of the logical paradox implied in the pursuit of truth—which we also seek.

IN THIS ISSUE

The pursuit of truth, according to John Dewey, is accountable to nothing and to no one not a part of that pursuit itself. For researchers in agricultural economics, this means that they are not obliged to account for their procedures and findings to just anyone; but it also means that they are held responsible for explaining their results to certain people—people who are a part of the pursuit.

When we use data and theory to deal with a problem and we reach particular conclusions which we claim to be true, to whom are we—as researchers—accountable? Gunnar Myrdal warned that the individual cannot be "left alone to manage his rationalizations as he pleases, without interference from the outside. His valuation will, instead, be questioned and disputed."

Who can dispute us and ask us questions about procedures and findings that we are obliged to anwer? Some who have thought about this question are clear about what goes wrong when we act as if we were accountable to those who are not a part of the pursuit of truth. We are not obliged to be slaves to the logic and theories of what John Maynard Keynes called "some defunct economist." Michael Polanyi warned against accepting conclusions imposed by authority—it is not so that "might makes right." Georg Wilhelm Friedrich Hegel said "the simple tendency of the naive mind is to accept with trustful conviction the truth which is publicly known." John Locke warned that to retain credibility we must not put "passion in the place of reason"—we cannot go off on our own merely concluding what we want to believe.

When we ask which of several roads might lead us to the truth we are not unlike Alice who, in her journey through Wonderland, looked up at the Cheshire Cat on a bough of a tree and asked: "Would you tell me, please, which way I ought to go from here?" "That depends," said the cat, "a good deal on where you want to get to." Alice said she did not much care—as long as she got somewhere. She knew she wanted to get home but did not know where that was. Similarly, we often know we want the truth but do not know what it is. The Cheshire Cat was not a part of Alice's search and did not try to influence her decision. But agricultural economists are often not so lucky. We are too often given advice by—and held accountable to—people who are not a part of that pursuit itself.

The articles in this issue question how we can know whether our research is leading down the path toward truth. Kost, in the first article, which examines a net foreign trade model, recognizes that validating a model is a subjective process—we have confidence in a model that somehow, intuitively, feels right. Another researcher working on the same problem is in a legitimate position to raise a question about validity if, to that researcher, the proposed model subjectively feels wrong.

Rodriguez and Kunkel, in the second article, which examines a model of the agricultural sector of the Philippine economy, mention a number of economic development models which simulate competitive markets. Another researcher working on the same problem might judge that competitive markets are not functioning in the public interest and that purposive government intervention needs to be modeled instead. Such a researcher is in a legitimate position to raise a question about the social values implied in the competitive model.

Authors of the first two articles make their main appeals for validity to data. They claim their models track the historical data well. Another researcher who finds that Kost's or Rodriguez and Kunkel's conclusions differ from other empirical evidence can legitimately question the empirical foundations of the models.

Baron, in an article on farm tenure agreements, and Salathe, in an article on the food stamp program, appeal to logic and economic theory for evidence that their conclusions are true. Each has reformulated economic theory in a manner which raises questions about the truth others have seen in the logical conclusions of earlier formulations.

The articles in this issue illustrate that pursuers of truth are accountable to those who are a part of that pursuit and who can, therefore, raise legitimate questions about the data, theory, value judgments, and subjective feelings which the proposed models draw upon.

CLARK EDWARDS

CONTENTS

MODEL VALIDATION AND THE NET TRADE MODEL

By William E. Kost*

In this article I discuss validation of structural economic models. I emphasize goodness-of-fit measures for historical simulations plus comparisons with alternative models. I then use these procedures to evaluate the world trade forecast modeling system being developed in the International Economics Division, ESCS.

The article discusses processes of validating structural forecasting models. It summarizes methods of evaluating the goodness of fit of model simulations over historical periods and methods of comparing the forecasting behavior of structural models with that of simple time series models. The net trade model provides a case study for these two validation processes.

Keywords

Validation
Modeling
Forecasting
International trade
Wheat
Coarse grain

VALIDATION

A common approach to analyzing economic issues involves developing a model that simulates economic behavior. This model becomes a proxy for reality. The model's behavior is then evaluated to provide insight into analyzing economic issues. Used this way, models have to represent reality accurately. One determines whether or not a model is good through the process known as validation. Determining the "goodness" of a model is a subjective process that involves using both economic and statistical criteria. One usually begins to construct and validate a model by defining the economic problem that model will analyze. This procedure restricts the model's size and scope to only relevant aspects of economic behavior.

Once the problem has been identified, an initial structural hypothesis can be proposed. General statements are developed concerning the form of the structural equations, the availability of data, structural shifts over time, and the signs and magnitudes of coefficients. An appropriate sampling and equation estimation procedure is defined and the preliminary model is estimated. These initial equations are evaluated on the basis of both the prior economic hypotheses and statistical, econometric criteria. In light of this evaluation, several equations may have to be made more accurate through an alternative equation specification (and/or possibly estimation procedure) that is also consistent with the set of hypotheses previously specified. In some instances this equation evaluation leads to rejection of the previously specified hypotheses. The prior hypothesis framework must then be redefined and new equations specified and estimated that will be consistent with the new hypotheses. This process may also lead to the rejection of the data base, which then requires the generation of a new data base that will lead to different, more acceptable model parameters.

An initial model is constructed with this process of hypothesis generation, data base construction, equation estimation, and equation evaluation. Only after these steps have been taken can we evaluate the behavior of the complete model. How does the complete model track within the historical period of the sample? How does it respond to shocks? How does the model forecast outside the period of the sample?

Simulation Methods

The purpose of model validation is to increase one's confidence in the ability of the model to provide useful information. Attention focuses on goodness of fit of the complete model (as opposed to goodness of fit of any single equation). Therefore, model validation continues throughout model construction and even into model use.

Tracking the model through the historical period of fit allows evaluation of interdependence between its equations. The lowest level of interdependence in any historical simulation is the residual check. Under a residual check simulation, all equations are assessed with all explanatory variables set at their actual values. For example, assume the model can be represented by a set of n equations:

$$Y_t = F(Y_t, Y_{t-1}, \ldots, Y_{t-i}, X_t, X_{t-1}, \ldots, X_{t-j}, \xi_t) \qquad (1)$$

*The author is an agricultural economist in the International Economics Division, ESCS.

where:

- n = the number of endogenous variables in the model,
- m = the number of exogenous variables in the model,
- t = the time period,
- Y = an n column vector of endogenous variables,
- i = the maximum number of lag periods on endogenous variables,
- X = an m column vector of exogenous variables,
- j = the maximum number of lag periods on exogenous variables, and
- ξ = an n column vector of errors.

A residual check simulation would be the solution of:

$$\hat{Y}_t = \hat{F}(Y_t, Y_{t-1}, ..., Y_{t-i}, X_t, X_{t-1}, ..., X_{t-j}) \quad (2)$$

for $\hat{Y}_t$. Estimates of model parameters, $\hat{F}$, and actual values for all right-hand variables are used in this calculation. This is equivalent to solving each equation independently. It provides a check on the accuracy of the solution algorithm. The residuals:

$$\hat{\xi}_t = Y_t - \hat{Y}_t \quad (3)$$

will be identical to those produced in the econometric estimation of F.

The next level of interdependence involves solving:

$$\hat{Y}_t = \hat{F}(\hat{Y}_t, Y_{t-1}, ..., Y_{t-i}, X_t, X_{t-1}, ..., X_{t-j}) \quad (4)$$

In this static simulation, all exogenous and lagged endogenous variables are set at actual values. This provides a series of simultaneous solutions for endogeneous variables, each for a single time period. Static simulation errors will typically be larger than those in a residual check as this simulation allows for interactions among current-period endogenous variables.

A dynamic simulation provides the highest level of interdependence. The dynamic simulation involves solving:

$$\hat{Y}_t = \hat{F}(\hat{Y}_t, \hat{Y}_{t-1}, ..., \hat{Y}_{t-i}, X_t, X_{t-1}, ..., X_{t-j}) \quad (5)$$

where only exogenous variables and the initial i period endogenous variables are set at actual values. The first time period simulated will have the same solution as the static simulation. The second time period will differ; its simulation will use values of the estimated lagged endogenous variables from the previous period, $\hat{Y}_{t-1}$, rather than the actual values, Y_{t-1}. The third time period simulated will use estimated endogenous variable values for the first two time periods, and so on throughout the simulation horizon.

The dynamic simulation furnishes a simultaneous solution that starts at an initial point in time, based on a set of initial conditions, then feeds on itself for additional inputs throughout the simulation time horizon. A dynamic simulation differs from a static simulation; it generates a single multiperiod simulation rather than a series of single-period simulations. All multiperiod forecasts of future behavior are dynamic simulations. These forecast simulations, of course, also require forecasted, rather than actual, values for the exogenous variables.[1] Dynamic simulation errors will typically be larger than those in a static simulation. Errors can be propagated throughout the system both by interactions among current-period endogenous variables and by interactions among current and lagged endogenous variables.

Each of the three simulations can be evaluated for goodness of fit. As a residual check simulation yields information identical to that from the econometric evaluation of individual equations, this article will focus on static and dynamic simulations.

Validating Multivariable Models

Problems arise in evaluation of models that simulate many endogenous variables simultaneously. Virtually no techniques exist for overall goodness-of-fit evaluation of multiple-response simulation models. One can sometimes circumvent this multiple-response problem, either by viewing a simulation with many responses as many simulations, each with a single response, or by combining several responses and treating the combination as a single response.

[1] A dynamic simulation that forecasts future behavior involves solving $\hat{Y}_t = \hat{F}(\hat{Y}_t, \hat{Y}_{t-1}, ..., \hat{Y}_{t-i}, \hat{X}_t, \hat{X}_{t-1}, ..., \hat{X}_{t-j})$.

The mathematician says that 2+2 is identicially equal to 4. The statistician says that 2+2 is approximately 4. The economist asks, "What kind of number are you looking for?"

Oral tradition

Wallace suggests that "if the question that promotes the research relates to a specific variable, the research should be keyed on that variable. Reliability of the model should be based upon how well the key variable is predicted" (*12*, p. 15).[2] By their nature, models contain several variables that are relatively unimportant. However, the problems for which models are typically used require more than one key variable. Wallace's approach narrows the range of focus but still leaves a subjective decision regarding a model's goodness of fit.

GOODNESS-OF-FIT MEASURES

Several goodness-of-fit measures are now presented for each endogenous variable. To the extent they are favorable, they increase one's subjective confidence in the model, and help evaluate changes in the model. A comparison of prechange and postchange simulations, in terms of these goodness-of-fit measures, provides information concerning the merit of the structural change. Five types of goodness-of-fit measures will be examined: errors, regression, correlation, inequality coefficients, and turning points.

Errors

Several alternative measures of simulation error can be calculated. They all measure the deviation of a simulated variable from the actual path. The simplest measure is mean error:

$$\text{Mean error} = \frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t) \quad (6)$$

where:

T = the number of periods simulated,

$\hat{Y}_t$ = the simulated level of the variable at time period t, and

Y_t = the actual level of the variable at time period t.

The mean error can be misleading. Large positive and negative errors offset each other and bias the mean error downward.

The mean absolute error (MAE) is defined as:

$$\text{MAE} = \frac{1}{T}\sum_{t=1}^{T}|\hat{Y}_t - Y_t| \quad (7)$$

The mean absolute error is not subject to the bias associated with the mean error.

Probably more frequently used in the literature is the root-mean-square (RMS) error:

RMS error =

$$\sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \quad (8)$$

This measure weights large errors more than the mean absolute error.

These three errors can best be evaluated relative to the average size of the variable. They, therefore, become more relevant expressed in percentage terms:

mean percentage error =

$$\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\hat{Y}_t - Y_t}{Y_t}\right) \quad (9)$$

mean absolute relative error (MARE) =

$$\frac{1}{T}\sum_{t=1}^{T}\left(\frac{|\hat{Y}_t - Y_t|}{Y_t}\right) \quad (10)$$

RMS percentage error =

$$\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\hat{Y}_t - Y_t}{Y_t}\right)^2} \quad (11)$$

In all cases, the smaller the error, the better the fit.

Regression

A linear regression of actual values of a variable on predicted values has been suggested by Cohen and Cyert (*1*, pp.112-127) as a method of testing goodness of fit:

$$Y_t = \beta_0 + \beta_1 \hat{Y}_t + \xi_t \quad (12)$$

$\hat{Y}_t$ would equal Y_t for all t in perfect models and the resulting regression is one with zero intercept ($\beta_0 = 0$) and unit slope ($\beta_1 = 1$). Parameters of the regression would be tested to see if they differed significantly from zero and one and if the ξ_t's are small.

Correlation Coefficient

Association between predicted and actual values for a variable can be measured by the correlation coefficient (R) or by R-square. R-square measures the proportion of

[2] Italicized numbers in parentheses refer to items in References at the end of this article.

the variation explained by a linear regression of predicted on actual values. A disadvantage of the R or R-square as the sole measure of goodness of fit is that perfect correlation only implies an exact linear relationship between predicted and actual values. For simulations to be unbiased, and therefore perfect, regression parameters of $\beta_0 = 0$ and $\beta_1 = 1$ must also exist.

Theil's Inequality Coefficients

Theil proposed the inequality coefficient as a measure for analyzing accuracy. Several definitions of the inequality coefficient exist in the literature. Even Theil presents different definitions at different points. The first inequality coefficient was proposed by Theil in *Economic Forecasts and Policy* (*11*, pp. 32-33).[3]

$$U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\hat{Y}_t^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}Y_t^2}} \qquad (13)$$

This inequality coefficient is bounded by zero and one. When $U = 0$, $\hat{Y}_t = Y_t$ for all periods, and a perfect simulation exists. When $U = 1$, either the model always predicts zero for nonzero actual values, or the model predicts nonzero values for actual values that are always zero, or negative proportionality exists between predicted and actual values. Unlike the correlation coefficient, this inequality coefficient penalizes a consistent bias in the simulation. However, again unlike the correlation coefficient, it is sensitive to additive transformation of variables.[4] When one is evaluating alternative variations of a single model, where general levels of endogenous variables remain relatively unchanged, this disadvantage is not a serious drawback. However, this version of the inequality coefficient may not be comparable across models.

To overcome sensitivity to an additive transformation, Theil proposed defining the inequality coefficient in terms of changes in a variable. The base from which all predicted and actual variables are measured is fixed, and comparisons can then be made across models. This inequality coefficient is defined as:

$$U_1 = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}((\hat{Y}_t - Y_{t-1}) - (Y_t - Y_{t-1}))^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_{t-1})^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t - Y_{t-1})^2}} \qquad (14)$$

[3]This is the definition used in the FEDEASY "Actfit" comparison of actual and predicted time series. FEDEASY refers to the set of linkules added to SPEAKEASY by the Federal Reserve System. SPEAKEASY is a software package widely used for analysis by ESCS economists.

[4]Adding a constant, k, to any set of predicted and actual values will reduce the value of this inequality coefficient by increasing the denominator and leaving the numerator of the fraction defining U unchanged.

$$\sqrt{\sum_{t=1}^{T}((\hat{Y}_t + k) - (Y_t + k))^2} \Big/ \sqrt{\sum_{t=1}^{T}(\hat{Y}_t + k)^2} + \sqrt{\sum_{t=1}^{T}(Y_t + k)^2}$$

$$= \sqrt{\sum_{t=1}^{T}(\hat{Y}_t + k - Y_t - k)^2} \Big/ \sqrt{\sum_{t=1}^{T}(\hat{Y}_t + k)^2} + \sqrt{\sum_{t=1}^{T}(Y_t + k)^2}$$

$$= \sqrt{\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \Big/ \sqrt{\sum_{t=1}^{T}(\hat{Y}_t + k)^2} + \sqrt{\sum_{t=1}^{T}(Y_t + k)^2}$$

$$< \sqrt{\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \Big/ \sqrt{\sum_{t=1}^{T}(\hat{Y}_t)^2} + \sqrt{\sum_{t=1}^{T}(Y_t)^2}$$

The Math-Econ make exquisite modls finely carved from the bones of walrus. Specimens made by their best masters are judged unequalled in both workmanship and raw material by a unanimous Econographic opinion. If some of these are "useful"—and even Econ testimony is divided on this point—it is clear that this is purely coincidental in the motivation for their manufacture.

Axel Leijonhufvud
"Life Among the Econ"

The U_1 inequality coefficient also ranges between zero and one with $U_1 = 0$ occurring when a perfect simulation exists. U_1 is always less than U as only the denominator changes from one formulation to the other.[5]

Theil proposed a third inequality coefficient in *Applied Economic Forecasting* (*10*, pp.26-29):

$$U_2 = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}((\hat{Y}_t - Y_{t-1}) - (Y_t - Y_{t-1}))^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t - Y_{t-1})^2}} \qquad (15)$$

This U_2 inequality coefficient ranges from zero to infinity. For a perfect simulation, when $\hat{Y}_t = \hat{Y}_t$ for all periods, $U_2 = 0$. A no-change forecast model, where $\hat{Y}_t = \hat{Y}_{t-1}$ for all periods, generates a U_2 inequality coefficient of 1. No upper bound on the U_2 inequality coefficient means that there can be a model that is worse than a no-change forecast model.

Regardless of the definition chosen for the inequality coefficient, the numerator remains unchanged. It is the RMS error defined in equation (8).

The square of the RMS error can be decomposed into several terms, each reflecting a different type of error:

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2 = (\bar{\hat{Y}} - \bar{Y})^2 + (s_{\hat{Y}} - s_Y)^2 + 2(1-r)\,s_{\hat{Y}}s_Y \qquad (16)$$

where:

$$\bar{\hat{Y}} = \frac{1}{T}\sum_{t=1}^{T}\hat{Y}_t$$

$$\bar{Y} = \frac{1}{T}\sum_{t=1}^{T}Y_t$$

$$s_{\hat{Y}} = \sum_{t=1}^{T}(\hat{Y}_t - \bar{\hat{Y}})^2$$

$$s_Y = \sum_{t=1}^{T}(Y_t - \bar{Y})^2$$

$$r = \frac{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - \bar{\hat{Y}})(Y_t - \bar{Y})}{s_{\hat{Y}}s_Y}$$

The first term is zero only when the means of actual and predicted variables are equal. Errors that lead to a positive value for this term can be interpreted as a bias or central tendency error. The second term is zero only when standard deviations of actual and predicted variables are equal. A positive value for this term can be interpreted as error due to different variation. The third term is zero only when the correlation coefficient between predicted and actual values is one. Therefore, a positive value for this term can be interpreted as an error due to different covariation. To compare different model decompositions, one should convert the three components to proportional terms by dividing each by their sum:

$$U(\text{bias}) = \frac{(\bar{\hat{Y}} - \bar{Y})^2}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \qquad (17)$$

$$U(\text{variation}) = \frac{(s_{\hat{Y}} - s_Y)^2}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \qquad (18)$$

$$U(\text{covariation}) = \frac{2(1-r)s_{\hat{Y}}s_Y}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \qquad (19)$$

This implies that:

$$U(\text{bias}) + U(\text{variation}) + U(\text{covariation}) = 1. \qquad (20)$$

---

[5]The equivalence of the numerator is demonstrated as follows:

$$\sqrt{\frac{1}{T}\sum_{t=1}^{T}((\hat{Y}_t - Y_{t-1}) - (Y_t - Y_{t-1}))^2} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_{t-1} - Y_t + Y_{t-1})^2} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2}$$

One expects U(bias) to be low. If it is large, the average errors are large and considerable bias exists in the simulation. Even if the inequality coefficient cannot attain its optimum level of zero, the most desired level for U(bias) remains zero. One would like U(variation) to be low. If U(variation) is high, predicted and actual values have unequal standard deviations. This might suggest that the model structure (or equation) underlying the variable in question is misspecified. The expectations regarding U(covariation) differ. It is unlikely that any model can generate simulations that are perfectly correlated with actual outcomes; therefore, one cannot expect U(covariation) to be low. As simulations will not all be perfect, the goal should be the lowest inequality coefficient possible with a decomposition showing U(bias) and U(variation) approaching zero and U(covariation) approaching one. With this type of decomposition, systematic error is minimized.

An alternative decomposition of the square of the RMS error, that is:

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2 = (\overline{\hat{Y}} - \overline{Y})^2 + \quad (21)$$

$$(s_{\hat{Y}} - rs_Y)^2 + (1 - r^2)s_Y^2$$

can be evaluated in relation to the regression equation defined in equation (12). The first term is the same as that in equation (16). A perfect simulation generates a regression equation with zero intercept and unit slope:

$$Y_t = \hat{Y}_t + \xi_t. \quad (22)$$

Because ξ_t has zero mean by definition, Y must equal $\hat{Y}$ and the first term of the decomposition becomes zero. Furthermore, the regression slope in equation (12) can be defined as:

$$\beta_1 = \frac{\sum_{t=1}^{T}(\hat{Y}_t - \overline{\hat{Y}})(Y_t - \overline{Y})}{\sum_{t=1}^{T}(\hat{Y}_t - \overline{\hat{Y}})^2} = \frac{rs_Y}{s_{\hat{Y}}} \quad (23)$$

For a perfect simulation, β_1 equals one and this second decomposition term also becomes zero. These three components can also be converted to proportional terms:

$$U(\text{mean}) = U(\text{bias}) = \frac{(\overline{\hat{Y}} - \overline{Y})^2}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \quad (24)$$

$$U(\text{regression}) = \frac{(s_Y - rs_Y)^2}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \quad (25)$$

$$U(\text{residual}) = \frac{(1 - r^2)s_Y^2}{\frac{1}{T}\sum_{t=1}^{T}(\hat{Y}_t - Y_t)^2} \quad (26)$$

with:

$$U(\text{bias}) + U(\text{regression}) + U(\text{residual}) = 1. \quad (27)$$

The objective is to generate a model with the lowest inequality coefficient possible for each variable; the decomposition should show U(bias) and U(regression) approaching zero and U(residual) approaching one. In fact, if the two decomposition terms differ significantly from zero, a linear correction factor[6] can be applied that will generate the desired decomposition.

Turning Point Errors

Another important goodness-of-fit measure is how well actual turning points are simulated during the historical period. Turning points are important because many economic time series exhibit positive serial correlation. For a model to be superior to a simple time trends model, it must predict turning points.

A simulation, with respect to turning points, has four possible outcomes: A turning point will actually exist and the model will either predict or not predict it; or no turning point will exist and the model will either predict or not predict one. These four possibilities are illustrated in the following diagram:

[6] The optimal linear correction factor to Y_t will be of the form $\hat{\beta}_0 + \hat{\beta}_1\hat{Y}_t$ where $\hat{\beta}_1 = rs_Y/s_{\hat{Y}}$ and $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{\hat{Y}}$.

ECONOMETRICS ANONYMOUS.
One of the major trends of the past decade has been the proliferation of redundant and useless econometric models and analyses. This new professional body has been formed to enable an economist, when he feels the urge to run multiple regressions far into the night, to telephone a fellow member of E.A. who will come over and sit up with him until the desire to regress passes.

Leonard Silk
"New Remedies for Economists"

		Predicted	
		Turning Point	No Turning Point
Actual	Turning Point	f_{11}	f_{12}
	No Turning Point	f_{21}	f_{22}

Each cell represents the frequency of each alternative. Perfect turning point forecasting implies $f_{12} = f_{21} = 0$; that is, no turning point errors. If f_{12} or f_{21} are not equal to zero, turning point errors are occurring. Expressing these errors in proportional terms provides a measure of turning point error.

A turning point error can be defined as:

$$\text{TP error} = \frac{f_{12} + f_{21}}{f_{11} + f_{12} + f_{21} + f_{22}} \quad (28)$$

A measure of error due to turning points missed is:

$$\text{TP}_M \text{ error} = \frac{f_{12}}{f_{11} + f_{12}} \quad (29)$$

A measure of error due to falsely predicted turning points is:

$$\text{TP}_F \text{ error} = \frac{f_{21}}{f_{21} + f_{11}} \quad (30)$$

Each of these measures ranges between zero and one; small values indicate good turning point simulations.

NET TRADE MODEL VALIDATION

The world trade forecast modeling system under development in the International Economics Division (IED), ESCS, centers on net trade models. The net trade model accounts for the interaction among major trading countries by commodity. Each commodity model is a system of export supply and import demand functions, by country, that are solved simultaneously for net trade (exports and imports) and world price levels. The net trade functions are specified as functions of own price, other commodity prices, production, income, population, and other demand shifters. Net trade models for individual commodities are linked through cross price variables. I evaluate the wheat and coarse grain net trade models here. These models were developed to support the USDA world trade/ U.S. export outlook process (*6*).

Static Simulation Results

A static simulation of wheat and coarse grain net trade models was performed over 1964-75. The data base for equation specification is, in general, 1960-75. The EEC wheat threshold price series used in the EEC-6 wheat consumption equation starts in 1964. This limitation shortens the simulation period to 1964-75.

Table 1 presents a summary of several measures of goodness of fit for each variable. Generally, the mean absolute relative errors and root-mean-square percentage errors are low, the slope coefficients of the regressions of actual on predicted values are close to one, the R-square coefficients are high, and the Theil inequality coefficients are low.

All U_2 inequality coefficients are substantially below one; therefore, this model proved to be significantly better than a no change forecast model. A decomposition of the inequality coefficients shows that errors are due primarily to differences in the covariation between actual and predicted values. All U(covariation) and U(residual) terms approach one and other components approach zero. Turning points for all variables are, in general, forecasted accurately.

In terms of the key variables concept suggested by Wallace, these models were developed to forecast U.S. agricultural commodity trade within the context of an integrated world agricultural commodity trade model. Therefore, U.S. exports and world price level forecasts become key variables. The MARE's for U.S. wheat (USWHEX) and coarse grain (USCGRX) exports are 4.7 and 4.6 percent, respectively. The corresponding RMS percentage errors are 5.9 and 5.5 percent. The MARE's for U.S. wheat (PXWHEUS) and corn (PXCORUSG) prices are 9.4 and 4.2 percent, respectively. The corresponding price RMS percentage errors are 12.0 and 5.3 percent.

TYPICAL RESULTS ARE SHOWN . . . the best results are shown.

CORRECT WITHIN AN ORDER OF MAGNITUDE . . . wrong.

Table 1 — Summary of goodness-of-fit statistics for each endogenous variable, 1964-75 static simulation

Model and variable[1]	Mean[2]	MARE	RMS percentage error	β_1	R^2	Theil's inequality coefficients			TP error
						U	U_1	U_2	
Wheat net trade model:									
ARWHEXC	2787	12.5	14.0	1.12	0.95	0.06	0.17	0.33	0.08
AUWHEXC	7109	6.6	7.4	1.08	.88	.04	.12	.25	.08
BRWHEM	2439	5.1	6.5	1.03	.93	.03	.08	.16	.08
CAWHECON	4522	1.9	2.2	.85	.87	.01	.27	.57	.33
CAWHEEK	15151	9.4	10.5	1.00	.93	.05	.24	.50	0
CAWHEX	11991	12.1	13.7	.82	.51	.07	.34	.71	.08
DEWHENM	1076	20.8	24.4	.90	.84	.11	.14	.26	.33
DEWHESTK	2493	13.8	16.6	.71	.53	.08	.27	.50	.17
EGWHENM	2682	3.6	4.1	1.00	.96	.02	.17	.33	.08
E6WHECON	29855	2.1	2.7	1.00	.75	.01	.35	.62	.25
E6WHEEK	5615	18.2	21.6	.49	.38	.10	.38	.76	.25
E6WHEMW	2620	11.7	14.1	1.03	.82	.07	.17	.32	.08
E6WHEXW	4876	4.5	5.7	1.09	.96	.03	.10	.19	.08
FRWHENX	5524	11.0	12.8	1.15	.91	.06	.18	.35	.17
INWHENM	4527	4.1	5.8	1.01	.99	.03	.02	.04	0
IRWHENM	483	24.1	28.4	1.01	.92	.10	.05	.09	.08
JPWHEM	4661	2.4	3.6	.90	.96	.02	.27	.52	0
KRWHEM	1288	6.8	8.4	1.09	.96	.04	.27	.57	0
LAWHENMC	1940	6.3	8.3	.97	.91	.04	.28	.52	.08
NAWHEM	1839	8.3	10.2	1.00	.94	.04	.16	.34	.25
PKWHENM	1317	11.7	14.6	.95	.77	.07	.20	.39	.17
PXWHEUS	2.34	9.4	12.0	.97	.94	.05	.24	.50	.17
RWWHENM	17155	3.7	4.8	1.03	.96	.02	.09	.19	.17
UKWHEM	4205	6.8	7.8	.88	.76	.04	.19	.39	.08
USWHEX	22776	4.7	5.9	1.04	.95	.03	.11	.21	.08
Coarse grain net trade model:									
ARCGRNX	5942	8.2	9.4	0.95	0.87	0.05	0.15	0.28	0.17
AUCGRNX	1599	13.3	15.7	1.04	.94	.07	.19	.41	.17
DECGRNM	4374	7.1	9.0	.71	.63	.04	.20	.40	.08
ESCGRNM	2891	12.2	15.1	1.09	.70	.07	.28	.52	.25
FRCGRNX	5167	9.6	12.4	1.05	.91	.06	.21	.43	0
ITCGRNM	5780	4.4	5.5	.84	.76	.03	.18	.35	.17
JPCGRNM	9773	5.1	6.5	.99	.95	.03	.25	.47	.08
PXCORUSG	1.79	4.2	5.3	1.00	.98	.02	.13	.24	.08
RWCGRNM	14265	6.8	8.5	1.00	.96	.04	.24	.46	.25
SACGRCON	5753	4.4	6.0	.82	.83	.03	.53	.80	.25
SACGRNX	1926	18.2	21.8	1.01	.89	.09	.07	.13	0
SVCGRNM	2166	20.5	29.8	.99	.98	.06	.07	.13	.33
THCGRX	1652	5.0	6.0	.95	.97	.03	.10	.20	.17
UKCGRNM	3710	5.6	7.1	.93	.53	.04	.34	.61	.17
USCGRX	26672	4.6	5.5	1.00	.98	.03	.12	.23	0

[1] Variables are defined in table 8.
[2] Quantities are 1,000 metric tons except for PXWHEUS and PXCORUSG which are in dollars per bushel.

THREE OF THE SAMPLES WERE CHOSEN FOR DETAILED STUDY . . . the results on the others didn't make sense and were ignored.

INTUITIVELY OBVIOUS . . . I don't understand it either.

Oral tradition.

Table 2 presents goodness-of-fit measures for the overall model. These measures are for the integrated wheat and coarse grain model and for both subcomponents of that model separately. To calculate the slope coefficient, R-square, Theil inequality coefficients, and turning point relative errors, I assume that the responses of the several variables can be combined and treated as the response of a single variable. This approach seems inappropriate for the MARE and RMS percentage error. For these measures, a simple average of respective individual endogenous variable measures is reported. Thus, the static simulation of the wheat and coarse grain model generally exhibits an 8.8-percent average MARE and a 10.9-percent average RMS percentage error. The regression slope coefficient, R-square, and Theil inequality coefficient, including its decomposition, all indicate unbiased forecasts and explain a major proportion of the variation in the actual variables throughout the historical period of fit.

Dynamic Simulation Results

A dynamic simulation of the wheat and coarse grain net trade model was performed over 1964-75. Table 3 presents goodness-of-fit measures for each variable. These results are similar to those from the static simulation. Some variables perform slightly worse, and others, slightly better.

Table 2 — Goodness-of-fit statistics, wheat and coarse grain net trade models, 1964-75 static simulation

Goodness of fit statistic	Wheat and coarse grain net trade model	Wheat net trade model component	Coarse grain net trade model component
Average MARE	8.8	8.9	8.6
Average RMS percentage error	10.9	10.6	11.4
β_1	1.00	1.00	1.01
R^2	.99	.99	.99
Theil inequality coefficients:			
U	.03	.03	.03
U_1	.05	.05	.09
U_2	.09	.09	.17
U(covariation)	.9973	.9987	.9831
U(residual)	.9999	.9985	.9910
Turning point errors:			
TP	.13	.13	.14
TP_M	.12	.10	.14
TP_F	.14	.13	.14

The U_2 inequality coefficient for South African coarse grain consumption (SACGRCON) exceeds one. This signifies that the dynamic forecast for this variable is significantly worse than a simple no-change model forecast. This poor performance is easily explained. South African consumption in the model is a function of lagged consumption. Between 1963 and 1964, actual coarse grain consumption increased 30 percent. Therefore, starting the simulation in 1964 creates a large forecast error, one carried through all periods of the simulation. The static simulation does not have this large error as the 1965 forecast depends on actual consumption levels in 1964 rather than on levels forecast for 1964. As South African coarse grain net exports (SACGRNX) are a function of consumption, these large consumption errors generate large net export errors. These errors would have been substantially reduced had any year other than 1964 been chosen for starting the dynamic simulation.

Korean wheat imports (KRWHEM) exhibit a U_2 inequality coefficient close to one. This simulation forecasts a more rapid rise in Korean imports throughout the mid sixties than actually occurred. Because actual growth was slow in the earlier years of the simulation, a no-change forecast, on the average, would have proved more accurate. However, imports did double during the sixties and the model picked up this phenomenon. The other goodness-of-fit measures indicate that the Korean wheat import simulation is satisfactory.

Table 3 — Goodness-of-fit statistics for each endogenous variable, 1964-75 dynamic simulation

Model and variable[1]	MARE	RMS percentage error	β_1	R^2	Theil's inequality coefficients			TP error
					U	U_1	U_2	
Wheat net trade model:								
ARWHEXC	12.5	14.0	1.12	0.95	0.06	0.17	0.33	0.08
AUWHEXC	6.7	7.8	1.20	.89	.04	.12	.23	.08
BRWHEM	5.4	6.8	1.03	.92	.03	.08	.16	0
CAWHECON	1.9	2.3	.77	.88	.01	.30	.63	.08
CAWHEEK	7.7	10.3	1.02	.94	.05	.23	.52	0
CAWHEX	13.0	14.4	.77	.47	.07	.36	.75	.17
DEWHENM	20.9	23.7	1.00	.84	.10	.12	.23	.33
DEWHESTK	12.2	16.5	.74	.51	.08	.25	.47	.08
EGWHENM	3.7	4.2	1.01	.96	.02	.18	.34	.08
E6WHECON	2.1	2.7	1.00	.75	.01	.35	.62	.25
E6WHEEK	14.4	19.2	.58	.41	.09	.30	.59	.25
E6WHEMW	12.5	15.3	1.07	.79	.07	.17	.33	.25
E6WHEXW	12.4	15.4	.92	.68	.07	.25	.44	.17
FRWEENX	9.8	13.3	1.12	.90	.06	.16	.30	.08
INWHENM	4.3	5.9	1.03	.99	.03	.03	.06	0
IRWHENM	29.4	36.1	1.01	.87	.13	.15	.27	.08
JPWHEM	3.0	4.0	.87	.96	.02	.29	.57	0
KRWHEM	8.6	11.4	1.11	.95	.05	.35	.92	0
LAWHENMC	6.5	8.6	.94	.91	.04	.30	.57	.08
NAWHEM	8.3	10.2	1.00	.94	.04	.16	.34	.25
PKWHENM	14.2	17.3	.89	.71	.09	.28	.52	0
PXWHEUS	13.2	15.3	1.06	.90	.07	.30	.62	.17
RWWHENM	5.8	7.0	1.12	.93	.03	.15	.32	.17
UKWHEM	7.5	8.4	.94	.71	.04	.22	.42	.17
USWHEX	5.4	6.5	1.10	.95	.03	.13	.23	.08
Coarse grain net trade model:								
ARCGRNX	9.6	10.8	0.98	0.82	0.05	0.17	0.31	0.08
AUCGRNX	12.5	14.9	1.06	.95	.06	.17	.35	.08
DECGRNM	7.1	9.0	.71	.63	.04	.20	.40	.08
ESCGRNM	12.1	14.8	1.10	.72	.07	.27	.51	.25
FRCGRNM	11.8	15.7	1.02	.86	.07	.22	.42	0
ITCGRNM	4.4	5.5	.83	.77	.03	.17	.33	.25
JPCGRNM	5.1	6.5	.99	.95	.03	.25	.47	.08
PXCORUSG	5.9	7.2	1.02	.97	.03	.18	.32	.17
RWCGRNM	8.6	11.3	.98	.94	.05	.31	.62	.33
SACGRCON	18.8	20.1	1.81	.88	.11	.78	1.99	.75
SACGRNX	86.7	93.2	.71	.86	.29	.33	.94	.25
SVCGRNM	25.9	32.4	.93	.99	.06	.11	.22	.25
THCGRX	5.0	6.0	.95	.97	.03	.10	.20	.17
UKCGRNM	5.7	7.3	.91	.51	.04	.35	.62	.25
USCGRX	7.7	8.9	1.00	.97	.04	.20	.38	.08

[1] Variables are defined in table 8.

The MARE's for U.S. wheat (USWHEX) and coarse grain (USCGRX) exports are 5.4 and 7.7 percent, respectively. The corresponding RMS percentage errors are 6.5 and 8.9 percent. The MARE's for U.S. wheat (PXWHEUS) and corn (PXCORUSG) prices are 13.2 and 5.9 percent, respectively. The corresponding price RMS percentage errors are 15.3 and 7.2 percent. Relative to the static simulation, these errors are larger, especially for U.S. coarse grain exports and wheat prices.

Table 4 presents several goodness-of-fit measures for the complete model. Generally, the dynamic simulation of the wheat and coarse grain model exhibits an 11.7-percent average MARE and a 17.5-percent average RMS percentage error. The other measures indicate that the dynamic model forecasts are unbiased and explain a major proportion of actual variation throughout the historical period of fit.[7]

Table 4 — Goodness-of-fit statistics, wheat and coarse grain net trade models, 1964-75 dynamic simulation

Goodness-of-fit statistic	Wheat and coarse grain net trade model	Wheat net trade model component	Coarse grain net trade model component
Average MARE	11.7	9.7	15.1
Average RMS percentage error	17.5	17.4	17.6
β_1	1.01	1.00	1.04
R^2	.99	.99	.98
Theil inequality coefficients:			
U	.04	.04	.05
U_1	.15	.15	.26
U_2	.27	.27	.62
U(covariation)	.9795	.9967	.8854
U(residual)	.9924	.9945	.9234
Turning point errors:			
TP	.15	.12	.21
TP_M	.12	.09	.16
TP_F	.16	.13	.22

Cross Simulation Comparison

The above evidence focuses on goodness of fit of a particular model simulation. Table 5 compares the goodness-of-fit measures across the three kinds of simulations discussed. Only two measures are broadly comparable across the residual check, static, and dynamic simulation: the R-square and the coefficient of variation. The R-square for the residual check simulation is derived from the ordinary least squares estimation procedure. The R-square for the static and dynamic simulation comes from the regression of actual on predicted values. In all three cases, I derive the coefficient of variation by dividing the standard error of the regression by the mean of the dependent variable. The residual check simulation measures in table 5 are not completely comparable to the static and dynamic simulations. The regression equations are generally fitted over slightly longer time periods. The data, however, broadly indicate behavior across simulations. The residual check generally performs somewhat better than the static simulation, and the static simulation performs somewhat better than the dynamic simulation. These results are expected; each simulation allows for an additional source of errors.

VALIDATION THROUGH COMPARISON WITH ALTERNATIVE MODEL SPECIFICATION

Validation questions essentially refer to a model's goodness of fit. The validation results discussed earlier are absolute measures of this. They all measure the degree of goodness of fit relative to an ideal: the perfect forecast model.

Another way to validate a model is to compare its specification with those of other models. Two types of simple models are good candidates for comparison. The first is the no-

[7] The high U_2 inequality coefficient for the coarse grain component is the result of the errors for South Africa explained previously.

Table 5 — Three simulations of wheat and coarse grain net trade models

Model and variable[1]	R-square			Coefficient of variation		
	Residual check	Static simulation	Dynamic simulation	Residual check	Static simulation	Dynamic simulation
Wheat net trade model:						
ARWHEXC	0.85	0.95	0.95	25	15	15
AUWHEXC	.80	.88	.89	11	8	9
BRWHEM	.88	.93	.92	9	7	7
CAWHECON	.84	.87	.88	3	2	3
CAWHEEK	.99	.93	.94	4	11	11
CAWHEX[2]		.51	.47		15	16
DEWHENM	.90	.84	.84	23	27	26
DEWHESTK	.76	.53	.51	1	18	18
EGWHENM	.97	.96	.96	5	5	5
E6WHECON	.75	.75	.75	3	3	3
E6WHEEK[2]		.38	.41		24	21
E6WHEMW	.88	.82	.79	16	15	17
E6WHEXW	.96	.96	.68	10	6	17
FRWHENX	.94	.91	.90	15	14	15
INWHENM	.98	.99	.99	7	6	6
IRWHENM	.92	.92	.87	33	31	40
JPWHEM	.98	.96	.96	4	4	4
KRWHEM	.93	.96	.95	1	9	12
LAWHENMC	.95	.91	.91	9	9	9
NAWHEM	.93	.94	.94	16	11	11
PKWHENM	.90	.77	.71	11	16	19
PXWHEUS[2]		.94	.90		13	17
RWWHENM	.93	.96	.93	1	5	8
UKWHEM	.92	.76	.71	5	9	9
USWHEX	.92	.95	.95	9	6	7
Coarse grain net trade model:						
ARCGRNX	0.87	0.87	0.82	13	10	12
AUCGRNX	.93	.94	.95	20	17	16
DECGRNM	.82	.63	.63	11	10	10
ESCGRNM	.81	.70	.72	21	17	16
FRCGRNX	.94	.91	.86	17	14	17
ITCGRNM	.88	.76	.77	10	6	6
JPCGRNM	.98	.95	.95	1	7	7
PXCORUSG[2]		.98	.97		6	8
RWCGRNM	.96	.96	.94	11	9	12
SACGRCON	.91	.83	.88	0	7	22
SACGRNX	.93	.89	.86	1	24	102
SVCGRNM	.98	.98	.99	19	33	36
THCGRX	.97	.97	.97	1	7	7
UKCGRNM	.66	.53	.51	8	8	8
USCGRX	.97	.98	.97	8	6	10

[1] Variables are defined in table 8.

[2] No measures are available for the residual check simulation. These variables are not econometrically estimated but derived from identity equations and the simultaneous nature of the net trade model.

An economist can tell you what will happen under any conditions. And his guess is liable to be just as good as anybody else's.

Will Rogers

change model, which assumes that next year's forecast will be the same as this year's actual level. The second type of model is a simple time trends model, where each endogenous variable is estimated as a function of time only. If the net trade model is no better than these relatively simple models, it should be rejected as a useful forecasting tool.[8] However, the net trade model proved superior to both alternatives.

The Theil U_2 statistic for the net trade model provides a comparison to a no-change forecast model. Except for the special case of South African coarse grain consumption, all statistics are below 1.0.

The time trends model provides an interesting comparison because time trends are popular with forecasters. Equations in a trend model can take numerous forms, but for this analysis, I chose a linear time trend. With annual data for 1960-75, far too few observations were available for developing even moderately sophisticated equations. When evaluating the linear trend results, in all but a few cases which demonstrated relatively rapid rates of exponential growth, there appeared to be no advantage to using other functional forms.

Table 6 presents linear trends model results and table 7 shows summary statistics for the net trade and the linear trends models. The no-change model/linear trends model comparison can be evaluated solely based on a U_2 inequality coefficient. For 7 of 40 variables (CAWHEEK, INWHENM, KRWHEM, PXWHEUS, PXCORUSG, RWCGRNM, and UKCGRNM), the U_2 statistic exceeds 1.0 for the linear trends model, and the no-change forecast is superior. U.S. wheat and corn prices are two variables for which a forecast of no change is better than a linear trend forecast. The average U_2 inequality coefficient is 0.84 for wheat and 0.80 for coarse grain. Thus, the linear trends model is more accurate than the no-change model. Table 7 indicates that the static simulation of the net trade model is superior to the linear trends model.[9] A more conclusive evaluation involves the net trade model dynamic simulation as the basis for comparison. Dynamic simulation results put a model in the worst possible light. The net trade model again performs better than the linear trends model (table 7).

Rather naive trend models were used. Trend model results could probably be improved if a commodity analyst was careful in choosing the appropriate trend equation for each variable rather than routinely choosing the linear form for all the variables. When one examines the time series plots for all the endogenous variables in the net trade model, it seems unlikely that any forecast from a reasonably simple time series model will be better than the net trade model's. The net trade model also provides the core upon which a detailed, structural world trade modeling system can be built. As it provides equal or better forecasts as well as a structural model framework, the net trade model seems superior to any time series approach.

CONCLUSION

The only true test of validity involves using the net trade model in an actual forecasting environment. However, how well the model represents one's perception of reality, in both structural and historical tracking senses, provides some preliminary validation. Although no definitive conclusions based on statistical theory can be drawn from such an analysis, impressions gathered from it can increase one's confidence in the model.

---

[8] The net trade model structure itself may be rejected as a forecast tool but may still be accepted as a valid modeling construct. It can still support a world modeling system framework that, with better and more complete country-sector detail, provides both a better forecast and has explicit structural integrity.

[9] Regressing actual values on predicted values generated by a linear trend does not provide a basis for comparison. When the ordinary least squares procedure minimizes the sum of squared deviations from the equation, an intercept of zero and a slope of one is assured. For the same reason, the decomposition of the inequality coefficient provides little information.

Table 6 — Summary of goodness-of-fit statistics for each endogenous variable, linear trends model

Model and variable[1]	MARE	RMS percentage error	R^2	Theil's inequality coefficients		TP error	Coefficient of variation
				U_1	U_2		
Wheat net trade model:							
ARWHEXC	40.0	52.6	0.04	0.46	0.88	0.31	56
AUWHEXC	15.7	19.1	.15	.40	.66	.12	20
BRWHEM	16.8	21.8	.06	.44	.71	.31	23
CAWHECON	3.0	3.5	.76	.35	.62	.31	4
CAWHEEK	28.9	35.6	.02	.59	1.38	.44	38
CAWHEX	17.3	20.0	.04	.45	.83	.31	21
DEWHENM	38.5	51.7	.09	.41	.70	.25	55
DEWHESTK	14.8	19.6	.03	.41	.67	.31	21
EGWHENM	9.2	11.4	.84	.39	.71	.19	12
E6WHECON	2.7	3.8	.63	.54	.95	.38	4
E6WHEEK	15.2	19.9	.02	.42	.71	.25	21
E6WHEMW	21.7	25.1	.58	.40	.70	.67	27
E6WHEXW	17.6	22.3	.65	.42	.72	.06	24
FRWHENX	17.4	22.9	.82	.40	.71	.25	24
INWHENM	40.3	46.3	.01	.60	1.36	.30	49
IRWHENM	63.4	85.1	.28	.44	.82	.25	91
JPWHEM	4.4	5.7	.94	.32	.54	.12	6
KRWHEM	20.7	23.7	.74	.55	1.16	.19	25
LAWHENMC	12.8	16.0	.73	.43	.81	.12	17
NAWHEM	29.4	32.7	.55	.44	.80	.19	35
PKWHENM	23.0	26.7	.05	.40	.68	.25	29
PXWHEUS	29.3	34.8	.36	.55	1.24	.40	31
RWHENM	14.4	19.3	.26	.48	.94	.38	21
UKWHEM	9.6	12.4	.20	.43	.74	.25	13
USWHEX	16.3	19.8	.37	.47	.89	.19	21
Coarse grain net trade model:							
ARCGRNX	18.2	22.6	0.54	0.40	0.63	0.25	24
AUCGRNX	37.9	44.1	.55	.45	.88	.25	47
DECGRNM	16.3	19.8	.22	.36	.60	.25	21
ESCGRNM	22.6	24.7	.71	.46	.88	.25	26
FRCGRNM	23.5	29.1	.74	.40	.68	.06	31
ITCGRNM	14.4	17.1	.48	.49	.98	.06	18
JPCGRNM	5.4	7.1	.98	.28	.52	.06	8
PXCORUSG	21.9	27.2	.61	.55	1.21	.40	29
RWCGRNM	25.8	31.3	.49	.55	1.25	.12	33
SACGRCON	2.9	3.9	.96	.30	.53	.31	4
SACGRNX	50.8	54.9	.09	.42	.71	.19	59
SVCGRNM	202.7	285.5	.60	.46	.74	.19	305
THCGRX	10.9	17.9	.83	.34	.55	.12	19
UKCGRNM	8.1	10.3	.14	.50	1.00	.31	11
USCGRX	20.2	25.2	.68	.48	.91	.31	27

[1]Variables are defined in table 8.

Table 7 — Average goodness-of-fit statistics, alternative model comparison

Goodness-of-fit statistic	Net trade model						Linear trends model		
	Static simulation			Dynamic simulation					
	Wheat	Coarse grain	Wheat and coarse grain	Wheat	Coarse grain	Wheat and coarse grain	Wheat	Coarse grain	Wheat and coarse grain
MARE	8.9	8.6	8.8	9.7	15.1	11.7	20.9	32.1	25.1
RMS percentage error	10.6	11.4	10.9	17.4	17.6	17.5	26.1	41.4	31.8
U_1	.19	.20	.20	.22	.25	.22	.45	.43	.44
U_2	.36	.38	.38	.43	.54	.47	.84	.80	.83
TP	.13	.21	.13	.12	.21	.15	.28	.24	.27

Table 8 — Definitions of endogeneous variables in net trade model

First Field

AF	Africa
AR	Argentina
AU	Australia
BR	Brazil
CA	Canada
DE	West Germany
EG	Egypt
ES	Spain
E6	European Economic Community-6
FR	France
IN	India
IR	Iran
IT	Italy
JP	Japan
KR	Korea
LA	Other South America
NA	North Africa
PK	Pakistan
RW	Rest of the world
SA	South Africa
SV	U.S.S.R.
TH	Thailand
UK	United Kingdom
US	United States

Second field

CGR	Coarse grain
COR	Corn
WHE	Wheat

Third field

CON	Consumption
EK	Ending stocks
M	Imports
MW	Imports excluding intraregional trade
NM	Net imports
NMC	Net imports, local crop year
NX	Net exports
STK	Stocks
X	Exports
XC	Exports, local crop year
XW	Exports, excluding intraregional trade

REFERENCES

(*1*) Cohen, Kalman J., and Richard M. Cyert. "Computer Models in Dynamic Economics." *Quar. J. Econ.* Vol. 75, Feb. 1961, pp. 112-127.

(*2*) Dhrymes, Phoebus J., E. Philip Howrey, Saul H. Hymans, Jan Kmenta, Edward E. Leamer, Richard E. Quandt, James B. Ramsey, Harold T. Shapiro, and Victor Zarnowitz. "Criteria for Evaluation of Econometric Models." *Annals of Econ. and Social Measurement*, 1/3 (1972), pp. 291-324.

(*3*) Howrey, E. Philip, Lawrence R. Klein, and Michael D. McCarthy. "Notes on Testing the Predictive Performance of Econometric Models." *International Econ. Rev.* Vol. 15, No. 2, June 1974, pp. 366-383.

(*4*) Johnson, Keith N., and Lawrence R. Klein. "Error Analysis of the LINK Model." Project LINK Working Paper Number 7, LINK Central, Econ. Res. Unit, Univ. Pennsylvania, Philadelphia, Oct. 1975.

(*5*) Kost, William E. "Validation Criteria for Models with More than One Endogenous Variable." Unpublished paper, May 1973.

(*6*) ________, Martin Schwartz, and Anthony Burris. "FDCD World Trade Forecast Modeling System." FDCD Working Paper, Econ., Stat., Coop., Serv., U.S. Dept. Agr., Aug. 1979.

(7) Labys, Walter C. *Dynamic Commodity Models: Specification, Estimation, and Simulation*. Lexington Books, D.C. Heath and Company, Lexington, Mass., 1973, pp. 215-240.

(*8*) Leuthold, Raymond M. "On the Use of Theil's Inequality Coefficients." *Am. J. Agr. Econ.*, Vol. 57, No. 2, May 1975, pp. 344-346.

(*9*) Pindyck, Robert S., and Daniel L. Rubinfeld. *Econometric Models and Economic Forecasts*. McGraw-Hill, New York, 1976, pp. 308-335.

(*10*) Theil, Henri. *Applied Economic Forecasting*. Rand-McNally, Chicago, 1966, pp. 1-36.

(*11*) ________. *Economic Forecasts and Policy*. North-Holland Publishing Company, Amsterdam, 1961, pp. 1-48.

(*12*) Wallace, Thomas D. "The General Problem of Spatial Equilibrium: A Methodological Issue." *Interregional Competition Research Methods* (ed. Richard A. King). Agr. Policy Inst. Series 10. School Agr. and Life Sci., N.C. State Univ., Raleigh, 1963.

MODEL VALIDATION AND THE PHILIPPINE PROGRAMMING MODEL

By Gil R. Rodriguez, Jr., and David E. Kunkel*

This research demonstrates the need and the procedure for testing sector programming models. It compares the model estimates of endogenous variables to carefully selected base period parameters. It uses an operational, static, deterministic, and highly aggregate programming model of Philippine agriculture as the framework. Alternative formulations of the Philippine model are also examined for possible errors in the consumption, production, and objective function data sets.

Keywords

Mathematical programming
Model validation
Philippines
Agricultural sector analysis
Development planning

The use of programming models to analyze the economic implications of supply and demand shifts for the agricultural sector of developing countries has increased significantly. Notable models are those by Duloy and Norton (*4*); Pomareda (*16*); Cappi, Fletcher, and others (*3*); Miller and others (*14*); and Heady (*9*).[1] These models use an objective function that incorporates supply and demand functions to simulate competitive market equilibrium.

Despite the substantial investment in technical skills and data-processing inputs, validation of sector programming models is rarely discussed explicitly. Nugent was the first analyst to test the reliability of programming models (*15*). His work can be summarized in two propositions:

1. If a market in the real world approximates a competitive condition closely, any deviation—for that market—of the results of a programming model from an existing observable empirical data base represents model specification errors.
2. If the programming model simulates a competitive market solution, but the real world situation being modeled has market imperfections, then these imperfection are likely responsible for some deficiencies in the predictive ability of the programming framework.

Recent works by Duloy and Norton (*4*) and Kutcher (*13*) have employed both propositions to validate the Mexican agricultural sector (CHAC) model. Later Shumway and Talpez (*17*) concentrated on the first proposition when examining the output of a model of major crops in California.

In this article, we use Nugent's first proposition to validate the optimal levels of production, exports, imports, and the shadow prices of commodities and resources. It is our principal objective to illustrate the validity tests conducted on a programming model of the Philippine agricultural sector known as MAAGAP.[2]

THE STRUCTURE OF THE PHILIPPINE MODEL

MAAGAP, a highly aggregate, static, and deterministic model, includes rice, corn, sugar, coconuts, vegetables, and livestock products that collectively accounted for about 90 percent of the total value of Philippine agricultural commodities in 1976. Detailed discussion of the data set appears in Gonzalez and Kunkel (*9*). MAAGAP was developed in Project ADAM (Agricultural Diversification and Markets) with the assistance of both Filipino and U.S. agricultural economists.

The MAAGAP model forms an important part of the agricultural policy analysis system within the Philippine Bureau of Agricultural Economics. The model has been used for several policy analyses, such as the fertilizer subsidy analysis and the evaluation of supply and demand

*Gil R. Rodriguez, Jr. is a Senior Economist in the Bureau of Agricultural Economics, Philippine Ministry of Agriculture. David E. Kunkel was formerly an ESCS resident consultant who is with the Foreign Agricultural Service, U.S. Department of Agriculture (USDA). The research reported was funded under Project ADAM, a joint USDA-Bureau of Agricultural Economics undertaking. Funding was provided by the Agency for International Development, the Philippine National Science and Development Board, The Philippine Council for Agricultural Research, and the Republic of Philippines Ministry of Agriculture. The authors wish to acknowledge comments from Jerry A. Sharples, Clark Edwards, and Mark Rosegrant. The views expressed here do not necessarily reflect those of the Philippine Ministry of Agriculture or USDA.

[1]Italicized numbers in parentheses refer to items in References at the end of this article.

[2]MAAGAP, a Filipino word which means alert, stands for Model Analysis of Agricultural Adjustments in the Philippines.

projections estimated by the National Economic Development Authority (NEDA). The most important papers published during the model development were by: Atkinson and Kunkel (*1*): Kunkel, Gonzalez, and Hiwatig (*11*); Gonzalez, Kunkel, and Alix (*8*); Ferrer (*6*); Atkinson and Kunkel (*2*); Foote (*7*); and Encarnacion (*5*). The objective function for the MAAGAP model is:

$$\max(Z) = \sum_j \left[\int_o^{C_j^i} P_j dC_j \right] + \sum_j v_j E_j - \sum_j u_j I_j - \sum_n c_n X_n - \sum_k w_k R_k - \sum_t f_t F_t - \sum_j g_j 0_j - \sum_m b_m M_m \qquad (1)$$

where the variables for this model are defined in table 1.

Equation (1) simply sums the areas under the demand curves and contains the value of exports minus the costs of imports, incidental production items, input supply, and feed-mixing processing. The objective function simulates a competitive market by using stepped demand functions.

The step demand functions are formed by grid linearization of:

$$C_j = \int_o^{C_j^i} P_j dC_j;$$

Thus, the ith step is the sum of the area under the demand curve up to the quantity C_j^i. The convex combination constraint allows only the corresponding quantity (C_j^i) to be sold. In the optimum solution the shadow price is:

$$P^* = \frac{\Delta \int P_j^* dC_j}{\Delta C_j},$$

Table 1 — Variables in the Philippine Programming Models

Endogenous		
P_j	=	f (C_jY) is the inverse demand function for the jth final product
C_j	=	domestic consumption of the jth product
E_j	=	quantity of the jth product exported
I_j	=	amount of the jth commodity imported
X_n	=	production levels of the nth production activity
R_k	=	amount supplied of the kth input
F_t	=	amount of the tth feed ration supplied
O_j	=	activity level of the jth final product transferred
M_m	=	activity level of the mth processing activity
$\Pi_{\ell j}$	=	shadow prices of the ℓth absolute land class used in production of the jth product
*	=	indicates equilibrium value
Exogenous		
Y	=	income level
v_j	=	export price of the jth product
u_j	=	import price of the jth commodity
w_k	=	input cost of the kth input supplying activity
f_t	=	unit cost of the tth feed-mixing activity
g_j	=	unit marketing margin of the jth final product
b_m	=	unit processing cost for the mth processing activity
c_n	=	miscellaneous cost of the nth production activity

All input-output coefficients are positive.

Thus, only one point on each demand, supply, or transformation function is validated.

where $\Delta \quad P_j dC_j$ is the change in the value of the objective function between steps, and ΔC_j is the change in the quantity demanded. Thus, by this formulation of the programming problem, *marginal price (shadow price) of output is equal to the average price*, or the intersection of the supply and the demand curve in a competitive market solution. Kunkel, Gonzalez, and Hiwatig (*11*) and Norton (*4*) provide illustrative examples.

Such an objective function implies the following individual behavioral assumptions:

1. Farmers are technically efficient and governed by profit-maximizing behavior.
2. Farmers are pricetakers in the input and commodity markets. The income variable appears in the demand function (P_j); income shifts are considered exogenous.[3] Because the Philippines is generally a price-taker in international markets, export (v_j) and import (u_j) prices are taken as given.

The product price function (P_j) does not contain any cross-price elasticity terms. They can easily be included through aggregation of commodities into composite groups. The formulation used allows substitution possibilities within a group but not across groups. Solutions which allowed substitution in the consumption set were found in computer runs not to be significantly different from ones that did not.

[3]The model does not capture the income impacts on the farmers' and the other sectors' expenditure patterns within a finite time period.

The objective function is maximized subject to a set of constraints defining production, processing, and marketing. These constraints are reported by Kunkel and others (*3*).[4] The model structure is shown in figure 1. Programming models for policy analysis are sensitive (particularly the shadow prices of fixed resources) to specification and measurement errors. The use of a programming framework imposes strong conditions on variables in the optimum solution. If the actual model specification used differs from the theoretical specifications of a perfectly competitive model, equilibrium shadow prices may be biased throughout. As shown by Kunkel, Gonzalez, and Hiwatig, the marginal revenue product of all resources used in each production process is equal to resource cost (*11*, p.6). Mathematically, this can be expressed as:

$$\lambda_{ijk} Q'_{ijk} = \gamma_j \qquad (2)$$

where λ_{ik} is the marginal cost of the *i*th product from the *k*th production process, Q'_{ijk} is the first derivative for the *k*th production process of the *i*th product and *j*th resource, and γ_j is the input cost of the *j*th resource.

Whenever price or quantities of inputs supplied or products demanded are fixed *a priori*, then neither $\lambda_{ik} = P_i^*$ nor $\gamma_j = w_j^*$ will hold.[5] For example, when resource levels are fixed, as with land in MAAGAP, the shadow price of land may differ from the actual competitive market price. To help detect any biases due to measurement and specification errors, the analyst must validate the model against a base period.

[4]This report is available from the authors on request. Overall, the MAAGAP model contained 158 rows and 504 columns (activities) for the 1976 base.

VALIDATING THE PHILIPPINE MODEL[6]

The validation procedure compared the MAAGAP results to actual base period levels for the set of endogenous variables (table 1). The base period, 1976, was chosen as being most representative of recent years. We are considering cross-section data and are not validating the ability of the model to capture turning points over a time path.

Thus, only one point on each demand, supply, or transformation function is validated.

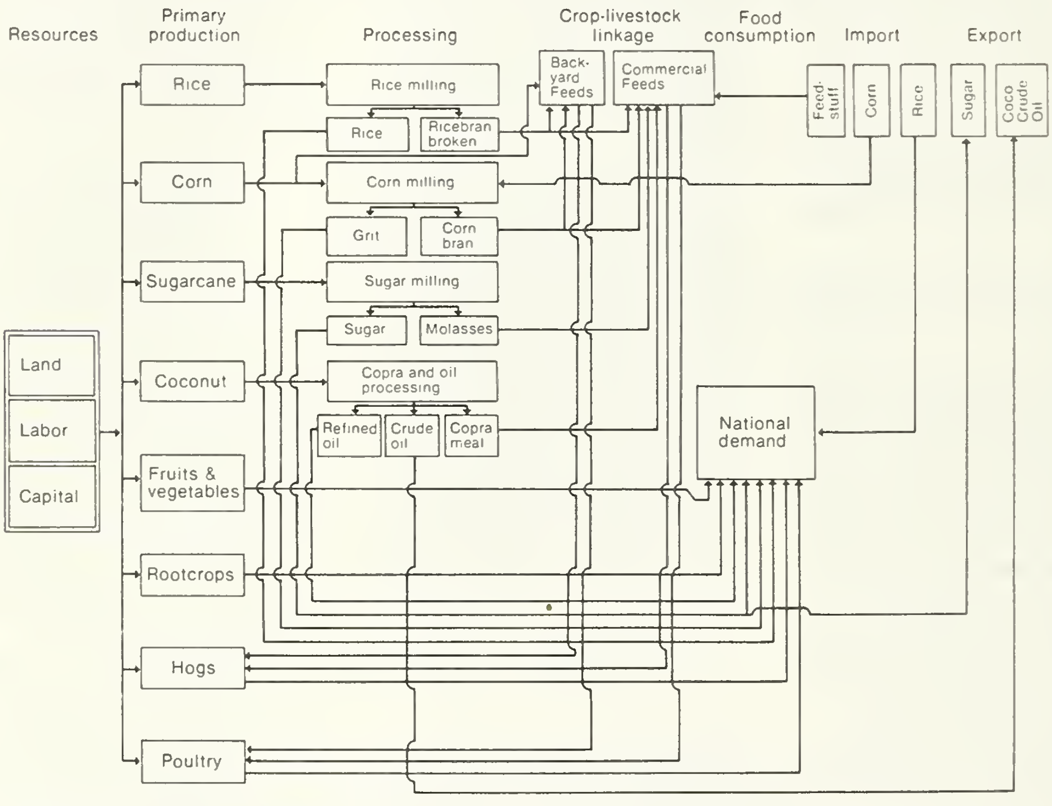
Data limitations made it difficult to determine some of the base period resource levels and prices ($R_k F_t$ and Π_{lj} in table 1). For example, land prices are acquisition costs, and the corresponding shadow prices are rates of return. To evade taxes, landowners usually undervalue prices on property not being sold. The major input which can be validated is the level of fertilizer use.

[5]Note that in this case P_j^* and w_j^* are market equilibrium prices on demand and input supply curves. See (*5*) for more detail.

[6]The validation tests performed in the Philippine model were influenced partly by the earlier work of Kutcher (*13*) on the consistency tests of the Mexico — Pacific Northwest Regional Model.

Figure 1

Schematic Diagram of the Adam National Model



The endogenous variables subjected to a close scrutiny were agricultural market price (P^*), imports (I_j), and production levels (X_m).

The first test involves a check on the production capacity (implicitly involving also the input-output coefficients) of the model. This is accomplished by treating final domestic commodity demand as perfectly elastic at fixed price levels. If a given commodity is partially or totally imported by the model (when, in fact, it is not imported in the base period), this implies an underestimate by the model of actual capacity. That is, the production vector may be too "expensive." The reverse holds true for "excessive" exports of a given commodity.

The second validation test entails redefining the model's objective function as the minimization of the costs of producing domestic output levels in the base period. This is accomplished by fixing the level of domestic and foreign demand for all products at base period levels. The shadow prices generated in the commodity balance equation are then marginal costs. To validate the model's assumption of a competitive structure, one then compares these with the base period market prices.

The third, and final test, compares the full model results with data for the base period. The models corresponding to these tests are called the

We found a high correlation between the crop prices estimated from the full model and the actual prices for 1976.

fixed demand, cost minimization, and full model.

In these validation tests, the following measures are used to judge how closely the model approximates the base period:

1. The correlation between the model-derived commodity outputs and prices and those observed for 1976.
2. A simple regression of the form,[7]

$$Y_o = a + b\, Y_m \qquad (3)$$

where Y_o is the observed value, and Y_m is the model-estimated value.

For this test, the model results and real world data on the various agricultural commodity outputs and prices will have a correlation of one (or, equivalently, the expected value E $(a) = 0$ and E $(b) = 1$ in equation (*10*) if the objective function, production, consumption, and constraints sets of the model are ideal).

3. The Information Inaccuracy Index, that is:

$$I(Y_o, Y_m) =$$

$$\sum_{i=1}^{n} Y_{io}\, [h(Y_{io}) - h(Y_{im})]$$

where Y_{io} =

$$Y_{io} / \sum_{i=1}^{n} Y_{io} \qquad Y_{io} > 0$$

$$Y_{im} =$$

$$Y_{im} / \sum_{i=1}^{n} Y_{im} \qquad Y_{im} > 0$$

$$h(Y_{io}) = \ell n_e(Y_{io})$$

$$h(Y_{im}) = \ell n_e(Y_{im})$$

The Information Inaccuracy Index was developed by Tilanus and Theil to evaluate the estimation errors of the endogenous variables in an input-output model. Their rationale in developing the index was that "... errors in less important variables are weighed less heavily than the same relative errors in the more important ones." A high value for the Information Inaccuracy Index (which does not have an upper and lower bound) indicates a deterioration of the estimation capability of the model. A perfect model would result in: $E(Y_o, Y_m) = 0$.

However, the critical value that separates "pass" from "fail" depends largely on the utility function of the researcher. A logical criterion which the researcher can use in selecting a critical value may depend on an awareness of the marginal returns from the model's improvement and the value of the marginal effort.

Crop production and price estimates from various alternative model formulations are compared with observed data in table 2. Table 3 gives the regression and correlation parameters used as indicators of goodness of fit for the linear model.[8] The linear regression results indicate two types of directional biases, as a and b are either less than or greater than 1.0. The first type (T_1 in figure 2) is that used for small values of the relevant base period data (Y_o); the model's estimations are biased upwards. The reverse is true for larger values of Y_o. The second type of bias (T_2 in figure 2) is one in which all the model's estimates are biased upward if the constant term is positive. As indicated by table 3, the full and fixed demand model's estimates of crop prices belong to the second bias type. This is not the case for the cost minimization model.

The full model's estimate of crop production is also T_2. However, the crop area and production estimates of the fixed demand model are of the first bias type. The latter type is also present in the full model's determined crop prices and in the cost minimization model's generated crop

[7]The regression form $\ell n Y_o = \ell n a + b \ell n Y_m$ was also estimated to determine nonlinear biases, but results were not significantly different from the linear case. A serious limitation arising from using equation (*10*) (or its log transform) is that formal statistical tests of significance cannot be applied to the regression parameters because the model estimates are not independent. Such parameters should merely be interpreted as informal measures of goodness of fit and model biases.

[8]Standard errors are given for informational purposes only and should not be used for formal statistical testing. See footnote 7.

Table 2 — Crop area, production, and prices of alternative model formulations

Crop	Actual base period	Area		Actual base period	Production		Actual base period	Prices		
		Full model	Fixed demand		Full model	Fixed demand		Full model	Fixed demand	Cost minimization
	--- *1,000 hectares* ---			--- *1,000 million tons* ---				--- *Pesos/kg.* ---		
Palay (rough rice)	3,579.3	4,198.0	4,173.2	6,159	6,705	6,710	0.94	1.01	1.09	1.18
Corn	3,257.0	3,169.3	3,144.0	2,767	3,119	2,960	.94	.52	.33	.50
Sugarcane	533.0	538.6	538.6	2,514	2,455	2,455	1.94	1.89	1.89	2.64
Coconut	2,521.2	2,387.0	2,387.0	10,662	8,619	1,730	1.85	[3]1.63	[3]1.63	[3]3.35
Banana	298.7	244.7	224.5	3,068	954	875	.41	.38	.38	.50
Cabbage	8.1	15.1	18.0	54	64	69	1.53	1.46	1.44	1.51
Pechay[1]	4.5	6.7	7.3	37	25	27	1.40	1.35	1.20	1.49
Tomatoes	21.0	21.9	26.6	153	79	96	2.04	1.53	1.56	1.63
Eggplant	16.2	26.5	32.1	82	99	120	.97	1.13	1.09	1.16
Camote[2]	192.3	196.0	243.3	781	687	745	.42	.64	.61	.61
Cassava[1]	118.0	150.5	163.2	621	464	503	.38	.33	.33	.38

[1] Leafy vegetable.
[2] Root vegetable.
[3] Copra equivalent price.

Note: 7.30 = $1.00

Source: Philippine Sugar Commission (PHILSUCOM), Bureau of Agricultural Economics (BAEcon), Philippine Coconut Authority (PCA).

prices. Judging from the standard errors for *b* and the correlation coefficient (*r*) given in table 3, the full model seems to perform better than the other model formulations for area and price but not for production.

The log linear regression results indicate the full and fixed demand model's estimates of crop areas have a nonlinear bias downward for small values of $\ell n Y_m$ and a nonlinear bias upward for large values.

We found a high correlation between the crop prices estimated from the full model and the actual prices for 1976. This supports the plausibility of assuming the competitive market structure of the Philippine model. However, for coconuts a large gap occurred in the cost minimization model (table 2). This gap can be attributed to data errors which had the following causes:

1. The conversion rate was overestimated; a rate of 4.5 per kilograms (kg.) of copra (coconut meat) was used.
2. The domestic coconut oil demand was overestimated by 65 percent due to a data error.
3. The coconut hectarage constraint was underestimated by 5.3 percent.
4. The export levels set up for coconut oil and copra may have been too high due to the absence of any stock adjustments.

The coconut data misspecification will likely affect the shadow prices of other major agricultural commodities, particularly sugarcane.

A general reason for the prices of the cost minimization model deviating from the actual prices is that, by dropping the first and second terms which allow market pricing of output from equation (1), we are utilizing the total model structure information less efficiently. Graphically, this means that if we disregard D_1 in figure 3, the probability of estimating the "true" market price (P_1) is low. If S_2 is the implicit supply function generated by the cost minimization model, the error in price estimation is the area abP_1P_2 (fig. 3).

The low model price for corn (0.50 pesos per kg.) compared with the observed price (0.94 pesos per kg.) can be attributed to a possible downward bias in the model's estimate of the cost of producing corn. The problem is partly caused by the

The full model predicts the crop production proportions well in comparison with the fixed demand framework.

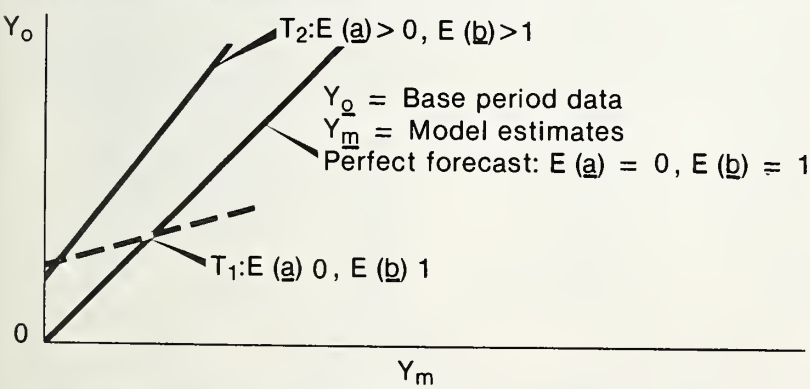
Table 3 — Regression results for actual versus model levels[1]

Result and model	a	b	r
Area:			
Full model	32.49	0.930 [2](.0373	0.9928
Fixed demand	1051.69	.941 (.45)	.5688
Production:			
Full model	148.80	1.086 (.0961)	.9665
Fixed demand	24.93	.938 (.109)	.9927
Prices:			
Full model	.00138	1.079 (.134)	.9370
Fixed demand	.0965	1.018 (.149)	.9157
Cost minimization	.4032	.56 (.120)	.8413

[1]Based on table 2.
[2]Numbers in parentheses are standard errors.

Figure 2

Illustration of Linear Directional Biases



difficulty of determining the appropriate spatial aspects of corn production vectors.

For the production capacity test, the fixed demand model solution registered 40,800 metric tons of commercial broiler imports. However, as no broilers were imported by the Philippines in 1976, the domestic commercial broiler production activities incorporated in the Philippine model may be too expensive; that implies an upward bias in the pricing of such activities. Comparing the export levels of coconut and sugar products with the base levels indicates an "over-capacity" for centrifugal sugar (1.720 million metric tons (mmt) versus 1.455 mmt) whereas the reverse holds for molasses (0.657 mmt versus 0.792 mmt) and copra meal (0.170 mmt versus 0.497 mmt).

Table 4 gives usage levels obtained from the model formulations. All three models overestimated the levels of fertilizer use, probably as a result of aggregation error because production vectors are based on farm survey data. However, the full model did well in predicting directions of change in fertilizer prices. The full and the fixed demand models performed better than the cost minimization model in predicting nitrogen and potash consumption in 1976.[9]

[9]Nitrogen is considered the most important fertilizer nutrient in the Philippines. Experiments conducted by the Bureau of Soils (BS), the Philippine Sugar Commission (PHILSUCOM), and the International Rice Research Institute (IRRI) show that most crops responded favorably compared with their reponse to potash and potassium.

Figure 3

Market Equilibrium in the Cost Minimization Model

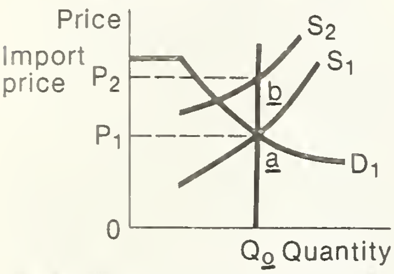


Table 4 — Fertilizer usage levels under alternative model assumption

Fertilizer	Model formulation			Actual 1976[1]
	Full model	Fixed demand	Cost minimization	
	1,000 metric tons			
Nitrogen (N)	178	180	195	152
Phosphorous (P_2O_5)	92	84	92	38
Potash (K_2O)	68	68	103	55

[1] Fertilizer and Pesticides Authority.

Although no formal level of significance can be attached to the information parameters provided in table 5, the full model predicts the crop production proportions well in comparison with the fixed demand framework. The full model incurs a relative information loss of −3.75 percent, compared with the fixed demand model's loss of −20.74 percent. All model types perform well in estimating the crop prices. The Information Inaccuracy Index (in absolute terms) ranges from 2.02 percent to 4.92 percent.

CONCLUSIONS

Programming models are rarely subjected to validation tests. In this article, we have shown that consistency checks on the shadow prices of programming solutions are useful because any misspecifications in resource constraints or prices will tend to affect shadow prices.

Tests of a programming model of Philippine agriculture revealed biases in the production, consumption, constraint, and objective function sets. Each model compared represented a unique theoretical structure from which to test for inconsistencies. The fixed demand model appeared to have significant biases in the crop production vectors for commercial broilers, corn, and copra meal. The wide disparity between the cost minimization model's coconut shadow prices and the actual price in 1976 helped to identify measurement errors.

The numerical measures used to judge how well a specific programming model approximate the Philippine agricultural conditions in 1976 were simple correlation coefficients, regression of actual versus model results, and the Information Inaccuracy Index. Based on these indices, the full model outperformed the others.

Validating resource usage and price levels of the MAAGAP model was limited to fertilizer use, a limitation dictated by the availability of the basic data. The three models' comparisons of the estimated fertilizer nutrients with actual 1976 levels indicate that all these models overestimated use. Nevertheless, the full and fixed model's yields of nitrogen and potash consumption levels were more accurate than those determined by the cost minimization model.

Table 5 — Information indices for evaluating the relative magnitude of estimation errors in the model types

Endogenous variables	Model type	Information Inaccuracy Index	Expected information content[1]	Relative information content[2]
Crop area	Full model	−0.01219	1.4811	0.82
Crop area	Fixed demand	−.0103	1.6654	3.75
Crop production	Full model	−.0624	1.6654	20.74
Crop production	Fixed demand	−.3454	1.6654	20.74
Crop prices	Full model	−.0202	2.2569	0.89
Crop prices	Fixed demand	−.0377	2.2569	1.67
Crop prices	Cost minimization	−.0492	2.2569	2.18

[1] Computed as:

$$H(Y_{io}) = \sum_{i=1}^{n} Y_{io} h(Y_{io})$$

[2] Computed as:

$$[I(Y_o Y_{ml}) H(Y_{io})] \times 100$$

REFERENCES

(*1*) Atkinson, L. J., and D. E. Kunkel. *High-Yielding Varieties of Rice in the Philippines, Progress of the Seed Fertilizer Revolution.* FAER-113. Econ. Res. Serv., U.S. Dept. Agr., Feb. 1976.

(*2*) Atienza, F. M., and D. E. Kunkel. "Determining the Economic Family Size Farm for Land Reform Areas," *Philippine J. of Agr. Econ. and Devt.* Vol. IV, No. 2, July 1974.

(*3*) Cappi, Condos, Leland Fletcher, and Roger Norton. "An International Trade Model for Some Latin American Countries," World Bank working paper, unpublished, 1976.

(*4*) Duloy, J. H., and R. D. Norton. "CHAC: A Programming Model for Mexican Agriculture," *Multi-Level Planning: Case Studies in Mexico*, (ed. L. Goreux and A. Manne). North Holland Publishing Company, Amsterdam, 1973.

(*5*) Encarnacion, J. "On Predicting Subsectoral Outputs and Prices Via an Optimizing Model." Discussion Paper No. 75-17, Univ. of the Philippines, School of Economics, Nov. 1975.

(*6*) Ferrer, J. B. "Alternative Methods in the Estimation of Demand for Selected Agricultural Products in the Philippines." Unpublished M. S. Thesis, Asian Social Institute, 1977.

(*7*) Foote R. "An Econometric Model for the Agricultural Sector of the Philippines, 1955-1969." *J. of Agr. Econ. and Devt.*, Los Banos, Philippines, Vol. VIII, No. 3, Nov. 1978, pp. 78-96.

(*8*) Gonzales, L. A., D. E. Kunkel, and J. C. Alix. "Selected Agricultural Policy Issues and the MAAGAP Model 1975-1977, *Agr. Econ. and Devt.*, Los Banos, Philippines, Vol. VIII, No. 3, Nov. 1978, pp. 26-53.

(*9*) Gonzales, Leonardo A., and D. E. Kunkel. "Data Base and Technology Set of the MAAGAP Model." *J. of Agr. Econ. and Devt.*, Los Banos, Philippines, VIII, No. 3, Nov. 1978, pp. 122-169.

(*10*) Heady, Earl. "Annual Report: Agriculture Sector Analysis in Thailand." Center for Agr. and Rural Devt., Feb. 1977.

(*11*) Kunkel, D. E., L. A. Gonzales, and M. Hiwatig. "Application of Mathematical Programming Models Simulating Competitive Market," *International Agr. Econ. Assoc., Occasional Paper* No. 1, Nov. 1977, pp. 3-19.

(*12*) Kunkel, David E., G. R. Rodriguez Jr., L. A. Gonzales, and J. C. Alix. "Theory, Structure and Validated Empirical Performance of MAAGAP: A Programming Model of the Agricultural Sector of the Philippines," *J. of Agr. Econ. and Devt.*, Los Banos, Philippines, Vol. VIII, No. 3, Nov. 1978, pp. 1-25.

(*13*) Kutcher, Gary. "A Regional Agricultural Planning Model for Mexico's Pacific Northwest." Forthcoming in *Programming Studies for Agricultural Policy*, by L. M. Bassoco, R. E. Norton, J. S. Solis, and L. M. Solos, Oxford Univ. Press.

(*14*) Miller, T., and R. Miller. "A Prototype Model to the U.S. Agricultural Sector." Unpublished USDA station paper, Colorado State Univ., 1976.

(*15*) Nugent J. B. "Linear Programming Models for National Planning: Demonstration of a Testing Procedure," *Econometrica*, Vol. 38, Nov. 6, Nov. 1970.

(*16*) Pomareda, C., and R. L. Simmons. "A Risk Programming Model to Evaluate Mexican Rural Wage Policies," *Operational Res. Quar.*, 1977.

(*17*) Shumway, R. C., and H. Talpaz, "Verification of Linear Programming Solutions with Emphasis on Supply Implications," *So. J. Agr. Econ.*, Vol. 9, No. 2, Dec. 1977.

(*18*) Tilanus, C. B., and H. Theil. "The Information Approach to the Evaluation of Input-Output Forecasts," *Econometrica*, Vol. 32, No. 4, Oct. 1965.

TRANSACTION COSTS, RISK AVERSION, AND CHOICE OF TENURE REVISITED

By Donald Baron*

INTRODUCTION

Is allocative efficiency affected by the type of leasing or employment contract that tenants and landlords negotiate on rental and owner-operated farmland? How do tenants and landlords determine their choice of contracts? This article examines why risk-averse tenants and landlords pick the type of contracts they do.

Until recently, the consensus among economists was that a tenant who manages his own farm would achieve greater efficiency by renting land for a fixed rental payment than by receiving a share of farm revenue. In other words, a fixed rental contract would allocate resources more efficiently than a share contract. Moreover, the efficiency of the fixed rental contract was viewed as equal to that of the fixed wage contract, under which a landowner/farm manager pays a tenant/employee a contractually fixed salary. Therefore, the fixed wage contract was also considered more efficient than the share contract (*1*, *3*, *4*, *6*, *9*, *13*, *18*).[1]

The share contract is thought to be less efficient in that the share tenant receives only a fraction of the marginal product of all variable inputs. He therefore employs fewer inputs than does either the fixed rental tenant or the fixed wage landlord. In this article, the reduction in input employment caused solely by the share tenant's receipt of only a fraction of the marginal product will be referred to as "shirking."

The related questions of how tenants and landlords choose leasing contracts and how these contracts affect efficient allocation of resources continue to divide economists. This article rejects answers suggested by transaction cost models developed by Cheung and by Ip and Stahl and argues that among risk-averse farmers, contract choice is determined by the relative intensities of tenant and landlord aversion to risk. The risk model examined here suggests that all contract forms—whether fixed rent, fixed wage, or crop share—can generally achieve the same allocative efficiency.

Keywords

Leasing contracts
Tenure
Allocative efficiency
Transaction costs

CHEUNG'S THEORY

Recently, a number of economists have challenged this traditional condemnation of the share contract. Cheung has argued that in a competitive economy rational landlords do not permit share tenants to determine unilaterally a level of variable input use that is less than the optimum under alternative tenure forms. They act to prevent shirking. Moreover, as the share contract enables risk-averse landlords and tenants to increase the utility of their incomes by sharing the risk of uncertain farm output and prices, risk-averse landlords and tenants always, Cheung stated, prefer the share contract as a method of counteracting uncertainty. Formally proving this latter argument, Sutinen further proved that the utility-maximizing share contract will actually achieve a greater allocative efficiency than either of the two nonshare contracts (*20*). Specifically, he proved that the risk-averse landlord who shares risk with his or her tenant under a share contract will employ more resources and produce more than under any nonshare contract (*19*, *20*).

Cheung and Sutinen recognized that their new theories of share contracting raised new questions (*7*, *19*). If landlords act to prevent shirking, and if share contracts can always disperse risk, why is share contracting not the only observed tenure form? One answer is that some landlords and tenant farm managers are either neutral as to risk or willing to gamble. A risk-neutral decisionmaker has no incentive to share risk and, therefore, factors unrelated to risk determine choice of contract. A risk-preferring landlord or farm manager has an incentive to assume the entire risk under a fixed wage contract, whereas the risk-preferring tenant will want to assume the entire risk under a fixed rental contract.

Studies by Wolgin, Wiens, and Moscardi and deJanvry (*22*, *21*, *14*) suggest, however, that at least in peasant agriculture, most small farmers are risk averse. Therefore, the important question is that which Cheung first addressed and that which is also the subject of this article: Why do risk-averse landlords and tenants in peasant agriculture so often choose fixed rental or fixed wage contracts, despite their obvious preference for risk sharing?

*The author is an agricultural economist with the Natural Resource Economics Division, ESCS.

[1] Italicized numbers in parentheses refer to items in References at the end of this article.

Cheung's answer was that choice of tenure primarily depends on the extent to which risk sharing and transaction costs offset each other. The share contract would indeed always be preferred by risk-averse peasants; however, it requires transaction costs that are substantially higher than costs under fixed wage and rental contracts. Peasant farmers choose the share contract only if these higher transaction costs are compensated for by gains they expect from risk sharing. Alternatively, they choose a fixed wage or rental contract if the higher transaction costs of the share contract would more than offset its risk-sharing advantages.

Sutinen suggested that transaction costs may sometimes be less under share contracts than under nonshare contracts, but he agreed with Cheung that, where transaction costs are "greater for share leasing . . . , a share lease may not be preferred even when risk and risk aversion exist" (*20*, p. 616). Sutinen further argued, however, that even if transaction costs are equal for all contracts, peasant farmers may still choose nonshare contracts because other more effective risk dispersal methods, such as crop insurance or future markets, may be available (*20*, pp. 616-617).

FOCUS OF INVESTIGATION

This article, as mentioned, focuses on farming regions in developing countries, where risk aversion is the norm. As farmers in these areas typically lack access to futures markets and private or government crop insurance, I assume that share contracting is the only feasible method available for sharing risk. I will argue that Cheung's transaction cost theory is an inappropriate explanation of tenure choice in this environment. I will then propose that choice of tenure among risk-averse farmers depends on differences between the intensities of landlord and tenant risk aversions. This theory will also confirm Cheung's view that allocative efficiency should generally be the same under all tenure forms. Finally, the analysis will determine which of the conflicting theories of tenure choice is most consistent with the available empirical evidence of transaction costs under different tenure forms.

INCONSISTENCIES IN THE CHEUNGIAN THEORY

Cheung argued that transaction costs are higher under share contracting than under either fixed wage or fixed rental contracting because more time must be devoted to negotiating and enforcing contract terms under the former. Share tenants and landlords must devote substantial time to negotiating contract terms that specify in great detail the duties both parties will perform. Moreover, they must agree on such terms as the "rental percentage, the ratio of nonland input to land, and the types of crops to be grown," whereas under "fixed rent[al] and wage contracts . . . , given the market prices, one party alone can decide how much of the other party's resources he shall employ and what crops shall be grown." In addition, share landlords, unlike fixed rental landlords, must devote substantial time to supervision to prevent tenants from shirking (7, pp. 67-68).

Cheung failed to explain exactly how the higher transaction costs of the share contract might operate to offset its risk-sharing advantages. Perhaps recognizing this shortcoming, Ip and Stahl proposed an explanation which views the typical landlord's labor supply curve as the basic measurement of transaction costs. Like Sutinen, Ip and Stahl believed that risk sharing by itself makes the allocative efficiency of the share contract higher than that of the nonshare contract. They argued, however, that, given their labor-leisure preferences, most landlords are unwilling to devote the time to contract negotiation and enforcement that is necessary to completely eliminate shirking by tenants. The less the tenant shirks, the more inputs the tenant will employ. Yet landlords will continue their efforts to prevent shirking only as long as the expected utility of their share of the additional output exceeds the disutility of their additional efforts (*12*, pp. 22-24).

The marginal disutility of landlords' work increases as the time they devote to contract negotiations and monitoring tenant activities increases. At the same time, the marginal product of the additional variable inputs tenants employ decreases as employment of resources increases. Therefore, the marginal utility of the landlord's share of

This article examines why risk-averse tenants and landlords pick the type of contracts they do.

output is likely to equal the marginal disutility of contract negotiation and supervisory work at a point where a significant amount of shirking by the tenant still occurs. This residual amount of shirking offsets the gains from risk sharing, according to Ip and Stahl.

Residual shirking may be a valid measure of the higher transaction costs of share contracting envisioned by Cheung. Unlike Cheung, Ip and Stahl were not concerned about why different landlords and tenants choose different tenure forms. They wanted instead to explain why empirical studies have shown that allocative efficiency is generally the same under all forms of agricultural tenancy (*2*, *5*, *10*, *11*, *17*), despite Sutinen's proof that share contracting can always allocate resources more efficiently because risks are shared. Ip and Stahl wanted to determine how the share tenant, who shirks despite landlord supervision, might still produce as much as the nonshirking owner-operator farmer who hires no wage labor. They suggested that the amount by which residual shirking reduces optimum output equals the amount by which risk sharing increases output under the share contract. In other words, net production is the same under the share contract as it is under owner cultivation (*12*, pp. 23-24).

Moreover, although they did not say so explicitly, Ip and Stahl intended this reasoning to explain why allocative efficiency should be the same under share contracting as it is under fixed rental and fixed wage contracting.[2] Like the owner-cultivator, the fixed rental tenant and the fixed wage landlord have no incentive to shirk. Yet, they produce no more than share tenants because their assumption of the entire risk has the same negative impact on production that residual shirking has under the share contract.

Ip and Stahl's treatment of residual shirking differs from Cheung's transaction cost theory in not viewing transaction costs as the key factor determining choice of tenure. Risk sharing and residual shirking offset each other to the point that share contracting is equally as efficient as the nonshare tenure forms, Ip and Stahl suggest. Then we are left with no explanation of why one tenure form may be preferred over another.

If one agrees with Cheung that transaction costs *do* determine choice of tenure, how do we then explain the evidence that all tenure forms use resources equally efficiently? One might argue that the share contract achieves maximum utility and is the preferred tenure form only when the output gain from risk sharing exceeds the loss attributable to residual shirking. Fixed rental and fixed wage contracts would be chosen when the loss from shirking exceeds the output gain from risk sharing. This explanation clearly implies that farms operating under share contracts should actually achieve a greater allocative efficiency—contrary to the empirical evidence—than farms operating under fixed wage or fixed rental contracts, because of the net gain attributable to risk sharing. Thus, the possibility that both residual shirking and risk sharing would determine choice of tenure contradicts the theory that all tenure forms are equally efficient.

Thus far no theory has been posed that can consistently answer both the choice of tenure issue and the allocative efficiency issue. To develop such a new approach, I first present a modified form of Sutinen's proof that, in the absence of transaction costs and other more effective risk sharing methods, the share contract will always be the utility-maximizing choice of tenure among risk-averse farmers, and it will always achieve a greater allocative efficiency than any nonshare contract. I will also propose an alternative choice of tenure theory, consistent with Ip and Stahl's findings of equal efficiency.

THE SUTINEN MODEL

Sutinen showed how the share contract can always achieve a greater expected utility and a greater allocative efficiency than the nonshare contracts. To do so, he derived the necessary conditions for the maximization of a risk-averse landlord's expected utility of income subject

[2] Some of the empirical studies cited by Ip and Stahl compare share contracting only with fixed rental and fixed wage contracting rather than with owner cultivation without wage labor. For example, in a study of farm tenure in Malaysia (cited by Ip and Stahl as an example of the "equal efficiency" school of thought (*12*, p. 23)), Huang showed that share tenants produced at least as much or more per acre as did both fixed wage landlords and fixed rental tenants (*11*, pp. 706-15).

to a constraint. This constraint was that the risk-averse tenant's expected utility of income from any particular tenure form had to equal the expected utility of the tenant's opportunity income (*20*).

In this article, I modify Sutinen's approach by maximizing both the tenant's and the landlord's expected utility. The person whose expected utility is being maximized is assumed to act as the decisionmaking party. When the landlord's expected utility is maximized, I assume that the tenant acts as an employee with no managerial responsibilities. Maximizing the landlord's expected utility will, therefore, be appropriate for determining the allocative efficiency of both the fixed wage contract and also any share contract that assigns all managerial responsibilities to the landlord and none to the tenant. Alternatively, I maximize the tenant's expected utility to determine the allocative efficiency of both the fixed rental contract and any share contract that assigns all managerial responsibilities to the tenant.

Maximization from the landlord's point of view is defined as maximization of the expected utility of the landlord's income, $\pi = r_l pqu - c(q) - \theta$. A condition is that the expected utility of the tenant's income, $\phi = r_t pqu + \theta$, must equal the expected utility of the income, I, which the tenant could earn by employing his or her assets elsewhere. The landlord's share of total revenue is r_l, and the tenant's share is r_t, where $r_t = 1 - r_l$. Total revenue is represented by pqu; u is a nonnegative random variable which accounts for variations in either the output price, p, or environmental factors, such as weather and disease, both of which effect output, q. As the expected value of u is a constant, the expected total revenue is simply pq E(u).[3] θ is a shift variable "which acts to adjust the (tenant's) expected income to a level where he is indifferent about employing his assets in this farming activity or elsewhere in the economy" (*20*, p. 615). C(q) equals the total variable costs of producing expected output q, with the first and second derivatives $C'(q)$ and $C''(q)$, both being positive. The decisionmaking landlord pays all variable costs, since one decision will be the amount of each variable input to employ. Both the landlord and tenant are assumed to have continuous, concave utility functions, $U(\pi)$ and $U(\phi)$, such that $U'(\pi)$ and $U'(\phi)$ are both positive while $U''(\pi)$ and $U''(\phi)$ are negative.

Following Sutinen's approach, I assign specific functional forms to the utility functions, $U(\pi)$ and $U(\phi)$, and a specific probability law to u. Let:

$$U(\P) = -e^{-\alpha \P} \text{ and } U(\phi) = -e^{-\beta\phi},$$

where α and β equal the absolute risk aversions of the landlord and tenant, respectively (*20*, p. 617). Also assume that u follows the gamma probability law, which defines the following probability density function:

$$f(u) = \frac{\Lambda^\rho}{\Gamma(\rho)} u^{\rho - 1} e^{-\Lambda u}$$

for $u \geqslant o$, with parameters $\rho = 1,2,...$ and $\Lambda > 0$ (*15*, p. 180).

In terms of the Lagrangian, the landlord maximizes:

$$L = E(- e^{-\alpha\pi}) + \lambda[E(- e^{-\beta\phi}) - E(U(I))]$$

or

$$L = - \int_0^\infty e^{-\alpha\pi} f(u)\, du + \lambda[- \int_0^\infty e^{-\beta\phi} f(u)\, du + \int_0^\infty e^{-\beta I} f(u) du].$$

Substitution of $\pi = r_\ell pqu - C(q) - \theta$, $\phi = r_t pqu + \theta$, and $f(u) = \Lambda^\rho u^{\rho - 1} e^{-\Lambda u}/\Gamma(\rho)$ yields:

$$L = - \Lambda^\rho[\alpha r_l pq + \Lambda]^{-\rho} e^{\alpha[c(q) + \theta]} + \lambda \left\{\Lambda^\rho [\beta r_t pq + \Lambda]^{-\rho} e^{-\beta\theta} - e^{-\beta I} \right\}$$

[3] E(u) $\gtreqless$ o are all possible. When E(u) > 1, marginal returns to random inputs, such as weather conditions, are increasing. When E(u) = 1, or E(u) < 1, marginal returns are constant or decreasing, respectively. Increasing marginal returns encourage decisionmakers to increase production (*8*, pp. 27-28). However, if decisionmakers are risk averse, this increase will be partly or completely offset by the negative impact that income variance has on production. The net effect of uncertainty and risk aversion on production depends on the specifications of the production and utility functions.

The choice of tenure theory I am proposing here suggests that the three tenure forms are equally efficient because landlords and tenants choose different tenure forms to adjust for differences in their attitudes toward risk.

The first order conditions are:

$$L_r = -\Lambda^{\rho}(-\rho)(\alpha pq)[\alpha r_l pq + \Lambda]^{-\rho-1} e^{\alpha[C(q)+\theta]} + \lambda \Lambda^{\rho}(-\rho)(-\beta pq)[\beta r_t pq + \Lambda]^{-\rho-1} e^{-\beta\theta} = 0 \qquad (1)$$

$$L_\theta = -\Lambda^{\rho}\alpha[\alpha r_l pq + \Lambda]^{-\rho} e^{\alpha[c(q)+\theta]} + \lambda \Lambda^{\rho}(-\beta)[\beta r_t pq + \Lambda]^{-\rho} e^{-\beta\theta} = 0 \qquad (2)$$

$$L_q = -\Lambda^{\rho}(-\rho)(\alpha r_l pq)[\alpha r_l pq + \Lambda]^{-\rho-1} e^{\alpha[C(q)+\theta]} - \Lambda^{\rho}(\alpha C'(q))(\alpha r_l pq + \Lambda)^{-\rho} e^{\alpha[C(q)+\theta]} + \lambda \Lambda^{\rho}(-\rho)(\beta r_t p)[\beta r_t pq + \Lambda]^{-\rho-1} e^{-\beta\theta} = 0 \qquad (3)$$

$$L_\lambda = \Lambda^{\rho}[\beta r_t pq + \Lambda]^{-\rho} e^{-\beta\theta} - e^{-\beta I} = 0 \qquad (4)$$

Using equations (1) and (2) we can calculate the optimum value for the landlord's share rate r as:

$$r_l = \frac{\beta}{\alpha + \beta}. \qquad (5)$$

Since $\alpha > 0$ and $\beta > 0$ (both landlord and tenant are risk averse), it is clear that $0 < r_l < 1$. Thus, Sutinen concluded that the landlord will always maximize expected utility by choosing a share contract rather than a fixed wage contract (*20*, pp. 615-17).

A similar conclusion can be derived for the tenant who maximizes the expected utility of income, $\phi = (r_t)pqu - C(q) - \theta$, subject to the condition that the expected utility of the nonmanagerial landlord's income, $\pi = rpqu + \theta$, equals the expected utility of the income, Y, which the landlord could earn by employing assets elsewhere.

In exponential form, maximizing the tenant's expected utility requires maximizing:

$$L = -\Lambda^{\rho}[\beta(r_t)pq + \Lambda]^{-\rho} e^{\beta[C(q)+\theta]} + \lambda\left\{\Lambda^{\rho}[\alpha r_l pq + \Lambda]^{-\rho} e^{\alpha(-\theta)} - e^{\alpha Y}\right\}$$

where:

$$E\left\{U(\phi)\right\} = \frac{-\Lambda^{\rho} e^{\beta[C(q)+\theta]}}{[\beta r_t pq + \Lambda]\rho}$$

and:

$$E\left\{U(\pi)\right\} = \frac{-\Lambda^{\rho} e^{\alpha(-\theta)}}{[\alpha r_l pq + \Lambda]^{\rho}}.$$

The first order conditions are:

$$L_r = -\Lambda^{\rho}(-\rho)(-\beta pq)[\beta r_t pq + \Lambda]^{-\rho-1} e^{\beta[C(q)+\theta]} + \lambda \Lambda^{\rho}(-\rho)(\alpha pq)[\alpha r_l pq + \Lambda]^{-\rho-1} e^{-\alpha\theta} = 0 \qquad (6)$$

$$L_\theta = -\Lambda^{\rho}(\beta)[\beta r_t pq + \Lambda]^{-\rho} e^{\beta[C(q)+\theta]} + \lambda \Lambda^{\rho}(-\alpha)(\alpha r_l pq + \Lambda]^{-\rho} e^{-\alpha\theta} = 0 \qquad (7)$$

$$L_q = -\Lambda^{\rho}(-\rho)(\beta r_t p)[\beta r_t pq + \Lambda]^{-\rho-1} e^{\beta[C(q)+\theta]} - \Lambda^{\rho}(\beta C'(q))[\beta r_t pq + \Lambda]^{-\rho} e^{\beta[C(q)+\theta]} + \lambda \Lambda^{\rho}(-\rho)(\alpha r_l p)[\alpha r_l pq + \Lambda]^{-\rho-1} e^{-\alpha\theta} = 0 \qquad (8)$$

$$L_\lambda = \Lambda^{\rho}[\alpha r_l pq + \Lambda]^{-\rho} e^{-\alpha\theta} - e^{-\alpha Y} = 0 \qquad (9)$$

Using equations (6) and (7), we can calculate the optimum value of the tenant's share rate as:

$$r_t = \frac{\alpha}{\alpha + \beta}. \qquad (10)$$

Again, as $\alpha > 0$ and $\beta > 0$, $0 < r_t < 1$. Therefore, the tenant will always maximize his or her expected utility by choosing a share contract rather than a fixed rental contract (*20*, pp. 615-617).

Let us now compare the optimum marginal costs achieved under each tenure form in equilibrium to see why Sutinen further concluded that the utility-maximizing share contract will always achieve a greater allocative efficiency than any nonmaximizing, nonshare contract. Consider first the managerial landlord who views the fixed wage and share contracts as the two major tenure alternatives. For the fixed wage contract, we use equation (2) to solve for λ, and substitute λ into equation (3) to derive:

$$C'(q) = P\left[\frac{E(U'(\pi)u)}{E(U'(\pi))}\right]_{r_l = 1}$$

$$= P\left[\frac{\rho}{\alpha pq + \Lambda}\right]_{r_l = 1} \qquad (11)$$

We use the same procedure to derive the optimum marginal cost under the landlord-managed share contract:

$$C'(q) = P\left[\frac{E(U'(\pi)u)}{E(U'(\pi))}\right]_{0 < r_l < 1}$$

$$= P\left[\frac{\rho}{\alpha r_l pq + \Lambda}\right]_{0 < r_l < 1} \qquad (12)$$

We derive similar equations for the managerial tenant who views the fixed rental and share contracts as his major tenure alternatives. We solve equation (7) for λ, substitute λ into equation (8) and derive the following values for the optimum marginal costs under fixed rental and tenant-managed share contracts, respectively:

$$C'(q) = P\left[\frac{E(U'(\phi)u)}{E(U'(\phi))}\right]_{r_t = 0}$$

$$= P\left[\frac{\rho}{\beta pq + \Lambda}\right]_{r_t = 0} \qquad (13)$$

$$C'(q) = P\left[\frac{E(U'(\phi)u)}{E(U'(\phi))}\right]_{0 < r_t < 1}$$

$$= P\left[\frac{\rho}{\beta r_t pq + \Lambda}\right]_{0 < r_t < 1} \qquad (14)$$

The bracketed expressions in equations (11) through (14)—the risk factors—measure the impact of uncertainty on the decisionmaker's optimum employment of resources. Because $\rho/\alpha r_l pq + \Lambda$ under the landlord-managed share contract exceeds $\rho/\alpha pq + \Lambda$ under the fixed wage contract, and $\rho/\beta r_t pq + \Lambda$ under the tenant-managed share contract exceeds $\rho/\beta pq + \Lambda$ under the fixed rental contract, it is clear that the optimum marginal cost is also higher under the utility-maximizing share contract than under any nonshare contract. Moreover, this higher marginal cost translates into a higher optimum production level, which measures the amount by which the allocative efficiency of the optimum share contract exceeds that of the nonshare contracts (*20*, pp. 617-619).

RISK AVERSION AND CHOICE OF TENURE

Now reconsider the question of why some risk-averse landlords and tenants forego the risk-sharing benefits of the share contract and select instead the fixed wage or rental contract. One explanation can be derived from an analysis of the inverse relationship in equations (5) and (10) between the optimum values of r_l and r_t and the relative values of the landlord and tenant's risk aversions. This relationship will be used to develop a new theory of tenure choice that does not rely on hypotheses (such as Cheung's) concerning the transaction costs of different tenure forms.

First, assume that a nonmanagerial tenant willing to let the landlord act as decisionmaker is likely to be either equally or more risk averse than the landlord. This assumption is important because equation (5) reveals that the lower the value of the landlord's risk aversion relative to the tenant's risk aversion, the greater the landlord's optimum share rate, r_l, and the lower the tenant's optimum share rate, r_t. Moreover, as the risk aversion of the tenant increases over that of the landlord

The greater the difference between the risk aversions of the landlord and tenant, the more closely the utility-maximizing contract will resemble a fixed wage or rental contract rather than a standard share contract.

and the optimum value of r_l increases towards $r_l = 1$, the optimum value of θ needed to satisfy the tenant's constraint equation (4) will increase towards $\theta = W$. Thus, the more closely will the utility-maximizing share contract resemble the fixed wage contract.

Eventually, the tenant's risk aversion will exceed the landlord's to the point that the landlord will actually maximize expected utility by choosing the fixed wage contract. The landlord will choose this rather than the standard share contract, under which r_l is significantly less than 1, and θ is significantly less than W.

A similar conclusion can be derived for the decisionmaking tenant. This tenant is likely to be as risk averse or less risk averse than the nonmanagerial landlord. Equation (10) shows that the lower the value of the tenant's risk aversion relative to the landlord's, the greater will be the tenant's optimum share rate, r_t, and the lower will be the landlord's optimum share rate, r_l. Moreover, as the risk aversion of the landlord increases over that of the tenant, and the optimum value of r_t increases towards 1, the optimum value of θ needed to satisfy the landlord's constraint equation (9) will increase towards $\theta = R$. Thus, the more closely will the utility-maximizing share contract resemble the fixed rental contract.

Eventually, the landlord's risk aversion will exceed the tenant's to the point that the tenant will maximize his expected utility by choosing the fixed rental contract. The tenant will choose this rather than a standard share contract, under which r_t is significantly less than 1, and θ is significantly lower than R.

The relationship between risk aversion and the expected utility of different tenure forms suggests that tenure choice depends on the relative values of landlord and tenant risk aversions. The greater the difference between landlord and tenant risk aversions, the more closely will the utility-maximizing contract resemble a fixed wage or rental contract rather than a standard share contract. A landlord who is much less risk-averse than the tenant and who chooses a fixed wage contract does so, not because transaction costs under share contracting will be prohibitively high, but because the landlord can maximize expected utility by assuming the entire risk. Similarly, less risk-averse tenant/farm managers who choose a fixed rental contract do so because they can expect to maximize utility by assuming the entire risk. On the other hand, a landlord and tenant who choose a standard share contract do so because the difference between their risk aversions is small enough to make risk sharing attractive.[4]

[4]On farms jointly managed by landlords and tenants, the standard share contract is likely to be the preferred tenure form because the sharing of management responsibilities is inconsistent with the assumption of the risk by only one party. Moreover, the difference between the risk aversions of landlords and tenants who are willing to share management responsibilities is likely to be small enough (perhaps zero) to make risk sharing under a standard share contract feasible. Indeed, the output share assigned to each party is likely to reflect the amount of management responsibilities each party has assumed.

RISK AVERSION AND EMPIRICAL STUDIES

This theory of tenure choice confirms empirical studies which demonstrate that all three tenure forms are equally efficient. Ip and Stahl's transaction cost model suggests that this is so because the greater efficiency of the share contract—as a result of risk dispersion—is offset by the residual shirking of the decisionmaking share tenant or landlord.[5] The choice of tenure theory I am proposing here suggests that the three tenure forms are equally efficient because landlords and tenants choose different tenure forms to adjust for differences in their attitudes toward risk.

I demonstrate this theory by comparing the risk factor value of each tenure form when it is the utility-maximizing, or "rationally chosen," contract. Let equation (11) now represent the optimum marginal cost achieved by the fixed wage contract only when tenant risk aversion exceeds landlord risk aversion to the point that the expected utility of the fixed wage contract is higher than that of the standard share contract under which r_l is significantly less than 1. Now let equation (12) represent the optimum marginal cost of the standard share

[5]Ip and Stahl were concerned only with share contracting on tenant-managed farms. However, their transaction cost model clearly predicts that residual shirking by the landlord decisionmaker will occur under share contracting on landlord-managed farms and that this shirking will also offset any increase in output induced by risk sharing.

contract only when the difference between the tenant's and the landlord's risk aversion is small enough (perhaps zero, in which case $r_l = 1/2$) to make the standard share contract the expected utility-maximizing tenure form.

To compare the values of the risk factors in these two equations, note that in both the risk factor depends on the landlord's risk aversion, α, as well as the landlord's share rate, r_l. This is important, because for any given value of β, α should be greater under the share contract, represented by equation (12), than under the "rationally chosen" wage contract, represented by equation (11). Moreover, as equation (5) indicates, the higher the value of α, the lower the optimum share rate, r_l. Thus, the amount by which α in equation (12) exceeds α in equation (11) determines the amount by which r_l in equation (11) exceeds r_l in equation (12).

I derive a similar relationship for the tenant-managed farm. Recall that the fixed rental contract is the rational choice when landlord risk aversion exceeds tenant risk aversion to the point that the fixed rental contract will have a higher expected utility than the standard tenant-managed share contract. Let equation (13) represent this utility-maximizing fixed rental contract, and let equation (14) represent the utility-maximizing standard share contract. As the excess α over β is greater under the fixed rental contract than under the standard share contract, β in equation (14) should exceed β in equation (13) for any given value of α. Moreover, as equation (10) indicates, the amount by which β in equation (14) exceeds β in equation (13) determines the amount by which r_t in equation (13) exceeds r_t in equation (14).

The inverse relationships between the risk aversion of the decisionmaking tenant or landlord and the rational choice of r_t and r_l may explain why all tenure forms are equally efficient. The different risk aversions and different optimum values of r_l and r_t characterizing each tenure form when rationally chosen have offsetting effects on the value of the risk factor. The inverse relationship between r_l and α suggests that among all decisionmaking landlords, the average amount by which r_l on farms operating under the fixed wage contract exceeds r_l on farms operating under the share contract approximates the average amount by which α under the share contract exceeds α under the fixed wage contract. Thus, the average values of both the risk factor and the optimum marginal cost, $C'(q)$, as shown in equations (11) and (12), are the same under fixed wage contracting as under share contracting. Similarly, the inverse relationship between r_t and β suggests that among all decisionmaking tenants, the average amount by which r_t on farms operating under the fixed rental contract exceeds r_t on farms operating under the share contract approximates the average amount by which β under the share contract exceeds β under the fixed rental. Thus, the average values of both the risk factor and the optimum $C'(q)$, as shown in equations (13) and (14), are the same under fixed rental contracting as under standard share contracting.

Empirical studies have shown that the three tenure forms are equally efficient, but not because of a trade off between the greater efficiency of share leasing as a response to risk and the lower transaction costs of fixed rental and fixed wage contracting, as Ip and Stahl suggest. The three tenure forms are equally efficient because, for any given level of transaction costs, the nonshare contract chosen when differences between landlord and tenant risk aversions are greater will generally achieve the same optimum marginal cost and the same total output as the standard share contract chosen when differences between landlord and tenant risk aversions are less.

TRANSACTION COSTS AND EMPIRICAL STUDIES

Both my theory of tenure choice and theory of allocative efficiency contradict Cheung's conclusion that transaction costs are necessarily and consistently higher under share contracting than under the other two tenure forms. I suggest that transaction costs may be the same for all tenure forms, because even in a zero transaction cost situation, share contracting as usually practised is not always the utility-maximizing or the most efficient tenure form.

Recent studies tend to support this conclusion. In his analysis of post Civil War farming in the South, Reid confirms Cheung's view that landlords under share contracts incur substantial transaction costs in negotiating detailed contractual terms and in monitoring tenant performance to prevent shirking on

Landlord and tenant farm managers will choose the standard share contract only if the difference between landlord and tenant risk aversions is small enough to make risk sharing attractive.

variable inputs. However, his evidence reveals that fixed wage and rental contracts were also "costly to negotiate and enforce" (*16*, p. 569).

Reid found that:

> . . . typical wage contracts required each laborer's attendance to an overseer and specified in detail daily work schedules and a renumeration schedule related to each worker's satisfaction of his contractual obligations. Rental contracts resembled their sharecropping counterparts in paying much attention to the details of land use and of maintenance duties. Landlords often placed specific contractual restraints upon renters to guard against deterioration of land and capital (instructions regarding drainage, type of plowing permissible, number and type of crops allowed, maintenance of fences and buildings, use of manures and fertilizers, prohibitions on stock grazing in clover or fallow). To insure that renters would pay their rents and honor their contracts, landlords often supervised their work as well (*16*, pp. 569-70).

Reid concludes that as

> . . . all agricultural production requiring cooperation among different factor owners necessitates costly negotiation and enforcement . . . , little plausibly differentiates the landlord's requisite transaction costs under self-cultivation with hired labor or under renting from his transaction costs under sharecropping (*16*, p. 570).

CONCLUSION

I have reviewed Cheung's theory that in a situation of zero transaction costs, risk-averse landlords and tenants will always maximize their utilities by choosing some form of share contract rather than a nonshare contract. I have also examined Sutinen's proof that the utility-maximizing share contract will always be more efficient than a nonshare contract as it will achieve a higher optimum marginal cost and a higher optimum expected output.

My principal conclusion here is that the greater the difference between the risk aversions of the landlord and tenant, the more closely the utility-maximizing contract will resemble a fixed wage or rental contract rather than a standard share contract. On landlord-managed farms, the expected utility of the fixed wage contract will increase relative to that of the standard share contract as the risk aversion of the tenant over that of the landlord increases.

Eventually, the increase will be so great that the fixed wage contract will be preferred to the standard share contract. Similarly, on tenant-managed farms, the expected utility of the fixed rental contract will increase relative to that of the standard share contract as the risk aversion of the landlord over that of the tenant increases. Eventually, the increase will be so great that the fixed rental contract will be the preferred tenure form. Thus, landlord and tenant farm managers will choose the standard share contract only if the difference between landlord and tenant risk aversions is small enough to make risk sharing attractive.

In this article, I have also developed an explanation of why empirical studies have shown all tenure forms to be equally efficient. I have suggested that landlord and tenant decisionmakers who rationally choose fixed wage and fixed rental contracts are generally less risk averse than those who choose share contracts. Therefore, even though the former assume the entire risk by themselves—that is, $r_l = r_t = 1$—they will still generally achieve the same optimum marginal cost and total output as do their share-contracting counterparts.

This conclusion and the choice of tenure theory on which it is based are clearly distinguished from Cheung's theory of tenure choice and Ip and Stahl's explanation of the equal efficiency findings by their rejection of the argument that transaction costs are necessarily higher under share contracts than under nonshare contracts. I suggest that, even in a world where transactions costs are equal for all contracts, differences in the risk aversions of landlords and tenants and in the amounts of risk they assume are enough to ensure the continued existence of a wide variety of equally efficient contract forms.

REFERENCES

(*1*) Adams, D. W., and N. Rask. "Economics of Cost-Share Leases in Less-Developed Countries." *Amer. J. Agr. Econ.*, Vol. 50, 1968, pp. 935-942.

(*2*) Ahmad, M. "Farm Efficiency Under Owner Cultivation and Share Tenancy." *Pakistan Econ. and Social Rev.*, Vol. 12, 1974, pp. 132-143.

(*3*) Bardhan, P. K., and T. N. Srinivasan. "Cropsharing Tenancy in Agriculture." *Amer. Econ. Rev.*, Vol. 62, 1972, pp. 777-795.

(*4*) Bhagwati, J. N. *The Economics of Underdeveloped Countries.* New York: McGraw-Hill Book Co., 1966.

(*5*) Bray, J. O. "Farm Tenancy and Productivity in Agriculture: The Case of the United States." *Food Res. Inst. Stud.*, Vol. 4, 1963, pp. 25-38.

(*6*) Castle, E. "Some Aspects of the Crop-Share Lease." *Land Econ.*, Vol. 28, 1952, pp. 177-179.

(*7*) Cheung, S. N. S. *The Theory of Share Tenancy*. Chicago: University of Chicago Press, 1969.

(*8*) Edwards, C. "Gambling, Insurance, and the Production Function." *Agr. Econ. Res.*, Vol. 30, 1978, pp. 25-28.

(*9*) Georgescu-Roegen, N. "Economic Theory and Agrarian Reform." *Oxford Econ. Papers*, Vol. 12, 1960, pp. 1-60.

(*10*) Hendry, J. B. "Land Tenure in South Vietnam." *Econ. Devt. and Cult. Change*, Vol. 9, 1960, pp. 27-44.

(*11*) Huang, Y. "Tenancy Patterns, Productivity, and Rentals in Malaysia." *Econ. Devt. and Cult. Change*, Vol. 23, 1975, pp. 703-718.

(*12*) Ip, P. C., and C. W. Stahl. "Systems of Land Tenure, Allocative Efficiency, and Economic Development." *Amer. J. Agr. Econ.*, Vol. 60, 1978, pp. 19-28.

(*13*) Issawi, C. "Farm Output Under Fixed Rents and Share Tenancy." *Land Econ.*, Vol. 38, 1957, pp. 74-77.

(*14*) Moscardi, E., and A. de Janvry. "Attitudes Toward Risk Among Peasants: An Econometric Approach." *Amer. J. Agr. Econ.*, Vol. 59, 1977, pp. 710-716.

(*15*) Parzen, E. *Modern Probability Theory and Its Applications.* New York: John Wiley & Sons, 1960.

(*16*) Reid, J. D., Jr., "Sharecropping and Agricultural Uncertainty." *Econ. Devt. and Cult. Change*, Vol. 24, 1976, pp. 549-576.

(*17*) Ruttan, V. W. "Tenure and Productivity of Philippine Rice Producing Farms." *Philippine Econ. J.*, Vol. 5, 1966, pp. 42-63.

(*18*) Sen, A. K. "Peasants and Dualism With or Without Surplus Labor." *J. Polit. Econ.*, Vol. 74, 1966, pp. 425-450.

(*19*) Sutinen, J. G. "An Economic Theory of Share Contracting." Ph.D. thesis, University of Washington, 1973.

(*20*) ________. "The Rational Choice of Share Leasing and Implications for Efficiency." *Amer. J. Agr. Econ.*, Vol. 57, 1975, pp. 613-621.

(*21*) Wiens, T. B. "Peasant Risk Aversion and Allocative Behavior: A Quadratic Programming Experiment." *Amer. J. Agr. Econ.*, Vol. 58, 1976, pp. 629-635.

(*22*) Wolgin, J. M. "Resource Allocation and Risk: A Case Study of Smallholder Agriculture in Kenya." *Amer. J. Agr. Econ.*, Vol. 57, 1975, pp. 622-630.

FOOD STAMP PROGRAM IMPACTS ON HOUSEHOLD FOOD PURCHASES: THEORETICAL CONSIDERATIONS

By Larry E. Salathe*

INTRODUCTION

The principal objective of the Food Stamp Program (FSP) is to promote the general welfare of the Nation's population by raising levels of nutrition among low-income households. To accomplish this objective, the Food Stamp Act authorizes the distribution of food coupons (stamps) to households which meet certain income eligibility requirements, thereby enabling these households to buy more food to improve their diets.

Numerous researchers have attempted to measure the impact of the FSP on participant households' food purchases. They concur that participation in the FSP increases household food purchases. But there are wide variations in the estimated magnitude of the program's impact. For example, estimates of the marginal propensity to spend on food at home from bonus food stamps range from 0.30 (*10*) to 0.72 (*6*).[1]

This article presents a theoretical framework for estimating empirically the impact of participation in the FSP on food purchased by household members for use at home. Previous studies by Southworth (*8*) and Mittlehammer and West (*4*) provided the basis for developing this framework.

The model for analyzing the impact of the Food Stamp Program on food purchased for use at home indicates that no continuous relationship exists between at-home food expenditures and income of food stamp participant households. As previous studies have not allowed for this fact, they may have measured the program's impact inaccurately. Elimination of the purchase requirement likely decreased food-at-home purchases by some participant households. However, elimination of the purchase requirement probably did not affect food-at-home purchases of food stamp households with incomes near the upper income eligibility bound.

Keywords

Food expenditures
Food Stamp Program
Income

PREVIOUS STUDIES

Previous studies have used indifference curves to analyze the theoretical implications of the FSP on houshold food purchase behavior. Prior to the work of Mittlehammer and West, these analyses focused primarily on explaining the level of participation and the FSP's impact on food-at-home purchases for a household with a given level of income. Little attention was given to explaining the FSP's impact over alternative levels of household income. Mittlehammer and West used indifference curves to analyze the impact of the FSP on household food-at-home purchases, given alternative levels of household income.

The theoretical framework presented here assumes some functional relationship exists between household food-at-home purchases and household income. Indifference curves are not examined explicitly. But if we assume households allocate their income optimally, the theoretical framework will produce the same results as would examining the FSP's impact with indifference curves.

Indifference curves have also been employed to explain nonparticipation of eligible households in the FSP. These analyses were conducted before the purchase requirement was eliminated and cannot explain nonparticipation under current FSP provisions. Furthermore, the theoretical framework presented here cannot explain wh households eligible for the FSP would not participate. Instead, it analyzes the impact of participation on household food-at-home purchase behavior.

THEORETICAL FRAMEWORK

Figure 1 presents the theoretical framework for analyzing the impact of the FSP on participant food-at-home purchases. Line AB represents the assumed functional relationship between household food-at-home expenditures and household income prior to participation in the FSP.[2]

*The author is an agricultural economist in the National Economics Division, ESCS. The helpful comments of William Boehm and Richard King are gratefully acknowledged.

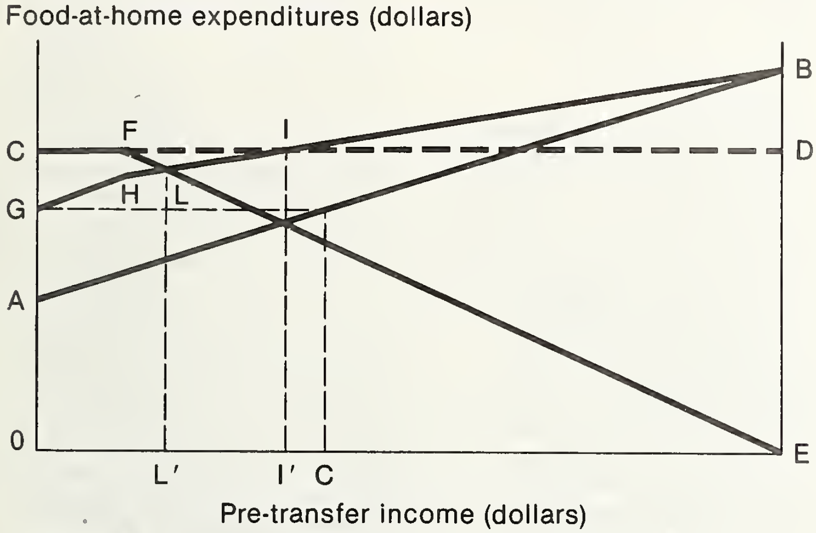
[1] Italicized numbers in parentheses refer to items in References at the end of this article.

[2] A linear relationship between income and food-at-home expenditures was assumed, but is not necessary to derive the results presented here. Figure 1 assumes that factors other than income, such as household size, are held constant.

This article presents a theoretical framework for estimating empirically the impact of participation in the FSP on food purchased by household members for use at home.

Figure 1

Impact of Food Stamp Program on Household Food-at-Home Purchases



CFE shows the relationship between the cash (face) value of food coupons the household is eligible to receive and the household's income.

Cash Transfer Program

Initially, let us assume households participating in the FSP receive the transfer as cash rather than food coupons. Under these conditions, eligible households need not spend their FSP transfer on food to consume at home. Instead, they can allocate the transfer between food at home and other goods in the same fashion as they would do with additional income.

A household with pretransfer income of 0 dollars would receive C dollars of cash by participating in the FSP. Assuming this household allocates the transfer between food at home and other goods in the same fashion as additional income, it would spend a total of G dollars on food at home after participation in a cash transfer program. Thus, this household would expand food-at-home purchases by G minus A dollars and increase other purchases by C minus G plus A dollars. Selecting successively higher pretransfer income levels, one can determine the upward shift in food-at-home purchases resulting from participation in a cash transfer program. GHB shows the relationship between household food-at-home purchases and household (pretransfer) income after participation in a cash transfer program.

Current Food Stamp Program

The Food Stamp Program distributes food coupons rather than cash to participant households. Assuming the marginal utility derived from food is positive (that is, a household desires to spend more on food than its income permits), a participant household will not spend less on food at home than the cash value of food coupons it receives. Thus, if a participant household's income is 0 dollars, it would receive C dollars worth of food coupons and increase its food-at-home purchases to C dollars after participation in the FSP. Purchases of other goods would be increased by A dollars, or the level of expenditure on food at home prior to participation in the FSP. Compared with participation in a cash transfer program, this household would expand food-at-home purchases by C minus G dollars and reduce other purchases by that same amount. In other words, distributing the transfer as food coupons rather than cash will cause this household to spend more on food at home and less on other items. This is because a transfer in the form of food coupons forces participant households to allocate at least the value of the transfer to food at home.

Analyzing successively higher income households reveals that households with incomes below L′ will spend more on food at home (and less on other items) if they receive coupons not cash. Households with incomes at or above L′ can allocate the same amount of income to other items as under a cash transfer program. Thus, CFLB defines the relationship between household food-at-home expenditures and household

Distributing the transfer as food coupons rather than cash will cause this household to spend more on food at home and less on other items.

income for FSP participants (fig. 1). The difference between CFLB and GHB denotes the increase in food-at-home purchases resulting when the transfer is in the form of food coupons rather than cash, at each level of household income.

Food Stamp Program with a Purchase Requirement

Prior to January 1, 1979, households participating in the Food Stamp Program were required to spend a specified amount of their income to receive their allotment of food coupons. The cash value of this allotment did not vary with household income. But the amount of income the household had to spend to receive this allotment increased as household income rose.

In figure 1, line CD represents the cash value of food coupons an eligible household could purchase. The difference between CD and CFE represents the amount of income the household must spend to obtain the allotment of food coupons at each level of household income. Under this program all participant households will spend at least C dollars on food at home, if the marginal utility derived from food at home is assumed to be positive. Thus, the relationship between food-at-home purchases and household income for participants in this program is given by CIB in figure 1. Households with incomes below I′ are forced to spend more on food at home (and less on other items) under this program than under a cash transfer program. However, the purchase behavior of households with incomes above I′ would be the same under all three programs. The theoretical framework also suggests that the impact of the FSP on food-at-home purchases depends on the income distribution of participants.

Alternative placements of AB, the income-expenditure relationship for households before participation, yield slightly different interpretations of the three programs' impacts on food at home purchases. For example, if AB is shifted upward by an amount equal to GC, then GHB would be equal to or above CD, the value of the food stamp allotment. In this case, elimination of the purchase requirement or adoption of a cash transfer program would not alter food purchases by food stamp households. Or, stated differently, a FSP with or without a purchase requirement would be no more effective in increasing food purchases than a cash transfer program providing the same benefits. If this situation exists, empirical estimates of the marginal propensity to spend on food from bonus food stamps and ordinary income would not be statistically different. But a number of empirical studies indicate that these marginal propensities to spend differ statistically. For example, studies by Benus, Kmenta, and Shapiro (*1*), by Hymans and Shapiro (*2*), by Smeeding (*7*), and West and Price (*10*) all indicate that the marginal propensity to spend on food from bonus food stamps exceeds that from ordinary income. Information in figure 1 coincides with these findings.

CHANGING THE VALUE OF FOOD COUPONS DISTRIBUTED

Suppose the cash value of food coupons distributed to participants was increased by a specified amount. Figure 2 analyzes the impact of such an increase on household food purchase behavior.[3] Let AB define the relationship between food-at-home purchases and income prior to participation in the FSP and let CFE represent the value of food coupons distributed at each level of household income. Then CFLB is the relationship between food-at-home purchases and household (pretransfer) income for participant households. Now let us assume the value of food coupons distributed is increased by C′ minus C dollars at each level of household income. Under these conditions C′F′L′B′ gives the relationship between food-at-home purchases and income for participant households. Notice that the difference between C′F′L′B′ and CFLB varies with income, or equivalently, that the impact of an increase in the value of food coupons distributed varies by household income.

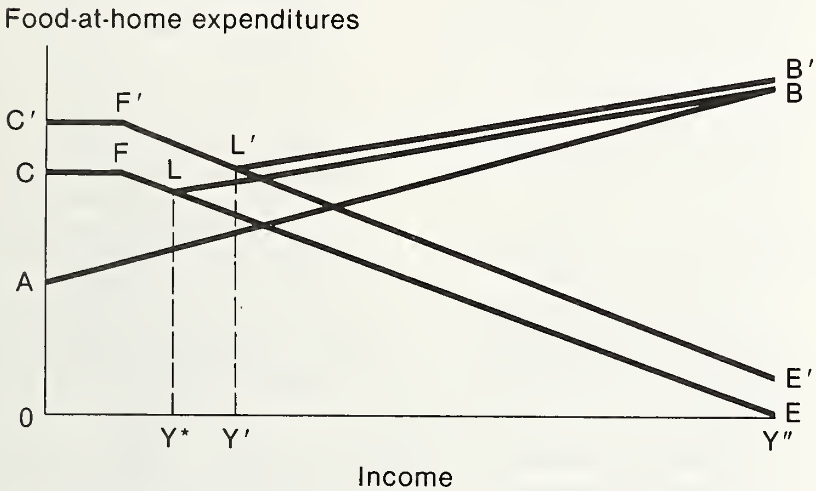
The effect of a $1 increase in the value of food coupons distributed can be shown to range between $1 and the marginal propensity to spend on food at home out of ordinary income. If household income is between 0 and Y*, a $1 increase in

[3] Figure 2 is for a FSP without a purchase requirement. A similar figure for a FSP with a purchase requirement can be easily derived.

A transfer in the form of food coupons should be more effective in increasing food purchases than would a cash transfer of the same value for very low income households.

Figure 2

Increase in Value of Food Coupons Distributed to Participant Households



the value of food coupons distributed will result in a $1 increase in food-at-home purchases. Between income levels Y* and Y′, the effect of a $1 increase in the value of food coupons distributed declines as household income increases. It ranges between $1 and the marginal propensity to spend on food at home out of ordinary income. Between income levels Y′ and Y″, the impact of a $1 increase in the value of food coupons distributed on household food-at-home purchases equals the response resulting from a $1 increase in household income.

IMPLICATIONS FOR ESTIMATION

Figure 1 indicates that the relationship between food-at-home expenditures and income for participant households is discontinuous, contrary to assumptions of past empirical studies. Spline functions could be used to capture the discontinuity between food-at-home expenditures and household income for FSP participant households (*9*). Alternatively, food stamp households spending no more than the cash value of food coupons received on food at home could be excluded from the total sample of participants. Both approaches require identifying FSP-participant households spending no more than the cash value of food coupons received on food at home. However, existing household survey data are inadequate for this purpose.

Another approach is to segment households into participants and eligible nonparticipants. Food-at-home purchase data for eligible nonparticipant households could be used to estimate the relationship between food-at-home purchases, household income, and other household characteristics prior to participation in the FSP. This relationship could provide estimates of participants' food-at-home purchases prior to participation in the FSP. A comparison of these estimates with data on actual food-at-home purchases of participants would provide an estimate of the FSP's impact on food-at-home purchases. This approach does not ignore the discontinuity between food-at-home purchases and household income for FSP participants. Thus, it should provide better estimates of the FSP's impact on household food-at-home purchases.

CONCLUSIONS

A transfer in the form of food coupons should be more effective in increasing food purchases than would a cash transfer of the same value for very low income households, based on this study's findings. For households with incomes at the upper income eligibility bound, a transfer in the form of food stamps would probably be no more effective than a cash transfer. In addition, a FSP containing a purchase requirement is likely to be more effective in increasing food purchases per dollar distributed among participants than one without such a requirement.

Increasing the value of food coupons distributed has impacts that vary depending on household income. For very low income households, a $1 increase in the value of food coupons received will increase food-at-home purchases by $1.

For very low income households, a $1 increase in the value of food coupons received will increase food-at-home purchases by $1.

For participant households with incomes at the upper eligibility bound, such an increase will likely result in an increase in food-at-home expenditures equal to the marginal propensity to spend on food at home out of ordinary income.

Previous estimates of the impact of the FSP on household food-at-home purchases may be misleading because earlier studies did not allow for the possibility that the relationship between participants' food-at-home expenditures and their income is not continuous. A household food-expenditure survey containing monthly food purchases or the value of food stamps used to purchase food would provide more accurate estimates of the overall impact of the FSP on household food purchases and also of the FSP's impact on food purchases of particular subgroups of participants. Analyses which segment households into participants and eligible nonparticipants should also provide more accurate estimates of the FSP's impact on household food-at-home purchases.

REFERENCES

(*1*) Benus J., J. Kmenta, and H. Shapiro. "The Dynamics of Household Budget Allocation to Food Expenditures." *Rev. Econ. and Stat.* Vol. 57, 1976, pp. 129-138.

(*2*) Hymans, S. H., and H. T. Shapiro. "The Allocation of Household Income to Food Consumption." *J. Econometrics*, Vol. 4, 1976, pp. 167-188.

(*3*) Madden, Patrick J., and Marion Yoder. *Program Evaluation: Food Stamps and Commodity Distribution in Rural Areas of Central Pennsylvania.* Pa. State Agr. Expt. Sta. Bull. 780, June 1972.

(*4*) Mittlehammer, R., and D. West. "Food Stamp Participation Among Low-Income Households: Theoretical Considerations of the Impact on the Demand for Food." *So. J. Agr. Econ.* Vol. 7, 1975, pp. 223-231.

(*5*) Neenan, Pamela H., and Carlton G. Davis. "Impact of the Food Stamp Program on Low Income Household Food Consumption in Rural Florida." *So. J. Agr. Econ.* Vol. 9, 1977, pp. 89-97.

(*6*) Reese, Robert B., J. Gerald Feaster, and Gary B. Perkins. *Bonus Food Stamps and Cash Income Supplements.* Econ. Res. Serv., U.S. Dept. Agr., Res. Rpt. 1034, Oct. 1974.

(*7*) Smeeding, T. "Measuring the Economic Welfare of Low-Income Households, and the Antipoverty Effectiveness of Cash and Noncash-Transfer Programs." Unpubl. Ph. D. disser., Univ. Wisc., Madison.

(*8*) Southworth, Herman. "The Economics of Public Measures to Subsidize Food Consumption." *J. Farm Econ.* Feb. 1945.

(*9*) Suits, Daniel B., Andrew Mason, and Louis Chan. "Spline Functions Fitted by Standard Regression Methods." *Rev. Econ. and Stat.* Vol. 70, 1978, pp. 132-139.

(*10*) West, Donald A., and David W. Price. "The Effects of Income, Assets, Food Programs, and Household Size on Food Consumption." *Am. J. Agr. Econ.* Vol. 58, 1976, pp. 725-730.

RESEARCH REVIEW
INDEX NUMBERS AND CHANGES IN FOOD PRICES

By R. McFall Lamm, Jr.*

INTRODUCTION

During periods of rapid inflation it is frequently asserted that fixed weight price indices overstate the magnitude of changes in the cost-of-living as consumption patterns change over time. This is an erroneous proposition, however, because most price indices are designed to measure changes in prices and not to serve as indicators of cost-of-living change. Even so, many fixed weight indices such as the Consumer Price Index (CPI) are often improperly utilized as cost-of-living measures.

A more relevant topic in the measurement of price change involves the appropriate type of weighting scheme to use in index number construction. Fixed base-period quantity weights (Laspeyres) are used most often because of ease of construction.[1] Variable current-period quantity weights (Paasche) are used infrequently since time series data on both quantities and prices are necessary for construction.

This note compares changes in food prices implied by alternative index number specifications. Laspeyres, Paasche, and Fisher's Ideal Index are calculated for a market basket of 42 basic foods over 1964-77. The results imply that the use of variable weight price indices does not lead to a significant restatement of food price change. For this reason, most U.S. Department of Agriculture (USDA) statistics based on Laspeyres indices (such as the market basket) would not be altered substantially if computed with variable weights.

ALTERNATIVE PRICE INDICES

Virtually all price indices that USDA currently uses are Laspeyres indices: Prices Received by Farmers, Prices Paid by Farmers, Market Basket Statistics, most food consumption and production indices, and Farm to Retail Price Spreads—all of which utilize fixed base-period weights. The Consumer and Producer Price Indices, constructed by the U.S. Department of Labor's Bureau of Labor Statistics (BLS) (*2*),[2] are also Laspeyres. The Implicit Price Deflator, constructed by the U.S. Department of Commerce's Bureau of Economic Analysis, is perhaps the most important Paasche index currently used on a large scale[3]. The major reason for the dominant use of Laspeyres indices is that they are the simplest to construct and maintain; only additional price data are required following a survey in the base year to determine quantity weights.

Laspeyres indices measure price change under the assumption that the same market basket of goods consumed in the base year is consumed in subsequent years. The Laspeyres index is calculated as follows:

$$L = \frac{\Sigma\, q_0 p_1}{\Sigma\, q_0 p_0}$$

where L is the value of the index, the q's are quantities of goods consumed, the p's are the corresponding prices, and the subscripts denote time period with 0 representing the base year.

The Paasche index allows quantity weights to change each year. It represents price changes under the assumption that the same market basket of goods consumed this year was consumed in the base year. The Paasche index is computed as follows:

$$P = \frac{\Sigma\, q_1 p_1}{\Sigma\, q_1 p_0}$$

Since household consumption patterns change from year to year, but only to a limited extent, actual changes in prices may lie somewhere between these two extremes. One index used to approximate partial adjustment in quantity weights is Fisher's Ideal Index, which is calculated on the basis of the Laspeyres and Paasche indices as follows:

$$F = (L \cdot P)^{1/2}$$

This index satisfies a weak set of the five conditions required of index numbers as proposed by Fisher (see Eichhorn (*3*) for a discussion).

*The author is an agricultural economist with the National Economics Division, ESCS.

[1] A true cost-of-living index requires that utility be constant over time. The Laspeyres index is the cost-of-living index corresponding to a fixed-coefficients utility function which allows no substitution among commodities—an unlikely representation of consumer behavior.

[2] Italicized numbers in parentheses refer to items in References at the end of this note.

[3] See (*1*) for a discussion of how the Implicit Price Deflator is constructed.

METHODOLOGY

So that the effects of using variable-quantity indices as measures of changes in food prices could be determined, Laspeyres, Paasche, and Fisher's Ideal indices were computed for a market basket of 42 foods sold at retail over the 1964-77 period. Quantity data were obtained from *Food Consumption, Prices, and Expenditures* and other USDA sources. Prices were obtained from series provided by BLS. All quantity statistics used were expressed on a per capita consumption basis in pounds, and prices were expressed as dollars per pound. The base period for prices was 1967, the same base as for the CPI. The base period for quantities was 1972, the first year of the most recent BLS expenditure survey used to obtain quantity weights for the CPI.

Table 1 lists the 42 foods included in the index number calculations. These foods represent most of the basic foods consumed domestically and more than 50 percent of consumer food expenditures (for food consumed at home) in each year of the study period. Some important foods are omitted, however, because of data limitations.

RESULTS

Table 2 presents computed index numbers for the 42-food market basket, as well as corresponding values of the CPI for food. Strikingly apparent is the close association between the Laspeyres and Paasche indices for the 42-food basket. The indices are highly colinear, with a correlation coefficient of 0.999. The same is true for the relationship between Fisher's Ideal Index, and the Laspeyres and Paasche indices. Further, the CPI for food increases more rapidly than do the 42-food market basket indices. This most likely occurs because the composition of the indices differ and because all food consumed away from home is reflected in the CPI for food.

The strong relationship between the three constructed indices emerges even more clearly if percentage changes are compared over time (table 3). Indications are that the indices are highly correlated with changes of similar magnitude in years of declining and increasing food prices. A substantial discrepancy does occur in 1977, however; the Laspeyres index increases 3.9 percent, while the Paasche index rises only 2.8 percent.

Although the Laspeyres and the Paasche indices are closely related for the 42-food market basket, it is not clear that this would be true for the major components of these indices. Thus, Laspeyres and Paasche indices were computed for meats, based on the beef and pork data from the 42-food market basket. The results of this exercise indicate that, as with the aggregate basket, both the Laspeyres and Paasche indices are highly colinear. The largest discrepancy is a 0.7-percent difference occurring in 1975 when meat prices rose 12.5 percent based on the Laspeyres index and 11.8 percent based on the Paasche index.

Table 1 — Market basket foods

Meats	Frozen broccoli	*Dairy products*
Beef	Potatoes	Eggs
Pork		Butter
Chicken	*Fruits*	Cheese
Turkey	Bananas	Evaporated milk
	Grapes	Fluid milk
Vegetables	Strawberries	Fluid lowfat milk
Frozen french fries	Canned pears	Ice cream
Tomatoes	Frozen orange juice	
Asparagus	Canned fruit cocktail	
Cabbage	Watermelon	*Other*
Carrots	Oranges	Rice
Celery	Grapefruit	Sugar
Lettuce	Apples	Wheat flour
Spinach		Roasted coffee
Canned peas	*Fats and oils*	Instant coffee
Canned tomatoes	Cooking oil	Tea
Dry beans	Margarine	

Table 2 — Food price indices, 1964-77

Year	Laspeyres	Paasche	Fisher's Ideal	CPI for food
		(1967=100)		
1964	94.1	94.0	94.0	92.4
1965	97.2	97.3	97.2	94.4
1966	102.1	102.0	102.1	99.1
1967	100.0	100.0	100.0	100.0
1968	103.6	103.5	103.6	103.6
1969	110.3	110.1	110.2	108.9
1970	114.3	114.2	114.2	114.9
1971	116.1	115.8	115.9	118.4
1972	123.1	123.1	123.1	123.5
1973	148.3	147.8	148.1	141.4
1974	168.0	167.9	168.0	161.7
1975	180.9	180.0	180.5	175.4
1976	178.0	177.6	177.8	180.8
1977	184.9	182.6	183.7	192.2

Table 3 — Changes in alternative food price indices, 1965-77

Year	Laspeyres	Paasche	Fisher's ideal	CPI for food
		Percent		
1965	3.3	3.5	3.4	2.2
1966	5.1	4.9	5.0	5.0
1967	−2.1	−2.0	−2.0	0.9
1968	3.6	3.5	3.6	3.6
1969	6.4	6.4	6.4	5.1
1970	3.6	3.7	3.6	5.5
1971	1.6	1.4	1.5	3.0
1972	6.0	6.3	6.2	4.3
1973	20.5	20.1	20.3	14.5
1974	13.3	13.6	13.4	14.4
1975	7.7	7.2	7.4	8.5
1976	−1.6	−1.3	−1.5	3.1
1977	3.9	2.8	3.3	6.3

CONCLUSION

Using Paasche variable quantity indices does not, based on these results, lead to a substantial restatement of food price change. This is true both for a broad basket of foods, as well as for a subgroup of related foods such as meats. For this reason, the large costs involved in collecting additional data to construct Paasche or other variable quantity indices may not be justified. These findings are similar to those of other studies indicating that computing the CPI for all items using the Paasche index would not lead to results significantly different from those obtained with the Laspeyres index.

In addition, stability over time characterizes domestic food consumption patterns. If this were not the case, changes in the Paasche and Laspeyres food price indices would be less highly correlated.

REFERENCES

(*1*) Backman, J., and M. Gainsbrugh. "Inflation and the Price Indices." Joint Economic Committee, U.S. Congress, 1966.

(*2*) Bureau of Labor Statistics. "The Consumer Price Index: Concepts and Content Over the Years." U.S. Dept. Labor, 1978.

(*3*) Eichhorn, Wolfgang. *Functional Equations in Economics.* Reading, Mass. Addison-Wesley, 1978.

INTRADAY COMMODITY PRICE MOVEMENTS

By Jitendar S. Mann*

INTRODUCTION

Sharp changes in commodity prices in the past few years have focused attention on the need to understand shortrun price movements better. The behavior of daily price changes has been widely studied.[1] This note analyzes the movements of intraday prices.

Research on intraday prices began in 1937 when Irwin tried to identify different kinds of trading on the floor of a commodity exchange. Irwin was interested in the impacts of speculation, manipulation, and movement trading on price changes (*5*). According to Irwin, there is a:

> tendency for the speculative operations to center approximately upon the price justified by the conditions existing at the time, whereas movement trading has no such a check. This difference accounts for the tendency of true speculation to stabilize prices and for the tendency of movement trading to widen price swings.

Working, studying the actual behavior of traders, tried to establish a relationship between price movements and the behavior of scalpers and day traders (*13*, *15*). He concluded that the following major difference existed between scalping and day trading: "In scalping, the interval between purchase and sale, or between a sale and a subsequent purchase, is ordinarily not more than a few minutes. In day trading, the interval may be an hour or more." Working characterized the scalper as one who stands ready to buy at 1/8 cent below the last price or to sell at 1/8 cent above it.[2] For successful scalping, any small price change should be followed by a price change in the opposite direction. Working called this tendency for price reversals "price jiggling." Recently, interest in intraday price movements has revived because of the problem of dual trading; that is, the floor brokers and future commission merchants trade for their own accounts as well as for those of customers.

Olson analyzed the participation of floor trades in intraday price movements of potatoes on the New York Mercantile Exchange (*11*).[3] There are several categories of trades in the commodity pit. Intraday price analysis may help to explain the activities of floor traders and scalpers. For example, too many reversals (in contrast to continuations) provide scalpers an opportunity to buy and sell on small price changes.

The data used in this analysis are from a computer system maintained by the Chicago Mercantile Exchange. Prices are collected by exchange pit clerks, who on hearing the price of a trade, a higher bid, or a lower offer than the prevailing market price, record the information on cards which in turn are time-stamped and given to keyboard operators to enter into the computer. If the market is not active and the price does not change, the last price is repeated after a specific time interval. However, if trading is extremely active, not all price changes are recorded. Thus, the prices are not strictly one price for each transaction. The prices analyzed here are for the July 1975 contracts of frozen pork bellies for all trading days during July 1975. Each observation during July 1975 has been included in the analysis.[4]

THE RANDOM WALK HYPOTHESIS

This study tests the hypothesis that intraday commodity prices behave like a random walk. In a random walk, a price series follows the stochastic process:

$$P_t = P_{t-1} + E_t$$

where E_t is an independent random variable with zero mean. Working's theory of anticipatory prices outlines the basic process for random walk in futures prices (*14*). In an efficient competitive market, price is determined by the actions of many traders, each acting based on expectations. Traders' expectations, in turn, are based on information from diverse sources. As prices reflect expectations, new information affects prices only to the extent that it differs from what was previously anticipated. The price-making mechanism starts with a specific opening price and adds to it in each interval

---

*The author is an agricultural economist in the International Economics Division, ESCS.

[1] Italicized numbers in parentheses refer to items in References at the end of this note.

[2] Working was referring to wheat.

[3] See also Niederhoffer and Osborne (*10*) for a study of intraday prices in stock markets.

[4] Analysis of the February 1976 contract for pork bellies, traded during the latter half of July 1975, gave results similar to those reported in this note.

a random factor, E_t, which encompasses the influence of all the new information available to generate the next price. All currently available information is incorporated into each successive price change. The best expected price for the next period is the current price; past price history is irrelevant.

Muth states that in the very short run "if the production and consumption flows are negligible compared with the speculative inventory level, the process approaches a random walk" (*9*).

In the very short run, when the demand and supply of contracts depend only on price movements, the price will follow a random walk. Let:

$$D_t = b \Delta P_t + G_t$$

$$S_t = g \Delta P_t + N_t$$

where D, S, and ΔP are respectively demand, supply and price change; *b* and *g* are constants; and G and r are random errors. Equating demand and supply, one obtains:

$$\Delta P_t = \frac{N_t - G_t}{b - g}$$

which is a random walk, a linear combination of independently distributed random variables.

The empirical interpretation of the random walk hypothesis is that the price differences are temporally independent. The price change following a given transaction is not influenced by the sequence of past price changes.

A more precise statement of the requirements for a "fair" market is provided by the Martingale hypothesis, which requires only that the conditional expectation of ΔP_t be zero.

$$E(\Delta P_t \mid \Delta P_{t-1}, \Delta P_{t-2} \ldots) = E(\Delta P_t) = 0$$

Successive ΔP_t may be drawn from different distributions as long as their means are zero. A "fair" market is defined as a market where no trader can profit from predicting price movements based on past observations.[5]

STATISTICAL ANALYSIS

To test the random walk hypothesis, I constructed a joint frequency distribution of consecutive price changes.[6] The results showed only 285 cases of a price change—142 declines and 143 rises—equal to the minimum of 2.5 cents per cwt. The price changes involving half-cents were generally less frequent. Price changes clustered around multiples of 5 cents. Price changes of 10 cents were more frequent than those of 5 cents.

To guarantee that there was no zero entry (table 1), I collapsed the frequency table of successive price changes and calculated a matrix of transition probabilities (table 2). This stochastic matrix gives the probabilities of each of the current price changes (items in boxheads of table 2), given a certain last price change (stub entries in table 2). The tendency for large changes to be followed by small changes is apparent when one examines the largest probability for each row. Table 2 shows, for example, that in all cases of a price decline between 12.5 and 10.0 cents, the ratio of another similar decline was 0.114 (=29/255).

This change pattern is highlighted by table 3, which has been abridged to give only the direction of change.[7] Although it appears that the number of positive and negative changes were equal, the tendency for changes of the same sign to follow each other is more frequent. A Chi-square test for a 2 × 2 table rejected the hypothesis that price changes were independent. The number of cases where a change was followed by a change in the same direction (continuity) was 1,933, whereas the number of reversals was 1,360. In the series as a whole, 41 percent of cases were reversals.[8]

Under the random walk hypothesis, the probabilities of a price change are not influenced by the past price changes; that is:

$$\text{Prob}(\Delta P_t = X \mid \Delta P_{t-1}, \Delta P_{t-2} \ldots) = \text{Prob}(\Delta P_t = X).$$

[5] See Samuelson (*12*).

[6] The pork belly prices are quoted in cents per pound. A contract equals 38,000 pounds. The minimum price change was 2.5 cents per cwt (*3*). This note uses cents per cwt.

[7] Cases with zero changes have been excluded from table 3.

[8] Working reported a tendency to reversals in intraday price movements for Chicago wheat (*13, 15*). He studied 143 series of 100 successive price changes during 1927-40 and found that 140 of the series had 65 percent or more reversals.

Table 1 — Frequency of successive price changes, frozen pork bellies, July 1975 contract*

Last price change (P(t−1)) (cents/cwt.)	This price change (P(t)) (cents per cwt.)							
	−15 and less	−12.5 and −10.0	−7.5 and −5.0	−2.5 to 2.5	5.0 and 7.5	10.5 and 12.5	15.0 and more	Total
				Number				
−15 and less	1	6	13	2	29	16	4	71
−12.5 and −10.0	6	29	75	15	111	13	6	255
−7.5 and −5.0	15	68	595	180	269	129	18	1,274
−2.5 to 2.5	5	32	149	283	187	24	4	684
5.0 and 7.5	23	93	295	177	633	95	13	1,329
10.5 and 12.5	19	21	122	27	87	35	6	317
15.0 and more	2	6	25	1	12	6	3	55
Total	71	255	1,274	685	1,328	318	54	3,985

*During July 1975.

Table 2 — Matrix of transition probability of successive price changes, frozen pork bellies, July 1975 contract*

Last price change (P(t−1)) (cents/cwt.)	This price change (P(t)) (cents per cwt.)								
	−15.0 and less	−12.5 and −10.0	−7.5 and −5.0	−2.5 to 2.5	5.0 and 7.5	10.5 and 12.5	15.0 and more	Marginals	Chi-square
−15.0 and less	0.014	0.084	0.183	0.028	0.408	0.225	0.056	0.018	42.00
−12.5 and −10.0	.024	.114	.294	.059	.435	.051	.024	.064	42.23
−7.5 and −5.0	.012	.053	.467	.141	.211	.101	.014	.320	162.84
−2.5 to 2.5	.007	.047	.218	.414	.273	.035	.006	.172	345.75
5.0 and 7.5	.017	.070	.222	.133	.476	.072	.010	.334	135.28
10.5 and 12.5	.060	.066	.385	.085	.274	.110	.019	.080	56.77
15.0 and more	.036	.109	.454	.018	.218	.109	.054	.014	22.48

*For July 1975.

This means that the probabilities in each set of entries for each stub item (each row) should be independent of each other. This hypothesis is tested by a Chi-square test recommended by Anderson and Goodman (*1*). The null hypothesis is that the probabilities in each row are equal to the marginal probabilities. The estimated value of Chi-square appears below the last boxhead items (column in) table 2. The null hypothesis (of independent rows) is rejected for each of the seven rows. A hypothesis that both rows and columns are independent was also rejected (results are not shown here).

Another null hypothesis tested was that conditional probabilities of a price change, given a past price change, are constant (and equal), which means that each probability in a row equals 1/7. This hypothesis was also rejected for each row.

Table 3 — Direction of change of successive changes, frozen pork bellies, July 1975 contract*

Last price change	This price change		
	Negative	Positive	Total
Negative	922	678	1,600
Positive	682	1,011	1,693
Total	1,604	1,689	3,293

*During July 1975.

CONCLUSION

Although the random walk hypothesis for price movements has been well accepted in academic circles,[9] the empirical work is based on analysis of daily price changes. An analysis of intraday price movements leads to a rejection of the random walk hypothesis. Niederhoffer and Osborne reach a similar conclusion from a study of intraday stock prices (*10*). Martell and Helms indicate that their analysis of intraday prices leads to conclusions different from those based on daily closing prices (*8*).

A possible explanation for lack of randomness in intraday price movements is that not all price changes result from discounting of information. The floor traders, who trade on their own account, make trades based on past and expected price movements. This behavior adds a nonrandom element to price changes.

---

[9]See (*4*).

REFERENCES

(*1*) Anderson, T. W., and Leo A. Goodman. "Statistical Inference about Markov Chains," *Ann. Math. Stat.*, Vol. 28, 1957, pp. 89-110.

(*2*) Cargill, Thomas F., and Gordon C. Rausser (1975), "Temporal Behavior in Commodity Futures Markets," *J. Finance*, Vo. 30, 1975, pp. 1,043-1,053.

(*3*) Chicago Mercantile Exchange, Market News Department. *Yearbook 1974-75.*

(*4*) Cootner, Paul H., *The Random Character of Stock Market Prices*, The MIT Press, Cambridge, Mass., 1964.

(*5*) Irwin, H. S. "The Nature of Risk Assumption on Organized Exchanges," *Am. Econ. Rev.*, Vol. 27, 1937, pp. 267-278.

(*6*) Kendall, Maurice G., and Alan Stuart. *The Advanced Theory of Statistics*, Vol. III, Hafner Pub. Co., New York, 1968.

(*7*) Mann, Jitendar S., and Richard G. Heifner. *The Distribution of Short-run Commodity Price Movements*, Tech. Bull. 1536. U.S. Dept. Agr., Econ. Res. Serv., 1976.

(*8*) Martell, Terrence F., and Billy P. Helms. *A Re-examination of Price Changes in Commodity Futures Markets*, Univ. Alabama (undated).

(*9*) Muth, John F. "Rational Expectations of the Theory of Price Movements," *Econometrica*, Vol. 29, 1961, pp. 315-335.

(*10*) Niederhoffer, Victor, and M. F. M. Osborne. "Market Making and Reversal on the Sotck Exchange, *JASA*, Vol. 61, 1966, pp. 897-916.

(*11*) Olson, Wayne L. *The Extent and Price Effects of Floor Trading in the New York Mercantile Exchange Potato Futures Markets*, U.S. Dept. Agr., Commodity Exchange Authority, 1970.

(*12*) Samuelson, Paul A. "Proof that Properly Anticipated Prices Fluctuate Randomly," *Ind. Man. Rev.*, Vol. 6, 1965, pp. 41-49.

(*13*) Working, Holbrook. "Price Effects of Scalping and Day Trading," *Proceedings of the Seventh Annual Symposium: Commodity Markets and the Public Interest*, Chicago, 1964, pp. 114-143.

(*14*) ________. "A Theory of Anticipatory Prices," *Am. Econ. Rev., Papers and Proceedings*, Vol. 48, 1958, pp. 188-199.

(*15*) ________. "Tests of a Theory Concerning Floor Trading on Commodity Exchanges," *Food Res. Inst. Studies*, Vol. 7, 1967, pp. 5-48.

CAQUEZA: LIVING RURAL DEVELOPMENT

Hubert Zandstra, Kenneth Swanberg, Carlos Zulberti, and Barry Nestel. International Development Research Centre, Ottawa, 1979, 321 pages $15.00.

*Reviewed by David W. Culver**

It is interesting and unusual to find a book that treats a range of development issues while telling the story of a small rural development project in Colombia. And the story is generally convincing.

The book and, to a significant extent, the project itself are products of the International Development Research Centre (IDRC) of Canada. In addition to publishing the book, IDRC provided the outside funding (U.S. $908,000) for the 5 years covered in the book as well as the expatriate advisors—initially two and later three. Of the four authors, three served as IDRC advisors on the project, while the other was an associate director of IDRC's Agriculture, Food, and Nutrition Sciences Division from 1970 to 1976, during which time he negotiated and managed IDRC's involvement in the Caqueza project.

The book is divided into five parts. Part 1 covers the experience and ideas on rural development drawn from various countries; the status of agricultural research and extension in Colombia; and characteristics of the Caqueza project area. Chapter 1 reviews four specific rural development projects which made "some significant progress . . . (on) the question of how to transform existing institutions so as to enable society to capture the economic gains implicit in new technological alternatives . . ." The four projects are Borgo Mozzano in Italy, the Comilla Project in East Pakistan (now Bangladesh), the Intensive Agricultural District Program in India and the Puebla Project in Mexico. Chapters 2 and 3 are helpful as background for the Colombia story but are not otherwise of general interest.

Part 2 describes the 5 years of experience in the Caqueza project. Each chapter is organized identically; the sections on organization and programming, research, dissemination, and evaluation are probably the most important. The closing section of each chapter, a brief "resùme of the year," is helpful. The mixture of frustration and hope, of personal and organizational conflict, the apparent resourcefulness and flexibility of project staff combine so as to keep the reader eagerly pressing to discover what lies ahead.

The two chapters in part 3 interpret the research activities of the project. The first discusses the evolving methodologies used and describes the work on understanding existing production systems. The second summarizes those experiments testing the value of recommended practices. An interesting point is the gradual recognition by members of the project staff of the importance of economic factors in the farmers' production decisions, including their adoption of new technology. Most staff members were trained in one of the biological sciences, especially agronomy. While that may seem logical for research and extension focused mainly on crops, the economist reader will probably be amused by the gradual process through which the staff learn about the role of prices and price variability.

Part 4 examines factors which relate to technology-adopting rates—risk, credit, marketing, training, and buffer institutions. The chapters on risk and credit are good. However, the chapter on marketing is weak—the low point of the book for this reader. For example, one of the discussions relates to a comparison between atomistic (competitive) markets—as with crops in Caqueza—and oligopolistic markets (pp. 236-237). One of its arguments is that "the atomistic market system is a dumping ground for the unemployed," with the clear implication that an oligopolistic system would be preferable. The authors do not discuss the impact of such a change on overall efficiency. They imply that since the system has many marketing agents, there is also overcapacity in transport equipment. And they offer the "corollary" that "returns to the marketing agents were below the opportunity cost for the value of the services that they rendered." They do not indicate the available alternatives nor calculate opportunity costs for these unwanted marketing agents. Nor do they recognize the apparent conflict with their earlier description of the marketing system as "a dumping ground for the unemployed." Perhaps there is good reason why the marketing plan developed by project staff failed.

Part 5 deals with measuring achievements. Although one must be cautious about how people evaluate their own work, as in this case, this section offers a reasonable view. Even with some bias allowed for, it is likely that the project was relatively successful.

Most of the book is pleasant to read. There are only a modest number of typographical or other mechanical irritants, although the use of abbreviations and acronyms is excessive. There is a rather lengthy bibliography but no index.

*The reviewer is an agricultural economist with USDA's Foreign Agricultural Service.

In Earlier Issues

For forty years Wesley C. Mitchell was a pathfinder in business-cycle research—the analysis of the processes of expansion, recession, contraction, and revival . . . There is considerable dissatisfaction on the part of some with the progress in Mitchell's approach of painstaking examination series by series, as well as the inadequacies of the historical approach toward the development of the comprehensive theory of business cycles. The viewpoint of the econometricians (is) that a system of structural equations can be developed which will describe the operations of the economy and the theory of business fluctuations . . . (One test) found that (an econometric) model fared no better than a "naive" model which simply extrapolated the value of each variable.

Nathan M. Koffsky
Vol. IV, No. 4, pp. 142, 143
October 1952

Many extremes have occurred in our economic situation and activity. There have been eras of great prosperity, severe depression, recovery, war, and postwar readjustment . . . Economic forecasting is difficult because of the large number of factors and the complexity of relationships that influence the economic system . . . Federal Government economic forecasts that relate to agriculture over a 30-year period received an accuracy evaluation score of 76 on a scale which ranged from 100 for perfect forecasts to 0 for totally wrong forecasts, with 50 for the expectation from pure guessing.

John D. Baker, Jr. and Don Paarlberg
Vol. IV, No. 4, pp. 105, 107, 114
October 1952

It is obviously a waste of time and money to address envelopes and to mail questionnaires to people who do not return them. The statistical aspects of the problem are even more serious. Individuals who receive mail questionnaires are more likely to fill them out and return them when they have had some previous personal contacts with the agency.

Cecil C. Smith
Vol. IV, No. 4, p. 126, October 1952

Progress toward the solution of the world food problem is a stern and urgent challenge to Western Civilization . . . The best immediate prospects are in areas of high present economic activity and well developed education in technology. But the long-time view is a different matter involving the much more difficult tropical lands.

Charles E. Kelley
Vol. IV, No. 4, pp. 135, 136
October 1952

*U.S. GOVERNMENT PRINTING OFFICE : 1980 O-310-945/ESCS-139

Suggestions for Submitting Manuscripts for *Agricultural Economics Research*

Contributors can expedite reviewing and printing of their papers by doing these things:

1. SOURCE. Indicate in a memorandum how the material submitted is related to the economic research program of the U.S. Department of Agriculture and its cooperating agencies. State your own connection with the program.
2. CLEARANCE. Obtain any approval required in your own agency or institution before sending your manuscript to one of the editors of *Agricultural Economics Research*. Attach a copy of such approval to the manuscript.
3. ORIGINALITY OF MATERIAL. It is our policy to print original material. We consider alternative treatments of material published elsewhere, but such treatments need to differ substantially from the original approach. When submitting your manuscript, identify, if applicable, related material either published or submitted for publication.
4. ABSTRACT. Include an abstract and at least three keywords when you submit your article. The abstract should not exceed 100 words.
5. NUMBER OF COPIES. Submit three good copies.
6. TYPING. Double space everything, including abstract and footnotes.
7. FOOTNOTES. Number consecutively throughout the paper.
8. REFERENCES. Check all references carefully for accuracy and completeness.
9. CHARTS AND OTHER ARTWORK. Use charts sparingly for best effect. Keep design as simple as possible to improve communication. Submit all artwork in draft rather than final form, accompanied by neatly prepared pages with essential data for replotting.

The Secretary of Agriculture has determined that the publication of this periodical is necessary in the transaction of public business required by law of this Department. Use of funds for printing this periodical has been approved by the Director, Office of Management and Budget, through February 28, 1985.

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 204202. $6.50 a year domestic; $8.25, foreign.

Microfiche copies are available from two different sources: Microfilming Corporation of America, 21 Harristown Road, Glen Rock, NJ 07452, 1974 on: $4.95 per year or per issue; Congressional Information Service, Inc., P.O. Box 30056, Washington, D.C. 20014, 1973: $5 (2 fiche), 1974 on: average is 3 fiche per year ($3.75 first fiche, $1.25 each remaining fiche.)

UNITED STATES DEPARTMENT OF AGRICULTURE
WASHINGTON, D.C. 20250

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF
AGRICULTURE
AGR 101
THIRD CLASS

U.S. MAIL