January 1983

A.E. Res. 83-7

A REVIEW OF EMPIRICAL TECHNIQUES FOR THE ANALYSIS OF COMMODITY INSTABILITY

BY

Susan E. Offutt and David Blandford

Department of Agricultural Economics
Cornell University Agricultural Experiment Station
New York State College of Agriculture and Life Sciences
A Statutory College of the State University
Cornell University, Ithaca, New York 14853

A REVIEW OF EMPIRICAL TECHNIQUES FOR THE ANALYSIS OF COMMODITY INSTABILITY

by

Susan E. Offutt and David Blandford*

INTRODUCTION

Instability in commodity markets raises questions of magnitude, causes, consequences, and control. Economic analysis seeks to provide answers to these questions. Often, it is the need for policy evaluation and guidance that motivates the empirical analysis of instability. In U.S. agriculture, for example, the instability problem is a salient one, as Burnstein points out.

> In a competitive market characterized by relatively inelastic supply and demand schedules and a low income elasticity of demand, small quantity changes in agricultural goods induce disproportionate changes in price. The recent increase in price fluctuations of agricultural commodities has provoked a re-examination of our social tolerance to price instability and has forced policymakers to consider the social desirability of governmental intervention to reduce price fluctuations. (p. 1)

A prerequisite for policy action is the characterization and measurement of instability, and the identification of sources and means of control. This bulletin describes and analyzes a number of empirical techniques which may be used to achieve these ends. It seeks to identify the limitations and the strengths of the various techniques, and to act as a guide for selecting the most appropriate for a particular application.

The bulletin is divided into three sections. The first considers single variable measures of instability, which can be used to answer such questions as:

- for several commodities, which displays the most instability in price, revenue, yield, acreage, or output; and,

- for a single commodity, which of price, revenue, etc. is most unstable.

---

* Susan Offutt is an assistant professor in the Department of Agricultural Economics, University of Illinois, Urbana, IL 61801. David Blandford is an associate professor in the Department of Agricultural Economics, Cornell University, Ithaca, NY 14853.

Each variable's time series is evaluated separately by a technique that describes its behavior without reference to functional or behavorial relationships to other variables. Examples of these univariate techniques would be moving averages and coefficients of variation. It is shown that the assessment of relative instability can be dependent upon the choice of indicator.

The methods of the next section seek to identify the sources of instability in variables that can be expressed as a function of other variables through multiplicative or additive identities. For example, output is the product of acreage and yield. Total output can be expressed as the sum of all individual farm or country outputs in a market. Empirical techniques are available to apportion the variance of say, revenue, into that attributable to the variance of price, of output, and to the interaction between the two. An extension of this methodology identifies the source of the variability in the underlying supply and demand schedules. The analysis in this section emphasizes the potential pitfalls involved in making policy recommendations based on such simplistic empirical decomposition of variability.

Once the sources of instability have been uncovered, there may exist a need to determine the means for its control in order to achieve some policy goal. For this purpose, the third section considers behavioral relationships which go beyond identities and provide information on causality through the use of multivariate regression and multiplier analysis. For example, if the analysis of the identity shows output to be responsible for more of the fluctuation in revenue than is price, and if the interaction between price and output is positive in direction, then policy actions to stabilize revenue should probably center on output. However, this is insufficient information with which to formulate policy because the causes of output fluctuation are not known. Variability might be largely due to stochastic weather effects, or it might be attributable to changes in production decisions in response to changes in prices of competing crops. Consequently, the choice of policy instruments may be difficult in the absence of this information. The section examines the means by which the more important, and controllable, sources of instability may be identified.

I. Single Variable Measures

Several different single variable measures or indicators of variation have been used in the literature on commodity instability. In a number of studies indicators have been used to determine the relative variability of commodity prices or value, and the results employed to evaluate or guide stabilization policies (Coppock; UN; UNCTAD). In most cases, the choice of indicator is treated as incidental, if it is addressed at all. This presents no problem if all indicators provide the same assessment of relative variability. However, if the conclusions drawn are sensitive to the measure used, then selection is an important consideration.

The comparison of the various measures is of more than academic interest since policy recommendations can be based on the results of empirical analysis. For example, McNicol states that "commodity prices are much less stable than the prices of manufactures... an impression that is confirmed by the coefficients of variation" (p. 16). He construes this result to provide support for a primary commodity price stabilization scheme. As seen below, it is entirely possible that this conclusion could be significantly altered should another indicator of instability be employed. In another application, Massell (1970a) argues that deviations of commodity export earnings from an exponential trend should be the basis for measuring their instability. He asserts that policymakers see constant annual percentage growth in earnings as an acceptable form of earnings' movement through time. However, it could be that, say, linear trend gives a better description of the series' behavior, if better is determined by the reduction of deviations between predicted and actual values. Since the magnitude of the "instability" problem is affected by how much of the variation is assumed to be systematic in origin the choice of trend can be an important consideration.

Knudsen and Parnes briefly considered the tradeoffs involved in the choice of alternative indicators. They used indices based on logarithmic variance, exponential trend and a moving average to rank instability in export earnings for fifty-three countries. The moving average as an indicator of instability yielded results substantially different from the other two. They state,

> If this study were concerned with the absolute level of instability, the differences...would be disturbing. What is low instability under one index is high under another; clearly, this would cause problems at the policymaking level. (p. 49)

In as much as the focus of their study was on explaining relative rather than absolute differences in instability between countries, Knudsen and Parnes concluded that high correlation of rankings among indices was sufficient for the purposes of their investigation.

In this section, a number of single variable measures previously employed to analyze commodity instability are discussed, and the degree to which they provide the same assessment of relative variability is evaluated. The analysis is conducted using acreage, yield, output, price, and revenue data for ten U.S. field crops over the period 1950 to 1977 (USDA).

The concept of instability and its measurement

An unambiguous definition of instability would provide the ideal starting point for the selection of an appropriate empirical indicator. Unfortunately, the concept of instability is nebulus because the perception of what constitutes unstable behavior is largely subjective. It is crucially dependent on who is evaluating the "instability" and what problems he/she views it to present. For example, from a producer's perspective only downward fluctuations in commodity prices may be viewed as a problem because of their effects on revenues, whereas from a consumer's perspective upward fluctuations may be the focus of concern because of their effects on expenditures. From a policymaker's perspective, upward and downward fluctuations in prices resulting from systematic changes in such factors as consumer income may be viewed to be acceptable, since these act as signals for resource allocation. However, fluctuations which are created by stochastic factors such as weather conditions may be viewed with concern.

Since much of the discussion on instability is directed towards the analysis of government stabilization policies, the perceptions of policymakers are particularly important. A given degree of variation in the price of a major commodity (however defined) may not be viewed in the same light as that for a minor commodity. Furthermore, in the case of a single commodity, a given degree of variation in price may not be viewed in the same way as the same degree of variation in revenue.

This brief discussion indicates that the definition of instability is complex. It is clear that variability and instability cannot necessarily be equated and that the measurement of instability requires that an implicit or explicit judgment be made as to what constitutes "acceptable" versus "unacceptable" variability. In many cases, for example, some type of trend is removed from the data before instability is measured, possibly on the grounds that such trend is predictable and does not therefore constitute instability. This clearly reflects some judgment on acceptable variability (that due to trend) and unacceptable variability (deviations from trend).[1/] Only rarely is the rationale behind the particular specification of trend adopted explicitly considered (e.g., Massell, 1970a). Gardner, for example, has stressed the need to be explicit about the rationale for the exclusion of trend and for the types of fluctuations to be treated as instability.

Analysis of Alternative Measures

The single variable measures considered in this section have all been used in previous empirical studies of instability (Offutt). They include percentage range and average period-to-period change measures, moving averages, a logarithmic index developed by Coppock, and several versions of the coefficient of variation. Their formulae and major characteristics are outlined in Table 1. A Fortran computer program to calculate the measures is given in the Appendix.

The measures in Table 1 were applied to acreage, yield, output, price and revenue data on ten U.S. field crops (barley, corn, cotton, oats, rice, rye, sorghum, soybeans, sugarbeets, and wheat) over the period 1950 to 1977.[2/] For each instability indicator, two kinds of rankings were compiled. On a cross-commodity basis, commodities were ranked in descending order of the level of

---

1/ Knudsen and Parnes use the permanent income hypothesis to decompose the variability of export earnings into "permanent" and "transitory" components. They then compute an index of instability using the sum of the squared transitory component normalized by permanent export earnings. This approach is extremely unusual in that it relies on a specific theory to separate acceptable from unacceptable variability.

2/ The moving average measure was calculated using 3 and 5 period lengths.

Table 1. Single Variable Measures of Instability.

Measure	Formula	Major Characteristics
Percentage Range (PR)	$PR = W_M - W_m$ where $W_M = MAX(W_2, \ldots, W_{n-1})$; $W_m = MIN(W_2, \ldots, W_{n-1})$; $W_t = \frac{\lvert V_t - V_{t-1} \rvert}{V_{t-1}} \times 100 \quad t=1, \ldots, n$	No account made for presence of trend; strongly affected by outliers.
Average Percentage Change 1 (APC1)	$APC1 = \frac{\sum_{t=2}^{n} \left\lvert \frac{V_t - V_{t-1}}{V_{t-1}} \right\rvert}{n - 1} \times 100$	Measures average period-to-period change; absolute value in numerator moderates influence of outliers; asymmetrical treatment of increases, which can be greater than 100%, and decreases, which cannot be.
APC2	$APC2 = \frac{\sum_{t=2}^{n} \left[\frac{V_t - V_{t-1}}{V_{t-1}} \right]^2}{n - 1} \times 100$	
APC3	$APC3 = \frac{\sum_{t=2}^{n} \left[\frac{V_t - V_{t-1}}{MAX(V_t, V_{t-1})} \right]^2}{n - 1} \times 100$	Measures average period-to-period change; squaring accentuates influence of outliers; use of $MAX(V_t, V_{t-1})$ in denominator corrects asymmetry in ACP1 and 2.
Moving Average (MA)	$MA = \frac{\sum_{t=r+1}^{n-r} \left\lvert \frac{V_t - V_{t*}}{V_{t*}} \right\rvert}{n + 1 - m}$ where $V_{t*} = \frac{\sum_{t-r}^{t+r}}{n}$; $r = (m - 1)/2$; m = period of moving average.	Coincides with trend increasing by constant amount each period, lies below if increasing by decreasing amounts, and above if by increasing amounts; uses only subset of data to determine trend value, fits data closely; generates oscillatory movement in residual; minimizes cyclic fluctuation, allows specification of length of average, can be symmetric, as shown, or backward looking.

Table 1. (Continued)

Measure	Formula	Major Characteristics
Coppock Index (CI)	$CI = \text{antilog}\left[\frac{1}{n-1}\sum_{t=1}^{n-1}\left[\log\left[\frac{X_t + 1}{X_t}\right] - m\right]^2\right]^{1/2}$ $m = \frac{1}{n-1}\sum_{t=1}^{n-1}\log\left[\frac{X_t + 1}{X_t}\right]$ X_t = log of data's first difference	Linear trend removed by taking first differences; taking logs of first differences moderates influence of outliers in squaring; interpreted as approximation of average period-to-period percentage variation net of trend; m depends only on first and last points in differenced series.
Coefficients of Variation (CV)	$CV(S) = \frac{\left[\frac{\sum_{t=1}^{n}(V_t - \bar{V})^2}{n}\right]^{1/2}}{\bar{V}}$ $\bar{V}$ = mean of V_t	Unit free measure of relative dispersion from variable's mean; no account for trend in mean; results in overstatement of variability relative to nontrending series; no constant range; squaring accentuates outlier effect.
	$CV(D) = \frac{\left[\sum_{t=1}^{n}\frac{\lvert V_t - \bar{V}\rvert}{n}\right]^{1/2}}{\bar{V}}$	Absolute value formulation modifies effects of outliers; more sensitive to difference between dispersed and more compact data series than CV(S); otherwise similar.
Standardized Coefficients of Variation (SCV)	$SCV(S) = \frac{CV(S)}{n-1}$	Division by maximum value of CV(S) gives constant range between zero and one; independent of length of series, facilitates comparison among variables.
	$SCV(D) = \frac{CV(D)}{2(1 - 1/n)}$	Division by maximum value of CV(D) gives constant range between zero and one, independent of length of series, facilitates comparison among variables.

Note: Coefficients of variation are frequently calculated using deviations from trend. In this $\bar{V}$ in the numerator is replaced by $\hat{V}_t$, the predicted value of V derived from the trend line.

instability exhibited in each of revenue, yield, output, acreage, and price. On an intra-commodity basis, the five variables (revenue, etc.) for each of the commodities were ranked in the same way. In this manner, 15 sets of rankings were obtained for each indicator.

The main objective of this empirical application was to discover whether or not the measures provide a consistent assessment of instability. This can be determined by comparing the rankings obtained in each of the two schemes. As a summary measure of the degree of agreement, Spearman correlation coefficients were computed for all relevant pairings. This nonparametric coefficient provides an index of the degree of similarity between two rankings of the same list of items. Its value ranges from positive unity, indicating complete agreement, to negative unity, indicating complete disagreement. Averaging the values of the Spearman coefficient across rankings (intra- and cross-commodity) and across pairs facilitates a general comparison of behavior among and between measures (Table 2).

In general, the first group of measures (PR, APC, MA, and CI, see Table 1 for the key to abbreviations) agree well among themselves in both ranking schemes, with an average Spearman coefficient of 0.81 (Table 2). The coefficients of variation had only a few cases of disagreement among themselves, due mainly to the differences in treatment of outliers. However, the agreement between the first group of measures and the coefficients of variation was fairly low, an average correlation of 0.41 by the Spearman coefficient. The discrepancy seems attributable to the influence of trend in a number of data series. Trend appeared to outweigh any other data characteristic when evaluated by the coefficients of variation. Variables with the strongest trend were ranked most unstable by the coefficients. The other measures seemed more sensitive to outliers and sawtooth-like data features, so jagged series were identified as most unstable by them, practically regardless of the presence of trend. Thus, the coefficients of variation identified as most unstable those data series with smooth but strong trend as opposed to nontrending series with significant negative serial correlation.

Individual measures had some idiosyncratic features which deserve mention. The PR measure recorded its lowest values for series with fairly constant percentage trend, such as yield, and was influenced strongly upward by outliers such as occurred toward the end of most price series (during 1973-1974). The average percentage change measures moderate the influence of these outliers and so agreed most closely with PR for smoother series. Although the moving average (fit with three and five period lengths) accounts for trend, it agrees fairly well with the PR and APC measures. The similarity is due to the flexibility of this average, which uses subsets of the data in determining trend values, such that it is influenced more strongly by outliers than, say, linear regression. This sensitivity to extreme values accounts for its agreement even with non-detrended measures.

The Coppock Index is supposed to yield a close approximation of the period-to-period percentage change adjusted for (linear) trend. Yet, curiously it agrees quite well with the average percentage change measures, which do not account for trend. Furthermore, the agreement is best (0.92 Spearman value)

Table 2. Average Spearman Correlation Coefficients.

	First Group of Measures Among Themselves [a]	First Group of Measures with Coefficients of Variation (as a group)
Intracommodity [b]		
Barley	.75	.52
Corn	.78	.69
Cotton	.77	.44
Oats	.71	-.15
Rice	.97	.75
Rye	.79	.09
Sorghum	.94	.55
Soybeans	.80	.84
Sugarbeets	.90	.91
Wheat	.82	.84
Average	.82	.55
Intercommodity [c]		
Revenue	.82	.07
Output	.68	.00
Acreage	.89	.03
Yield	.87	.51
Price	.63	.12
Average	.78	.15
Global Average	.81	.41

a/ Percentage range, average percentage change (3 variants), moving average with 3 and 5 period lengths, and the Coppock index (see Table 1 for formulae).

b/ Revenue, output, acreage, yield, and price were ranked for each commodity separately. In this case a coefficient of |0.90| or higher is significantly different from zero at the 5% level (two-tailed test).

c/ For one variable, all ten commodities are ranked. In this case a coefficient of |0.64| or higher is significantly different from zero at the 5% level (two-tailed test).

for the cross-commodity ranking of yield, the variable that generally displayed the most trend. That adjustment for trend has no apparent effect on the rankings is an anomalous result. The sensitivity of CI to the particular period chosen, pointed out by Knudsen and Parnes, was demonstrated. The expectations part of the measure, m, depends only on the first and last observations; changes in the period often had dramatic effects on the ranking of a variable. For example, when the 1977 observation was dropped and CI recomputed, cotton fell from the third to the tenth most unstable in a cross-commodity ranking of output. This sensitivity makes CI an unreliable measure.

Coefficients of variation (in particular the measure CV(S) given in Table 1) derived from trend lines, rather than deviations around the arithmetic mean. Therefore, all 50 data series were subjected to both linear and exponential detrending by least squares regression. Based on examination of the coefficients of multiple correlation (R^2's) for these equations, "best" estimates of the coefficient of variation CV(S) were chosen. If both R^2's were less than 0.6, the non-detrended coefficient was selected; if one or both were greater than 0.6, the higher of the linear or exponential was chosen. In this fashion, a best estimate list of coefficients of variation was developed. This list was then compared with coefficients from the non-detrended data. The Spearman coefficient between the two lists was only about 0.20. Here again, the lack of agreement can largely be attributed to the influence of trend in the mean. Those series which the non-detrended coefficient of variation (CV(S)) identified as being most unstable very often fell in ranking once trend was removed, as the systematic change in mean inflated the value of the non-detrended coefficient.

A comparison of the PR, APC, and MA measures with the best estimate CV(S) shows, on the average, little concurrence between rankings (Table 3). Note the disparity across commodities and by measures. The lesson appears to be that a judicious accounting for trend can produce rankings radically different from those obtained when trend is ignored. The results of this application should eliminate any remaining skepticism as to the dependence of the characterization of instability through single variable measures on the choice of empirical technique.

Implications

As evidenced above, the treatment of trend is perhaps the paramount conceptual and empirical issue in the application of single variable measures. Whether or not trend should be regarded as instability depends on the context of the analysis; however, recognition of the existence of trend should always be made inasmuch as it influences a measure's empirical evaluation of data series. Commonly, some coefficient of variation is applied to the residuals of a series net of trend. Residuals from the moving average should not be used for this purpose because they tend to be serially correlated as do those from differencing.3/ The coefficients provide unbiased estimates of varia-

---

3/ This is the major problem of using a coefficient of variation based upon differenced series e.g. Tintner's variate difference method (see Offutt).

Table 3. Spearman Correlation Coefficients: Best Estimate CV(S) Rankings Compared to PR, APC, MA and CI Measures.

	PR	APC1	APC2	APC3	MA3	MA5	CI
Revenue	0.36	-0.05	0.31	-0.08	-0.07	-0.02	0.08
Acreage	0.90	0.88	0.92	0.84	0.90	-0.26	0.94
Output	0.44	0.72	0.72	0.68	-0.09	-0.13	0.90
Price	0.36	0.61	0.58	0.45	-0.45	-0.22	0.58
Yield	0.40	0.75	0.70	0.68	0.59	0.84	0.68

Note: For ten pairs, a correlation of $|0.64|$ or higher is significant at the 5% level (two-tailed test) and one of $|0.78|$ or higher is significant at the 1% level.

variability only for random series. For this reason, it is most often the residuals from linear regressions which are used, on the assumption that they are random once deterministic trend has been removed.

Should regression residuals not be random, as indicated perhaps by the Durbin-Watson statistic, stochastic process models can be employed to account for the remaining oscillatory movements if data series are sufficiently long. Integrated autoregressive moving average (ARIMA) models, as discussed by Box and Jenkins, can account for the deterministic and oscillatory parts of a time series and leave a random residual for which a coefficient of variation can be calculated. However, the identification and estimation of these models can be difficult and time-consuming, so the use of the residuals from linear regression can be considered an acceptable approximation to randomness for most purposes.

The dissimilarity in the rankings demonstrates that it is unlikely that all single variable measures will provide the same assessment of relative instability and results will be dependent on the particular measure chosen. Since the determination of what type of behavior constitutes instability is subjective, it is not possible to advocate unequivocally the use of any one measure. However, some general guidance can be given in making the selection.

The first step should always be to plot the data under investigation; this will reveal the presence of trend or outliers which, as indicated above, can markedly affect a measure's performance. An understanding of each measure's characteristics can then be used to determine which one might be most appropriate. While it is probably advisable in any case to compute several of the single variable measures for purposes of comparison, some can be eliminated from consideration. Because of its limitations in identifying the effects of trend, the percentage range measure seems too simple to be of much use. The Coppock Index, due to its sensitivity to the period of the data series, might also be excluded, especially since techniques such as regression can also account for trend with much less computational burden.

The average percentage change and moving average measures may have applicability in some situations. The former may be useful, for example, when some idea of the absolute average yearly change in a variable is of importance, as opposed to an index of relative dispersion from a mean value, as obtained from the coefficient of variation. The coefficients are more useful for relative comparisons. The flexibility of the moving average and its use of only a subset of the data in the calculation of trend values may have appeal, particularly if one is attempting to represent a policymaker's expectations. These measures can be computed in a straightforward fashion and provide a useful comparison to the coefficients of variation.

The use of the coefficients of variation on detrended data is probably suitable for most purposes. A coefficient of variation can be applied to the results of the regression to yield a measure of instability. While the coefficient which uses the sum of squared residuals is probably most easily obtainable, that which uses absolute deviations may be preferable. This is because such a form can distinguish widely dispersed data from that which is more compact. The sum of squared residuals will not be as sensitive to the absolute value of the distance of the data points from the fitted line; this feature may be of significance in a study of instability in which absolute as well as relative distances are important. Standardization of the coefficients is desirable if comparisons are to be made across data periods of different length.

II. The Decomposition of Variability Using Identities

For a number of variables, instability has been investigated using identities. Total output is the sum of all individual outputs within a market or country. Revenue is the product of price and output, and output the product of acreage and yield. These additive and multiplicative identities can be used to apportion variance into that attributable to each of the component variables and to their interaction.

Another approach, based on the identities, relates instability to its source in supply and demand fluctuation. The variance of gross revenue is attributed to movement in supply and demand schedules rather than just price and output. A related procedure attempts to identify instability in supply or demand by examining the covariance between commodity price and output.

Additive identities

Many commonly used commodity variables are aggregates. For example, national production is the sum of the production in individual regions. World consumption is the sum of the consumption in individual countries. It may be of interest to examine the source of variability in such aggregates in order to understand, or perhaps predict, fluctuations. While the techniques to be discussed do not encompass an explicit forecasting procedure, the implicit assumption is that past variability is a guide to that of the future. This view presupposes no or very little change over time in the underlying structural relationships among the variables. Below, the discussion will focus on whether the technique can provide any clue to the constancy or nature of these fundamental relationships.

If $x_1, \ldots, x_n$ are random variables with a multivariate normal distribution and finite variances $\sigma_1^2, \ldots, \sigma_n^2$; and $S_n = x_1 + \ldots + x_n$; then

$$(1) \quad \mathrm{Var}(S_n) = \sum_{k=1}^{n} \sigma_k^2 + 2 \sum_{j,k} \mathrm{cov}(x_j, x_k),$$

the last sum extending over each of the $\binom{n}{2}$ pairs (x_j, x_k) with $j < k$. Normalizing by division by $\mathrm{var}(S_n)$

$$(2) \quad 1 = \frac{\sum_{k=1}^{n} \sigma_k^2}{\mathrm{Var}(S_n)} + 2 \frac{\sum_{j,k} \mathrm{cov}(x_j, x_k)}{\mathrm{Var}(S_n)} .$$

The first set of terms can be considered the direct contribution of each individual component variable to the total variability of the aggregate S. The second set of terms may be considered as the contribution of the interaction of pairs of variables.

Rourke applied this formulation to the variability of world coffee production. He concluded that Brazil directly contributes 86 percent of total variability in year-to-year changes in world coffee production. Further, the interaction between Brazil and other countries accounts for another 11.84 percent of the total. The remaining 2.03 percent is due to the separate and combined influences of the other seven major and all other exporters. Rourke infers that changes in both productive capacity and in yield from existing capacity contribute to the year-to-year variability in production. However, the variance decomposition has nothing more to contribute in the way of explanation since it cannot provide information as to the cause of fluctuations in, say, Brazil's production. Similarly, knowing that the interaction between Brazil and Colombia accounts for 5.03 percent of total variability is not necessarily helpful. There is no obvious way to decide whether the correlation is spurious or indicative of a structural relationship. In order to deal with the issue of causality, Rourke examines the contribution of a two year bearing cycle on the annual production changes. However, the cycle is but one of many variables which could affect or explain fluctuations. An alternate procedure might be to decompose the production relation using the multiplicative relationship between acreage and yield.

Multiplicative identities

Interest in decomposing the variance of a multiplicative identity dates to the early 1950's. Foote, Klein and Clough and Meinken dealt with the question of the relative importance of yield and acreage changes in production variability. The results of Foote, et al. were obtained by determining the average year-to-year changes in yield and acreage as a percentage of average acreage over a sample period and then summing the two. The average annual changes in yield and in acreage were then expressed as a percentage of this sum. As Sackrin notes, the method's "drawback is that it fails to equate strictly changes that take place in acreage and yield with changes in production" (p. 136). Meinken's procedure is similar and susceptible to the same criticism.

Sackrin proposes an alternative approach, which entails expressing the multiplicative identity as an additive one in terms of natural logarithms. The sum of the coefficients on acreage and yield obtained when production is regressed on these variables is exactly one, so that this decomposition does account for all the change in output. However, as Burt and Finley (1968) point out, there are two objections to Sackrin's procedure. First, the expression of the variables in log form complicates interpretation. Second, and more important, is its failure to account for the statistical dependence between acreage and yield through an interaction term.

Consequently, Burt and Finley (1968) advanced another method for decomposing the variance of a multiplicative identity. While Goodman published the same results eight years earlier, Burt and Finley appear to have been responsible for promoting the method among agricultural economists. Burt and Finley's procedure rests on a Taylor's series expansion; any nonlinear function may be handled in this way, but Burt and Finley concentrate on the multiplicative identity, as does the rest of the literature related to agricultural issues. This interest is in large measure due to the importance, for farmers and policymakers, of revenue, the product of price and quantity sold, and also of production, the product of acreage and yield. Much of the empirical work on variance decomposition was stimulated by a desire to investigate the nature of these relationships.

The variance expression is derived by writing the identity in terms of a Taylor's series expansion. For $y = x_1x_2$, the expansion is

$$(3) \quad y = \mu_1\mu_2 + (x_1 - \mu_1)\mu_2 + (x_2 - \mu_2)\mu_1 + (x_1 - \mu_1)(x_2 - \mu_2),$$

$$\text{and } E(y) = \mu_1\mu_2 + \operatorname{cov}(x_1, x_2),$$

where μ_1 and μ_2 are the arithmetic means of x_1 and x_2. Then, the variance of y is given by

$$(4) \quad \operatorname{var}(y) = E\{y - E(y)\}^2,$$

$$= E\{(x_1 - \mu_1)\mu_2 + (x_2 - \mu_2)\mu_1 + (x_1 - \mu_1)(x_2 - \mu_2) - \operatorname{cov}(x_1x_2)\}^2,$$

$$= A + B + C + D + E + F,$$

where $A = \mu_2^2\operatorname{var}(x_1)$,

$$B = \mu_1^2\operatorname{var}(x_2),$$

$$C = 2\mu_1\mu_2 \operatorname{cov}(x_1x_2),$$

$$D = E\{(x_1 - \mu_1)(x_2 - \mu_2) - \operatorname{cov}(x_1x_2)\}^2,$$

$$E = 2\mu_1 \cdot E(x_1 - \mu_1)(x_2 - \mu_2)^2,$$

$$F = 2\mu_2 \cdot E(x_1 - \mu_1)^2(x_2 - \mu_2).$$

Burt and Finley (1968) explain the significance of the six terms in the variance:

> The first two terms (A,B) are the direct effects of x_1 and x_2, and the term (C) is a first order interaction effect.[1] The fourth term (D) is the variance of the covariance product about the covariance parameter, is necessarily positive, and is neutral for purposes of interpretation. The last two terms (E,F) are higher order interactions. Since

> these last three terms have their origin in the second degree terms of the Taylor's series, we would expect them to be relatively unimportant but in some sets of data they might give trouble. (p. 737)

For any more than a two variable decomposition, the number of terms in the variance formula increases dramatically (e.g., to 81 in the three variable case).

The derivation given above assumes that the two variables x_1 and x_2 are not independent. Goldberger (1970) points out that Burt and Finley were in error in stating that the variance formula reduces to the sum of the first two terms, A + B, in the event of independence. In fact, "independence implies that joint moments factor into products of univariate moments," so that

$$(5)\quad \text{var}(y) = \mu_2^2\text{var}(x_1) + \mu_1^2\text{var}(x_2) + \text{var}(x_1)\text{var}(x_2)$$

In applications of the decomposition to revenue or production identities, the most realistic assumption would seem to be lack of independence between price and quantity or acreage and yield.

For ease of computation, it might be desirable to approximate the variance of the function with a few terms of the expansion. In the two variable case, Burt and Finley (1968) suggest that an appropriate route would be the use of the conventional asymptotic approximation, given by the first three terms, A + B + C, of the var(y) expansion. They state that "we would expect the first interaction term (C) to dominate the higher order terms in most situations" (p. 737). Following this reasoning, the terms D, E, and F would be dropped from the computations. The accuracy of this approximation depends on the size of each variable's mean and variance as well as their covariance.

In defending the approximation, Burt and Finley (1968) state that the results will be satisfactory if the "individual means are large relative to their respective variances" (p. 155). They cite Goodman, who shows that the relative inaccuracy of the approximation for the case in which the two variables are independent is likely to be small if either variable's coefficient of variation is small. For the case Burt and Finley consider, in which the two variables are not independent, the possible source of inaccuracy is not so easily seen.

The accuracy of the approximation when the variables are not independent depends on the size of the higher order interaction terms, D, E, and F. In this case, not only the magnitudes of each variable's coefficient of variation are factors affecting accuracy but also the magnitude of the joint product moments. Without knowledge of the distributions of the underlying variables x_1 and x_2, a statement about the significance of these higher order moments cannot be made. Therefore, the assertion that the linear interaction term C can be expected to dominate the others may be true only for a limited number of situations. If, for example, x_1 and x_2 are normally distributed, the higher product moments will not exist (Kendall and Stuart). However, in practice, there would remain the possibility that the sample moments would be nonzero even if the normality assumption were valid asymptotically. Therefore, the safest course would seem to be to calculate all the terms in the exact variance formula. For the two variable case, the work is not that burdensome, as Burt and Finley themselves point out. For the three variable case, the expression is significantly longer, but for evaluation with the aid of computers, not prohibitively so.

The lack of clear cut guidelines for handling the interaction terms means that the decomposition is most useful when all terms but A and B, the direct effects, are small. Even if the linear interaction term dominates the others, care must be taken in allocating its effects to each variable. Houck, for example, divides the interaction term C equally between the two variables. The logic behind such an approach is not evident in the absence of a theory about the nature of the relationship between the two variables. Essentially, the difficulty with combining or ignoring terms in the variance formula is that there is no basis for it in the logic of the derivation of the expression. At best, there is a loss of accuracy in the results, and, at worst, bias.

Nevertheless, potentially useful information can be obtained from the decomposition. The relative contributions of price and quantity to revenue variation are of interest because of the implications for stabilization policy. A price stabilization scheme, for example, will not help stabilize revenue, and may actually destabilize it, if the source of variation is quantity sold. Or, if price and quantity seem directly responsible for about equal amounts of the variance and if the linear interaction term is strongly negative, stabilization of one or the other might again destabilize revenue.

In considering any combination of characteristics of the terms, the question of the underlying structural relationship arises. For example, price variability may be a result of a number of supply-demand relationships. McKinna considers the case of a volatile supply curve moving along a fixed, inelastic demand curve, when this is not known by the researcher.

> The decomposition procedure would indicate that price variability was the important contributor to revenue instability. Yet the underlying cause of revenue instability is shifts in supply and a program aimed at stabilizing the supply function... would probably contribute more toward stabilizing revenue than a price stabilization programme. (p. 62)

However, should this be the case, the decomposition alone will not indicate it. Since several different kinds of underlying relationships could produce similar behavior in the variance components, the allocation procedure suffers from a potentially serious limitation. The information contained in the interaction terms is invaluable in adducing the appropriate policy response, but is not recoverable in its calculated form. The decomposition procedure can be most usefully regarded as a diagnostic tool in analyzing the sources of a variable's instability. Because it is essentially nonstructural, however, the information it provides cannot substitute for knowledge of the fundamental behaviorial relationships which drive the system. Consequently, policy recommendations based solely on the indications of the decomposition run the risk of being wrong.

The Burt and Finley variance decomposition was applied to the revenue data for the ten U.S. field crops used in the previous section. Both a two variable decomposition, price multiplied by output to obtain revenue, and a three variable, with revenue as the product of price, yield, and acreage, were performed. The results are reported in Tables 4 and 5. The Fortran computer program employed is given in the Appendix.

The major purpose of this discussion is to examine the performance of the Burt and Finley method from a primarily statistical or computational viewpoint. However, it is interesting to note the economic implications of the results obtained. The two variable decomposition indicates that in eight of ten commodities, quantity contributes less variance to revenue than price, a result that coincides with the notion of inelastic supply within a season. Turning to the three variable case (Table 5), the decomposition shows price to be the most variable component for eight of ten commodities. In general, then, the results are consistent with the idea that price tends to be the most unstable variable in commodity markets. In order to determine whether more specific information might be gained and to gauge its reliability, a closer look at the results is in order.

One of the comments made above about the Burt and Finley method concerned the advisibility of checking the magnitudes of the higher order interaction terms rather than applying the approximation right away. Burt and Finley (1968) state, "... we would expect the first order interaction term to dominate the higher-order terms in most situations" (p. 737). Looking at the two tables, the ambiguity of "domination" becomes evident. In both the two and three variable cases, there appears to be no consistent relationship between the first and higher order interaction terms. Furthermore, in assessing the size of these effects, it is difficult to say whether just the sum or the absolute value of the sum of the higher order terms should be used in judging dominance. The higher order terms are as much as 14 times the first order and as little as one hundredth its size. Looking over the results, it is difficult to see where the dominant line might logically be drawn; its placement would be arbitrary in any event. A secondary point is that while the first and higher order terms are frequently of the same sign, examination of the expansion formulae yields no reason as to why this might be so. Furthermore, it is difficult to say what the higher order terms mean in an economic sense, so there is no a priori reasoning to apply in deciphering them. For a large number of the results, the higher order terms are too large to be ignored.

The decompositions were also run using deviations from linear trend as data. Burt and Finley (1970) noted "a substantial reduction in the error of the approximate formula when variance was measured around a trend instead of around the mean of the time series" (p. 168). That is, the data used were the residuals between the actual value in a period and that predicted by the

Table 4. Burt and Finley Two Variable Decomposition of Revenue.[a]

Commodity	Revenue Variance	Direct[b] Quantity A	Direct Price B	Quantity-Price Interaction C	D	E	F	Sum
Barley	32502.82	5691.85	38636.72	-9207.82	689.36	-3794.89	487.59	32502.81
		(0.1284)	(0.8716)	(-0.2077)	(0.0156)	(-0.0856)	(0.0110)	
Corn	13907854.40	2705677.92	4528501.81	3212915.18	352711.04	1997879.79	1110167.83	13907853.57
		(0.3740)	(0.6260)	(0.4441)	(0.0488)	(0.2762)	(0.1535)	
Cotton	427580.94	179340.96	350150.67	-6565.87	15936.89	-105201.45	-6080.29	427580.91
		(0.3387)	(0.6613)	(-0.0124)	(0.0301)	(-0.1987)	(-0.0115)	
Oats	29226.12	41903.59	80040.38	-69831.90	8882.55	-47449.63	15681.13	29226.12
		(0.3436)	(0.6564)	(-0.5727)	(0.0728)	(-0.3891)	(0.1286)	
Rice	94839.40	20560.36	25624.39	25622.09	2729.57	14401.70	5900.20	94838.40
		(0.4452)	(0.5548)	(0.5548)	(0.0591)	(0.3118)	(0.1278)	
Rye	36.99	106.39	166.71	-187.90	15.66	73.66	9.79	36.99
		(0.3869)	(0.6104)	(-0.6880)	(0.0573)	(-0.2697)	(0.0359)	
Sorghum	250795.28	93263.28	70943.82	31502.23	8297.48	32859.63	13928.63	250795.27
		(0.5680)	(0.4320)	(0.1918)	(0.0505)	(0.2001)	(0.0848)	
Soybeans	8259452.01	1790143.97	1380459.02	2298793.57	439360.50	1339243.70	1011450.77	8259451.52
		(0.5646)	(0.4354)	(0.7250)	(0.1386)	(0.4224)	(0.3190)	
Sugarbeets	52465.78	8736.95	24130.52	12826.79	860.22	5888.10	23.20	52465.77
		(0.2658)	(0.7342)	(0.3903)	(0.0262)	(0.1791)	(0.0007)	
Wheat	2982885.17	469929.51	1004317.59	621618.25	89673.27	531127.98	266218.39	2982884.99
		(0.3188)	(0.6812)	(0.4217)	(0.0608)	(0.3603)	(0.1806)	

[a] The number in parentheses is the ratio of the term to the sum of the direct effects.

[b] See the text for the algebraic expression for each term, A, B, C, D, E, and F.

Table 5. Burt and Finley Three Variable Decomposition of Revenue.[a]

Commodity	Revenue Variance	Direct Acreage[b]	Direct Yield	Direct Price	Acreage-Yield-Price Interaction I	II	III	RESIDUAL
Barley	32502.82	9401.58	8027.45	40360.05	-10002.45	-21540.96	11811.13	-5643.99
		(0.1627)	(0.1389)	(0.6984)	(-0.1731)	(-0.3712)	(0.2044)	(-0.0977)
Corn	13907854.40	354015.22	3056442.89	4646933.27	-955493.57	838511.25	2234343.08	3733101.43
		(0.0439)	(0.3793)	(0.5767)	(-0.1186)	(0.1041)	(0.2773)	(0.4633)
Cotton	427580.94	498335.17	128088.28	378347.88	-366649.14	28587.22	-18003.41	-221125.09
		(0.4960)	(0.1275)	(0.3766)	(-0.3649)	(0.0285)	(-0.0179)	(-0.2201)
Oats	29226.12	109109.93	17856.35	89765.31	-77054.95	-85734.91	12711.95	-37427.56
		(0.5034)	(0.0824)	(0.4142)	(-0.3555)	(-0.3956)	(0.0587)	(-0.1727)
Rice	94836.40	6325.89	8681.99	25037.10	4364.22	12364.64	11045.46	11045.46
		(0.1580)	(0.2168)	(0.6252)	(0.1090)	(0.3150)	(0.2758)	(0.6685)
Rye	36.99	99.48	77.34	178.65	-80.30	-186.46	11.07	-53.78
		(0.2799)	(0.2176)	(0.5026)	(0.2512)	(0.5246)	(0.0311)	(-0.1513)
Sorghum	250795.28	25037.56	54502.09	65201.77	39717.91	1029.59	22177.49	43218.86
		(0.1730)	(0.3765)	(0.4505)	(0.2744)	(0.0071)	(0.1532)	(0.2980)
Soybeans	8249452.01	1097682.29	89531.67	1265773.95	550300.67	1773094.99	327121.68	3155946.28
		(0.4475)	(0.0365)	(0.5160)	(0.2243)	(0.7228)	(0.1334)	(1.2866)
Sugarbeets	52465.78	5199.35	858.37	23417.96	2806.34	8239.49	4294.84	7649.42
		(0.1764)	(0.0291)	(0.7945)	(0.0952)	(0.2795)	(0.1457)	(0.2595)
Wheat	2982885.17	193766.39	323280.60	1015831.53	-88476.61	466262.67	90139.84	982080.59
		(0.1264)	(0.2109)	(0.6627)	(-0.0577)	(0.3042)	(0.0588)	(0.6407)

[a] The number in parentheses is the ratio of the term to the sum of the direct effects.

[b] An explanation of the algebraic expression for each term is given in Appendix II.

regression. In the original formula, the μ's were the arithmetic means of the original raw data. In their 1968 paper, Burt and Finley used deviations from trend with the means set equal to the trend value in the final period.

As previously mentioned, the accuracy of the approximation is improved when the coefficients of variation of the variables are small. Burt and Finley apparently viewed detrending as a way of reshaping the data to fit this criterion. The results reported in the 1968 paper do show an apparent improvement in accuracy. On closer examination, however, the significance of this new decomposition must be called into question.

The essential problem is that the identity of the original decomposition no longer holds when the means in the formula are replaced by trend values and the data used are deviations from trend. The variance sum so obtained can be equal to, greater, or less than the variance of the dependent variable. Table 4 shows that, for raw data, the identity holds exactly. However, when the two and three variable detrended versions for the 10 field crops were run, in two cases, cotton and oats, the variance of revenue calculated from the decomposition exceeded the variance found from the observations themselves. If the judgement about the improvement in accuracy of the approximation is made by comparing the sizes of the higher order interaction effects, then the detrending appears to do its job. Burt and Finley overlooked the error because they either did not compare their detrended results to the original variance of the dependent variable or had data which coincidentally satisfied the identity.

The fact that the detrended decomposition used here and that of Burt and Finley used different means still does not solve the essential problem that using deviation from trend destroys the identity. McKinna, who employed the technique in a study of Australian potato revenue, spent some time discussing the selection criteria for means.

> In principle, the results could be quite sensitive to this choice, especially if trends in x_1 and x_2 are in opposite directions. Moreover, the choice of means is not obvious. While some would argue that the formulae should be evaluated at the midpoint of the time period (in a sense the midpoint is the most 'representative' year), Burt and Finley evaluate the formula using means corresponding to the computed trend values for the last year of the time period. A case can be made for this choice as well. The last year in the time period is the year nearest to the future period in which the policies are to be implemented. (p. 60)

However, the choice of means is a moot point. Computing the variance using the trend residuals and some arbitrarily chosen (from a statistical viewpoint) mean yields results that have no apparent meaning. The premise of expansion of the total variance around the arithmetic means no longer holds. While computation of the variance of the dependent variable by the same technique might yield an identity, it is difficult to see what the interpretation of that quantity might be. The advantages of the detrended decomposition lie in its improving the accuracy of the approximation through reduction of the higher order interaction terms and in its allowing recognition that the means of the data series may be functions of time. However, the use of the detrending technique is not defensible because of the destruction of the underlying identity. Nevertheless, given that accuracy and trend are omnipresent concerns, can anything be done about them?

The concern over accuracy is precipitated by the frequent appearance of higher order interaction terms. Burt and Finley attribute these to the systematic movement due to trend in the variables. As discussed earlier, those terms cannot be further broken down by the decomposition technique. Furthermore, although detrending seems to eradicate them, it destroys the identity upon which the Taylor's series expansion is based. Consequently, the decomposition is valid only on raw data. If the higher order terms are large, it is unfortunate, but nothing more can be done correctly to reduce or eliminate them. The decomposition is not capable of providing the desired information.

The second concern, over data which may be trending, cannot be addressed the way Burt and Finley suggest, by selecting the appropriate values for the means. This concern over trend is indicative of the type of problem discussed in the previous section, that is, the determination of what part of a variable's variance is of interest. The decomposition of the raw data will not shed much light on trend effects. For example, it may be possible to examine a series and see that output trends strongly, price does not, and therefore any trend in revenue would appear to be contributed by output. However, the Burt and Finley procedure will not reveal this, for it gives a quantitative estimate of output's influence, not a qualitative one. So, large higher order interaction terms show that the two variables together act on the dependent variable but do not tell in what proportion or fashion.

Nevertheless, the Burt and Finley decomposition could be used to obtain information on the relative importance of the variables over time. The optimal use of the technique would be with cross section/time series data. When the mean varies with time, as is the case with trend, the variance should be computed around the value of the mean in each time period. If the decomposition were run on cross section data for each period, the results could be examined for evidence of systematic change through time in the relative contributions of the variables. The use of this approach ensures that the identity holds and allows analysis of the effects of nonstationarity in the means.

Analysis of identities with reference to underlying structural relationships

Attempts to relate underlying supply and demand relationships to behavior of components in the variance expression have come from two sources. One has its genesis in an interest in improving the usefulness of the Burt and Finley method in policy formulation by explicitly incorporating the effect of elasticity and intercept coefficients of the supply and demand functions. By redefining the form of the price and quantity variables in terms of supply and demand expressions, the same method of expansion can be used to derive a formula for the exact variance of gross revenue. The other method of interest is used in the literature concerned with the welfare and revenue effects of international stabilization schemes for primary commodities. The distributional impact of the schemes is, under certain assumptions, dependent on the source of instability in supply or demand schedules. Research has concentrated on examining the inferences about the source of price instability given the characteristics of an interaction term between price and quantity.

R. R. Piggott has advanced a method which modifies the Burt and Finley expansion so that the variance of gross revenue is decomposed into supply and demand rather than price and quantity components. Piggott reiterates the criticisms of the Burt and Finley method given in the last section, emphasizing "the great dangers of implementing price stabilization policies when the root cause of the fluctuating prices remains unknown" (p. 148). The point of his procedure is to "(uncover) the historical pattern of <u>supply and demand variability</u> underlying a particular pattern of revenue instability" (p. 148).

At the outset, the use of Piggott's method requires a formulation and estimation of a product's supply and demand functions. In his article, these relationships are assumed to be linear with fixed slopes over time. The assumption of linearity in the demand and supply curves is crucial to the rest of the procedure. Nonlinearity would greatly complicate the derivation of the variance expression. In agricultural economics, sufficient grounds exist for questioning the realism in assuming not only linearity but also constant slopes and additive disturbance terms as shifters. The restrictive nature of these assumptions would seem to limit the applicability of the method to commodity markets.

The transformation of the price and quantity variables is accomplished by solving for their equilibrium values. From these, there results an expression for gross revenue (GR) equal to equilibrium price times equilibrium quantity. Piggott obtains an expression for the variance of GR which is partitioned into a demand effect (DE), a supply effect (SE), and an interaction effect (I), i.e., var (GR) = DE + SE + I. DE and SE are the direct and interaction effects of the demand and supply intercept shifters, respectively.

The interaction term (I) is handled differently than under the Burt and Finley procedure. Piggott calculates I equal to the difference between GR and (DE + SE), where GR is calculated from the actual data, not from the derived formula. He argues that if I is insignificant compared to DE and SE, nothing is lost by its omission. However, if these terms are indeed significant, the problem of their interpretation is no closer to being solved than with the Burt and Finley method. In some respects it is worse since the components are not identified separately.

Even the included components of the demand and supply components are hard to explain in economic terms. Each is comprised of magnitudes attributable to the variance, skewness, and kurtosis of the distribution of the historical set of intercepts (Piggott, p. 150). However, Piggott does not explain what, for example, the skewness or kurtosis of these distributions imply for the variance of gross revenue. These quantities can be viewed as a correction to the variance of the intercepts, but it would be more interesting to know what a non-symmetrical or flat-topped distribution implies for the behavior of the structural system. Once again, the decomposition produces higher order moments which may be large, but which have no clear theoretical interpretation.

Under certain conditions, the formula for the variance of GR simplifies. Fortunately, these special cases are often of interest to agricultural economists. Simplification results if either supply is perfectly price inelastic or demand is perfectly price elastic. Piggott shows that when both these conditions are met, the variance expression collapses to that of Burt and Finley, in which price and demand shifts are equivalent as are quantity and supply shifts.

The main advantage of the Piggott over the Burt and Finley procedure lies in its identification of the source of variability in supply and demand rather than price and quantity. Such a formulation provides some measure of assurance that one has not been misled by price and quantity movements that can mask the real roots of instability. Nevertheless, even if the source of instability has been identified as supply and/or demand fluctuations, policy recommendations do not obviously follow. The blind spot in the analysis is its inability to identify the composition of the forces which are shifting the curves. While the decomposition can indicate whether price or production stabilization schemes are appropriate, it cannot identify controllable variables. That is, knowing that an inelastic and volatile supply curve is the main source of revenue fluctuation does not imply that one has any idea what causes the shifts or the rigidity in the function.4/

A related part of the literature has been concerned with identifying supply or demand fluctuations as the source of price variability. The focus is on price movement because of its effects on revenue and also because, in an international context, price stabilization is a more viable policy mechanism than quantity adjustment. This is not to say that variability in production is not important, only that the existing literature concentrates on price variability.

As with the Burt and Finley technique, the goal is the partitioning of variance, but because the relationship is not directly expressible as an identity, some other approach to the decomposition must be developed. There exists a theoretical rationale for the derivation of supply and demand expressions which lend themselves to a type of analysis of variance or correlation.

4/ Most of the theoretical literature on the effects of price stabilization to which this and the technique discussed below relate, assumes that shifts in supply and demand are random and normally distributed (Turnovsky). If this were so then controllability would not be an issue. Unfortunately, when these simple methods are applied to time series data, there is no guarantee that the disturbances identified are stochastic or normally distributed.

In particular, interest has been in determining what the implications of price variability due to supply and demand shifts are for revenue stabilization and maximization. The currency of the topic is in large part due to the concern over the fluctuation and level of export revenue in developing countries. The theoretical framework for this investigation is found in the works of Waugh, Oi, and Massell (1970b), which deal with the effects of price stabilization schemes on producer and consumer welfare and their revenue or expenditures.

Brook, Grilli, and Waelbroeck develop a framework for empirical analysis of the revenue effects of a buffer stock price stabilization scheme on developing nations. The welfare effects, defined in terms of producer and consumer surplus, are not directly comparable to the income effects, defined in terms of total revenue or expenditure. It is possible for any one group of consumers or producers to have positive gains in economic surplus but an increase in expenditures or a reduction in revenue. Nevertheless, in choosing a first approximation criterion for choice, Brook, et al. argue that the income effect is a reasonably good qualitative indicator of the net benefits or losses to LDCs. Quantification of the combined welfare and income effects can be done empirically on a country by country basis. The commodity analysis they employ, which is less costly and time consuming, is expected by them to yield satisfactory proxy results.

Accordingly, Brook et al. are interested in determining for which commodities a price stabilization scheme generates larger income for the developing nations. Their study includes 17 primary commodities for which LDCs are significant net exporters or importers. To determine the income effect, the source of each commodity's price instability in supply or demand must be determined (see Massell, 1970b). Then, if supply is the source, LDCs as net exporters would experience an increase in revenue under price stabilization. If the source of a commodity's price fluctuation is in demand shifts, LDCs as net importers would gain through a decrease in expenditures.

The empirical work is carried out using a methodology for identifying the source of a commodity's price instability in supply or demand shifts. Brook, et al. set out by postulating that, for each of the 17 internationally traded primary commodities they study,

(6) $q = s = \alpha p + x \qquad (\alpha \geq 0),$

(7) $q = d = -\beta p + y \qquad (\beta \geq 0),$

where

s = quantity supplied,
d = quantity demanded,
q = quantity traded,
p = price,
α and β are positive constants,
x and y are random variables with contemporaneous moment matrices σ_{xx}, σ_{xy}, σ_{yy}.

Further, the variables are all expressed in terms of deviations from trend (four specifications were fit for each commodity). The authors derive the above equations so that x and y contain the influences of all other independent variables which may appear in the supply and demand relations and act to shift the intercepts. Slopes are held to be constant over time for the detrended data. Solving for the equilibrium values for price and quantity yields

$$(8)\quad p = \frac{y - x}{\alpha + \beta}$$

$$(9)\quad q = \frac{\alpha y + \beta x}{\alpha + \beta}.$$

In order to determine the source of price instability, Brook, et al. rely on the sign of covariance between price and quantity, given as

$$(10)\quad \sigma_{pq} = \frac{\beta\sigma_{xx} + (\beta - \alpha)\sigma_{xy} + \alpha\sigma_{xy}}{(\alpha + \beta)^2} = \theta\sigma_{pp},$$

$$\text{where } \sigma_{pp} = \frac{y^2 - 2xy + x^2}{(\alpha + \beta)^2}$$

The covariance is taken as the difference between the mean value of revenue in a nonstabilized and a stabilized market. They explain,

> Since $\sigma_{pp} > 0$ (the variance of price), from the sign of θ (the regression of observed quantity deviations from trend on price deviations from trend), it is possible to determine whether $\sigma_{pq} \gtrless 0$. If $\theta > 0$, the income effect is favorable to consumers (importers) since their expenditures in a stabilized market are smaller than in an unstable market and unfavorable to producers (exporters) since their revenue is lower with stable prices than with unstable prices. Vice versa if $\theta < 0$, the income effect is favorable to producers (exporters) and unfavorable to consumers (importers). (p. 21)

These implied relationships between the covariance and underlying demand and supply shifts are not immediately obvious. Positive covariance implies that price and quantity move in the same direction, behavior associated with demand fluctuations. Price and quantity move in opposite directions when supply shifts. The sign of σ_{pq} gives the correct indication of the source of instability when one of the schedules is held constant while the other fluctuates. The theoretical work on which Brook, et al. base their method does make this assumption of constancy when assessing the distribution of income effects. However, in applying their empirical methodology Brook, et al. make no corresponding stipulation. When using real world data it is probably reasonable to assume that both supply and demand have shifted over time. Thus, Brook, et al. expect the sign of σ_{pq} to reveal the dominant source of price instability.

The sign of σ_{pq} will give the correct indication of the source of instability only if certain assumptions about the underlying supply and demand structures are met. Porter rearranges the covariance expression to show how its sign depends on the values of the variances of the random shifters x and y and also on the price elasticities of demand and supply, α and β. Using Brook, et al.'s notation the term can be written

$$(11)\quad \sigma_{pq} = \frac{\left(\frac{\sigma_y}{\sigma_x}\right)^2 + \left(\frac{\beta}{\alpha} - 1\right)\left(\frac{\sigma_y}{\sigma_x}\right)^2\left(\frac{\sigma_{xy}}{\sigma_x\sigma_y}\right) - \frac{\beta}{\alpha}}{(\beta + \alpha)^2}$$

Only if α and β are close to the same value will the Brook, et al. conclusions be correct. To see this, assume $\alpha = \beta$ so that the expression collapses to

$$(12)\quad \sigma_{pq} = \frac{\left(\frac{\sigma_y}{\sigma_x}\right)^2 - 1}{(\beta + \alpha)^2}$$

where σ_y^2 is variance of random demand shifts

σ_x^2 is variance of random supply shifts.

Then, if $\sigma_y > \sigma_x$, the term $\left(\frac{\sigma_y}{\sigma_x}\right)^2 > 1$, and σ_{pq} will indeed be positive and demand will be correctly identified as the dominant source of price fluctuation. Porter shows that unless the elasticities are equal, it is possible to obtain negative values of σ_{pq} even when demand is in fact the dominant source of instability, and vice versa. In any case, high values of $\sigma_{pq}/\sigma_p\sigma_q$, the correlation coefficient, do not necessarily imply σ_y greatly exceeds σ_x.

Porter's findings cast doubt on the Brook, et al. technique. They stated,

> Knowledge of model structures... is not necessary to ascertain the income effect of price stabilization. The income effect depends on observable variables and it can be ascertained directly from them. (p. 20)

Their technique is not, however, reliable unless the random variances and price elasticities taken on specific values. If these values were to be estimated, then presumably the Piggott decomposition could be employed to assign definitively the source of price instability in demand, supply, or both. If neither schedule is the dominant source, the welfare and income gains are indeterminate and must be calculated on a country by country basis.

General remarks on application

Some general remarks on the merits and limitations of this general set of procedures are in order. Comments are made on the level of aggregation in the data, application of the techniques to other than revenue or output relations, the implications of the linearity assumptions, and the ubiquitous detrending problem.

First, use of the decomposition techniques on aggregate data does not necessarily allow extension of the results to the farm or individual level. Those units whose variance (be it in revenue, production, etc.) follows the same pattern as that of the aggregate will experience the desired effects of a stabilization policy. Those whose variances differ from the aggregate norm will either lose or gain depending on the nature of the divergence. The fact that relationships implied by aggregate data do not always hold at the microeconomic level would be of particular interest to a policymaker who is concerned with the distribution of the effects of any policy over the population.

The decomposition of the multiplicative revenue identity has received much attention because of its importance to both policymakers and individuals in the economic system. It is fortunate that the relationship can be expressed in terms of supply and demand function variables as this formulation is often helpful for purposes of interpretation. However, in the case a partitioning of output variance into, say, acreage and yield components, the underlying structural relationships are not so directly obtained. Furthermore, the decomposition of functions other than identities is considerably more complex in computation and interpretation. Hence, only certain types of relationships can be easily analyzed by these techniques. However, this is really only a problem if one believes that economic relationships are nonlinear and contain disturbance terms with nonadditive and non-normal properties.

It is also important to realize just how pervasive the linearity assumption is and how restrictive is its nature. Evidence of nonlinearity in economic systems appears frequently. Little discussion is given to the consequences of nonlinearity because they can be extremely complex. Nevertheless, one cannot ignore the possibility that an erroneous assumption about linear relationships could damage the validity of the results. Similarly, the issue of non-normality in residual distributions (which are the data to which most of the methods are applied) is equally complex and surely deserves more attention than it gets (Blandford and Lee).

The problem of incorporating information about trend in a variable cannot be handled in these decomposition techniques. The single variable measures explicitly describe this facet of behavior, although they were not able to relate the movement of one variable to that of others. Consequently, it seems best to calculate both kinds of descriptions, univariate and the multivariate decomposition, in an empirical study. The information each provides is complementary.

III. Explanatory Models of Instability

The previous section included a discussion on the work of Piggott, whose aim was to "(uncover) the historical pattern of supply and demand variability underlying a particular pattern of revenue instability" (p. 148). His supply and demand schedules had price as the only explanatory variable, all other influences being included in the intercept. Combining these simplified schedules in the Burt and Finley decomposition allowed Piggott to determine what percentage of revenue variability was attributable to supply, demand, and their interaction. Any more detail on the forces shifting these curves could not be incorporated, even if it were known, because of the resulting complexity of the terms in the decomposition. Because the orientation was toward past variability, there was no explicit way to examine possible future patterns, unless the assumption were made that the historical behavior would continue unaltered.

When the identification of possible means of control of instability becomes an objective of empirical analysis, its sources must be known in greater detail. In addition, it is useful to be able to predict the effects that manipulation of a chosen source variable would have on instability. Regression analysis provides a framework for formulating and estimating the causal relationships among variables which supply the necessary detail and the capacity to analyze possible future behavior. Nevertheless, instability cannot be measured as it was previously. The variance of the dependent variable is divided into that which is deterministic - explained by the regression - and that which is stochastic - the unexplained residual. The explanation is now in terms of other variables and does not provide an explicit univariate description as before, when single variable measures could capture the nature of trend. The problem of interaction can be reflected in multicollinearity between regressors, and quantifying its effects is often difficult. The earlier methods allowed a cardinal ranking of instability among variables or among its sources. Oats revenue, for example, could be determined to be twice as variable or unstable as oats production, using a measurement index. With regression, the sources of instability can only be ranked ordinally in terms of importance, since the effects of multicollinearity do not allow a precise measure of the contribution of each. There is therefore a tradeoff in moving from one form of empirical analysis to another. Regression, while not providing as simple a method of measurement *per se*, does allow an examination of the sources and means of control of instability not previously possible.

An advantage of regression analysis, especially when applied to simultaneous equation models, is that both the relativity and dynamism of instability can be incorporated in the empirical model. That is, causal relationships can be formulated to include linkages to other variables in the system. Instability is not perceived in isolation, but relative to behavior elsewhere, which serves as a benchmark in defining what types of variability constitute instability. Furthermore, instability occurs over several periods or units of time. The single variable measures captured this intertemporal aspect through their isolation of trend, but the variance techniques of the third chapter did not reveal it explicitly. Systems of simultaneous difference equations can be used to capture the feedbacks which occur among time periods and to produce a description of the time path of a variable's movement.

Single equation analysis

A single equation which functionally relates the behavior of a dependent to that of one or more explanatory variables can be constructed using a priori knowledge about causal relationships. In the previous section, production was viewed as the product of acreage and yield. Using regression analysis, production can be related to other variables which do not appear in the identity. Yield effects can be broken down into the influence of weather, fertilizer prices, technology, etc. This more detailed information can then be useful in policy formulation that involves selection of appropriate instruments to reduce instability in production. The estimated coefficients from the regression equation can be employed to identify those sources which seem most important.

Firch (1977) advanced a method of variance analysis using a single linear regression equation. He proposes two models for analyzing variance and explains,

> The first model is appropriate for the analysis of the resulting variance when two series A and B are summed. In this case it can be shown that the variance of the combined series (A + B) is the following function of the variance of the individual series and their covariance.

(13) $\sigma^2_{A+B} = \sigma^2_a + \sigma^2_B + 2\sigma_{AB}$

The final term in (13) can be translated into an equivalent but more meaningful form,

(14) $\sigma_{AB} = \Gamma_{AB}\sigma_A\sigma_B$

> where σ_{AB} is the correlation coefficient between A and B ($\sigma_{AB}/\sigma_A\sigma_B$). (p. 325)

The net effect of the series B is given by the last two terms of (13); Firch suggests standardizing these two terms by division by σ_A to find the net change due to series B as a proportion of the variance of the original series, A.

Before looking at Firch's second model, it is necessary to point out that this result for the variance of the series A + B is valid only if the assumption is made that A and B have finite variance. Furthermore, standardization by the total variance, σ^2_{A+B}, rather than by σ^2_A would seem intuitively more straightforward than the one Firch suggests.

Firch introduces his second model as a generalization of the first, when it is impossible to include explicitly all relevant variables and the relationship between the included variables is linear, not just a simple sum. The model can be written initially as

(15) $Y = c + \beta_1 X_1 + \beta_2 X_2 + u$

where u is an error term representing the influence of excluded variables. Then, he writes the variance of Y as

(16) $\sigma^2_Y = \beta^2_1 \sigma^2_{X_1} + \beta^2_2 \sigma^2_{X_2} + 2\beta_1 \beta_2 \sigma_{X_1 X_2} + \sigma^2_u.$

Firch argues that

> ...it is possible to 'explain' the variance of one series by the variances and covariance of two other series and the residual variance of the error term. All of the information needed to make the allocation of the variance as in (16) is obtained from the least squares estimation of the regression coefficients of (15). (p. 326)

This approach is applied to data in which

Y = percentage year-to-year change in deflated cash receipts from marketing of farm products,

X_1 = percentage year-to-year change in deflated total national income, and

X_2 = percentage year-to-year change in the index of farm output.

All data are expressed in terms of first differences so that intercept term c is the trend value of Y.

In effect, Firch allocates the contribution to total variance in Y to each of the independent variables, whose variances are weighted by the square of their respective estimated regression coefficients. Note that the first two terms in (16) are identical to the numerator of the R^2 coefficient for multiple regression. The interesting aspect of Firch's method is that it takes into account the covariance between X_1 and X_2, which the conventional R^2 does not. Earlier, this interaction term was seen to have significance in many situations.

In many applications, the assumption of independence of the regressors is not warranted. On the other hand, the possibility of total linear dependence is ruled out, since the matrix of regressors would then be singular and estimation of the parameters impossible. Nevertheless, the intermediate case, referred to as intercorrelation, when $0 < \sigma_{AB}/\sigma_A\sigma_B < 1$, must be addressed. Severe multicollinearity between regressors results in the estimators of the parameters being inefficient though not biased. This inefficiency poses potentially serious problems in the use of Firch's model. If σ_{AB} is large, the variance of β_1 and β_2 may be large. Since these coefficients are used as the weights in the variance allocation formula, imprecise estimates can produce misleading results, dependent upon the characteristics of the particular sample used.

The existence of bias in the coefficient estimates cannot be completely ruled out without considering the nature of the excluded variables. If relevant variables have been omitted, and if those excluded are correlated with those included, the regressors and the disturbance term will be related, so that the least squares estimator is not consistent or even asymptotically unbiased. This bias further compounds the difficulties of using the coefficients as weights.

The essence of the problem is that the technique of multiple regression does not deal particularly well with the problem of multicollinearity. In practice, the degree to which intercorrelation influences the quality of the parameters' estimates can be difficult to assess. Consequently, Firch's approach suffers from its susceptibility to the multicollinearity complication. The severity of this deficiency, however, is largely dependent on the specification of the regression equation. Firch's own application to an equation containing national income and farm output may be suspect if one is concerned about the implications of multicollinearity. But might there be relationships for which the Firch technique would be appropriate?

The answer to the question is that if the structural equation specified has all the relevant variables in it, and these are believed to be generally uncorrelated, then the Firch method is satisfactory in terms of the validity of the coefficients as weights. Relative independence between the regressors implies a small covariance term. Therefore, Firch's decomposition is not of much interest, and one might as well proceed with the analysis using the conventional coefficient of determination. If, on the other hand, the regression coefficients as weights are not required, as in the case of identities, the logical approach would be a decomposition like that of Rourke or Burt and Finley as discussed in the previous section. But, as was seen, these latter techniques give rise to results that can be difficult to interpret.

For many relationships in economics, it is not realistic to expect that one can precisely delineate structural relationships, and further to assume that the designated regressors are unrelated. Therefore, asking the regression to provide the required information is akin to making stone soup, since the product cannot be much better than the input. But, if one is willing to settle for a less precise kind of information, there remains hope. From the procedure outlined below, one can expect to get a reasonable indication of the relative importance of independent variables' contributions to the dependent variable's variance. The operative word is relative.

Begin by writing in implicit fashion a simultaneous linear system that includes the particular variable of interest. From this specification, one can arrive at a reduced form expressible in a single linear equation for that variable. The point of the exercise is to determine which of the reduced form variables (X_i) appear to be important in relation to the others in the explanation of the dependent variable's (Y) movement. As Goldberger (1964) explains, for this purpose,

> The simple determination coefficient of Y on X_j, R^2_{jy}, is clearly inadequate. The sheer size of the coefficient b_j (the regression coefficient) is no measure of importance, since the size can be changed at will by changing the units of measurement of the variable. We may think of using the effect on Y of a typical or 'equally likely' change in each variable as a measure of importance; if Δ_j is the typical change in X_j, then $b_j\Delta_j$ is the typical effect on Y induced by X_j, and we may like to say that X_j is more important than X_k if $b_j\Delta_j > b_k\Delta_k$. (p. 197)

Beta, or path, coefficients can provide the measure of typical changes as represented by the sample standard deviation of the regressors. By dividing each variable by its standard deviation and using these values in the regression, the Beta coefficients are the ordinary coefficients so obtained. As Goldberger says, "the moments of these standardized variables are in fact correlation coefficients of the original variables" (p. 198). That is, the j, k th element of X'X, where X is the matrix of standardized variables, is $\Sigma x_j x_k / \sqrt{\Sigma x_j^2}\,\sqrt{\Sigma x_k^2}$, which is the square root of the determination coefficient $R^2_{j,k}$.

In this fashion, one has arrived at a crude sort of multiplier analysis. The square of the Beta coefficient can be interpreted as the direct contribution of that variable to R^2_Y. However, the sum of the squared Beta coefficients will not add to R^2_Y because of the covariance between regressors which cannot be disaggregated. Nevertheless, the Beta coefficients for each variable can be ranked according to the contribution of each independent variable to the variance in the dependent variable.

Ultimately, the best estimates of the reduced form coefficients are obtained when the full simultaneous equation model has been estimated. The problem of multicollinearity between regressors appearing in the reduced form is alleviated somewhat if they appear in different structural equations. In this case, estimates of their coefficients might be relatively good. Derivation of the reduced form coefficients subject to the restrictions implied by the structural specification avoids the dangers inherent in estimating the unrestricted coefficients directly. The next section discusses the estimation and interpretation of the reduced form coefficients in a full simultaneous model.

Multiplier analysis

The analysis of an equation system proceeds following the specification and estimation of the structural model. It is assumed, therefore, that the elements of the coefficient matrices are known and believed to be relatively "good" estimates. The notation to be used in what follows is introduced by considering the formulation of the general structural model.

$$(17)\quad \Gamma y_t + \beta_1 y_{t-1} + \beta_2 x_t + u_t = 0$$

where

- y_t is an n×1 vector of current endogenous variables
- y_{t-1} is an n×1 vector of lagged endogenous variables
- x_t is an m×1 vector of exogenous variables
- u_t is an n×1 vector of stochastic disturbances (with some elements zero corresponding to equations which are identities)
- Γ is an n×n matrix of coefficients on current endogenous variables
- β_1 is an n×n matrix of coefficients on lagged endogenous variables
- β_2 is an n×m matrix of coefficients on exogenous variables

For simplicity of exposition, it is assumed that the system involves lags of no more than first order. It can be shown that any system of higher order difference equations can, by means of a suitable transformation, be rewritten as a first order system (Chiang, p. 602).

Assuming Γ is a nonsingular matrix, the system can be solved for the current endogenous variables in terms of the residual and all predetermined variables (which can include lagged endogenous and both current and lagged exogenous variables). Premultiplying by Γ^{-1} and moving all terms but y_t to the right hand side (RHS) yields

$$(18)\quad y_t = (-\Gamma^{-1}\beta_1)y_{t-1} + (-\Gamma^{-1}\beta_2)x_t + (-\Gamma^{-1}u_t)$$

$$= \Pi_1 y_{t-1} + \Pi_2 x_t + v_t.$$

This is called the reduced form. The endogenous variables are now seen as explicit functions of all the other predetermined variables in the system. Contained in the Π_2 matrix are elements which are usually referred to as impact multipliers. These coefficients give the change in an endogenous variable, y_{it}, due to a ceteris paribus one unit change in an exogenous variable, x_{jt}, and may be understood as a partial derivative, $\frac{\partial y_{it}}{\partial x_{jt}} = \pi_{(i,j),t}$, to denote the (i,j)th element of Π_2. Consequently, the total number of multipliers is equal to the product of the number of endogenous and exogenous variables.

In addition to a matrix of current exogenous variables, the structural and thus the reduced form may contain one of lagged exogenous values. In this case, however, the corresponding matrix of reduced form coefficients cannot be interpreted as the effect of $x_{j,t-n}$ on $y_{i,t}$. This is essentially due to the simultaneity of the model, since if $x_{j,t-n}$ had been different so would $y_{i,t-n}$. Therefore, the effect of a unit change in $x_{j,t-n}$ on $y_{i,t}$ is the sum of effects of $x_{j,t-n}$ on $y_{i,t}$ directly and indirectly of the effects of $x_{j,t-n}$ on the values of all the y_{t-n}, which in turn affect $y_{i,t}$. To find the values of these delay multipliers it is necessary to transform (18). However, it should be noted that ultimately it is the matrix of lagged endogenous variables that gives the system its dynamic character. The effects of changes in current endogenous variables will be transmitted through the lagged endogenous variables as explained above.

Multipliers can be derived if the system of equations is stable, i.e., given the values of the exogenous variables and disturbances, the endogenous variables will converge to equilibrium values (Labys, p. 169). A description of the conditions required for stability can be found in Goldberger (1959). In the present study, stability is assumed for the purposes of discussion; however, the determination must always be made empirically for any estimated equation system.

The endogenous variables of a stable system can be written as a function of the exogenous variables and disturbance terms alone. Multiplier analysis is deterministic in that it ignores the possible influence of stochastic equation residuals. From this final form, as it is called, a number of different multipliers may be derived depending on the assumption made about the duration of the one unit change in the exogenous variable. The change may be for one period, after which the variable returns to its original level, or may be sustained over an infinite period. In either case, multipliers which describe the contemporaneous (impact), cumulative or delay (interim), and long run (final) effects of the exogenous change can be found. This discussion focuses on the use of impact multipliers, those most commonly used in economic analysis. For a discussion of the analytical uses of the other multipliers see Offutt.

The initial specification and estimation of the structural form is crucial in multiplier analysis based on the reduced form coefficient matrices, which are themselves linear functions of the structural coefficients. The stability of the system is dependent in the same way on the values of the structural estimates. So, while the reduced form yields considerable information on system dynamics, this information is only as good as the structural specification which underlies it.

One point about the use of multipliers needs to be made. The influence of an exogenous upon an endogenous variable depends not only on the size of the associated multiplier but also on the movement in the exogenous variable as well. Goldberger (1959) proposes a measure which accounts for the size of the multiplicand as well as the multiplier.

The contribution made by a predetermined variable X_i to the statistical explanation of the change in an endogenous variable Y_i in year t is defined as

(19) $\mu_{ijt} = \pi_{ij} X_{jt}$

where π_{ij} is the appropriate impact multiplier and X_{jt} is observed change in X_j from year t-1 to year t. Goldberger argues that "some type of sample average value of the μ_{ijt} would provide a summary measure of the importance of (X_j) in explaining (Y_i)" (p. 72). He proposes that the sum of absolute values of annual changes in X_j be used:

(20) $\mu_{ij} = \pi_{ij} \sum_t |X_{jt}|$.

The introduction of the Goldberger measure of the contribution of an exogenous variable to the change in an endogenous variable, μ_{ij}, leads to a consideration of the ways in which multiplier analysis might be applied in the context of the analyses of instability. Up to this point, the discussion has considered the formal derivation of the multiplier values; it now examines how assumptions upon which the derivations are based influence empirical application. The discussion begins with an examination of the usefulness of the μ_{ij} and takes up the issues concerning the time period over which multiplier analysis is applied.

From (20), μ_{ij} is seen to have as its multiplicand the sum of absolute values of the annual changes in the exogenous variable X_j. This formulation is prefered to other possible measures because the changes in X_j may be positive in some years and negative in others when measured in deviations from its mean or some other base value. Presentation of the net value of these changes could obscure important characteristics of X_j's behavior relative to the endogenous variable Y_i. However, the aggregate term $\sum_t |X_{jt}|$ does not provide a time profile of changes in X_j.

Given that the μ_{ij} have been computed for all relevant combinations of the endogenous and exogenous variables (as indicated by the reduced form equations), it should be a simple matter to compare these coefficients and to determine, for any Y_i, which X_j are important in terms of explanation of Y_i's movement. This subset of exogenous variables is called a "simplified reduced form" by Goldberger (1959, p. 73). These simplified equations can then be used in making rough predictions for the Y_i. However, in an instability study, the ranking of the X_j obtained from the μ_{ij} may be of equal interest.

Once the exogenous variables are ranked, it is instructive to consider the characteristics of and relationships among these variables. For example, for policy purposes, it may be useful to identify which of these exogenous variables are controllable, or at least susceptible to the influence of the policymaker. Dependent upon the variable's ranking, its manipulation may be a viable course of action in attempting to influence the movement of the endogenous variable.

In this context, it should be remembered that the exogenous variables themselves may be intercorrelated, due either to systematic or to spurious relationships. Systematic relationships should be explicitly incorporated in the structural model. In a correctly specified model, the regressors should be approximately orthogonal. However, apparently spurious correlation may still appear and cannot be accommodated by changing the model specification. While it may be difficult to distinguish spurious from causal relationships it is advisable to examine the correlation matrix for regressors which may exhibit multicollinearity. Multicollinearity can be a problem because it can result in imprecise coefficient estimates and also because it may confuse the selection of controllable variables.

The values of the multipliers are derived based on the relationships extant during the sample period. So, care must be taken in extrapolating the results outside that period. Furthermore, multipliers are derived from a *ceteris paribus* one unit change in the level of an exogenous variable. While these conditions are virtually never duplicated in the real world, the usually tacit assumption is that the multiplier values nevertheless retain some validity. Finally, the calculation of multipliers is only valid for linear models. Where non-linearities in variables or parameters exist a linear approximation may be used but this will not necessarily be acceptable. The alternative to the analytical multiplier technique is simulation using the structural model, to be discussed below.

Simulation analysis

Simulation techniques can be used to examine the influence of changes greater than one unit or of a random nature in exogenous variables, as well as the effects of the stochastic component of behavioral relationships. A major advantage of simulation is that it is not limited to linear models. This procedure can best be explained with reference to the Adelmans' examination of the Klein-Goldberger (K-G) model of the United States economy. Their purpose was to

> ...learn whether the (Klein-Goldberger model) really offers an endogenous explanation of a persistent cyclical process. We should like to learn whether the system is stable when subjected to single exogenous shocks, what oscillations (if any) accompany the return to the equilibrium path and what is the response of the model to repeated external and internal shocks. (p. 597)

The Adelmans' interest in explaining cyclic behavior is employed here for illustrative purposes; the techniques are more generally applicable.

In order to study the K-G model's dynamic properties, the Adelmans simulated the model over one hundred years. After some minor revisions of the structural model, the exogenous variables were extrapolated, many by fitting linear trends to post-war data. These values, along with as many lagged endogenous values as needed, were then used to solve the model using the values of the endogenous variables found in time t to compute the next set in time t + 1. The Adelmans found that the behavior of the system was monotonic and essentially linear, with no evidence of an internally generated cycle. By the eighth year, the system "was essentially on its long run equilibrium path" (p. 604).

Thus, the Adelmans examined the question of system stability using simulation rather than the analytical techniques to which the earlier discussion alluded. The projected values of the endogenous variables over time represent a moving equilibrium but the explicit expressions for these paths are difficult to derive analytically because some of the exogenous variables are extrapolated as nonlinear functions of time. The Adelmans never deal explicitly with the multiplier matrices, since their primary interest is in the response of the system to movements larger than the one unit change postulated in the multiplier concept.

The Adelmans noted that while their simulation of the K-G model predicted monotonic behavior in the endogenous variables, actual values exhibited cyclic behavior. They therefore considered the possibility that an exogenous shock to the system might produce more realistic results. To test this idea, the real magnitude of federal outlays, an exogenous variable, was reduced significantly from its extrapolated value in the ninth year (when the system was presumed on its equilibrium path) but was returned to its extrapolated path in subsequent years. It was found that although this change resulted in marked displacement from the equilibrium path for some 30 years, it did not produce the observed cyclical behavior.

In a sense, then, the Adelmans simulated multiplier analysis. The shock they employed, however, was much greater in magnitude than either the one time only or once and for all unit change postulated in multiplier analysis. They remark,

> While it is obvious that such a discontinuity in an exogenous variable is basically equivalent to a change in initial conditions, it is equally obvious that the response of a dynamic system to large displacements may be quite different from its behaviour under small perturbations. (p. 604)

It is not possible to examine the effects of greater than one unit displacements in exogenous variables within the framework of conventional multiplier analysis. As a partial derivative, the multiplier coefficient gives the value of change in the endogenous variable for an infinitesimally small change in the exogenous variable. Conventionally, infinitesimally small is interpreted as a one unit change in the exogenous variable. However, one unit may actually be quite a large change, depending on the scale of the variable. For example, wheat support price is often defined in terms of dollars per bushel but seldom would we expect that price to change by one dollar. Thus, the unit of measure does not seem infinitesimally small compared to the typical variation in support price. So, simulation allows consideration of the effects of a more realistically sized shock. A researcher would be able to identify those which seem relatively more important in influencing the behavior of endogenous variables.

The Adelman study also considers two more possible sources of cyclic behavior. The effects of random shocks superimposed on the extrapolated values of the exogenous variables are called Type I shocks. The addition of a random disturbance term to the empirically fitted equations provides what is referred to as a Type II shock. Their purpose was to determine "whether or not the introduction of relatively minor uncorrelated perturbations into the Klein-Goldberger structure (would) generate cyclical fluctuations analogous to those observed in practice" (p. 606).

Because the smooth extrapolation of the exogenous variables over time is not likely to produce realistic time paths, the imposition of random shocks seems a logical way of allowing for nontrend movement. These Type I shocks were constructed in the following fashion,

> ...define the value of an exogenous variable y_t at time t as its trend value y_t plus the shock term sy_t, and assume that sy_t has a Gaussian distribution with a mean of zero. In order that the shocks inflicted upon the system be of a more or less realistic magnitude at all times, we evaluate the standard deviation of sy_t over that portion of the data for

> which our least squares fit was made, and, for our subsequent calculations, we maintain the ratio of standard deviation of sy_t to y_{t-1} at a value independent of time. (p. 607)

For the K-G model, the Adelmans found that the imposition of Type I shocks did not produce cyclic behavior.

Given the original and extrapolated values of the exogenous variables, the Adelmans found it was impossible to induce cyclic behavior in the model whether by large displacements in one variable's value or by the imposition of random shocks on the extrapolated time paths of all exogenous variables simultaneously. However, Type II shocks (changes in the error terms of system equations) represent another potential source of cyclic or oscillatory behavior in a stable dynamic system.

The Type II, unlike the Type I, shock represents a source of cyclical behavior which is internal to the model. This is in contrast to the discontinuous shock and Type I shock, which applied to changes in the values of the exogenous variables defined as being determined outside the system. Therefore, their influence on the system's behavior is externally generated.

In general, the error term may be attributed to two sources. First, the specification error can contribute to the term. If X^* is the true variable, but some quantity $X = X^* + v$ is actually measured, the random component v will contribute to the error as would the influence of omitted variables. Second, economic relationships can rarely be expected to hold exactly, so that the inclusion of a random component in the functional form is as important in specifying the relationships among variables as the choice of the variables themselves.

In deriving the Type II shocks, the Adelmans assumed that, because there appeared to be no a priori reason why errors from the different sources should be correlated, the random error terms could be assumed to be normally distributed. The terms are further presumed to have mean zero and their standard errors are calculated in the same fashion as were Type I's. The terms were also drawn in the same way.

For the K-G model, the Adelmans found the imposition of Type II shocks did indeed result in cyclical model behavior. They found that the overall predicted behavior corresponded well to that of business cycle theory and empirical findings.

Should Type II, or stochastic, shocks appear to be important in explaining movement in one or a system of endogenous variables, what assumptions can be made about the genesis of the residual terms? As previously discussed, the terms may have their source in mis-specification (omitted variables, measurement error) or in inherent noise or truly random events (the prime example being weather). The ability to differentiate among these possible sources would aid in assessing the nature of instability (e.g., random movement in real world prices is not due to omitted variables).

Realistically, one cannot hope to disentangle these causes with the regression technique. However, some rough separation might be possible at least theoretically if not empirically. For example, in an equation which attempts to explain variation in yield, one might expect the weather to exert an influence. If the yield equation does not contain a proxy for weather, metereological influences would most likely be contained in the error term. Unfortunately, finding some logical explanation of the contents of the error term is not usually even this straightforward.

Results and comments on the applications

An empirical demonstration of all the methodological questions discussed in this chapter is beyond the scope of this study. As an alternative, selected techniques are illustrated with respect to an existing model. There was no attempt to specify and estimate an original dynamic, simultaneous equation model and some of the techniques examined in the section, such as the Adelmans' Type I and II shocks, were not analyzed. The rest of this section reports on what was learned in trying to use an econometric model in the analysis of instability.

Mo's model of the United States wheat sector, described in a 1968 USDA bulletin, was chosen for the analysis. For the period 1928 to 1964, Mo estimated a six equation recursive model explaining U.S. wheat consumption, including inventory demand in the form of government and commercial stocks, and U.S. exports, but specifying production exogenously. The equations, their estimated coefficients, and variable definitions are given in Table 6. Values for the short and long run multipliers from the model were also given by Mo who also determined that the system was stable.

Using Mo's results, three rankings of the importance of sources of instability for each endogenous variable were computed. The rankings were made using the Beta or standardized coefficients, the impact multipliers, and the Goldberger μ measure - the impact multiplier times the sum of the absolute values of annual changes in the exogenous variable. These results are reported in Tables 7, 8, and 9.

The Beta coefficients (Table 7) are based on the structural equations. For each endogenous variable, the explanatory variables are ranked in order of importance as indicated by the absolute value of the computed coefficient. These rankings are often quite different from those which would be obtained if the ordinary regression coefficients alone were considered. The differences in standard deviation or "typical variability" among the regressors make the rankings of the standardized coefficients, which account for this, different.

In contrast, the impact multipliers and the μ coefficients (Tables 8 and 9, respectively) are based on the reduced form equations, which express each endogenous variable as an explicit function of all the exogenous variables. Consequently, the composition of the sources of variability in an endogenous variable is often quite different than that inferred by simply

Table 6. Mo's Wheat Model.

Note: The figures in parentheses below the estimated coefficients are the standard errors of the estimates, and R is the estimated coefficient of multiple correlation. All estimates are OLS.

Farm Price and Support Relation

$$P_t = 0.1492 + \underset{(0.0448)}{0.9189}\, P_{st} + \underset{(0.0014)}{0.0108}\, K_t P_{fot}$$

$R = 0.97$

Food Consumption Relation

$$q^*_{ht} = 1.1989 - \underset{(0.0678)}{0.2284}\, P_t + \underset{(0.0042)}{0.0077}\, P_{ct} + \underset{(0.2254)}{1.6005}\, G(I_t)$$

$R = 0.97$

Feed Consumption Relation

$$q_{ft} = -137.8420 - \underset{(37.4650)}{143.7966}\, P_t + \underset{(0.5804)}{1.6302}\, P_{fot} + \underset{(0.8894)}{1.7860}\, L_t$$

$R = 0.88$

Government Inventory Relation

$$C_{gt} = -182.9923 + \underset{(78.0566)}{115.6075}\, P_{st} + \underset{(0.0913)}{0.1806}\, \overline{K}_t \overline{D}_{t-2} O_t + \underset{(0.0974)}{0.7446}\, C_{gt-1}$$

$R = 0.94$

Commercial Inventory Relation

$$C_{ct} = 200.2999 - \underset{(24.5510)}{64.4016}\, P_t - \underset{(0.0270)}{0.0422}\, C_{gt} + \underset{(0.1538)}{0.3635}\, C_{ct-1}$$

$R = 0.84$

Export Relation

$$q_{Et} = 433.5437 - \underset{(80.3589)}{112.09799}\, q^*_{ht} + \underset{(0.0695)}{0.0967}\, (C_{ct-1} + C_{gt-1}) + \underset{(0.1361)}{0.6494}\, q_{Et-1}$$

$R = 0.93$

Table 6. (Continued)

VARIABLE DEFINITIONS

Endogenous

P_t = average wheat price received by farmers in time t ($/bu)

q^*_{ht} = domestic per capita use of wheat for food in time t (bu per capita)

q_{ft} = domestic use of wheat for feed in time t (mil. bu)

C_{gt} = government wheat inventory at the end of time t (mil. bu)

C_{ct} = commercial wheat inventory at the end of time t (mil. bu)

q_{Et} = total U.S. exports of wheat in time t (mil. bu)

Exogenous

P_{st} = average wheat support price at time t ($/bu)

K_t = 1, if no price support program at time t

= 0, otherwise

P_{fot} = farm price index of other feed grains (corn, oats, barley, and sorghum) at time t (1957-59 = 100)

P_{ct} = consumer price index at time t (1957-59 = 100)

I_t = per capita disposable income at time t ($ per capita)

$G(I_t)$ = a nonlinear transformation of variable I_t

L_t = grain consuming animal units of livestock fed annually at time t (mil. units)

D_t = 1, during World War II

= 0, otherwise

$\overline{K}_t$ = 1, if there is a government price support program at time t

= 0, otherwise

$\overline{D}_t$ = 1, during World War II

= 0, otherwise

O_t = total U.S. wheat production at time t (mil. bu)

Table 7. Beta Coefficients for Mo's Wheat Model.

Explanatory Variable	Endogenous Variable P_t	q^*_{ht}	q_{ft}	C_{gt}	C_{ct}	q_{Et}
P_{st}	0.7927			0.1255		
$K_t P_{fot}$	0.3761					
P_{ct}		0.3601				
$G(I_t)$		1.1090				
P_{fot}			0.6036			
L_t			0.2303			
$\overline{K}_t \overline{D}_{t-2} O_t$				0.0712		
P_t		0.2799	0.8549		0.3777	
C_{gt-1}				0.7446		
C_{gt}					0.1968	
C_{ct-1}					0.3635	
q^*_{ht}						0.2226
q_{Et-1}						0.6494
RANKING	P_{st}	$G(I_t)$	P_t	C_{gt-1}	P_t	q_{Et-1}
	$K_t P_{fot}$	P_{ct}	P_{fot}	P_{st}	C_{ct-1}	q^*_{ht}
		P_t	L_t	O_t	C_{gt}	

Table 8. Impact Multipliers for Mo's Wheat Model.

Explanatory Variable	Endogenous Variable					
	P_t	q^*_{ht}	q_{ft}	C_{gt}	C_{ct}	q_{Et}
P_{st}	0.9189	-0.2099	-132.1347	115.6075	-64.0573	23.5268
$K_t P_{fot}$	0.0108	-0.0025	-1.5530		-0.6955	0.2765
P_{ct}		0.0077				-0.8632
$G(I_t)$		1.6005				-179.4127
P_{fot}			1.6302			
L_t			1.7860			
D_t			159.4989			
$\overline{K}_t \overline{D}_{t-2} O_t$				0.1806	-0.0076	
RANKING	P_{st}	$G(I_t)$	D_t	P_{st}	P_{st}	$G(I_t)$
	$K_t P_{fot}$	P_{st}	P_{st}	$\overline{K}_t \overline{D}_{t-2} O_t$	$K_t P_{fot}$	P_{st}
		$K_t P_{fot}$	L_t		$\overline{K}_t \overline{D}_{t-2} O_t$	P_{ct}
		P_{ct}	P_{fot}			$K_t P_{fot}$
			$K_t P_{fot}$			

Table 9. Goldberger's μ Measure for Mo's Wheat Model.

Explanatory Variable	Endogenous Variables P_t	q^*_{ht}	q_{ft}	C_{gt}	C_{ct}	q_{Et}
P_{st}	33.9993	-7.7663	-4888.9839	4277.4775	-3520.2392	870.4916
$K_t P_{fot}$	2.0304	-0.4700	-291.9640		-130.7540	-51.9971
P_{ct}		0.6430				-28.9462
$G(I_t)$		2.2919				
P_{fot}			894.9798			
L_t			364.4512			
D_t						
$\overline{K}_t \overline{D}_{t-2} O_t$				577.8279	-24.3161	
RANKING	P_{st}	P_{st}	P_{st}	P_{st}	P_{st}	P_{st}
	$K_t P_{fot}$	$G(I_t)$	P_{fot}	$\overline{K}_t \overline{D}_{t-2} O_t$	$K_t P_{fot}$	P_{ct}
		P_{ct}	L_t		$\overline{K}_t \overline{D}_{t-2} O_t$	$K_t P_{fot}$
		$K_t P_{fot}$	$K_t P_{fot}$			$G(I_t)$

considering the explanatory variables in the structural equations. The rankings again change from the impact multipliers to the μ coefficients, which take into account the size of the annual changes in the exogenous variables.

According to the μ coefficient, the wheat support price p_{st} is the most important determinant of the behavior of each of the endogenous variables. However, p_{st} does not figure so prominently in the other two rankings. If stabilization of farm price p_t were a policy objective, then the support price would be the obvious choice of policy instrument. However, the likely pervasive effects of its manipulation are not revealed unless the results of the μ coefficients are considered. One of the objectives in employing regression techniques is to examine how schemes for controlling instability in a variable could affect the rest of the system in which it is embedded. The use of the coefficients based on the reduced form from the simultaneous model facilitate this investigation. The use of standardized regression coefficients, which are derived from single equations, does not incorporate these linkages.

The rankings obtained from the three coefficients do not enable a precise apportionment of the historical variance in the endogenous variable. For one thing, they ignore the possible influence of random events, embodied in the residuals of the regression equations. More significantly, the coefficients for each exogenous variable are derived on the assumption that all the other exogenous variables are held constant. Historically, all these are likely to have fluctuated. Therefore, the relationships among exogenous variables, their interaction, are ignored in deriving those coefficients. The net effect of a change in any one exogenous variable will depend on the behavior of all the others and the nature of their intercorrelation. Thus, we obtain a ranking which does not take these possible interactions into account. As illustrated in the Burt and Finley decomposition, these interaction effects can be significant. The inability to account precisely for these interaction effects is the main drawback of this analytical approach.

Simulation, on the other hand, provides a way to examine the effects on an endogenous variable when more than one exogenous variable changes. The influence of changes in the stochastic residual can also be evaluated. Coupled with the analytical results of the coefficients discussed above, simulation can be used to look at those scenarios which are of most interest. For Mo's model, the analytical results indicate that manipulation of the support price would have significant impact not only on farm price but on other endogenous variables. Since the unit of measure of support price is dollars per bushel, simulation could be used to investigate the effects of smaller than one dollar changes, which would seem to be more realistic. The value of other variables, such as government stocks, could be simultaneously changed, and the net effect on farm price and other endogenous variables evaluated.

These demonstrations of methodology are simple compared with some of the techniques the section has discussion. However, it is still possible to see that both the structural specification of the model and the type of post-estimation analysis pursued can affect the conclusions derived from the model.

In this sense, methodologies based on simultaneous equation systems can be criticized for their ambiguity. However, it would seem that they still have the advantage of reflecting the relativity and dynamism implicit in even the most general concept of instability.

A model which is to be used for dynamic and/or stochastic simulation analysis must be built with this in mind. The structural specification should be done carefully since the subsequent analysis depends on it. The exact combination of techniques applicable in a particular case depends on the nature of the system and its instability. When random events or shocks, such as weather effects, play a large role in determining fluctuations in endogenous variables, then deterministic multiplier analysis may only be of limited use. In the process of model specification, the characteristics of the system should become clear so that the choice of techniques can be made with them in mind.

IV. Summary and Conclusions

Summary

The objective of this bulletin has been to appraise the use of empirical techniques in the analysis of instability. No attempt was made to define instability beyond a recognition that it involves some degree of variability in an economic quantity. What portion of this variability is regarded as "unacceptable" or unstable will depend, at least partly, on subjectivity. Key variables associated with agricultural revenue, specifically, acreage, yield, output, and price, provided the framework for the application of the empirical methodology.

In making the appraisal, empirical techniques were divided into several broad categories. The division was made on the grounds that no one technique provides all the information necessary for an analysis of instability. The general phenomenon of instability has different facets, and information on one or more might be required. This bulletin specifically considered methods which (1) characterize unstable behavior, (2) identify sources of instability, and (3) provide an evaluation of means for its control. Specific empirical methodologies were separated according to the type of information they provide. The first section dealt with single variable measures used to describe and quantify unstable behavior. The next section considered the use of identities to allocate instability to component variables. The final section examined the use of regression techniques to identify further the sources of instability and to evaluate means for its control.

Conclusions

Two major conclusions follow from this investigation. First, there must be explicit recognition of which aspect of instability a particular technique evaluates. This way, the limitations of any one approach are made clear. Significant complementarity and tradeoffs exist among empirical methodologies. Second, once a technique has been selected, its application must be consistent with its underlying assumptions. These assumptions directly influence the empirical results derived; they may imply restrictions which are unacceptable or untenable in particular cases. By considering them explicitly, limitations and biases can be made clear. Such clarity is essential since it is not possible to define instability unambiguously.

The characterization of unstable behavior was pursued through the use of univariate measurement techniques. These consider a single variable in isolation; no connection with other variables is incorporated. The importance of the assumptions and properties of different measures was demonstrated through their application to data on ten U.S. field crops. For several variables, e.g., yield and acreage, no unambiguous determination of which was the most unstable either across crops or within crops could be made. The rankings obtained were dependent on the measure employed. Differences in results emerged because of differences in assumptions and sensitivities of the measures. Thus, an analysis which uses one of these measures could well be challenged by one which employs some other. Selection can only be made on the basis of which measure reflects the view of instability most compatible with the aim of the study. For example, an analysis concerned with movements net of long term trend will require a different measure than one concerned with year-to-year variability. The criteria a measure uses to judge instability, must be stated explicitly in the analysis; and this can only be done if the analyst understands the assumptions which underlie it.

The first step toward decomposing the sources of instability by relating the behavior of the variable of interest to movements in other variables. A method of variance decomposition relying on identities, such as revenue as the product of price and output, was examined. Total variance was allocated among its sources; the distinction between deterministic and stochastic or trend and non-trend movement could no longer be made despite the fact that this is frequently assumed to be possible. It was concluded that the error in applying the decomposition to detrended data has arisen because of the lack of consideration of the underlying assumptions of the Taylor's series. In a similar way, reliance on the covariance between price and quantity to identify the source of instability in supply or demand was found to be warranted only under certain restrictive assumptions.

The next section considered the use of regression techniques to examine the sources of instability and to evaluate means for control. Although this approach allows the sources of instability to be evaluated in more detail than is the case with the decomposition of identities, it does not permit the same ease of measurement of instability. The use of simultaneous equation econometric models allows both the dynamic aspect of instability to be reflected, as in the univariate measures, and its relativity, as embodied in the variance decomposition of identities. An indication of the relative importance of sources of instability and the probable consequences of their manipulation can be gained from regression analysis. Depending on the assumptions made about the nature of change in behavior, simulation or multiplier analysis can be used to gauge the effects of control on unstable variables.

The progression from the single variable measures to multivariate regression represents a considerable increase in computational burden. However, this burden is balanced by a gain of information. Whether or not the increase in complexity is warranted depends on the objectives of the study at hand. However, it is always necessary to delineate the limitations of the approach chosen. This bulletin is intended to assist in an assessment of those limitations.

APPENDIX

Computer Programs for the Analysis of Instability

SVIM (PROGRAM TO DERIVE SINGLE VARIABLE INSTABILITY MEASURES)

Description

Derives single variable measures of variability defined in Table 1. Input is in the form of an unlimited number of sets of variables each composed of up to 10 variables each with up to 30 observations per variable.

Input

Card	Cols	
1	1-2	Number of variables in current variable set
	3-4	Number of observations per variable
	5-80	Title for variable set.
2	1-8	Name of 1st variable in set
	9-18	Name of 2nd variable in set
	"	"
	"	"
	72-80	Name of 10th variable in set.
3	1-80	Data format, e.g. (10F8.3).
4 et seq	Data punched in form specified by 3, in time series form by variable.	
Final	Blank to terminate job <u>or</u> to process further variable sets repeat cards 1-4 above.	

```
C
C PROGRAM TO DERIVE SINGLE VARIABLE INSTABILITY MEASURES
C
C VERSION SEPTEMBER 1982
C
C PROGRAM HANDLES AN UNLIMITED NUMBER OF VARIABLE SETS EACH
C COMPOSED OF 'NSET' VARIABLES WITH 'N' OBSERVATIONS PER VARIABLE
C
C STORAGE ALLOWS FOR 10 VARIABLES PER SET AND 30 OBSERVATIONS PER
C VARIABLE
C
C D. BLANDFORD, DEPARTMENT OF AGRICULTURAL ECONOMICS, CORNELL UNIVERSITY
C ITHACA NY 14853
C
      REAL*8 DATA(30),XA(30),XB(30),XC(30),A,AB,AC,RN1,RN2,V1,V2,CV1,
     1CV2,SCV1,SCV2,R2,F,DW,DSTO(30,10)
      INTEGER*2 N,NSET,NN,SW,PCT
      COMMON XA,A,AB,AC,V1,V2,CV1,CV2,SCV1,SCV2,R2,F,DW,N,SW
      INTEGER*4 CA(19),CB(20)
      DIMENSION FMT(20)
C
C READ NUMBER OF VARIABLES IN SET, NUMBER OF OBSERVATIONS,
C LABEL FOR THE SET, DATA FORMAT AND DATA
C
    1 READ(5,1000)NSET,N,(CA(I),I=1,19)
      IF(NSET.EQ.0) GOTO 50
      NL=NSET*2
      READ(5,1100)(CB(I),I=1,NL)
      READ(5,1100)(FMT(J),J=1,20)
      READ(5,FMT)((DSTO(I,J),I=1,N),J=1,NSET)
      NB=1
      NE=2
C
C COMPUTE SET OF MEASURES VARIABLE BY VARIABLE
C
      PCT=0
      DO 100 I=1,NSET
      IF(PCT.NE.3) GOTO 105
      WRITE(6,1010)(CA(J),J=1,19)
      PCT=0
      GOTO 106
  105 WRITE(6,1050)(CA(J),J=1,19)
      PCT=PCT+1
  106 DO 110 J=1,N
  110 DATA(J)=DSTO(J,I)
      WRITE(6,1150)(CB(J),J=NB,NE)
      NB=NB+2
      NE=NE+2
C
C
C    PERCENTAGE RANGE (PR)
C
      DO 200 J=2,N
      A=DATA(J)-DATA(J-1)
      XC(J)=A/DATA(J-1)
      XA(J)=DABS(XC(J))
      IF(A.GE.0.0) GOTO 220
      XB(J)=XA(J)
      GOTO 200
```

```
  220 XB(J)=A/DATA(J)
  200 CONTINUE
      A=XA(2)
      AB=XA(2)
      DO 230 J=3,N
      IF(XA(J).LT.A)A=XA(J)
      IF(XA(J).GT.AB)AB=XA(J)
  230 CONTINUE
      AC=(AB-A)*100.0
      WRITE(6,1200)AC
C
C   AVERAGE PERCENTAGE CHANGE (APC1-3)
C
      A=0.0
      AB=0.0
      AC=0.0
      DO 300 J=2,N
      A=A+XA(J)
      AB=AB+XC(J)*XC(J)
  300 AC=AC+XB(J)
      RN1=N-1
      A=A/RN1*100.0
      AB=AB/RN1
      AC=AC/RN1*100.0
      WRITE(6,1300)A,AB,AC
C
C   3 AND 5 PERIOD MOVING AVERAGES (MA)
C
      AB=0.0
      NN=N-1
      DO 400 J=2,NN
      XA(J)=0.0
      JJ=J-1
      JJJ=J+1
      DO 410 K=JJ,JJJ
  410 XA(J)=XA(J)+DATA(K)
      XA(J)=XA(J)/3.0
      XB(J)=DABS((DATA(J)-XA(J))/XA(J))
  400 AB=AB+XB(J)
      RN2=N-2
      AB=AB/RN2*100.0
      AC=0.0
      NN=N-2
      DO 420 J=3,NN
      XA(J)=0.0
      JJ=J-2
      JJJ=J+2
      DO 430 K=JJ,JJJ
  430 XA(J)=XA(J)+DATA(K)
      XA(J)=XA(J)/5.0
      XB(J)=DABS((DATA(J)-XA(J))/XA(J))
  420 AC=AC+XB(J)
      RN2=N-4
      AC=AC/RN2*100.0
      WRITE(6,1400)AB,AC
C
C   COPPOCK'S INDEX (CI)
C
      DO 500 J=1,N
```

```
  500 XA(J)=DLOG(DATA(J))
      A=0.0
      DO 510 J=2,N
      XB(J)=XA(J)-XA(J-1)
  510 A=A+XB(J)
      A=A/RN1
      AB=0.0
      DO 520 J=2,N
      RN2=(XB(J)-A)
  520 AB=AB+RN2*RN2
      AB=AB/RN1
      AB=DEXP(DSQRT(AB))
      WRITE(6,1500)AB
C
C   VARIANCE (N WEIGHTED), AVERAGE ABSOLUTE DEVIATION, AND
C   UNSTANDARDIZED/STANDARDIZED COEFFICIENTS OF VARIATION (CV,SCV)
C   FOR RAW, LINEARLY DETRENDED, AND EXPONENTIALLY DETRENDED DATA
C
      A=0.0
      AB=0.0
      AC=0.0
      RN1=N
      DO 600 J=1,N
      XA(J)=J
      XB(J)=DATA(J)
      XC(J)=DLOG(DATA(J))
      A=A+XA(J)
      AB=AB+XB(J)
  600 AC=AC+XC(J)
      A=A/RN1
      AB=AB/RN1
      AC=AC/RN1
      DO 610 J=1,N
      XA(J)=XA(J)-A
      XB(J)=XB(J)-AB
  610 XC(J)=XC(J)-AC
      CALL CVR(XB,AB)
      WRITE(6,1600)V1,V2,CV1,SCV1,CV2,SCV2
      WRITE(6,1610)
      SW=0
      CALL DETR(XB)
      WRITE(6,1600)V1,V2,CV1,SCV1,CV2,SCV2
      NN=N-2
      WRITE(6,1620)R2,NN,F,DW
      WRITE(6,1630)
      SW=1
      CALL DETR(XC)
      WRITE(6,1600)V1,V2,CV1,SCV1,CV2,SCV2
      WRITE(6,1620)R2,NN,F,DW
      WRITE(6,1640)
  100 CONTINUE
      GOTO 1
   50 CONTINUE
C
 1000 FORMAT(2I2,19A4)
 1010 FORMAT('1','INSTABILITY MEASURES FOR ',19A4)
 1050 FORMAT(' ','INSTABILITY MEASURES FOR ',19A4)
 1100 FORMAT(20A4)
 1150 FORMAT(' ','VARIABLE...',19A4)
```

```
 1200 FORMAT(' ','PERCENTAGE RANGE (PR) =',F7.2)
 1300 FORMAT(' ','AVERAGE PERCENTAGE CHANGE APC1 =',F7.2,'  APC2 =',
     1F7.4,'  APC3 =',F7.2)
 1400 FORMAT(' ','MOVING AVERAGE MEASURES MA3 =',F7.2,'  MA5 =',F7.2)
 1500 FORMAT(' ','COPPOCK INDEX (CI) =',F7.4)
 1600 FORMAT(' ','VARIANCE = ',D14.7,'  AVERAGE ABSOLUTE DEVIATION = ',
     1D14.7/' ','COEFFICIENTS OF VARIATION....V(S) = ',F6.2,'  S(S) = ',
     1F6.2,'  V(D) = ',F6.2,'  S(D) = ',F6.2)
 1610 FORMAT('0','LINEARLY DETRENDED DATA')
 1620 FORMAT(' ','R-SQUARED =',F7.4,2X,'F(1,',I2,') =',F9.4,2X,'DURBIN-W
     1ATSON D STATISTIC =',F7.4/)
 1630 FORMAT(' ','EXPONENTIALLY DETRENDED DATA')
 1640 FORMAT(' ')
      STOP
      END
C
C CALCULATION OF COEFFICIENTS OF VARIATION
C
      SUBROUTINE CVR(Z,M)
      REAL*8 Z(30),M,XA(30),A,AB,AC,V1,V2,CV1,CV2,SCV1,SCV2,R2,F,DW,RN
      INTEGER*2 N,SW
      COMMON XA,A,AB,AC,V1,V2,CV1,CV2,SCV1,SCV2,R2,F,DW,N,SW
      V1=0.0
      V2=0.0
      DO 100 I=1,N
      V1=V1+Z(I)*Z(I)
  100 V2=V2+DABS(Z(I))
      RN=N
      V1=V1/RN
      V2=V2/RN
      CV1=DSQRT(V1)/M
      CV2=V2/M
      SCV1=CV1/DSQRT(RN-1.0)
      SCV2=CV2/(2.0-2.0/RN)
      CV1=CV1*100.0
      CV2=CV2*100.0
      SCV1=SCV1*100.0
      SCV2=SCV2*100.0
      RETURN
      END
C
C LINEAR AND EXPONENTIAL DETRENDING
C
      SUBROUTINE DETR(Z)
      REAL*8 Z(30),XA(30),A,AB,AC,V1,V2,CV1,CV2,SCV1,SCV2,R2,F,DW,C,D,
     1E(30),G(30),RN
      INTEGER*2 N,SW
      COMMON XA,A,AB,AC,V1,V2,CV1,CV2,SCV1,SCV2,R2,F,DW,N,SW
      C=0.0
      D=0.0
      DO 100 I=1,N
      C=C+Z(I)*XA(I)
  100 D=D+XA(I)*XA(I)
      V1=C/D
      IF(SW.GT.0) GOTO 50
      V2=AB-V1*A
   50 IF(SW.EQ.0) GOTO 60
      V2=AC-V1*A
   60 DO 110 I=1,N
```

```
      E(I)=V1*XA(I)
      G(I)=E(I)
  110 E(I)=Z(I)-E(I)
      C=0.0
      D=0.0
      DO 130 I=1,N
      C=C+V1*Z(I)*XA(I)
  130 D=D+Z(I)*Z(I)
      R2=C/D
      C=0.0
      D=0.0
      DO 140 I=1,N
      C=C+(V1*V1)*(XA(I)*XA(I))
  140 D=D+E(I)*E(I)
      RN=N-2
      F=C*RN/D
      C=0.0
      D=E(N)*E(N)
      N1=N-1
      DO 150 I=1,N1
      D=D+E(I)*E(I)
      RN=(E(I+1)-E(I))
  150 C=C+RN*RN
      DW=C/D
      IF(SW.GT.0) GOTO 170
      DO 180 I=1,N
  180 E(I)=(AB+Z(I))-(V2+V1*(XA(I)+A))
  170 IF(SW.EQ.0) GOTO 190
      DO 200 I=1,N
  200 E(I)=DEXP(AC+Z(I))-DEXP(V2+V1*(XA(I)+A))
  190 CALL CVR(E,AB)
      RETURN
      END
```

BFD (Burt and Finley Decomposition)

Description

Derives 2 and/or 3 variable decomposition of variance in random identities using Burt and Finley's method AJAE 50, 1968, 734-744 for row and linearly detrended data. The program will accept multiple data sets of up to 30 definitional (Y) variables with up to 30 observations per variable.

Input

Card	Cols	
1	1-80	Job title
2	1-2	Number of definitions to be decomposed in current data set.
	3-4	Blank
	5-80	Data format, e.g. (10F8.3) or (F15.4)
3	1-8	Name of first definitional variable in data set (Y)
	9-10	Number of definitional variables entered (see explanatory note below).
	11-12	Number of observations per variable
	13-14	02 = 2 variable decomposition only required 03 = 3 variable decomposition only required 23 = both 2 and 3 variable decomposition required.
4 et seq		Data punched in time series form as per format specified on card 2 (for order see note below).
Final		Blank to terminate job or to process further variable sets repeat cards 1-4 above.

Explanatory Note

For a two variable decomposition only 02 should be specified in columns 9-10 of card 3. For decomposition of the revenue identity (revenue = output × price) the data will be in the form of revenue followed by output; price will be generated by the program. For decomposition of the output identity (output = acreage × yield) the data will be output followed by acreage harvested; yield will be generated by the program.

For a three variable decomposition only 03 should be specified in columns 9-10 of card 3. For decomposition of the revenue identity (revenue = acreage × yield × price) the data must be revenue, followed by output, followed by harvested acreage. Both yield and price will be generated internally by the program.

For both a two and three variable decomposition of the revenue identity 23 should be specified in columns 9-10 of card 3. Data should be entered in the order specified for the three variable decomposition specified above.

```
C
C PROGRAM TO DERIVE 2 AND 3 VARIABLE DECOMPOSITION OF VARIANCE IN
C RANDOM IDENTITIES USING BURT AND FINLEY'S METHOD (AJAE 50, 1968,
C 734-744) FOR RAW AND LINEARLY DETRENDED DATA
C
C VERSION SEPTEMBER 1982
C
C PROGRAM WILL ACCEPT MULTIPLE DATA SETS OF UP TO 30 Y VARIABLES,
C WITH UP TO 30 OBSERVATIONS PER VARIABLE
C
C D. BLANDFORD, DEPARTMENT OF AGRICULTURAL ECONOMICS, CORNELL UNIVERSITY
C ITHACA NY 14853
C
      REAL*8 DATA(30,5),DDMV(32,5),ST2(60,11),STD2(60,11),ST3(60,11),
     1STD3(60,11),T(32),STAT(30,8)
      INTEGER*2 NPROD,NOBS,NVAR,ASK,IND2,IND3,NNOBS(30),I,C2,C3
      INTEGER*4 CARD(20),NAME(30,2),NST02(30,2),NST03(30,2)
      DIMENSION FMT(19)
      READ(5,400)(CARD(I),I=1,20)
      WRITE(6,410)(CARD(I),I=1,20)
      WRITE(6,500)
      WRITE(6,510)
   95 READ(5,420)NPROD,(FMT(I),I=1,19)
      IF(NPROD.EQ.0) GOTO 100
      DO 1 I=1,60
      DO 1 J=1,11
      ST2(I,J)=0.0
      STD2(I,J)=0.0
      ST3(I,J)=0.0
    1 STD3(I,J)=0.0
      DO 2 I=1,8
      DO 2 J=1,NPROD
    2 STAT(J,I)=0.0
      IND2=0
      IND3=0
      C2=0
      C3=0
      DO 80 I=1,NPROD
      READ(5,430)(NAME(I,J),J=1,2),NVAR,NOBS,ASK
      NNOBS(I)=NOBS
      IF(ASK.EQ.2.OR.ASK.EQ.23)IND2=IND2+1
      IF(ASK.EQ.2.OR.ASK.EQ.23)C2=C2+1
      IF(ASK.EQ.3.OR.ASK.EQ.23)IND3=IND3+1
      IF(ASK.EQ.3.OR.ASK.EQ.23)C3=C3+1
      DO 5 J=31,32
    5 T(J)=0.0
      DO 6 J=1,NOBS
      T(J)=J
    6 T(31)=T(31)+T(J)
      T(31)=T(31)/NOBS
      DO 7 J=1,NOBS
      T(J)=T(J)-T(31)
    7 T(32)=T(32)+T(J)**2
      DO 10 J=1,30
      DO 10 K=1,5
      DATA(J,K)=0.0
   10 DDMV(J,K)=0.0
      READ(5,FMT)((DATA(J,K),J=1,NOBS),K=1,2)
      IF(NVAR.EQ.3)READ(5,FMT)(DATA(J,3),J=1,NOBS)
```

```
      IF(NVAR.EQ.2)GOTO 20
      DO 15 J=1,NOBS
   15 DATA(J,4)=DATA(J,2)/DATA(J,3)
   20 DO 25 J=1,NOBS
   25 DATA(J,5)=DATA(J,1)/DATA(J,2)
      DO 30 J=31,32
      DO 30 K=1,5
   30 DDMV(J,K)=0.0
      DO 35 J=1,5
      DO 40 K=1,NOBS
   40 DDMV(31,J)=DDMV(31,J)+DATA(K,J)
      DDMV(31,J)=DDMV(31,J)/NOBS
      DO 45 K=1,NOBS
   45 DDMV(K,J)=DATA(K,J)-DDMV(31,J)
      IF(NVAR.EQ.2.AND.J.EQ.3.OR.NVAR.EQ.2.AND.J.EQ.4)GOTO 35
      DO 50 K=1,NOBS
   50 DDMV(32,J)=DDMV(32,J)+DDMV(K,J)**2
   35 CONTINUE
      DDMV(32,1)=DDMV(32,1)/NOBS
      IF (ASK-3)55,60,55
   55 CALL BF2V(DDMV,NOBS,ST2,C2)
      DO 56 J=1,2
   56 NSTO2(IND2,J)=NAME(I,J)
      IF(ASK-3)65,65,60
   60 CALL BF3V(DDMV,NOBS,ST3,C3)
      DO 61 J=1,2
   61 NSTO3(IND3,J)=NAME(I,J)
   65 CALL DETR(DDMV,T,STAT,NOBS,ASK,I)
      IF(ASK-3)70,75,70
   70 CALL BF2V(DDMV,NOBS,STD2,C2)
      C2=C2+1
      IF(ASK-3)80,80,75
   75 CALL BF3V(DDMV,NOBS,STD3,C3)
      C3=C3+1
   80 CONTINUE
      IF(IND2.EQ.0)GOTO 85
      WRITE(6,450)
      CALL DPRINT(ST2,C2,NSTO2)
      WRITE(6,460)
      CALL DPRINT(STD2,C2,NSTO2)
   85 IF(IND3.EQ.0)GOTO 90
      WRITE(6,470)
      CALL DPRINT(ST3,C3,NSTO3)
      WRITE(6,480)
      CALL DPRINT(STD3,C3,NSTO3)
      CALL SPRINT(NAME,STAT,NNOBS,NPROD)
   90 CONTINUE
      GOTO 95
  100 CONTINUE
  400 FORMAT(20A4)
  410 FORMAT('0',20A4)
  420 FORMAT(I2,2X,19A4)
  430 FORMAT(2A4,3I2)
  450 FORMAT('1','TWO VARIABLE DECOMPOSITION OF VARIANCE - RAW DATA'///)
  460 FORMAT('1','TWO VARIABLE DECOMPOSITION OF VARIANCE - LINEARLY DETR
     1ENDED DATA'///)
  470 FORMAT('1','THREE VARIABLE DECOMPOSITION OF VARIANCE - RAW DATA'//
     1/)
  480 FORMAT('1','THREE VARIABLE DECOMPOSITION OF VARIANCE - LINEARLY DE
```

```
      1TRENDED DATA*////)
  500 FORMAT('0','CALCULATED USING RAW DATA MEANS')
  510 FORMAT('0','NOTE'/' ','FIRST ROW OF DECOMPOSITION = VARIANCE COMPO
      1NENTS'/' ','FOR 2 VARIABLES RESIDUAL IS ESSENTIALLY REDUNDANT BUT
      1CONTAINS CALCULATED SUM OF ZERO CROSS-PRODUCT TERMS'/' ','FOR 3 VA
      1RIABLES RESIDUAL IS THE SUM OF ALL NON-LINEAR TERMS'//' ','SECOND
      1ROW OF DECOMPOSITION = PERCENTAGE ATTRIBUTION USING DIRECT COMPONE
      1NTS AS DIVISORS'//' ','ERROR (FIRST ROW) IS COMPUTED WITH RESPECT
      1TO CALCULATED VARIANCE (SUM)'/' ','ERROR (SECOND ROW) USES Y VARIA
      1NCE')
       STOP
       END
C
C TWO VARIABLE DECOMPOSITION
C
       SUBROUTINE BF2V(X,N,S,M)
       REAL*8 X(32,5),S(60,11),MEM(4,4),A
       INTEGER*2 N,M
       DO 5 I=1,4
       DO 5 J=1,4
     5 MEM(I,J)=0.0
       A=1.0/N
       DO 10 I=1,N
    10 S(M,4)=S(M,4)-A*X(I,2)*X(I,5)
       DO 15 I=1,N
       S(M,1)=X(I,2)*X(31,5)
       S(M,2)=X(I,5)*X(31,2)
       S(M,3)=X(I,2)*X(I,5)
       DO 15 J=1,4
       DO 15 K=1,4
    15 MEM(J,K)=MEM(J,K)+A*S(M,J)*S(M,K)
       S(M,1)=X(32,1)
       S(M,2)=MEM(1,1)
       S(M,3)=MEM(2,2)
       S(M,4)=MEM(1,2)+MEM(2,1)
       S(M,5)=MEM(3,3)+MEM(4,4)+MEM(3,4)+MEM(4,3)
       S(M,6)=MEM(2,3)+MEM(3,2)
       S(M,7)=MEM(1,3)+MEM(3,1)
       S(M,8)=MEM(1,4)+MEM(4,1)+MEM(2,4)+MEM(4,2)
       S(M,9)=0.0
       DO 20 I=2,8
    20 S(M,9)=S(M,9)+S(M,I)
       S(M,10)=S(M,1)-S(M,9)
       S(M,11)=0.0
       DO 25 I=2,4
    25 S(M,11)=S(M,11)+S(M,I)
       S(M+1,11)=DABS(S(M,1)-S(M,11))/S(M,1)
       S(M,11)=DABS(S(M,9)-S(M,11))/S(M,9)
       A=0.0
       DO 30 I=2,3
    30 A=A+S(M,I)
       DO 35 I=2,7
    35 S(M+1,I)=S(M,I)/A
       S(M+1,8)=0.0
       RETURN
       END
C
C THREE VARIABLE DECOMPOSITION
C
```

```
      SUBROUTINE BF3V(X,N,S,M)
      REAL*8 X(32,5),S(60,11),MEM(11,11),A
      INTEGER*2 N,M
      DO 5 I=1,11
      DO 5 J=1,11
    5 MEM(I,J)=0.0
      A=1.0/N
      DO 10 I=1,N
      S(M,8)=S(M,8)-A*X(I,3)*X(I,4)*X(31,5)
      S(M,9)=S(M,9)-A*X(I,3)*X(I,5)*X(31,4)
      S(M,10)=S(M,10)-A*X(I,4)*X(I,5)*X(31,3)
   10 S(M,11)=S(M,11)-A*X(I,3)*X(I,4)*X(I,5)
      DO 20 I=1,N
      S(M,1)=X(I,3)*X(31,4)*X(31,5)
      S(M,2)=X(I,4)*X(31,3)*X(31,5)
      S(M,3)=X(I,5)*X(31,3)*X(31,4)
      S(M,4)=X(I,3)*X(I,4)*X(31,5)
      S(M,5)=X(I,3)*X(I,5)*X(31,4)
      S(M,6)=X(I,4)*X(I,5)*X(31,3)
      S(M,7)=X(I,3)*X(I,4)*X(I,5)
      DO 20 J=1,11
      DO 20 K=1,11
   20 MEM(J,K)=MEM(J,K)+A*S(M,J)*S(M,K)
      S(M,1)=X(32,1)
      S(M,2)=MEM(1,1)
      S(M,3)=MEM(2,2)
      S(M,4)=MEM(3,3)
      S(M,5)=MEM(1,2)+MEM(2,1)
      S(M,6)=MEM(1,3)+MEM(3,1)
      S(M,7)=MEM(2,3)+MEM(3,2)
      S(M,8)=0.0
      S(M,9)=0.0
      DO 25 I=1,11
      DO 25 J=1,11
   25 S(M,9)=S(M,9)+MEM(I,J)
      DO 30 I=2,7
   30 S(M,8)=S(M,8)+S(M,I)
      S(M,11)=DABS(S(M,9)-S(M,8))/S(M,9)
      S(M+1,11)=DABS(S(M,1)-S(M,8))/S(M,1)
      S(M,8)=S(M,9)-S(M,8)
      S(M,10)=S(M,1)-S(M,9)
      A=0.0
      DO 35 I=2,4
   35 A=A+S(M,I)
      DO 40 I=2,8
   40 S(M+1,I)=S(M,I)/A
      RETURN
      END
C
C LINEAR DETRENDING
C
      SUBROUTINE DETR(X,T,S,N,C,I)
      REAL*8 X(32,5),T(32),S(30,8),COV(4),B(4),SUM
      INTEGER*2 N,C,I
      DO 5 J=1,4
      COV(J)=0.0
    5 B(J)=0.0
      DO 10 J=1,4
      DO 15 K=1,N
```

```
   15 COV(J)=COV(J)+X(K,J+1)*T(K)
   10 B(J)=COV(J)/T(32)
      K=1
      DO 20 J=1,7,2
      IF(C.EQ.2.AND.J.EQ.3.OR.C.EQ.2.AND.J.EQ.5)GOTO 20
      S(I,J)=B(K)**2*T(32)/X(32,K+1)
      SUM=B(K)*COV(K)
      S(I,J+1)=SUM*(N-2)/(X(32,K+1)-SUM)
   20 K=K+1
      DO 25 J=1,4
      DO 25 K=1,N
   25 X(K,J+1)=X(K,J+1)-B(J)*T(K)
      RETURN
      END
C
C PRINTING OF DECOMPOSITION
C
      SUBROUTINE DPRINT(S,IND,NSTO)
      REAL*8 S(60,11)
      INTEGER*2 IND
      INTEGER*4 NSTO(30,2)
      WRITE(6,110)
      II=1
      DO 10 I=1,IND,2
      WRITE(6,120)(NSTO(II,J),J=1,2),(S(I,J),J=1,11)
      II=II+1
   10 WRITE(6,130)(S(I+1,J),J=2,8),S(I+1,11)
  110 FORMAT(' ','VARIABLE',4X,'VARIANCE Y',5X,'TERM 1',5X,'TERM 2',5X,
     1'TERM 3',5X,'TERM 4',5X,'TERM 5',5X,'TERM 6',3X,'RESIDUAL',8X,
     1'SUM',8X,'YV-SUM',2X,'ERROR'//)
  120 FORMAT('0',2A4,1X,E13.7,7(1X,E10.4),1X,E12.6,1X,E11.5,F7.4)
  130 FORMAT(' ',22X,7F11.4,25X,F7.4)
      RETURN
      END
C
C PRINTING OF REGRESSION DIAGNOSTIC STATISTICS
C
      SUBROUTINE SPRINT(N,S,NN,NP)
      REAL*8 S(30,8)
      INTEGER*2 NN(30),NP
      INTEGER*4 N(30,2)
      WRITE(6,200)
      WRITE(6,210)
      DO 10 I=1,NP
   10 WRITE(6,220)(N(I,J),J=1,2),(S(I,J),J=1,8),NN(I)
  200 FORMAT('1','DIAGNOSTIC STATISTICS FOR LINEARLY DETRENDED VARIABLES
     1')
  210 FORMAT('-','Y VARIABLE',18X,'X1',27X,'X2',27X,'X3',27X,'X4'/10X,
     14(10X,'R-SQUARED',2X,'F(1,N-2)'),4X,'N')
  220 FORMAT('0',1X,2A4,1X,4(8X,F10.4,F10.2,1X),2X,I2)
      RETURN
      END
```

REFERENCES

Adelman, Irma and Frank L. Adelman. "The Dynamic Properties of the Klein-Goldberger Model." Econometrica 27(1959):596-625.

Blandford, David and Seon Lee. "Quantitative Evaluation of Stabilization Policies in International Commodity Markets." Am. J. Agr. Econ. 61(1979): 128-134.

Box, G. E. P. and Jenkins, G. M. Time Series Analysis: Forecasting and Control. San Francisco: Holden Day, 1970.

Brook, E. M., E. R. Grilli, and J. Waelbroeck. Commodity Price Stabilization and the Developing Countries: The Problem of Choice. World Bank Staff Working Paper No. 262, 1977.

Burnstein, Harlan. Welfare Implications of Instability in Agricultural Commodity Markets. Dept. Agr. Econ. and Rural Soc., Pennsylvania State University, October 1977.

Burt, Oscar R. and Robert M. Finley. "Statistical Analysis of Identities in Random Variables." Am. J. Agr. Econ. 50(1968):734-744.

_________. "On the Statistical Analysis of Identities: A Reply," Am. J. Agr. Econ. 52(1970):155-156.

Chiang, A. C. Fundamentals of Mathematical Economics. New York: McGraw Hill, 1974.

Coppock, J. D. International Economic Instability. New York: McGraw Hill, 1962.

Firch, Robert S. "Sources of Commodity Market Instability in U.S. Agriculture." Am. J. Agr. Econ. 59(1977):162-169.

Foote, Richard J., John W. Klein, and Malcolm Clough. The Demand and Price Structure for Corn and Total Feed Concentrates. USDA Tech. Bull. No. 1061, 1952.

Gardner, B. "Instability in U.S. Agriculture - Discussion." Am. J. Agr. Econ. 59(1977):185-187.

Goldberger, Arthur S. Econometric Theory. New York: John Wiley and Sons, 1964.

_________. Impact Multipliers and Dynamic Properties of the Klein-Goldberger Model. Amsterdam: North-Holland, 1959.

_________. "On the Statistical Analysis of Identities: Comment." Am. J. Agr. Econ. 52(1970):154-155.

Goodman, Leo A. "On the Exact Variance of Products." J. Am. Stat. Assoc. 55(1960):708-713.

Houck, James P. "Some Aspects of Income Stabilization for Primary Producers." Aust. J. Agr. Econ. 17(1973):200-215.

Kendall, Maurice and A. Stuart. The Advanced Theory of Statistics Vol. III. London: Charles Griffin and Co., Ltd., 1966.

Knudsen, O. and A. Parnes. Trade Instability and Economic Development. Lexington, Mass.: Lexington Books, 1975.

Labys, Walter C. Dynamic Commodity Models: Specification, Estimation and Simulation. Lexington, Mass.: Lexington Books, 1973.

Massell, B. F. "Export Instability and Economic Structure." Am. Econ. Rev. 69(1970):618-630.

__________. "Some Welfare Implications of International Price Stabilization." J. Pol. Econ. 78(1970):404-417.

McKinna, David A. "The Nature and Extent of Instability in Annual Revenue from Certified Seed Potato Production in Victoria." M.S. Thesis, Monash University, 1976.

McNicol, D. Commodity Agreements and Price Stabilization. Lexington, Mass.: Lexington, 1978.

Meinken, Kenneth W. The Demand and Price Structure for Wheat. USDA Tech. Bull. No. 1136, 1955.

Mo, William Y. An Economic Analysis of the Dynamics of the United States Wheat Sector. USDA Tech. Bull. No. 1395, 1968.

Offutt, S. E. "The Empirical Analysis of Instability with Applications to Agriculture." M.S. Thesis, Cornell University, 1980.

Oi, W. "The Desirability of Price Instability Under Perfect Competition." Econometrica 29(1961):58-64.

Piggott, R. R. "Decomposing the Variance of Gross Revenue into Demand and Supply Components." Aust. J. Agr. Econ. 22(1978):145-157.

Porter, Richard C. "On Placing the Blame for Primary Product Instability." Intl. Econ. Rev. 11(1970):175-178.

Rourke, B. E. "Short-range Forecasting of Coffee Production." Food Res. Inst. Studies in Agr. Econ., Trade and Dev. No. 3 (1970):197-214.

Sackrin, S. M. "Measuring the Relative Influence of Acreage and Yield Changes on Crop Production." Agr. Econ. Res. 9(1957):136-139.