AGRICULTURAL ECONOMICS RESEARCH

JULY 1979 VOL. 31 NO. 3

Banning pesticide use on cotton...p.3

Agricultural Economics Research

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

• • •

When most people develop interest in a problem, they can usually soon articulate the problem and also an answer. But they frequently fail to come to grips with the reasons which explain the problem. Researchers, however, spend their working lives gaining familiarity with reasons—often at the expense of being articulate about specific issues or answers. This dichotomy may seem odd, but it apparently has existed for a long time. Joseph Schumpeter, in his *History of Economic Analysis*, had it in mind when he distinguished between economic analysis and economic views. He recognized the contribution of the ancient Greeks to analysis, but thought that, through most of history, we have had more to say about economic issues and answers than about economic reasons.

Schumpeter defined economic analysis as the development of an intellectual procedure that can clarify economic problems. The analyst has a command of techniques. Part of the genius of Western civilization can be attributed to the latitude given to people with analytical, research-oriented minds, even though, as Ed Bishop pointed out in the October 1976 issue of this journal, the history of academic freedom has been uneven. The articles in this issue involve research methods, the use of which can be important in arriving at reasons; but their importance may be lost on people who are concerned only with issues and answers. Yet, if the reasons are not right, the answers may not be right either, and the issues will continue to be unresolved.

Lack of access to good data often limits our search for reasons. Charles Sisson explains how two incomplete data sources can be merged to form what he calls a synthetic data file. Under certain conditions, the technique produces a file which contains information not in the separate sources. This increases the amount of information that can be extracted from secondary data sources and does so at a reasonable cost.

Budgeting allows researchers to search for reasons by weighing the consequences of alternative actions. Weisz, Miller, and Quinby use a computerized form of budgeting which they call stochastic simulation to compare the trends in prices and quantities of a farm commodity with what would be likely to happen after a change in agricultural technology. Here, the change is a ban on the use of the pesticide toxaphene on cotton.

When a researcher divides the elements of one data series, say income, by another, say price, the process is called deflation. The resulting variable, real income in this example, often is believed to be more appropriate for analysis than the original variables. Bell, Roop, and Willis examine the statistical properties of deflation and find it is not a technique to be used casually. Deflation can influence our discernment of reasons because it affects tests of significance, such as the correlation coefficient and t-ratio. It can change the sign for a regression coefficient. Take warning, because deflation can be used as one more means of lying with statistics.

CLARK EDWARDS

Contents

The Secretary of Agriculture has determined that the publication of this periodical
is necessary in the transaction of the public business required by law of this Department.
Use of funds for printing this periodical has been approved by the Director,
Office of Management and Budget, through February 28, 1980.

For sale by the Superintendent of Documents, U.S. Government Printing Office,
Washington, D.C. 20402. $6.50 a year domestic; $8.25, foreign.

The Synthetic Micro Data File: A New Tool for Economists

By Charles A. Sisson*

Detailed data files required to fill many economic models are not available. Direct construction of a needed file often proves to be prohibitively expensive. The author of this article poses one alternative: to synthesize a file by merging two or more existing ones that, between them, contain the needed information. For example, consider a researcher who wishes to know how economic variables affect sociological behavior and who has found one file with economic and demographic information and another with sociological and demographic information. By matching demographic characteristics, the researcher can construct a synthetic file to use in analyzing the relationship of economic and sociological characteristics.

Keywords:

Data
Synthetic data files
Research methods

Economists require detailed information about the characteristics of the economy to formulate a rational economic policy. The more complete the micro files they employ in their research, the more confident they can be in their policy recommendations. National income accounts and other summary figures cannot provide precise enough detail about the interactions at the micro level that are the foundation of economics (*9*, *17*, and *13*).[1] As Wassily Leontief noted:

> The time is past when the best that could be done with large sets of variables was to reduce their number by averaging them out or what is essentially the same, combining them into broad aggregates; now we can manipulate complicated analytical systems without suppressing the identity of their individual elements (*9*, p. 6).

Unfortunately, our ability to process large masses of data has exceeded our means to generate empirical bases for hypothesis testing and policy formulation. Ideally, all micro-based studies would stem from a carefully chosen sample of the population that included relevant income, expenditure, tax, and demographic factors. The direct approach to constructing such a file—collecting a sample—is prohibitively expensive. The impacts of many important policy changes are thus estimated by gross approximation. Yet it is obviously self-defeating to use macro subtotals to examine policy changes that have impacts only through their influence on individuals.

Seeking to improve their methodology within the formidable constraints barring construction of a true micro file, researchers have turned to synthetic micro files as a practical and improved base for their policy prescriptions. Synthetic micro files are not a true sample of the population. They are formed by a matching or merging of two different micro files that, between them, contain information about the desired variables. This technique might be useful, for example, if one had a microeconomic source of demographic characteristics for a specific socioeconomic group and another microeconomic source of information on their economic status, but wished to know how economic variables affected sociological structure. Suppose a researcher had a microeconomic data file, such as the agricultural census, and wished to extend its usefulness without creating a new microeconomic data file. The researcher might consider enlarging its applicability by "merging" it into a second, appropriate data file to create a file that would supply information on the missing relationships in the first file.

In this article, I examine merits and shortcomings of synthetic micro data files, review examples of such files, and explain procedures for constructing them. Results should be useful to each researcher in economics.

Synthetic data files are no panacea. Although they may be a useful research tool, they have shortcomings. They may produce the data base for useful and varied micro studies, but results are contingent on the appropriateness of the matching process. Although the creation of synthetic data files can significantly save money over the expense of designing and collecting a microeconomic survey, their construction requires huge investments of human and computer time, and patience. Studies employing such files can be costly. Also, it takes 2 or 3 years to collect the data and another year or more before the file can be constructed. If results from studies using such a file are to be more than an historical exer-

---

*At the time this article was written, the author was an economist with the National Economic Analysis Division, ESCS. He is now with the International Monetary Fund, Washington, D.C.

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

Detailed data files required to fill many economic models are not available. Direct construction of a needed file often proves to be prohibitively expensive. The author of this article poses one alternative: to synthesize a file by merging two or more existing ones that, between them, contain the needed information.

cise, one must presume that the relations depicted have not changed over the intervening span of years.[2]

Several questions arise: How is a synthetic micro file constructed? How relevant are these files? How big can a synthetic data file be—that is, if it makes sense to merge data files A and B to form C, does it make sense to merge C and some other data file D to form E? What types of synthetic data files have been constructed? What are their relative merits?

THE BROOKINGS MERGE FILE[3]

The following indepth view of the Brookings MERGE file should make the abstract concepts of synthetic micro files more understandable.

The Brookings MERGE file synthetically links individual records from two sources: The U.S. Internal Revenue Service (IRS) tax file for 1966, which contains individual Federal income tax returns; and the 1967 Survey of Economic Opportunity (SEO) data file, which samples the total U.S. population through field interviews. Both files reference individual family income for calendar year 1966, but each contains information not found in the other. The IRS tax file contains more complete tax information; the SEO file, more complete demographic information.

The Brookings MERGE file links information from the 87,000 individual records in the IRS tax file with the 30,000 household records in the SEO file. Each family record in the SEO file was examined to determine if any member of the family would be expected to have filed a tax return in 1966. If so, tax information from the return judged most likely to have been filed by that individual was added to the demographic information in the sample record. The process optimizes a "distance function" after certain basic criteria for a match are satisfied. How this is done is explained below.

The Brookings MERGE file is not perfectly synthetic. Low-income records from the SEO file have no tax data associated with them because no tax was paid. High-income tax file records in the final version of the MERGE file provide no demographic records because tax information alone is available for families with incomes above $30,000.[4] It is the vast middle range of records that have been artificially linked by the matching process.

Once an individual from a SEO family record is judged likely to have filed a tax return, the matching process consists of finding a return in the IRS tax file that closely represents the actual tax return that he or she would have filed. First, a set of "cells" or "equivalence classes" is constructed to serve as first-stage, or rough sort. All IRS tax returns occupying the same cell as the supposed (constructed) tax return from the SEO record are compared in a second-stage sort, based on income. Finally (ideally) a match is determined based on a distance function. Each match is randomly determined from all returns that fall within a standard income range of the SEO record.

The "equivalence classes" are formed on the basis of four criteria: (1) type of return filed—single, joint, surviving spouse, head of household; (2) age of the household head or spouse—65 years old or over; (3) number of exemptions—1, 2, 3, 4, 5 or more; and (4) reported pattern of income—major and minor sources of income (in absolute terms). Four classes of income were considered: wages, business, farm, and property.

If the cells as defined had been interpreted strictly, there would have been 1,420 different categories. So many cells would have been left empty that many SEO records would have been impossible to match, and still more would have yielded improbable income matches. Accordingly, the 1,420 original cells were collapsed to 74 somewhat densely populated cells.

The original criterion for the second stage, or major income source match, was the major income for the

---

[2] In the short run, this is usually a reasonable assumption. Joseph Pechman stated at the National Tax Association-Tax Institute of America Symposium held in Washington, D.C., July 10, 1975: "The 1966 MERGE file shows the gist of income and tax distributions even today."

[3] The material in this section is drawn from (*11* and *14*, pp. 84-92, unless otherwise noted.

[4] Groups with high and those with negative incomes are not really separate as returns with substantial losses are generally filed by wealthy persons. See (*11*, p. 335).

There may be a means of eliminating some bias in the SEO survey. Personal surveys are notorious for underreporting income for those at high income levels. See (*14*, p. 85).

SEO unit, plus or minus 2 percent.[5] All returns in the acceptable income range were subjected to a final sieve, a "consistency score," to help narrow the choice. Hitherto unused information was used to effect the most suitable match. Six criteria were used, each with different weights. If the tax return matched the related characteristic on the SEO record, it was awarded points.

Tax return	SEO record	Points
Home mortgage interest deduction on property tax return	Home ownership or debt (or house value in farm value)	12
Interest or dividend income	Interest or dividend income or ownership of stocks, bonds, and others	8
Farm income	Farm income or farm assets or debt	10
Business income	Business income or business assets or debt	10
Rental income or real estate property tax deduction	Rental income or real estate assets or debt	9
Nonzero capital gains income[1]	Dividends or interest on stocks, bonds, and others	8

[1] Capital gains equal to zero on the tax return and earnings from property in the SEO file are consistent.

The return with the highest consistency score was not necessarily matched to the SEO record. Any return in the highest 25 percent of those for which consistency scores were calculated was equally likely to be selected, if it scored 25 points or more. This procedure was sufficient to make most of the matches (97 percent).

SEO records that could not be matched by this technique were reprocessed, and an iterative process was begun. The income acceptability range was widened by 1 percent,[6] consistency scores for eligible tax returns were recalculated, and a match was determined based on the same consistency criteria. Records that failed this test were reprocessed six times, or until a match was determined; each time the acceptable income range was widened 1 percent. Records still lacking a match—0.5 percent of the 28,643 returns—were hand matched.

THEORETICAL BASIS FOR SYNTHETIC MICROECONOMIC DATA FILES

Constructing a synthetic, merged file generally involves merging two samples whose overlap is insignificant. Certain variables, denoted by the vector X, appear in both samples. Other variables, represented by the vector Y, appear in one sample; others, Z, appear in the second sample.[7] The ideal is a single sample with information on the joint distribution F(X,Y,Z). As this does not exist, an artificial one must be generated. This construction is a special case of the following general problem: Given samples from two marginal distributions of a joint distribution, estimate the joint distribution and generate a sample from it.[8] The difficulty is estimating the joint distribution.

The problem involves so many variables that it is difficult to conceptualize. Graphical presentation is also difficult. Two partial views that may assist the reader are in figures 1 and 2. Figure 1 shows the crux of the matter: The joint distribution of X and Y and the joint distribution of X and Z are known, but the joint distribution of Y and Z is unknown. If X and Y are single variables, the joint distribution of X and Y

[5] Limitations on the total amount of error excluded impossible (at lower limits) or overgenerous (at higher limits) margins of error.

[6] With corresponding increases in the maximum and minimum amounts that were permissible.

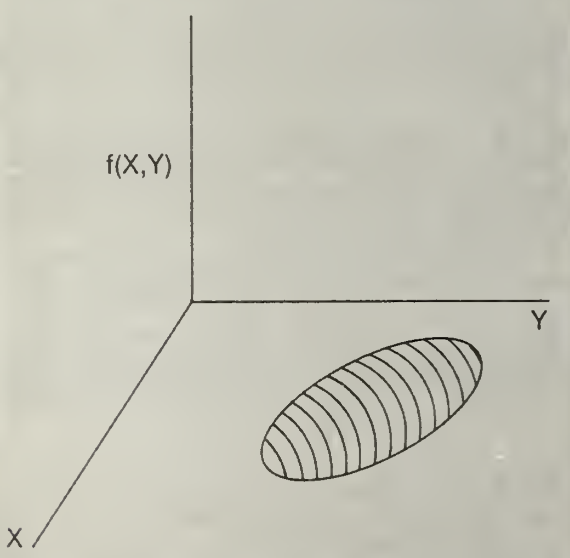
[7] This notation, which relies on (*21*), is standard in the literature.

[8] This approach relies on (*21*).

FIGURE 1
Distribution of Variables in Joint Sample of (X,Y,Z)

	X	Y
X	Known	Known
Z	Known	Not Known

FIGURE 2
Joint Distribution of X and Y



in 3-space might look as depicted in figure 2. The joint distribution F(X,Y,Z) occupies 4-space and cannot be represented here. It can, however, be defined in terms of conditional probabilities:

$$F(X,Y,Z) = F(Y/Z,X) \cdot F(Z/X) \cdot F(X)$$

The latter two terms in this expression are known. However, as the joint distribution of Y and Z is not known, F(Y/Z,X) is unknown. If we could estimate F(Y/Z,X), the joint distribution F(X,Y,Z) could be computed from the above equation.

As a first step, let us suppose that X is a k-dimensional variable; furthermore, let us suppose that it is divided into k-dimensional cells k(X) small enough for the distribution of Y to be essentially independent of X within each cell.[9] Then within each cell:

$$F(Y/X) \simeq F(Y)$$

and

$$F(Y/Z,X) \simeq F(Y/X)$$

so

$$F(X,Y,Z) \simeq F(X) \cdot F(Z/X) \cdot F(Y)$$

As the information needed to estimate F(Z/X) is in the original data file, the joint distribution F(X,Y,Z) can be estimated.

PROBLEMS OF APPLICATION

In theory, it is possible to calculate F(X,Y,Z) from knowledge about the joint distributions F(X,Y) and F(X,Z), but clearly there may be problems applying this technique. First, a means must be derived for determining when the cells k(X) are small enough to consider Y independent of X, within cells. Such a sieve need not be of uniform dimension with X. In fact, it would be better if the dimensions varied to reflect the density of the data. The more dense the data within cells, the better the estimate of the distribution F(Y/X) will be; however, the

[9] If the cells are densely filled, F(Y/X) = F(Y) can be estimated directly. If the cells are sparsely populated, regression techniques can be used in conjunction with some smoothness assumptions.

The more dense the data within cells, the better the estimate of the distribution F(Y/X) will be; however, the cells must be small enough to justify the assumption of independence between X and Y

cells must be small enough to justify the assumption of independence between X and Y.

As the data will vary in density, this trade off can be handled by varying the cell size as conditions warrant. Nonetheless, the sieve must be composed of cells in which X and Y are independent. A means of testing this assumption is required.

A first step in this direction has been proposed by Nancy and Richard Ruggles (*16*, pp. 360-362). They suggest a chi-squared test of the hypothesis that the samples in a cell came from different universes, and a second (correlation) test to evaluate the relative importance of these differences before making necessary adjustments. These tests do provide some basis for hypothesis testing, but I am unconvinced that they are reliable (*20*, p. 397).

An alternative to testing the validity of the synthetic micro file would be to introduce restrictions into the matching process deliberately. One could ignore some common variables in the two files (let us call these X') during the creation of the synthetic file, and estimate the mean square error of the artificial matches for X_1 to their actual value. This approach, however, would lead to inefficiencies in the actual matching process as some information instrumental in making the union would be deliberately sacrificed. At this point, the Heisenberg Uncertainty Principle becomes a factor.

For hypothesis testing to be meaningful, the matching cells must be densely occupied—otherwise the testing procedure cannot have statistical validity. This condition might be expected in the core (or central portion) of a large matched file, but it would require a large file. Even then, the fringe (or outer portions) of the file will be too scattered. For example, suppose X_1 is individual incomes and Y_1 is tax liability. At high values of X_1, the individuals will be too diverse to allow sufficient expansion of cell intervals along Y_1 to encompass a large number of wealthy individuals while maintaining the premise that Y_1 and X_1 are independent.

Obviously, the fewer the outliers and the more dense the data, the more technically correct the finished synthetic file will be. Thus, the more valid it is to "stack" this file with another—that is, to use this file as a basis for another synthetic file. However, this should not imply that outliers are "bad." They represent valuable information, and if their absence implies a "better" file, it is only "better" in a statistical sense related to ease of matching when pyramiding artificial files. A file that lacks outliers may be unrepresentative of the population and it may be quite misleading.[10] One reason statisticians square the distance from the regression line to sample points is to give greater weight to "extraordinary" points, and it behooves the synthetic file builder to be aware of the informative value of outlying points.

Yet outliers do pose a special problem for the type of matching technique proposed here. Their existence implies that some regions in the sample are so sparsely populated that some cells will lack match records. In practice, the samples—even for the largest files—will have cell-vacancy problems. The practical solution is to combine some of the X variables, which thereby collapses some of the cells (*16*, p. 357). The Brookings MERGE file, for example, has 74 nonincome classifications instead of the 1,000-plus first envisioned. If these cells had not been telescoped, many original cells in the grid would have produced obvious mismatches or no match at all.

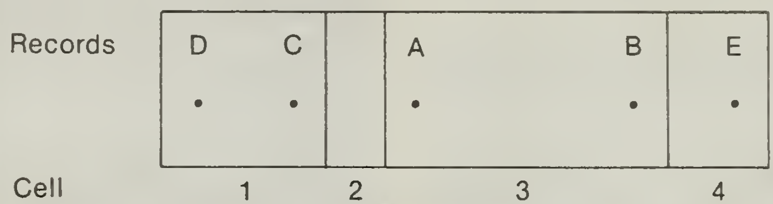
This difficulty is usually treated by using a metric technique, generally a distance function. Distance functions rank possible matches by their "closeness" to the record to be matched, not by whether they occupy the same cell. If a cell technique were strictly applied, only sample records sorted into the same cell would be linked. This situation might lead to some matched cells that were unnecessarily diverse.

Figure 3 depicts such a case for a single variable X_1. Sample A is a record for file (X,Z), and it is to be matched with a record from file (X,Y). The records B, C, D, and E are the leading candidates for matching. What is the best match for A? The cell technique would signify B should be chosen, because it occupies cell 3 in common with A. However, the distance function would rank B as the next to poorest choice of the four possibilities. The "closest" record to A is C, and the distance criterion would indicate it should be chosen, regardless of the difference in cells. Which is the best?

Following the premise that the cells were constructed such that the conditional probability of Y is not independent of X between cells, B is the best choice. But if

[10] The problem occurs particularly with income and expenditure distributions, which are generally skewed.

FIGURE 3
Individual Records in a Cell Structure



B did not exist and point E were the alternative to C and D, which should then be chosen? These are usual circumstances with a limited sample, and here the cell-matching technique breaks down completely. Of course, if the cells have been chosen to reflect Y's independence of X, the question is one of minimizing damage. Given that an ideal match is impossible, which is the best match? If one is willing to assume that the distribution will not be markedly asymmetrical, choosing the record that is closest or among those closest is a reasonable standard.[11]

SPECIFIC EXAMPLES OF SYNTHETIC MICRO DATA FILES

Several existing synthetic files have used variants of the distance function concept. Researchers creating the Canadian Survey of Consumer Finances-Family Expenditure Survey synthetic data file (SCF-FES) used multivariate analysis to determine their distance function ranking. The variables Y_i and Z_i were regressed individually on all the variables X. The explanatory power of the various X_i in the regressions on Y_i and Z_i was used to determine the variable's weight in the distance function. If a variable X_j had high partial R-squares for a wide range of the Y_i and Z_i variables, it was considered a crucial element in the matching process. Records from the two files that had similar values for that variable were awarded a relatively large number of points toward qualifying for matching. A variable X_k having low partial R-squares for most Y_i and Z_i variables was considered relatively inconsequential to a good match. It was assigned either a low or zero point contribution for the matching criteria. Matches between the two files were little influenced by the correspondence of these variables. Pairs of records with high match scores were linked (*1*).

The Brookings MERGE file uses a more *ad hoc* approach to the distance function. The relative importance of the X variables in distinguishing a good match was predetermined on what were considered reasonable grounds rather than by quantitative analysis (*11*, *12*, and *14*).

The Bureau of Economic Analysis (BEA), U.S. Department of Commerce, developed its synthetic file, a matching of the Current Population Survey and the Tax Model for the year 1964, in a similar manner to the Brookings effort. However, instead of using the Brookings technique of sampling for matches, the BEA file involved a one-to-one match between the two files. Each tax record was assigned to a unique population record

[11] It does, however, entail an implicit assumption of independence between the Y and Z, given X. See (*21*, p. 343).

The final match between two records (from different files) is determined by the closeness of the match between the corresponding X_i and by the preselected weights.

by matching records having the same rank order within broad *a priori*-determined equivalence classes. Each cell was defined so that it has the same weighted number of records from each file, which avoids the issue of improper population aggregates extrapolated from individual records.[12]

The synthetic file built by Nancy and Richard Ruggles under the auspices of the National Bureau of Economic Research (NBER) uses a distance function that is less arbitrary in cutting across previously defined cell structures. The Ruggles' success in adhering closely to the prescribed sampling structure is due chiefly to one advantage: more data. The files described earlier involved matching files on the order of 50,000 records each. The NBER file matches the 1970 Public Use Sample with the Social Security Longitudinal Employer-Employee Data file; each has 2 million records. These files are so large, in fact, that they not only permitted the use of a cell-structure technique but also eliminated explicit use of a distance function of the type employed by other researchers. Metric calculations take up computer time; and one must consider efficiency when processing 2 million records. A cell technique is not only theoretically more desirable, in this case it is practical.[13]

THE WEIGHTING PROBLEM

The NBER synthetic file technique does not necessarily approach optimal efficiency. The cell dimensions may well be larger than they should be. As noted, this will be especially true in the fringe, or outer portions of the file. Collapsing the cells in these portions of the file implies a distance function of a rudimentary sort. It is important to recognize the significance of the various weights or, as is the usual case, points assigned to each variable X_i in the distance function. The final match between two records (from different files) is determined by the closeness of the match between the corresponding X_i and by the preselected weights.

[12] This is known as the alignment problem. For a discussion of this problem, see (*14*, p. 88). For a discussion of the BEA synthetic file, see (*3* and *2*).

[13] (*16*, pp. 370-371). For a general review of the NBER data file, also see (*17*).

Obviously, as concepts about which variables are most important to a "good" match change and as those decisions are reflected in a different weighting scheme, the synthetic file changes. Certain records that would have been considered matches will no longer be considered satisfactory and will be dropped, and others that would have been considered unfit will now be linked.

The importance of the weighting scheme finally adopted lies in its determination of the accuracy of the file. Whether or not these weights are determined empirically, as Alter and the Ruggles did, or theoretically, as Okner did, there is considerable subjectivity in the final determination. It is also true that the file will have relative strengths and weaknesses according to the use that is made of it. The file is usually designed for general purposes, and the weights are chosen to provide a mean between conflicting goals. Ideally, all common variables X are in correspondence before a match is determined, but generally, only the most basic variables will approximate each other. For example, total income is generally such a basic variable, and records from two different files need to have very similar incomes to be eligible for matching. However, the source of that income is less important, and greater margins of error for individual sources of income are consistent with a "good" match.

The points awarded toward a match reflect an income source's importance and determine the trade off *vis-a-vis* other aspects of the X sector. In a specific study of farm taxation, for example, better results would be obtained if farm income were emphasized as a basic variable in the matching process. There would then be a greater likelihood that the (X,Y) and (X,Z) records would both have farm income, and the file quality for such a specific purpose would be improved. Given the expense of the matching process, however, it is more reasonable to construct a multipurpose file and use it for specific tasks rather than to construct a special file for each research task.

OTHER APPROACHES

Only one means of effecting file links—matching—has been considered although it is not the only process available. Perhaps the most important alternative to

Synthetic data files, when properly constructed, can provide more conclusive answers to policy questions than other more traditional approaches.

matching is a regression technique.[14] One file can be used to define the functional relationship between the common variables (X) and the disjunct variables it contains (Z). This relationship can then be used to append estimated Z values to the information in the second file, using the X values in file (X,Y) as a basis for the imputation (*16*, p. 354). That is, Z = f(X) would be estimated from the first file, and the X values in file (X,Y) would be used as the basis for calculating (X,Y,f(X)).

There are, however, several deficiencies in this approach. Perhaps the most grievous of the econometric problems (regressions imply their existence) is equation specification. The relationship between X and Z is unlikely to be well known—if it were, there would be little need for the first data file—and this relationship is even less likely to be linear throughout the domain. Thus it is extremely unlikely that the true joint distribution (X,Y,Z) can be well approximated by (X,Y,f(X)) (*20*, p. 395). A second major problem might be multicollinearity. A complex set of economic information, such as budget outlays, usually has highly inter-related components. Separate estimates of each outlay would lead to inconsistent estimates of the aggregate.[15] Another likely problem is heteroskedasticity. Many econometric studies using cross-sectional data find a changing variance in the disturbance term (7, p. 214).

For these and other reasons, regression analysis seems an inappropriate alternative to matching. From a methodological standpoint, it is inferior because it fails to produce the original variance of the data set when the imputations are generated. The regressions always assign mean values, whereas a matching process reproduces the distribution of variables in the original set over repeated imputations (7, p. 214). However, the basis for choice is hardly as one-sided as has sometimes been claimed. The Ruggleses have espoused a matching process over a regression technique because

> . . . for matching purposes no specific functional relationship need be determined in advance. Nonlinear relationships will automatically be handled as efficiently as linear relationships, without explicit recognition that the relationships are nonlinear (*16*, p. 354).

If this observation were pertinent, it would be sufficient cause to rely exclusively on matching techniques; however it misses the point. Under the simplest conditions, when Y and Z are independent, the two techniques give identical results (although the regression equation does not reproduce the original variance in the data set), but when there are interdependencies and nonlinearities, the two techniques differ. Sims has neatly summarized the problem:

> To justify a matching procedure one requires an assumption that the regression relation giving the conditional distribution of (say) X as a function of X is *constant* and *a fortiori* that the mean of Y is a constant conditional on X. This is a much stronger requirement than the assumption that the conditional mean of Y be linear in X (*20*, p. 395).

A matching technique does have greater flexibility than a regression, but the assumptions necessary for its success are more stringent.

CONCLUSIONS

Synthetic data files, when properly constructed, can provide more conclusive answers to policy questions than other more traditional approaches. However, in practice, several expedients are used which entail a loss in validity. Further, a micro data file cannot provide quick, rough approximations on exceedingly broad topics; summary tables still have their place for quick estimates and general guides. A microfile is more unwieldy and complicated, but—if well constructed—capable of precise calculations and it can be used for a range of topics. Although it would be possible to

[14] Other choices might be averaging or interpolation techniques.

[15] The Ruggleses consider this property to exemplify the superiority of a matching process, arguing that it is a simpler and more satisfactory way of transferring complete sets of budget information from observations in one sample to observations in another (*16*, p. 354). Admittedly, it does retain the integrity of each set of information, but it should be possible to improve on such a naive estimating approach. The regression analysis could be modified by including a constraint to produce consistent answers. See (7, pp. 155-159).

pyramid this process and to create synthetic files, the conditions necessary for this to be practical would prevent the file from being of more than academic interest.

Some rough guidelines for construction of a synthetic file can be these: If the functional relationships linking two of the three variable sets X, Y, and Z are well-known and the data are scattered, regression analysis would probably be superior to matching. Otherwise, matching is the best strategy. In matching, the more dense the data, the more closely one can approximate the conditions of the ideal scheme; therefore, larger data sets are preferred. Of course, a researcher may find that the relevant variables are only available in small samples—a situation which prevents him from constructing a more accurate synthetic file.

This may seem inconclusive. The ambiguity stems from the nature of the matching problem. As Sims has noted:

> . . . there is no way to avoid "subjective" use of economic theory in deciding when a match is bad. In their (the researchers') eagerness to avoid "subjective" assumptions about the nature of the distribution they are estimating, matchers have been letting the computer make foolish assumptions for them (*20*, p. 397).

The problem may not be easily answered, but we need to solve it because synthetic data do offer the promise of a better understanding of economic relationships.

In Earlier Issues

Studies of relationships between agriculture and the rest of the economy must continually weigh the conveniences of aggregation against losses of relevant detail. At one extreme are simple models which treat all agriculture as one enterprise selling a single composite product. But the diversity of conditions within agriculture generally forces us to frame price and production programs in terms of individual commodities. Modern techniques of analysis, such as the input-output or "interindustry relations" approach of Leontief and the "linear programming" methods of Dantzig, Koopmans and others, are creating a demand for more accurate data. . . . Electronic computers can handle the formidable calculations required for such studies, but the accuracy of the final results must depend on that of the basic data. For this reason, agricultural economists should take an active interest in the interpretation, application, and further development, of the interindustry relations approach most recently exemplified by the Bureau of Labor Statistics study of the U.S. economy in 1947. As time goes on, we need to supplement the input-output approach with one that permits us to use, among other things, our knowledge of demand and supply curves for agricultural commodities. Conceptually, this leads us into a very large system of simultaneous equations—a sort of "econometric map" of the agricultural economy in the framework of total economic activity. Our single-equation demand analyses, and sub-models of moderate complexity, would be as useful as ever. But the over-all model would force upon us a keener awareness of the nature of the approximations we were making, and of the variables or sets of economic relationships that we were assuming constant.

Karl A. Fox
and Harry C. Norcross
AER, Vol. IV, No. 1,
Jan. 1952, pp. 13 and 21

BIBLIOGRAPHY

(*1*) Alter, Horst E. "Creation of a Synthetic Data Set by Linking Records of the Canadian Survey of Consumer Finances with the Family Expenditure Survey 1970." *Annals of Economic and Social Measurement* 3 (Spring 1974): 373-394.

(*2*) Budd, Edward C. "Comments." *Annals of Economic and Social Measurement* 1 (Summer 1972): 349-354.

(*3*) ___________ . "The Creation of a Microdata File for Estimating the Size Distribution of Income." *Review of Income and Wealth 17* (December 1971): 317-334.

(*4*) Diesing, Paul. *Patterns of Discovery in the Social Sciences.* Chicago: Aldine-Atherton, 1971.

(*5*) Georgescu-Roegen, N. "Economic Theory and Agrarian Economics." *Oxford Economic Papers* 12 (Feb. 1960): 1-40.

(*6*) Goldsmith, Selma F. "Appraisal of Basic Data Available for Constructing Income Size Distributions." In *Conference on Research in Income and Wealth*, Vol. 13, pp. 267-377. New York: National Bureau of Economic Research, 1951.

(*7*) Johnston, J. *Econometric Methods.* New York: McGraw-Hill, 1972.

(*8*) Klein, Lawrence R. *A Textbook of Econometrics.* Evanston, Illinois: Row, Peterson and Company, 1953.

(*9*) Leontief, Wassily. "Theoretical Assumptions and Nonobserved Facts." *The American Economic Review* 61 (Mar. 1971): 1-7.

(*10*) Morgenstern, Oskar. *On the Accuracy of Economic Observations.* Princeton: Princeton University Press, 1973.

(*11*) Okner, Benjamin A. "Constructing a New Data Base from Existing Microdata Sets: The 1966 MERGE File." *Annals of Economic and Social Measurement* 1 (Summer 1972): 325-342.

(*12*) ___________ . "Data Matching and Merging: An Overview." *Annals of Economic and Social Measurement* 3 (Sept. 1974): 347-352.

(*13*) ___________ . "Reply and Comments." *Annals of Economic and Social Measurement* 1 (Summer 1972): 359-362.

(*14*) Pechman, Joseph A., and Benjamin A. Okner. *Who Bears the Tax Burden?* Washington, D.C.: The Brookings Institution, 1974.

(*15*) Peck, Jon K. "Comments." *Annals of Economic and Social Measurement* 1 (Summer 1972): 347-348.

(*16*) Ruggles, Nancy, and Richard Ruggles. "A Strategy for Merging and Matching Microdata Sets." *Annals of Economic and Social Measurement* 3 (Spring 1974): 353-371.

(*17*) Ruggles, Richard, and Nancy Ruggles. "The Role of Microdata in the National Economic and Social Accounts." *Review of Income and Wealth* 21 (June 1975): 203-216.

(*18*) Shoup, Carl S. *Quantitative Research in Taxation and Government Expenditure.* New York: National Bureau of Economic Research, 1973.

(*19*) Siegel, Sidney, *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill, 1956.

(*20*) Sims, Christopher A. "Comment." *Annals of Economic and Social Measurement* 3 (Spring 1974): 395-397.

(*21*) ___________ . "Comments." *Annals of Economic and Social Measurement* 1 (Summer 1972): 343-345.

(*22*) ___________ . "Rejoinder." *Annals of Economic and Social Measurement 1 (Summer 1972): 355-357.*

(23) Theil, Henry. *Linear Aggregation of Economic Relations.* Amsterdam: North-Holland Publishing Co., 1954.

The Stochastic Effects of a Ban on Toxaphene Use on Cotton

By Reuben N. Weisz, Ronald R. Miller, and William Quinby*

A ban on toxaphene use in control of the cotton budworm-bollworm would increase the average price of cotton as well as its price instability. It would decrease the level and increase the variability of cotton yield and production. Such a ban would also decrease the expected value but not the variance of exports. A Monte Carlo economic simulation model was used to evaluate stochastic impacts of pesticide regulation. This methodology should be applicable also in future technology impact policy analyses.

Keywords

Pesticide
Policy
Technology assessment
Simulation
Risk

Decisionmakers are considering explicitly the concept of risk whenever they behave as if they know the probability distribution of the consequences of the decision they have made. Policymakers, asking for the opportunity to play "Jimmy the Greek" in recent years, have asked research analysts to evaluate the full range of outcomes associated with a policy decision and to determine the probability associated with each level of outcome (*18*).[1]

Analysts of pesticide policy, however, have evaluated changes in levels of costs and yields but have failed to recognize the aggregate stochastic impacts of pesticide regulations (see *17*, for example). Because the dominant rationale for using pesticides may be to minimize risk, policymakers should evaluate the farm level and aggregate impacts of pesticide policies on risk. Numerous microeconomic studies have stressed the importance of incorporating information on risk into the analysis of technological changes for a firm (see *9*, for example).

We present a methodology that can be used to incorporate risk implications into evaluations of pesticide policies, and we use this approach to evaluate a ban on the use of toxaphene for the budworm-bollworm on cotton. This general approach should be applicable to other impact assessment studies.[2]

GENERAL METHODOLOGY

The National Agricultural Policy Simulator, POLYSIM, was the economic model used, particularly its cotton yield and acreage equations. The model can be made stochastic, which allows the decisionmaker to evaluate the statistical characteristics associated with the consequences of alternative policies.

Overview of POLYSIM

Ray and Moriak provide an overview of the deterministic version:

> The POLYSIM model was constructed differently from most simulation models to attain the desired policy analysis capability. The model makes full use of the forecasted data as a reference baseline. . . . POLYSIM simulates the effects of policy specifications that differ from those assumed in the baseline while holding all other supply and demand shifters the same. The model thus focuses on the interaction of supply and demand responses that result from specified changes in policy variables (*8*).

The model contains the following commodities: feed grains, wheat, soybeans, cotton, cattle and calves, hogs, sheep and lambs, chicken, turkeys, eggs, and milk.

POLYSIM was developed at Oklahoma State University by Richardson and Ray through cooperative agreements with the National Economics Division (then the

*Reuben N. Weisz and William Quinby are economists with the Natural Resource Economics Division, ESCS. Ronald R. Miller, an economist with the Commodity Economics Division, ESCS, while this research was conducted, is now with the U.S. Department of Energy.

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

[2] The sole purpose of this article is to present a methodology for technology impact policy analyses. Our main point is to show a useful method of evaluating the outcome of policy action. That alternative is a range of outcomes with the probability of occurrence of each outcome attached. The base line numbers presented are not intended, nor should they be construed, to represent a forecast of future cotton prices.

Because the dominant rationale for using pesticides may be to minimize risk, policymakers should evaluate the farm level and aggregate impacts of pesticide policies on risk.

Commodity Economics Division), ESCS. Documentation appears in several technical bulletins (*14, 15, 13*). ESCS economists have modified the model. The ESCS baseline used here is available (*18*).

Simulation Procedure

The user begins a simulation by changing one or more of the variables contained in the model's baseline. For example, a pesticide policy analysis could be conducted by changing the base values of one or more of the following variables:

- Crop variables—
  Exogenous change in yield per harvested acre;
  Exogenous change in variance of yield per harvested acre; and
  Exogenous change in variable production expense per harvested acre.
- Livestock variables—
  Exogenous change in production; and
  Exogenous change in nonfeed variable production cost.

The actual values of the variables are computed outside the model by the analyst and are inputed into POLYSIM.

The simulation procedure begins by shocking the model with the relevant changes in the cotton acreage and yield equations. We obtain simulated cotton statistics (tables 2-6) for 5 years, which we compare with the baseline statistics. The model may be viewed as an automated accounting routine that traces the initial effects on production through subsequent effects on price, use, and farm income, for each of the 11 commodity groups and for agriculture in the aggregate.

Role of Elasticities

Direct and cross-commodity supply and demand elasticities determine the magnitude and direction of endogenous variables' deviations from the baseline values. The elasticities used were derived by Ray and Richardson from many sources—subjective judgments of commodity specialists, a survey of the literature, and direct estimation based on recent data.

Although each commodity in the model has a unique set of parameters assigned to it, a large degree of similarity exists among the mathematical functions assigned to each commodity. We now describe the cotton yield equations we modified for use in our study (figs. 1 and 2).

Yield Equation

Simulated cotton yield in a given year, *t*, is calculated by adjusting the baseline yield in response to the following (fig. 1):

1. The change between simulated cotton price and baseline price in the previous year;
2. The change between the current simulated prices paid index and the index implied by the baseline;
3. The change between simulated cotton harvested acreage and baseline acreage;
4. A long-term adjustment coefficient which allows current adjustment in yield to reflect the behavioral, capital, and investment inertia of past decisions;
5. A shift, due to a change in pesticide or other policy, of the level of the production function; the exogenous change in yield in the simulated year *t*; and
6. A shift, due to a change in the variance of yield.

In the Ray and Richardson version, the analyst could select only one of the following yield options, and assume other things were equal for the other determinants of yield:

Option A: Deterministic, price-responsive yield equation (contains items 1 through 4, above);
Option B: Strictly exogenous yields (considers item "5", only); and
Option C: Strictly stochastic yields (considers item "6", only).

Option A contained no provision for incorporating a shift in the production function (item "5") or a change in the variance of yield (item "6") which may result from a change in technology. Therefore, the deterministic, price-responsive, yield equation is inadequate for pesticide regulation impact analysis.

However, the alternative of using a predetermined yield, Option B, is inappropriate, too. Whenever a prespecified yield is inserted into the model it overrides the feedback loops between price and yield. As the existence of the feedback effects is the main

Figure 1—Stochastic Cotton Yield Equation

Simulated cotton lint yield in pounds per harvested acre$_t$ = Baseline cotton lint yield in pounds per harvested acre$_t$ * [1.0 + (Elasticity of cotton yield wrt cotton price * % change in cotton price from baseline$_{t-1}$)

+ (Elasticity of cotton yield wrt change in prices paid for inputs * % change in prices paid index from the value implicit in the baseline data$_t$) − [Marginal effect of a change in cotton harvested acreage upon cotton yield/acre *

(Simulated cotton harvested acreage$_t$ − baseline cotton harvested acreage$_t$)] + ΔYLD/HA$_t$ + (1.0 − longrun adjustment factor)

* (simulated cotton yield$_{t-1}$ − ΔYLD/HA$_{t-1}$ − baseline yield$_{t-1}$)

+ (Normally distributed deviation from the baseline yield per harvested acre.)

Note: * = multiplied by.
wrt = with respect to.

reason for using POLYSIM, a strictly exogenous yield was not appropriate.

Option C, the strictly stochastic yield equation, also is inappropriate, because the dependency of yield on items "1" through "5" is ignored. These factors are all relevant for pesticide policy analyses.

In our methodology, all six factors have a simultaneous impact on yield. For example, in the first year simulated, t, the exogenous changes in the expected value and the variance of yield will result in a simulated yield different from the baseline yield. This yield differential results in the simulated year t production, and, hence, prices that differ from the baseline values. In the subsequent year, $t + 1$, the simulated yield is modified by the exogenous variables as well as the difference between the simulated and baseline values of the price variable for the preceding year. The change in simulated cotton yield in year $(t + 1)$ with respect to (abbreviated as wrt) a percentage change in cotton price in year t is computed by multiplying the percentage change between simulated and base figures of cotton price by the elasticity of cotton yield to obtain its own price. Other endogenous variables for cotton and other commodities in year $(t + 1)$ are affected through direct and cross elasticities with the price of cotton. In this manner, POLYSIM traces out the effects within and between time periods and the feedbacks among endogenous variables.

The marginal effect of a change in cotton harvested acreage upon the yield per harvested acre was obtained from Evans and Bell and incorporated into the POLYSIM yield equation (*6*). They showed a negative relationship between yield and harvested acreage because increases in cotton acreage involve bringing marginal land into production. Similarly, decreases in cotton acreage result in higher average yields because marginal land moves out of cotton production.

Acreage Equation

The cotton harvested acreage equation also is driven by initial policy shocks and subsequent price feedbacks (fig. 2). The simulated cotton acreage is a function of:

1. Prices of cotton and competing crops,
2. Production costs,
3. Expected yields, and
4. A long-term adjustment coefficient.

Net returns from cotton relative to those from competing crops influence acreage decisions. A negative relationship exists between acreage and AVOC, the sum of the average variable and opportunity costs of growing cotton (*6*). Evans and Bell calculate AVOC as a function of the prices, costs, and yields of competing crops, as well as the variable cost and yield per harvested

Figure 2—Cotton Harvested Acreage Equation

$$\text{Simulated cotton harvested acreage million acres}_t = \text{Baseline cotton harvested acreage million acres}_t * \Bigg[1.0 + \left(\text{Elasticity of cotton acreage wrt cotton price} * \text{\% change in cotton price from baseline}_{t-1}\right)$$

$$+ \left(\text{Elasticity of cotton acreage wrt corn price} * \text{\% change in corn price from baseline}_{t-1}\right) + \left(\text{Elasticity of cotton acreage wrt soybean price} * \text{\% change in soybean price from baseline}_{t-1}\right)$$

$$+ \left(\text{Elasticity of cotton acreage wrt wheat price} * \text{\% change in wheat price from baseline}_{t-1}\right) + \left(\text{Elasticity of cotton acreage wrt Prices Paid Index} * \text{\% change in Prices Paid Index from baseline}_{t-1}\right)\Bigg]$$

$$- \left[\text{marginal effect (in million acres/dollars per pound) of a change in cotton variable cost per pound upon cotton harvested acreage} * \Delta VPE/HA_t \div \text{Expected yield per acre}_t\right]$$

$$+ \left(\text{1.0-longrun adjustment factor}\right) * \left(\text{calculated cotton acreage}_{t-1} - \text{baseline cotton acreage}_{t-1}\right)$$

Where $\Delta VPE/HA_t$ = Exogenous change in cotton variable production expense per harvested acre (Dollars/acre);

$$\text{Expected yield per acre}_t = \left(\frac{YLD_{t-1} + YLD_{t-2} + YLD_{t-3}}{3}\right) + \Delta YLD/HA_t - \left(\frac{\Delta LD/HA_{t-1} + \Delta YLD/HA_{t-2} + \Delta YLD/HA_{t-3}}{3}\right);$$

YLD_t = simulated yield per harvested acre (Pounds/acre); in year *t*.

and where $\Delta YLD/HA_t$ = exogenous change in yield in year *t*.

Note: * = multiplied by.
wrt = with respect to.

acre of cotton. This relationship has been incorporated into the POLYSIM cotton acreage equation, so that the cotton acreage response curve (response to cotton price) can shift to the left as AVOC increases because of a pesticide ban.

Procedure for Making POLYSIM Stochastic

The procedure for making POLYSIM stochastic has been described by Ray and Richardson (*13*):

> A deterministic model can be made stochastic by drawing values for selected variables. The impact of the drawn values on the model endogenous variables are estimated with the simulator. By repeating the process a large number of times and recording the values of the output variables, experimental probability distributions are developed for the endogenous variables in the model.

In the original version of the model, the yield and export demand equations are bypassed when the model is run stochastically. In our version of the model, we add a normally distributed deviation from the baseline to the yield and export equations.

Several probability distributions for yields and exports are available to the analyst who wishes to use POLYSIM in the stochastic mode. We assumed that the yields and exports of each of the model's four crops are distributed as a correlated multivariate. Table 1 shows the default variance-covariance matrix for this option. This matrix, calculated by Richardson and Ray from detrended data for 1960-74 for average national values of crop yields and exports, is used to develop a stochastic baseline.

EMPIRICAL ANALYSIS

At the time of our study, USDA and Environmental Protection Agency researchers were developing partial budget and yield estimates of the impacts of this pesticide ban. Preliminary data developed for the Federal/State Assessment Team on Toxaphene (*17*) give the following average U.S. results: In response to a ban, average U.S. cotton yield per harvested acre would

Table 1—Default variance-covariance matrices for feed grain, wheat, soybean, and cotton yields and exports[1]

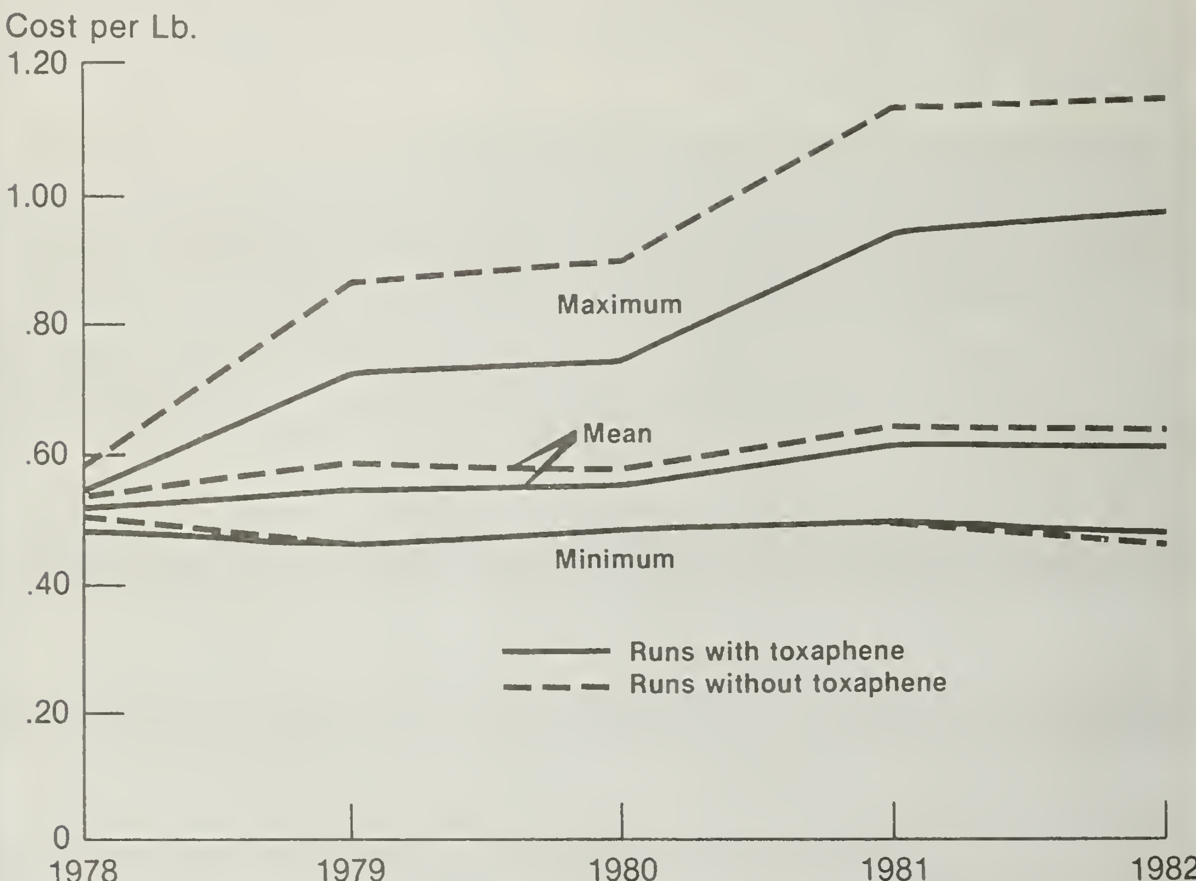
Commodity	Variance-covariance for crop yields				Variance-covariance for crop exports			
	Feed grains	Wheat	Soy-beans	Cotton	Feed grains	Wheat	Soy-beans	Cotton
	Tons/ acre	*Bushels/ acre*	*Bushels/ acre*	*Pounds/ acre*	*Million tons*	*Million bushels*	*Million bushels*	*Million net bales*
Yields:								
Feed grains	0.028	0.068	0.079	−0.304	0.588	8.760	2.890	0.051
Wheat		1.470	.498	.353	2.396	32.471	24.467	.064
Soybeans			.878	−2.275	1.876	37.421	19.512	.124
Cotton				102.567	−8.926	207.665	−79.504	4.921
Exports:								
Feed grains					23.365	529.668	82.816	2.100
Wheat						21,386.810	578.938	77.710
Soybeans							1,831.926	11.613
Cotton								.953

[1] These matrices were obtained by calculating the variances and covariances from detrended data for average national values of crop yields and exports, 1960-74 (*13*, p. 125).

A ban could raise cotton prices by 3.1 cents per pound for an average yield year and push prices up 15-20 cents per pound in a poor yield year. Thus, cotton could cost over $1 per pound at 1978 price levels, if toxaphene is banned.

Figure 3

Stochastic Cotton Prices



decrease by 16.5 pounds of lint. Standard deviation of yield would increase 2.7 pounds per acre, and variable cost per acre would go up $2.86. Because we have "massaged" preliminary data, our results are not official, yet they are reasonable estimates.

To perform this analysis, we chose a Monte Carlo simulation approach rather than a pure, point estimate method as the probability of achieving a given point is zero. Our approach presents ranges of estimates and the corresponding probabilities of their occurrence.

A Monte Carlo simulation generated the random deviations in the yield and export equations. A single 5-year simulation run would illustrate only one of the infinite possible combinations. To evaluate the relative frequency of possible outcomes, we ran two sets of simulations in each of which a 5-year (1978-82) Monte Carlo sequence of events was simulated 300 times.

The first set simulated a stochastic baseline, one with no user input changes but which includes by default the effect of the random shocks on baseline yields and exports. This set, the base, represents a stochastic view of the economy in the absence of a ban.

Before the 300 iterations of the second run, without toxaphene, were performed, the levels and variance of cotton yield and the variable expense per harvested acre were modified by the partial budget and yield results presented earlier. This is referred to as the simulated run.

Graphic Display of Results

Figure 3 summarizes the base and simulated cotton prices for each of the 5 years. The figure shows the maximum (top line), mean (middle line), and minimum (bottom line) observation recorded in each time period with (solid line) and without (dotted line) toxaphene.

The proposed policy's cost and yield shocks reduced output. The negative impact on production resulted in a wider range of cotton prices without toxaphene than with it. The annual maximum, mean, and minimum values without toxaphene are higher than the corresponding values with it. A ban could raise cotton prices by 3.1 cents per pound for an average yield year and push prices up 15-20 cents per pound in a poor yield year. Thus, cotton could cost over $1 per pound at 1978 price levels, if toxaphene is banned.

Hypothesis Testing

Numerical results for selected variables are illustrated in tables 2 through 5. In each case we test the hypothesis that no significant difference occurs in the results of the POLYSIM runs with and without toxaphene.

Comparison of Means

The reduction in average production from 11.09 to 10.86 million bales due to the pesticide ban pushes cotton prices an average of 3.1 cents per pound above the baseline (table 2). The higher level of cotton prices reduces the effect of the 16.5 pounds per acre exogenous change in yield per acre to –13.4, 81 percent of the initial value. Higher prices reduce exports from 4.5 to 4.4 million bales. The t' tests indicate that there is less than a one out of 1,000 probability that the differences

Table 2—Average 5-year mean for 1978-82

Item	Baseline	Simulated	t' statistic	Level of significance
Cotton price (dollars per pound)	0.564	0.595	9.83	0.001
Cotton yield (pounds per acre)	481.252	467.876	–27.03	.001
Cotton exports (million bales)	4.509	4.373	–3.64	.001
Cotton acreage (million acres)	11.064	11.144	4.81	.001
Cotton production (million bales)	11.091	10.861	–12.42	.001
Soybean price (dollars per bushel)	5.331	5.347	.40	N.S.

Note: N.S. means not significant.

in cotton prices, yields, exports, acreage, and production with and without toxaphene could have been obtained by chance alone.[3] However, there is no significant difference between 5-year mean soybean prices. Similar data on annual mean values appear in table 3.

Table 3—Annual means

Item	Base-line	Simu-lated	t' sta-tistic	Level of signifi-cance
Cotton price (dollars per pound):				
1978	0.511	0.535	19.71	0.001
1979	.543	.585	7.47	.001
1980	.550	.580	5.25	.001
1981	.609	.641	3.92	.001
1982	.606	.632	3.23	.005
Cotton acreage (million acres):				
1978	10.80	10.70	([1])	([1])
1979	11.164	11.173	.69	N.S.
1980	11.136	11.304	4.27	.001
1981	11.109	11.276	4.07	.001
1982	11.112	11.269	3.03	.005
Soybean price (dollars/bushel):				
1978	4.333	4.333	.02	N.S.
1979	4.704	4.717	.44	N.S.
1980	5.113	5.142	.58	N.S.
1981	6.185	6.222	.50	N.S.
1982	6.320	6.321	.01	N.S.

Note: N.S. means not significant.
[1] The stochastic variation was initiated in 1978. Thus, acreage in 1978 was nonstochastic because it was based on 1977 yields.

Comparison of Standard Deviations

Table 4 provides the 5-year standard deviations and table 5, the annual standard deviations of selected variables, with and without toxaphene. An F-test is used to test the hypothesis that there is no significant difference in the results of simulation runs with and without toxaphene. The hypothesis was not rejected for the 5-year standard deviations of cotton exports and soybean prices. Nor was it rejected for the annual values of cotton exports and soybean prices. However, F-tests for the other variables indicate that there is a significant difference between the baseline and simulated standard deviations. In these cases, there is a less than a one out of 100 probability that these results could have been obtained by chance alone.

Over the 5-year period, the increased variance on yield results in a significant (0.01 level) increase in the variability of the values of cotton price, acreage, and production. The standard deviation of price increased from 7.5 to 9.4 cents a pound. The standard deviation of harvested acreage increased from 390,000 to 510,000 acres. The standard deviation of production increased from 440,000 bales to 570,000 bales.

Table 4—Average 5-year standard deviation for 1978-82

Item	Base-line	Simu-lated	F sta-tistic	Level of signifi-cance
Cotton price (dollars per pound)	0.075	0.094	1.57	0.01
Cotton yield (pounds per bushel)	12.226	14.763	1.46	.01
Cotton exports (million bales)	1.008	1.039	1.06	N.S.
Cotton acreage (million acres)	.393	.512	1.69	.01
Cotton production (million bales)	.439	.569	1.68	.01
Soybean price (dollars/bushel)	1.094	1.110	1.03	N.S.

Note: N.S. means not significant.

[3] It cannot be assumed *a priori* that variances of the two means are identical. The t' statistic of Cochran (*3*) is used. As the number of observations for each sample is identical for our study, the Cochran approach here means simply to calculate $t' = t$ and adjust the degrees of freedom from $2(n-1)$ to $(n-1)$. Cochran's t' statistic is slightly more conservative than the solutions of (*1*, *2*, and *7*).

The input data for our stochastic POLYSIM simulation indicated a ban on toxaphene would decrease the level and increase the variability of yield and production.

Table 5—Annual standard deviation

Item	Baseline	Simulated	F statistic	Level of significance
Cotton price (dollars per pound):				
1978	0.012	0.017	1.99	0.01
1979	.058	.078	1.80	.01
1980	.059	.081	1.90	.01
1981	.085	.110	1.68	.01
1982	.083	.109	1.75	.01
Cotton acreage (million acres):				
1978[1]				
1979	.140	.163	1.35	.01
1980	.375	.490	1.71	.01
1981	.444	.554	1.56	.01
1982	.573	.694	1.47	.01
1978	.242	.242	1.00	N.S.
1979	.357	.361	1.02	N.S.
1980	.589	.602	1.05	N.S.
1981	.901	.932	1.07	N.S.
1982	1.226	1.251	1.04	N.S.

Note: N.S. means not significant.

[1] The stochastic variation was initiated in 1978. Thus, acreage in 1978 was nonstochastic because it was based on 1977 yields.

In Earlier Issues

A good textbook is the most reliable tool of a teacher. It often becomes the blueprint for a course, and in many instances the success or failure of a teacher working under the pressures of a heavy teaching load is dependent on the thoroughness of organization and presentation in the text material.

D. B. DeLoach
AER, Vol. III, No. 4,
Oct. 1951, p. 135

Future research should evaluate the impact on net farm income of changes in the variability of yields, costs and prices received.

Comparison of Variances and Frequency Distributions

The major stochastic shock in this analysis is the change in the variance of cotton yield that is induced by the proposed pesticide regulation. The variances of yield, then, deserve closer scrutiny. As table 6 indicates, the variance without toxaphene is over twice the value of the historical variance and 146 percent of the value of the stochastic baseline.

Table 7 illustrates the frequencies of alternative cotton prices. These indicate that a toxaphene ban would shift the expected frequencies of cotton prices from lower to higher values and expand the range of likely values from 51.1 cents to 69.1 cents. The mean and standard deviation both increase. The chi-square statistic, 276, measures the difference between baseline and simulated frequencies. This indicates that the probability that the price frequencies without toxaphene do not differ from those with it is close to zero.

Table 6—Variances on cotton yield per harvested acre

Source	Variance		
	Pounds per acre squared	*Percent of detrended values*	*Percent of stochastic baseline*
From detrended data for average national values, 1960-74	[1] 102.567	100	
From stochastic baseline, 1978-82	[1] 149.475	146	100
From stochastic without toxaphene run, 1978-82	217.946	212	146

[1] As described in the text, we used the POLYSIM default option for the probabilistic assumptions when constructing the stochastic baseline. An analysis of historical crop yield data indicates the presence of heteroskedasticity. Future 1978-82 variance of yield will likely differ from that observed in 1960-74 but why the actual difference occurred in our study remains unclear.

Table 7—Frequencies of cotton prices for 1978-82

Price interval (cents per pound)	Frequency of occurrence	
	Baseline	Simulated
45-50	[1] 240	[1] 102
50-55	559	506
55-60	313	336
60-65	205	246
65-70	106	140
70-75	39	62
75-80	19	50
80-85	9	20
85-90	3	17
>90	[2] 7	[3] 21

[1] The minimum baseline and simulated value recorded was 46.0 cents per pound. [2] The maximum baseline value recorded was 97.1 cents per pound. [3] The maximum simulated value recorded was 115.1 cents per pound.

Note: Chi-square equals 276. Other nonparametric statistical tests could have been applied to the empirical results of this Monte Carlo experiment. For example, the Kolmogorov-Smirnov test also rejects the common distribution hypothesis. Our examples only suggest the types of analysis that could be performed with such data.

CONCLUSIONS

The input data for our stochastic POLYSIM simulation indicated a ban on toxaphene would decrease the level and increase the variability of yield and production. The POLYSIM analysis indicates that a ban would also increase the average price of cotton as well as its price instability. It would decrease the expected value but not the variability of exports. Data on crops other than cotton would not be affected significantly.

The farm income part of POLYSIM is weak so we did not examine this component in detail. At the farm level, a pesticide ban will affect the variability of costs and yields. Future research should evaluate the impact on net farm income of changes in the variability of yields, costs, and prices received.

The stochastic method used in our policy analysis allows examination of a range of possible outcomes and assigns probabilities to alternative outcomes. In past pesticide policy analyses with the deterministic version of POLYSIM, we have used commercial econometric models to evaluate the consumer price implications of POLYSIM's results. A commercial econometric model could be used in conjunction with stochastic POLYSIM in future studies.

REFERENCES

(*1*) Aspin, A. A. "Tables for Use in Comparisons Whose Accuracy Involves Two Variances, Separately Estimated." *Biometrika.* 43:203, 1956.

(*2*) Behrens, W. V. *Landwirtschaftliche Jahrbücher.* 68:807, 1929.

(*3*) Cochran, W. G. "Approximate Significance Levels of the Behrens-Fisher Test." *Biometrics*, 20:191,1964.

(*4*) Davison, Cecil W., and Milton H. Ericksen. "Alternative Economic Settings for Agriculture." In *Agricultural-Food Policy Review.* Econ. Res. Serv., U.S. Dept. Agr. ERS-AFPR-1, Jan. 1977, pp. 12-29.

(*5*) Evans, Sam. "Regional U.S. Cotton Acreage Response." In *Cotton and Wool Situation.* Econ. Res. Serv., U.S. Dept. Agr. CWS-11, July 1977, pp. 20-24.

(*6*) Evans, Sam, and Tom Bell. "How Cotton Acreage, Yield, and Production Respond to Price Changes." In *Agr. Econ. Res.* Vol. 30, No. 2, Apr. 1978.

(*7*) Fisher, R. A., and F. Yates. *Statistical Tables.* 5th ed. Oliver and Boyd, Edinburgh, 1957, tables VI, VI_1, and VI_2.

(*8*) Moriak, Theo F. "Implications of Uncertainty in Yields and Exports." U.S. Dept. Agr., unpubl. Aug. 1977.

(*9*) Moscardi, Edgardo and Alain de Janvry. "Attitudes Toward Risk Among Peasants: An Economic Approach." In *Am. J. Agr. Econ.* Nov. 1977, pp. 710-716.

(*10*) Ray, Daryll E., and Earl O. Heady. *Simulated Effects of Alternative Policy and Economic Environments on U.S. Agriculture.* Ctr. Agr. and Rural Devlpt., Iowa State Univ. CARD rpt. 46T. Mar. 1974.

(*11*) Ray, Daryll E., and Theo F. Moriak. "POLYSIM: A National Agricultural Policy Simulator." In *Agr. Econ. Res.* Vol. 28, No. 1, Jan. 1976. pp. 14-21.

(*12*) Ray, Daryll E. "U.S. National Agricultural Projection and Policy Models." Paper presented at U.S.-U.S.S.R. Agr. Business Seminar, Nov. 16-17, 1977, Harvard University. Professional paper P-433, Okla. Agr. Expt. Sta., 1977.

(*13*) Ray, Daryll E., and James W. Richardson. *Detailed Description of POLYSIM.* Okla. State Univ., and U.S. Dept. Agr., Tech. Bull., T-151, Stillwater, Okla., Dec. 1978.

(*14*) Richardson, James W., and Daryll E. Ray. *User's Manual for UPDATE—An Auxiliary Program for POLYSIM.* Agr. Expt. Sta., Okla. State Univ., Res. Rpt. P-726. Nov. 1975.

(*15*) Richardson, James W., and Daryll E. Ray. *User's Manual for the National Agricultural Policy Simulator (POLYSIM). Agr. Expt. Sta., Okla. State Univ., Res. Rpt. P-727. Nov. 1975.*

(*16*) Trickett, W. H., B. L. Welch, and G. S. James. "Further critical values for the two-means problem." *Biometrika* 36:290, 1949.

(*17*) U.S. Department of Agriculture. "Biological Estimates and Partial Budgets Related to the Use of Toxaphene and Its Alternatives." Draft of supplement to report, "Assessment of Toxaphene in Agriculture," Sept. 9, 1977. Prepared by USDA State Assessment Team on Toxaphene in cooperation with Environmental Protection Agency. Coordinated by Office of Environmental Quality Activities, U.S. Dept. Agr. 1977.

(*18*) ———— . Transcript of remarks by Secretary of Agriculture, Bob Bergland, at meeting with USDA staff sponsored by Org. of Professional Employees of Dept. Agr., Wash., D.C., Feb. 8, 1977.

(*19*) ———— . "November 1977 Baseline: Intermediate Outlook Projections for Research Purposes." Econ., Statis., Coop. Serv., Commod. Econ. Div. working paper, Feb. 1978.

Deflating Statistical Series: An Example Using Aggregate U.S. Demand for Textile End-Use Categories

By Thomas M. Bell, Joseph M. Roop, and Cleve E. Willis*

Analysts frequently adjust price, income, or other data to eliminate the influence of inflation or differences in size. The authors of this article examine economic and statistical reasons for deflating time-series and cross-sectional data prior to estimating demand relations. Signs and magnitudes of regression coefficients change when aggregate demand equations for textiles are estimated from time-series data. Questions of heteroskedasticity, multicollinearity, and homogeneity are addressed. The demand equations are disaggregated by end-use category—apparel, household, and industrial demand.

Keywords:

Deflation
Demand analysis
Econometrics
Textile demand

Analysts often deflate data on prices, income, and other variables to eliminate the effects of inflation or household size in demand analyses. In time-series analyses, for example, they frequently deflate consumption by population, and investment by volume of sales. In cross-sectional studies, household income is often deflated by size of household and sales by size of firm.

Our purpose here is to present some reasons for deflating statistical series and to demonstrate the results —namely, that signs and magnitudes of regression coefficients change—when we use aggregate demand equations and time-series data for textiles.

ECONOMIC REASONS FOR DEFLATING

Variables that shift demand functions must be used if we are to isolate price-quantity relationships (*19*).[1] To measure consumer demand from time-series data, Foote divides shift variables into four classes: (1) consumer income or other measures of the general level of demand on a national basis; (2) the general price level; (3) supplies or prices of competing products; and (4) population (*6*, p. 27).

Let us focus on Foote's second category, the general price level. Assume that demand is homogeneous of degree zero for all prices and income, as economic theory suggests. We impose this assumption by deflating each price and income variable by the general level of prices.[2] We express demand for commodity y as:

$$y = b_0 + b_1 X_1 + b_2 X_2 + u \qquad (1)$$

where

X_1 = own price
X_2 = consumer income
b_i = unknown parameters, and
u = error term

The use of real income and relative prices is the "Marshallian" method; alternatively, we could use normalized prices (*9*) or a "mixed" demand curve specification (*15*). Utility theory requires that competing- and complementary-good prices be included. We ignore them here to simplify the presentation. However, this argument precedes the functional form.

From economic theory, price of y relative to other commodity prices influences consumption of y; thus X_1 should be the relative price of y. This is the verbal statement of homogeneity in the multigood world.

Operationally, we obtain a measure of the relative price of y by deflating its absolute price by an index of other prices. The relevant price becomes $^*X_1 = X_1/K$, and the relative income measure is $^*X_2 = X_2/K$. The original variables are expressed in nominal terms and K is an index of the general price level, such as the consumer price index (CPI).

Deflating by an index that contains the price of the dependent variable makes the resultant regression coefficients subject to bias. Bias is also introduced when the index is included as a separate variable.

*Thomas M. Bell is senior commodity analyst with Merrill Lynch, and Joseph M. Roop is senior economist with Evans Economics, Inc. Both Bell and Roop were with ESCS when this article was prepared. Cleve E. Willis is a professor in the Department of Food and Resource Economics, University of Massachusetts.

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

[2] We assume that (a) the good is relatively unimportant in the consumer's budget or that (b) the price movements of substitutes correspond approximately with the general price level.

Analysts frequently adjust price, income, or other data to eliminate the influence of inflation or differences in size.

This has led some investigators to construct special index numbers which eliminate the price(s) of the good(s) included in the analysis and to use these to deflate the price variables included in the study. Obviously, the unadjusted measure of the general price level should generally be used to deflate the income variables.

Brennan suggests that it is often better to deflate by some other index than that of all prices (*2*, p. 379). In the demand analysis for an agricultural commodity, Brennan deflates the price in question by an index of agricultural prices only. Further, if only one other commodity is a strong substitute for the good in question, it may be desirable to deflate the own price by the price of that substitute. No unambiguous rule for the choice of the appropriate index can be given. This choice is determined by the investigator's judgment and knowledge of the behavior of the subject being studied and the economic theory involved.

STATISTICAL OR ECONOMETRIC REASONS FOR DEFLATING

Some econometric considerations affect the choice of a deflator. Karl Pearson's early work on ratios having a common denominator showed that correlations between ratios can reflect spuriously high estimates of the relationship between the numerators (*13*). Kuh and Meyer showed that correlation among deflated series may also be spuriously low (*12*). The question of spurious correlation does not arise, of course, if the maintained hypothesis is in ratios. Kuh and Meyer further demonstrated two necessary and sufficient conditions for the correlation of ratios to yield correct estimates of the undeflated partial correlations. These are that (1) the coefficient of variation (the ratio of the standard deviation to the mean) of the deflating variable is small, and (2) the variables deflated are linear homogeneous functions of the deflator (*12*, p. 405). The degree of bias depends on the relative size of $r(X,Y)$ compared to $r(X,Z)$ and $r(Y,Z)$, where Z is the deflator and $r(\;)$ is the correlation operator. Hence, when cross-sectional data are deflated because of size, economic relationships will probably approximate the homogeneity requirement, so the ratio estimates generally should not be seriously biased.

Another focus of attention is the spherical attributes of the residuals. If the usual homoskedasticity is assumed for the undeflated series, deflating leads to heteroskedasticity because deflation of the included variables transforms the error term. The assumption of homoskedasticity is seldom appropriate for undeflated cross-sectional data.

Small observations are typically associated with small variances and large observations with large variances. David and Neyman demonstrated that least squares produce efficient unbiased estimates only if the residual sum of squares to be minimized is appropriately weighted (*4*). That is, we assume the usual Markov assumptions are met and the variance of the conditional distribution of the dependent variable is a weighted average of the unknown population variance (with weights w). Then, the most efficient unbiased estimate of the regression parameters (b_i) are produced by minimizing $(w(y - X_b))^t(w(y - X_b))$. The matrix w is diagonal with elements $w_i{-}½$; in the homoskedastic case, w is an identity matrix. This derivation of b is an Aitken generalized least squares estimator (*10*, p. 214).

Suppose the simple deflation is such that $w_i = D_i^2$ (D_i the deflator). Deflating yields efficient, unbiased estimators when the undeflated residuals are heteroskedastic. At worst, deflation will usually be superior to assuming falsely that a constant unitary weight is appropriate (see *12*, p. 407; *3*; *11*). A further advantage is that extreme observations will have less effect on the estimation.

Deflating statistical series to achieve a favorable specification of an econometric model may be appropriate for both theoretical and empirical reasons. Deflating an otherwise spherical relationship may induce heteroskedasticity although cross-sectional data frequently need to be deflated. Multicollinearity is also affected—most of the consequences are wellknown (*10*, p. 160; *14*, pp. 46-52; and *18*, pp. 127-128). The actual size and comparisons of $r_i(\;)$ determine the magnitude of the bias of partial correlation introduced when deflation is used.

AN EXAMPLE [3]

We now apply the deflating method by considering demand for three categories of textiles by end uses (apparel, household, and industrial). Each end use includes four fiber types (noncellulosic, cellulosic, cotton, and wool), with estimates of fiber content of purchases expressed in cotton equivalent pounds. Main explanatory variables of these categories are nominal disposable income, the end-use price index, an implicit deflator for all goods except the end use in question, and population. (Fiber types and classifications appear in appendix table 1; data are in appendix table 2. Problems of quality changes and aggregation are ignored.)

In functional form, the demand for each end-use category is expressed as:

$$w_i - \tfrac{1}{2}$$

$$Q_i = f(\text{CPI}_i, \text{PD}_i, \text{INCN}, \text{POP});$$

$$i = \text{A, H, I; end-use category}) \qquad (2)$$

where:

[3] This example does not illustrate the heteroskedasticity arguments. For a summary of tests for homoskedasticity, see (*10*, pp. 214-221).

Table 1—Correlation matrix

Variable	QA	QH	QI	CPIA	CPIH	WPII	PDA	PDH	PDI	INCN
QA	1.00	0.97	0.60	0.83	0.73	0.67	0.81	0.82	0.84	0.87
QH	.97	1.00	.50	.92	.83	.79	.89	.89	.91	.93
QI	.60	.50	1.00	.24	.12	.10	.21	.22	.25	.29
CPIA	.83	.92	.24	1.00	.97	.94	.99	.99	.99	.99
CPIH	.73	.83	.12	.97	1.00	.98	.99	.99	.98	.97
WPII	.67	.79	.10	.94	.98	1.00	.96	.96	.95	.94
PDA	.81	.89	.21	.99	.99	.96	1.00	1.00	1.00	.99
PDH	.82	.89	.22	.99	.99	.96	1.00	1.00	1.00	.99
PDI	.84	.91	.25	.99	.98	.95	1.00	1.00	1.00	1.00
INCN	.87	.93	.29	.99	.97	.94	.99	.99	1.00	1.00
POP	.95	.94	.48	.88	.80	.74	.88	.88	.90	.91
QAPC	.99	.94	.66	.76	.64	.59	.74	.74	.76	.80
QIPC	−.41	−.48	.44	−.64	−.67	−.62	−.68	−.67	−.66	−.64
QHPC	.97	1.00	.52	.91	.81	.77	.87	.88	.89	.92
INCNPC	.86	.93	.28	.99	.97	.95	.99	1.00	1.00	1.00
INCRA	.96	.99	.46	.94	.87	.82	.92	.93	.94	.96
INCRH	.96	.99	.46	.94	.87	.82	.93	.93	.94	.96
INCRI	.94	.98	.41	.96	.90	.86	.95	.95	.95	.97
CPIAR	−.76	−.75	−.23	−.80	−.83	−.80	−.88	−.88	−.88	−.87
CPIHR	−.94	−.94	−.45	−.88	−.80	−.75	−.89	−.89	−.91	−.91
WPIIR	−.92	−.88	−.47	−.79	−.71	−.61	−.80	−.80	−.83	−.83
CV	.21	.46	.07	.17	.15	.14	.23	.23	.22	.48

* Variable definitions in addition to those in table 2 are:

QAPC=QA/POP
QIPC=QI/POP
QHPC=QH/POP

INCNPC=INCN/POP
INCRA=INCNPC/PDA
INCRH=INCNPC/PDH
INCRI=INCNPC/PDI

CPIAR=CPIA/PDA
CPIHR=CPIH/PDH
WPIIR=WPII/PDI

Q_i = quantity consumed of the *i*th end-use category of total fibers (millions of pounds);

CPI_i = consumer price index for the *i*th end-use category (1967=100); WPI used for industrial category;

PD_i = implicit deflator for all except the *i*th good in question;

INCN = nominal disposable income (billion dollars);

POP = U.S. population (millions).

The population impact can be removed by deflating the quantity and income variables by population, and the real income and relative price impacts can be arrived at by deflation of own price and per capita income by the appropriate deflators.[4] We consider first the impact of converting to per capita measures. The simple correlation coefficients appear in table 1.

We first compare simple correlation coefficients between raw and deflated series, and then compare simple correlation coefficients of deflated series with partial correlation coefficients of raw series to determine the extent and magnitude of the bias among these measures.[5]

[4] Approximate because the implicit deflator does not contain all prices.

[5] The relationship between partial correlation coefficients and regression coefficients is discussed in (*10*, pp. 61-65, 132-135; and *18*, pp. 131-138). See also (*5*, pp. 192-197).

and coefficients of variability *

POP	QAPC	QIPC	QHPC	INCNPC	INCRA	INCRH	INCRI	CPIAR	CPIHR	WPIIR
0.95	0.99	−0.41	0.97	0.86	0.96	0.96	0.94	−0.76	−0.94	−0.92
.94	.94	−.48	1.00	.93	.99	.99	.98	−.75	−.94	−.88
.48	.66	.44	.52	.28	.46	.46	.41	−.23	−.45	−.47
.88	.76	−.64	.91	.99	.94	.94	.96	−.80	−.88	−.79
.80	.64	−.67	.81	.97	.87	.87	.90	−.83	−.80	−.71
.74	.59	−.62	.77	.95	.82	.82	.86	−.80	−.75	−.61
.88	.74	−.68	.87	.99	.92	.93	.95	−.88	−.89	−.80
.88	.74	−.67	.88	1.00	.93	.93	.95	−.88	−.89	−.80
.90	.76	−.66	.89	1.00	.94	.94	.95	−.88	−.91	−.83
.91	.80	−.64	.92	1.00	.96	.96	.97	−.87	−.91	−.83
1.00	.90	−.58	.93	.91	.96	.96	.95	−.87	−.99	−.97
.90	1.00	−.30	.95	.79	.92	.92	.89	−.68	−.89	−.87
−.58	−.30	1.00	−.45	−.64	−.54	−.54	−.57	.67	.59	.56
.93	.95	−.45	1.00	.92	.98	.98	.97	−.72	−.93	−.87
.91	.79	−.64	.92	1.00	.96	.96	.97	−.86	−.91	−.82
.96	.92	−.54	.98	.96	1.00	1.00	1.00	−.80	−.95	−.90
.96	.92	−.54	.98	.96	1.00	1.00	1.00	−.81	−.95	−.90
.95	.89	−.57	.97	.97	1.00	1.00	1.00	−.81	−.94	−.87
−.87	−.68	.67	−.72	−.86	−.80	−.81	−.81	1.00	.88	.82
−.99	−.89	.59	−.93	−.91	−.95	−.95	−.94	.88	1.00	.96
−.97	−.87	.56	−.87	−.82	−.90	−.90	−.87	.82	.96	1.00
.08	.14	.08	.40	.40						

The results show that considerable care is needed in appraising the relative price impact in a statistical or economic sense . . .

Briefly stated, the partial correlation coefficient r(X,Y.Z) may be defined as the measure of the linear relationship between X and Y after the linear relationship of Z on both X and Y is considered. This partial coefficient squared is the product of two regression coefficients:

> $r(X,Y.Z) = \sqrt{b_1 \,.\, b_2}$, choosing the sign of either coefficient where:
> b_1, b_2 are derived from the regression equations $X = a_1 + b_1 Y + c_1 Z$ and $Y = a_2 + b_2 X + c_2 Z$.

Consequently, the sign and relative importance of partial correlation coefficients resemble standardized regression coefficients (that is, coefficients adjusted for their standard errors). However, the two sets differ mathematically. Partial correlation measures interdependence between two variables, whereas regression involves the notion of dependent and independent variables.

Zero order results are important in increasing or decreasing multicollinearity. Thus, we examine the question of spurious correlation by looking at the correlation coefficients between undeflated series (such as Q_i and INCN, hereafter $r(Q_i,\text{INCN})$) relative to the correlations between deflated series—$r(Q_i\text{PC},\text{INCNPC})$. For apparel, $r(Q_A\text{PC},\text{INCNPC}) = 0.793$, while $r(Q_A,\text{INCN}) = 0.87$, $r(Q_A,\text{POP}) = 0.95$, and $r(\text{INCN},\text{POP}) = 0.91$. Coefficients of variability (CV's) are $CV(Q_A) = 0.21$, CV(INCN) = 0.48 and CV(POP) = 0.08. Relaxing somewhat the rigid assumptions in (*8*) (such as $CV(\text{POP})/CV(Q_A) = CV(\text{POP})/CV(\text{INCN})$ and $r(Q_A,\text{POP}) = r(\text{INCN},\text{POP})$), the correlation between the ratios can vary, depending on:

$$CV(\text{POP})/CV(Q_A) \gtrless 2\,r(Q_A,\text{POP}) \ldots 0.08/.21 < 2(0.95)$$

$$CV(\text{POP})/CV(\text{INCN}) \gtrless 2\,r(\text{INCN},\text{POP}) \ldots 0.08/0.48 < 2(0.91)$$

In both cases, we can infer decreased correlation between the ratios relative to the raw series. Results are similar for household use, but inconclusive for industrial use. The above results affirmed our inferences for apparel, as the correlation of the ratios is less than that of the undeflated series.

Proceeding to higher orders, we examine the question of partial versus ratio correlation by a test that the deflated series and the deflator are linear and homogenous, which is equivalent to testing the null hypothesis that the intercept of one on the other is zero. Simple regressions produce the following results (t-statistics below coefficients):

$$Q_A = -5334.73 + 45.06\ \text{POP} \qquad (3a)$$
$$(-8.77)\ (14.38)$$

$$Q_H = -14069.4 + 89.02\ \text{POP} \qquad (3b)$$
$$(-9.75)\ \ (12.00)$$

$$Q_I = 133.68 + 5.50\ \text{POP} \qquad (3c)$$
$$(3.02)\ (2.41)$$

$$\text{INCN} = -2564.36 + 16.21\ \text{POP} \qquad (3d)$$
$$(-8.17)\ (10.31)$$

The null hypothesis that the intercept is zero is rejected in all cases.

The coefficient of variation of population is relatively small (0.08). The degree of bias of $r(\text{INCN},Q_i.\text{POP})$ is given by comparing $r(\text{INCN},Q_i)$ with $r(\text{INCN},\text{POP})$ and $r(Q_i,\text{POP})$. Note that the simple correlation between income and population is 0.91 while $r(\text{INCN},\text{POP})$ and $r(Q_i,\text{POP})$ are high in all cases except the industrial use category.

Following Kuh and Meyer (*12*), we conclude, based on the degree of homogeneity between the deflator and the variable to be deflated, that deflating by population is likely inappropriate. Yet, the coefficient of variation for population is low (0.08). This evidence is perplexing. Judgment dictates that we deflate by population and examine the real income and relative price effects.

The simple ratio correlations $r(\text{INCNPC}/\text{PD}_i, \text{CPI}_i/\text{PD}_i)$ are −0.80, −0.95, and −0.87 for apparel, household, and industrial use respectively. The correlation between these ratios is not only less than that between the raw series $r(\text{INCNPC},\text{CPI}_i)$ (0.99, 0.90, and 0.94, respectively), but now differs in sign.

When we examine ratio versus partial correlation, the regression tests for linear homogeneity result in:

$$CPI_A = \underset{(9.79)}{28.72} + \underset{(26.9)}{89.81}\, PD_A \qquad (5a)$$

$$CPI_H = \underset{(14.5)}{38.19} + \underset{+\ (26.2)}{78.62}\, PD_H \qquad (5b)$$

$$CPI_I = \underset{(8.82)}{43.78} + \underset{(13.2)}{75.85}\, PD_I \qquad (5c)$$

$$INCNPC = \underset{(-17.3)}{-2.04} + \underset{(42.6)}{5.73}\, PD_A \qquad (6a)$$

$$INCNPC = \underset{(-18.4)}{-1.97} + \underset{(46.6)}{5.68}\, PD_H \qquad (6b)$$

$$INCNPC = \underset{(-21.5)}{-2.26} + \underset{(50.0)}{6.09}\, PD_I \qquad (6c)$$

leading to the rejection of the null hypothesis of linear homogeneity. Additionally the coefficients of variation are large. We would thus expect comparatively large bias in $r(CPI_i, INCNPC.PD_i)$ to be introduced when deflating by the implicit deflators. The correlation coefficient $r(CPI_i,INCNPC.PD_i)$ is large relative to the simple correlations $[r(CPI_i,PD_i)$ and $r(INCN,PD_i)]$ of the undeflated series; for example, the apparel results give 0.995 relative to 0.987 and 0.991.

We illustrate further, in quantity dependent form, four equations each for apparel and household use and three for industrial textile use in table 2. The first three equations are deflated by population and by the appropriate implicit price deflator. The double log demand function is used for reasons discussed by Sato (*14*).[6] Additionally, a price dependent demand specification is used for apparel.

The results show that considerable care is needed in appraising the relative price impact in a statistical or economic sense, as evidenced by the positive sign and low t-ratios when a homogeneous degree zero specification obtained by deflation is used (see equations A1, H1, I1 in table 2).

A typical solution to this problem has been to estimate the demand function in price dependent form (equation A6, table 2). However, the earlier analysis revealed $r(INCN,CPI_A.PD_A) = 0.6$, compared with $r(INCN/PD_A,CPI_A/PD_A) = -0.8$. Thus, we expect a negative regression coefficient for $INCN/PC_A$. (Actually, Q_APC should be included and higher order partial correlations computed, but these were ignored to simplify the argument. The regression analysis gives expected results and the low Durbin-Watson statistic indicates misspecification. Equation A7 with correction for autocorrelation (using Cochrane-Orcutt procedure) remains suspect (see *8* for a discussion of this problem). Equation A5, using the deflator as a separate variable, gives $R^2 = 0.999$. However, the multicollinearity makes confidence in the magnitude of the coefficients impossible.

Consequently, the second group of equations uses the price deflator (PD_i) as a separate variable. It is difficult to arrive at meaningful elasticity calculations because this functional form, although mathematically identical, is not zero homogeneous. For example, by assuming that the estimate of real income elasticity (1.56) is unbiased in equation A2, we calculate the relative price impact adding in the impact of the price deflator on the CPI (0.18 = 1.74 − 1.56) to the own-price coefficient (−0.57 + 0.18 = −0.39). The r matrix for the data in this group of equations exhibits very high multicollinearity, however, and the resulting problems mentioned by Johnston (*10*) and Rao and Miller (*14*) cast some doubt on this set of equations.

As an alternative, we estimated a third set of equations using restricted least squares. This restriction simply requires finding the vector B which minimizes $u^tu = (y - XB)^t(y - XB)$, subject to the restriction that $g = RB$.[7] The restricted coefficients, b^*, are calculated as:

$$b^* = (X^tX)^{-1}[X^ty + R^t[R(X^tX)^{-1}R^t]^{-1}(g - RB)],$$

[6] Sato points out how, in the double log form, the parameters estimated reflect the Slutsky-Hicks relation used in determining income and price elasticities.

[7] $R = [0,-1,1,1]$, $B = [b_0, b_1, b_2, b_3]$, $g = [0]$.

Our example demonstrates that both the magnitude and signs of regression coefficients may change because of deflating

Table 2—Regressions for retail demand for textile fibers, 1955-76

Equation	Dependent variable	Independent variable				Equation statistics			
		Ln c	*Ln* (CPI/PD)	*Ln* ((INCN/POP)/PD)		R^2	DW	SEE	
A1	*Ln*(Q_A/POP)	−0.37 [1] (−0.22)	0.44 (1.35)	0.95 (7.88)		0.87	1.05	0.05	
H1	*Ln*(Q_H/POP)	−5.25 (−1.44)	0.92 (1.39)	2.95 (8.18)		0.97	1.15	0.07	
I1	*Ln*(Q_I/POP)	1.57 (0.85)	0.23 (0.68)	−0.14 (−0.67)		0.33	1.73	0.07	
		Ln c	*Ln* CPI	*Ln* (INCN/ POP)	*Ln* PD				
A2	*Ln*(QA/POP)	3.65 (2.60)	−0.57 (−1.89)	1.56 (10.39)	−1.74 (−5.83)	0.95	1.82	0.037	
H2	*Ln*(Q_H/POP)	−3.07 (−1.00)	0.35 (0.60)	3.39 (10.31)	−4.31 (−5.06)	0.98	1.51	0.057	
I2	*Ln*(Q_I/POP)	1.27 (0.80)	0.10 (0.36)	0.53 (1.76)	−1.25 (−2.01)	0.52	2.36	0.06	
A3	*Ln*(C_A/POP)	2.55 (18.17)	−0.34 (−5.19)	1.56 (10.36)	−1.90 (−8.89)				
H3	*Ln*(Q_H/POP)	0.43 (2.07)	−0.30 (−3.38)	3.09 (15.42)	−3.39 (−11.88)				
I3	*Ln*(Q_I/POP)	3.46 (12.80)	−0.27 (−2.48)	0.24 (1.10)	−0.51 (−1.59)				
		C	CPI	INCN/POP	PD				
A4	Q_A/POP	45.58 (9.15)	−0.19 (−2.79)	59.60 (7.53)	14.99 (−8.76	0.92	1.51	0.76	
H4	Q_H/POP	45.87 (13.63)	−0.27 (−3.35)	21.76 (10.91)	−75.06 (−4.70)	0.99	1.74	0.82	
		*Ln*C	*Ln*(Q_A/POP)	*Ln*(INCN/POP)	*Ln*PD				ρ
A5	*Ln*(CPI_A)	4.625 (35.30)	−0.10 (−1.74)	0.35 (2.50)	0.25 (1.24)	0.99	1.08	0.01	0.78
		C	*Ln*(Q_A/POP)	*Ln*(INCN/ POP/PD_A)					
A6	*Ln*(CPI_A/PD_A)	4.80 (16.87)	0.20 (1.33)	−0.46 (−3.52)		0.67	0.33	0.04	
A7	*Ln*(CPI_A/PD_A)	4.37 (15.35)	−0.01 (−0.25)	0.16 (0.71)		0.93	0.50	0.02	0.96

[1] *t*-statistic.

which differs from the unrestricted estimator by a linear function of g – RB. (For a summary of this restricted estimation procedure, see *17* and 7.) One restriction was imposed on each equation so that, after ignoring the intercept, the absolute value of the price and income coefficients would equal the absolute value of the implicit deflator coefficient.[8] The resulting price elasticities were –0.34, –0.30, and –0.27 for apparel, household, and industrial use, respectively. Corresponding income elasticities were 1.56, 3.09, and 0.24. The real income and relative price impacts appear to be more reasonable and the t-statistics are larger. Although there is no guarantee, it appears likely this restriction will assure the "correct" signs, because of the dominance of the income and price deflator coefficients relative to the price coefficient. Two categories of textile demand, household and industrial use, exhibited positive signs on price when equation set 2 was used. Note, however, that the restrictions reversed the signs. It might be necessary to use inequality-restricted least squares to achieve the desired results.

If a homogeneous degree-zero demand function is desired, with the functional form exhibited by equation set 2, then R = (0 1 1 1), and equation set 1 results;

$$ln(\text{QAPC}) = -0.37 + 0.44\ ln(\text{CPIA})$$

$$+ 0.995\ ln(\text{INCNPC})$$

$$- 1.39\ ln(\text{PDA}) \qquad (7)$$

[8] This restriction does not imply homogeneity.

which is identical to equation A1 except for the t-statistics which are adjusted for degrees of freedom.

CONCLUSIONS

We developed our argument for partial correlation coefficients, but it can be used for regression coefficients. Deflating, for whatever reason, may have substantial impacts whether one deflates to maintain fidelity with the hypothesis formulated, as a preference for a particular functional form, to remove heteroskedasticity, or to improve what otherwise might be a severe multicollinearity problem.

Our example demonstrates that both the magnitude and signs of regression coefficients may change because of deflating. We simply call attention to these consequences as a reminder to those working with numbers. The crude restrictions used to obtain "reasonable" estimates of the parameters suggest it may be appropriate to use some form of restricted estimation in conjunction with deflated series, if nothing more than as a check on the results.

REFERENCES

(*1*) Bell, Thomas M., K. Kobe, and S. Evans. "The Supply and Demand for Textile Fibers with Emphasis on Cotton." Forthcoming in Proceedings, Beltwide Cotton Conference, National Cotton Council, Memphis, Tenn., 1979.

(*2*) Brennan, M. J. *Preface to Econometrics*. Southwestern Publ. Co., Cincinnati, 1965.

(*3*) Brown, J., and others. "A Study of Index Correlations." *J. Royal Stat. Soc.*, 77:317-46, 1914.

(*4*) David, F. N, and J. Neyman. "Extensions of the Markoff Theorem of Least Squares." *Stat. Res. Memoirs*, 2:105-16, 1938.

(*5*) Ezekiel, M., and K. A. Fox. *Methods of Correlation and Regression Analysis.* John Wiley & Sons, Inc., New York, 1966.

(*6*) Foote, R. F. "Analytical Tools for Studying Demand and Price Structures." U.S. Dept. Agr., Agr. Hdbk. 146, Washington, D.C., 1958.

(*7*) Goldberger, A. S. *Econometric Theory*. John Wiley & Sons, Inc., New York, 1964.

(*8*) Granger, C. W. J., and P. Newbold. "Spurious Regressions in Econometrics." *J. Econometrics*, 2:111-120, 1974.

(*9*) Houthakker, H. S. "Additive Preferences." *Econometrica*, 28:244-257, 1960.

(*10*) Johnston, J. *Econometric Method*. McGraw-Hill, 2nd ed., New York, 1972.

(*11*) Katona, G., and others. *Contributions of Survey Methods to*

Econometrics. Columbia University Press, New York, 1954.
(*12*) Kuh, E., and J. Meyer. "Correlation and Regression Estimates When the Data Are Ratios." *Econometrica*, 23:400-416, 1955.
(*13*) Pearson, K. "On a Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs." *Proc. Royal Soc. London*, 60, 1897.
(*14*) Rao, P., and R. L. Miller. *Applied Econometrics.* Wadsworth Publishing Co., Inc., Belmont, Calif., 1971.
(*15*) Samuelson, P. "Using Full Quality to Show That Simultaneous Additive Direct and Indirect Utilities Implies Unitary Price Elasticity of Demand." *Econometrica*, 33: 781-796, 1965.
(*16*) Sato, K. "Additive Utility Functions with Double-Log Consumer Demand Functions." *J. Pol. Econ.* 80:102-124, Jan.-Feb. 1972.
(*17*) Theil, H. "On the Use of Incomplete Prior Information in Regression Analysis." *J. Am. Stat. Assoc.* 38:404-414, 1963.
(*18*) Walters, A. A. *An Introduction to Econometrics.* Macmillan, 2nd ed., London, 1970.
(*19*) Working, E. J. "What Do Statistical Demand Curves Show?" *Qtr. J. Econ.* 41:212-235, Feb. 1927.

Appendix table 1—End use categories and fiber classifications

Apparel	Household	Industrial	Fiber classifications
Blouses	Bedspreads	Abrasive products	Cotton
Coats	Blankets	Artist canvas	Wool
Diapers	Curtains	Automotive upholstery	Cellulosic
Dresses	Draperies	Awnings	Staple (rayon)
Jackets	Mattresses and pads	Bags	Yarn (acetate)
Jeans	Pillowcases	Bookbindings	Noncellulosic
Pajamas	Pillow ticking	Electrical insulation	Staple
Rainwear	Quilts	Flags and banners	Polyester
Robes	Sheets	Industrial hose	Nylon
Shirts	Tablecloths and napkins	Life jackets	Olefin
Sport clothes	Thread	Luggage and handbags	Acrylic
Suits	Towels and washcloths	Machinery belts	Yarn
Sweaters	Upholstery	Rope, cordage, and twine	Polyester
Work clothes		Sleeping bags	Nylon
		Tents	Olefin
		Umbrellas	Glass
		Wall covering fabric	

Appendix table 2—Data for demand analysis

Year	Q_A	Q_H	Q_I	CPI_A	CPI_H	WPI_I	POP	INCN	PD_H	PD_H	PD_I
	- - - Million pounds - - -			*(1967=100)*			*Million*	*Billion dollars*	*(1972=1)*		*(1972=1.0)*
1955	2,483.26	1,707.65	2,326.68	88.9	91.9	98.7	165.93	273.41	0.64	0.63	0.63
1956	2,471.38	1,583.73	2,324.50	89.8	93.5	98.7	168.90	291.25	.65	.64	.64
1957	2,364.43	1,477.77	2,189.51	90.6	94.4	98.8	171.98	306.92	.67	.67	.66
1958	2,316.00	1,491.36	2,046.97	90.4	92.9	97.0	174.88	317.13	.69	.68	.68
1959	2,708.23	1,757.25	2,425.59	90.5	93.2	98.4	177.83	336.12	.70	.70	.69
1960	2,675.22	1,710.82	2,212.84	91.5	94.5	99.5	180.67	349.37	.72	.71	.70
1961	2,675.69	1,728.87	2,184.19	92.0	95.0	97.7	183.69	362.90	.72	.72	.71
1962	2,918.79	1,955.51	2,361.10	92.1	94.9	98.6	186.54	383.88	.74	.73	.72
1963	2,942.74	2,048.74	2,458.49	93.0	95.0	98.5	189.24	402.76	.75	.74	.73
1964	3,091.70	2,310.81	2,565.80	93.8	95.3	99.2	191.89	437.03	.76	.75	.74
1965	3,583.38	2,905.44	2,315.55	94.5	96.0	99.8	194.30	472.16	.77	.76	.76
1966	3,752.54	3,155.55	2,577.50	96.2	97.3	100.1	196.56	510.40	.79	.79	.78
1967	3,686.21	3,300.67	2,416.73	100.0	100.0	100.0	198.71	544.55	.81	.81	.80
1968	3,876.47	3,742.45	2,701.16	105.7	103.7	103.7	200.71	588.14	.84	.84	.84
1969	3,789.34	3,892.59	2,643.24	111.9	106.9	106.0	202.68	630.43	.88	.88	.88
1970	3,786.05	3,928.15	2,436.06	116.3	109.2	107.2	204.88	685.94	.92	.92	.92
1971	4,144.49	4,769.00	2,441.27	119.9	111.6	108.6	207.05	742.81	.96	.96	.96
1972	4,427.74	5,180.08	2,516.04	122.3	113.6	113.6	208.85	801.30	1.00	1.00	1.00
1973	4,478.98	5,760.54	2,705.51	126.5	116.2	123.8	210.41	901.70	1.06	1.06	1.06
1974	4,015.12	4,828.67	2,334.63	135.7	131.5	139.1	211.90	984.60	1.18	1.18	1.18
1975	3,940.60	4,730.89	2,154.34	140.6	141.4	137.9	213.56	1,084.40	1.27	1.27	1.28
1976	4,317.90	5,285.50	2,479.47	144.9	148.3	148.0	215.14	1,185.80	1.34	1.34	1.35

Q_A, Q_H, Q_I = Quantity demanded by category (million pounds). End use percentages calculated from National Cotton Council of America data and applied to total domestic consumption figures from the Economics, Statistics, and Cooperatives Service.

CPI_A, CPI_H, WPI_I = Consumer Price Index of Apparel minus footwear, Consumer Price Index of Textile Housefurnishings and Producer Price Index of Textile Products and Apparel respectively, Bureau of Labor Statistics, U.S. Department of Labor.

POP = Total U.S. population (million), U.S. Bureau of the Census.

INCN = Nominal personal disposable income (billion dollars), U.S. Department of Commerce.

PD_A = Price deflator of services, durables, food, gasoline and oil and other nondurables, U.S. Department of Commerce.

PD_H = Price deflator of services, nondurables, auto and parts of other durable goods, U.S. Department of Commerce.

PD_I = Price deflator of nondurables and services, U.S. Department of Commerce.

Research Review

INDUCED INSTITUTIONAL INNOVATION

by Vernon W. Ruttan*

The interpretation of technical and institutional change as endogenous, rather than exogenous, to the economic system is a relatively new development in economic thought. In a book published in the early seventies, Hayami and I demonstrated that technical change in agriculture had been induced along an efficient path that was consistent with physical and human resource endowments and relative factor prices (*5*).[1] The fact that much of the technical change had been produced by public sector institutions turned our attention to the problem of institutional innovation. Binswanger and I, along with several colleagues, have further refined and tested the theory of induced technical and institutional change (*1, 13*). My purpose here is to summarize our theory.

INSTITUTIONAL INNOVATION DEFINED

A distinction is often made between institutions and organizations. Institutions are usually defined as the behavioral rules that govern patterns of action and relationships. Organizations are the decisionmaking units—families, firms, or bureaus. This appears to be a distinction without a difference. What an organization accepts as a rule is the product of tradition or decision by another organization—a court, a labor union, or a religion, for example.[2]

We have found it useful to define the concept of institution broadly to include that of organization. Institutional innovation refers to change in the actual or potential behavior or performance of existing or new organizations; in the relationship between an organization and its environment; or in the behavioral rules that govern the patterns of action and relationships in the organization's environment.

This definition is sufficiently comprehensive to include changes in the market and nonmarket institutions which govern product and factor market relationships, ranging from the organized commodity market institutions to the patron-client relationships in traditional societies. It includes changes in public and private organizations designed to discover new knowledge and disseminate it to farmers; to supply inputs such as water, fertilizer and credit; or to modify market behavior through price support, procurement, or regulation. It encompasses changes which occur as a result of the cumulative effect of the private decisions of individuals as well as those which occur as a result of group action.

SOURCES OF DEMAND FOR INSTITUTIONAL INNOVATION

The demand for institutional innovation may arise out of the changes in relative factor endowments and relative factor prices associated with development. Douglass C. North and Robert P. Thomas (*11, 12*) explain the economic growth of Western Europe between 900 and 1700 primarily in terms of changes in the institutions governing property rights. These institutional changes were, in their view, induced by the pressure of population on increasingly scarce resource endowments. Theodore W. Schultz (*16*), focusing on more recent economic history, has identified the rising economic value of man during the process of economic development as the primary source of institutional change.

The suggestion that changing resource endowments and changing factor price ratios induce institutional innovation is consistent with contemporary experience in developing countries. In Thailand, the ruling elite shifted its major economic base from property rights in man to property rights in land as land prices rose relative to the price of labor in the latter half of the 19th century (*4*). Patron-client obligations were modified in favor of tenants and landless laborers in Indonesia between 1868 and 1928, a period of generally rapid economic growth. Since the late twenties, land prices there have risen relative to wage rates and the balance has again shifted in favor of landowners (*19*). Rising land prices associated with higher yielding rice varieties and increased population pressure in the Philippines are inducing changes in land tenure and harvest sharing arrangements. Farmers who acquired leasehold and security of tenure rights under the land reform are now finding it profitable to sublease part of their land under share tenure arrangements. Both owners and tenants are requiring laborers to carry out weeding operations to establish their right to share in

*Professor, Department of Agricultural and Applied Economics, University of Minnesota.

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

[2] According to Frank H. Knight, ". . . the term 'institution' has two meanings. . . . One type . . . may be said to be created by the 'invisible hand.' The extreme example is language, in the growth and changes of which deliberate action hardly figures; . . . law is in varying . . . degree of the same kind. The other type is of course the deliberately made, of which our Federal Reserve System and this (American Economic) Association itself are examples. With age, the second type tends to approximate the first" (*7*, p. 51).

The suggestion that changing resource endowments and changing factor price ratios induce institutional innovation is consistent with contemporary experience in developing countries.

harvest operations (*9*).

New income streams result from the efficiency gains associated with technical change or improvements in institutional performance. The partitioning of these income streams among factors of production and among social or economic classes is a major source of institutional innovation. According to the Ricardian model of distribution, the gains would flow to owners of the factors with relatively inelastic supplies. However, the primary function of capturing new income streams by the suppliers of inelastic factors—the factors that constrain growth—is establishing a claim on the increased production. As a result, advances in technology can prompt attempts to redefine property rights or change the behavior of market institutions so as to modify the partitioning of the new income streams. Much of the history of farm price support legislation in the United States, from the mid-twenties to the present, can be interpreted as a struggle between agricultural producers and the rest of society to capture the new income streams resulting from technical progress in agriculture.

The demand for institutional change may also shift because cultural endowments change. Even under conditions of unchanging demand, however, institutional change may arise from improvements in society's capacity to supply institutional innovations—that is, as a result of factors which reduce the cost of institutional change.

SOURCES OF SUPPLY OF INSTITUTIONAL CHANGE

Neither the institutionalist nor analytical schools in economics have adequately addressed the issue of the supply of institutional innovation. The older institutional tradition treated institutional change as primarily dependent on technical change. There is a tendency within modern analytical economics either to abstract from institutional change or to treat it as exogenous. Neither North-Thomas nor Schultz, whose insights assisted us in our theory building, has suggested a theory of the supply of institutional change.

The sources of institutional and technical change are similar. Just as the supply curve for technical change shifts as a result of advances in knowledge in science and technology, the supply curve for institutional change shifts as a result of advances in knowledge in the social sciences and related professions (law, administration, social services, and planning).

For example, research leading to quantification of commodity supply and demand relationships can contribute to more efficient supply management, food procurement, and food distribution programs. Research on the social and psychological factors affecting the diffusion of new technology can lead to greater effectiveness of agricultural credit and extension services. Research on alternative land tenure institutions or on the organization and management of group activities in agricultural production can increase equity in access to political and economic resources and increase resource productivity in rural areas.

Institutional change is not, however, primarily dependent on formal research. Technical change occurred prior to the invention of the research laboratory or experiment station. Similarly, the innovative efforts of politicians, bureaucrats, entrepreneurs, and others may cause institutional change.

The timing or pace of institutional innovation may be influenced by external contact or internal stress. If we were satisfied with the slow pace of technical and institutional change that characterizes trial and error, we would not institutionalize research in either the natural or the social sciences.

TOWARD A THEORY OF INDUCED INSTITUTIONAL CHANGE

A persistent dualism pervades much of the work of economists and other social scientists who study the historical and institutional dimensions of development. Either institutional change depends on technological change, or technological change depends on institutional change (*1*, pp. 328-333). Dissent over priority is unproductive. Technical and institutional change are highly interdependent and must be analyzed together.

The sources of *demand* for technical and institutional change are essentially similar. A rise in the price of labor relative to other factors induces technical changes designed to permit the substitution of capital for labor. It simultaneously induces institutional changes which enhance labor productivity and expand the worker's control of employment conditions. A rise in the price of land induces technical changes designed to release the constraints on production. This price rise also induces institutional changes that lead to redefinition of property rights in capital and land.

The new income streams generated by technical change and by increased institutional efficiency alter the relative demand for products and encourage more profitable innovations. The new income streams also induce

The sources of demand *for technical and institutional change are essentially similar..*

further institutional changes designed to alter the distribution of income.

The sources of *supply* of technical and institutional change are essentially similar too. Advances in knowledge in science and technology reduce the cost of new income streams generated by technical change. Similarly, advances in knowledge in the social sciences and related professions reduce the cost of the new income streams generated by improved institutional efficiency.

A large body of literature in economics on institutional change exists, but the radical and reformist thrust of much of it has impeded analytical considerations. When positive analysis has been applied to the issue of institutional change, it has often focussed on such grand themes as revolution rather than the less dramatic incremental changes in institutional performance or the gradual cumulation of institutional innovations.

The theory or model of induced institutional innovation can produce testable hypotheses regarding (a) alternative patterns of institutional change over time for a particular society and (b) divergent patterns of institutional change among countries at a particular time.

This theory requires clarification of the conceptual relationships among resource endowments, cultural endowments, technological change, and institutional change as they relate to agricultural development. The theory also calls for careful testing of those relationships against past and present experience. However, the methodology available for testing the induced institutional change hypothesis is not yet so rigorous as the econometric tests that confirm the strength of the induced technical change hypothesis.

Case studies provide an important method for testing the hypothesis. As our knowledge of the economic forces conditioning the rate and direction of institutional change improves, we will likely face policy issues similar to those in the areas of induced technical change. For example, price distortions bias current technology choices. The distortions also affect the availability of future technologies. Downward movements in the wheat-fertilizer or rice-fertilizer price ratio have delayed the development of fertilizer-responsive wheat and rice varieties by more than a generation in some countries. In several countries, exchange rate bias and direct subsidies have caused premature mechanization. These market distortions also bias the direction of institutional innovation and change.

As a next step, we need to identify and describe economic and political markets which turn latent sources of demand and supply of institutional change into effective innovations. These forces frequently operate through relatively imperfect or poorly organized markets or through nonmarket channels. Resources that shift the demand and supply of institutional change are distributed unequally among individuals and institutions.

This brief introduction cannot explore these issues in greater detail. In *Induced Innovation: Technology, Institutions, and Development*, we have tried to explore some ways in which economic and political market structure affect the demand for and supply of institutional innovation and bring about actual change (*1*, pp. 342-357).

Neither technical nor institutional change can be treated as entirely endogenous to the economic system. There is an autonomous thrust toward the accumulation of knowledge. Nature imposes constraints on what can be discovered. Changes in the evolution of ideas influence the path of institutional change. We reject the view that human activity can be partitioned neatly into distinct sets of "economic" and "noneconomic" activity. Rather, both economic and noneconomic sources of behavior may condition any dimension of human activity. Some of the goals that societies set are not achievable immediately through any combination of technological and institutional change. Thus, technical and institutional change cannot be fully explained by resource endowments.

No general theory of institutional innovation or economic development, that would have operational value, is feasible. It will be sufficient if we can demonstrate that changes and differences in resource endowments, reflected through either market or nonmarket forces, do influence the direction of technical and institutional innovation.

In Earlier Issues

. . . a successfully functioning and efficient economy consists of its natural resources and the people who use them; per capita income and therefore levels of living are resultants of the ratio between these two factors. This ratio can be changed in the long run only by producing economically more valuable products of resident natural resources and dividing the income from them among fewer people.

Carl C. Taylor
AER, Vol. IV, No. 1,
Jan. 1952, p. 26

The sources of supply *of technical and institutional change are essentially similar too.*

REFERENCES

(*1*) Binswanger, Hans P., and Vernon W. Ruttan. *Induced Innovation: Technology, Institutions, and Development.* Johns Hopkins University Press, Baltimore, 1978.

(*2*) Commons, John R. *The Economics of Collective Action.* Macmillan Co., New York, 1950.

(*3*) Davis, Lance E., and Douglass C. North. *Institutional Change and American Economic Growth.* Cambridge University Press, New York, 1971.

(*4*) Feeney, David Harold. *Technical and Institutional Change in Thai Agriculture, 1880-1940.* Unpublished Ph.D. thesis, University of Wisconsin-Madison, 1976.

(*5*) Hayami, Yujiro, and Vernon W. Ruttan. *Agricultural Development: An International Perspective.* Johns Hopkins University Press, Baltimore, 1971.

(*6*) Horowitz, Irving Louis. *Three Worlds of Development.* 2nd ed. Oxford University Press, New York, London, Toronto, 1972.

(*7*) Knight, Frank H. "Institutionalism and Empiricism in Economics," *Am. Econ. Rev.* 42, May 1952, pp. 45-55.

(*8*) Krueger, Anne O. "The Political Economy of the Rent Seeking Society," *Am. Econ. Rev.* 64, June 1974, pp. 291-303.

(*9*) Kukuchi, Masao, Luisa Maligalong, and Yujiro Hayami. "Evolution of Land Tenure System in a Laguna Village." The International Rice Research Institute, *Los Banos,* The Philippines, 1977.

(*10*) Lerner, Abba P. *The Economics of Control.* Macmillan, New York, 1944.

(*11*) North, Douglass C., and Robert P. Thomas. "An Economic Theory of the Growth of the Western World," *Econ. History Rev.* 23, 2nd Series, 1970, pp. 1-17.

(*12*) __________ . *The Rise of the Western World.* Cambridge University Press, London, 1973.

(*13*) Ruttan, Vernon W. "Technology and the Environment." *Am. J. Agr. Econ.* 53, No. 5, Dec. 1971, pp. 707-717.

(*14*) __________. "Institutional Innovation." *Distortions of Agricultural Incentives,* ed. T. W. Schultz. Indiana University Press, Bloomington, 1978.

(*15*) Samuelson, Paul A. *Foundations of Economic Analysis.* Harvard University Press, Cambridge, Mass., 1948.

(*16*) Schultz, Theodore W. "Institutions and the Rising Economic Value of Man." *Am. J. Agr. Econ.* 50, Dec. 1968, pp. 1113-1122.

(*17*) Seckler, David. *Thorstein Veblen and the Institutionalists.* Macmillan Co., London, 1975.

(*18*) Shubik, Martin. "Beyond General Equilibrium." Discussion Paper 417. Cowles Foundation for Research in Economics at Yale University, New Haven, Jan. 14, 1976.

(*19*) White, Benjamin. "Agricultural Involution: A Critical Note." *Agro-Economic Survey,* Jakarta, Indonesia, Oct. 1974. Mimeographed paper.

(*20*) Zingler, E. K. "Veblen vs. Commons: A Comparative Evaluation." *Kyklos* 27, 1974, pp. 322-344.

A NEW SET OF SMALL-FARM GUIDELINES

By David Brewster and Thomas A. Carlin*

Secretary Bergland has stated that it is USDA's policy "to encourage, preserve and strengthen the small farm as a continuing component of American Agriculture." The Department, he said, will provide assistance "which will enable small farmers and their families to expand the necessary skills for both farm and nonfarm employment to improve their quality of life."[1] The Secretary has also established a Policy Committee on Small Farm Assistance that includes the Assistant Secretaries and the Director of Economics, Policy Analysis, and Budget; Assistant Secretary Alex Mercure serves as Committee Chairman.

The new small-farm policy committee has directed that the following criteria be used to identify those families aided by USDA's small farm effort. The families so designated should:

- Operate farms by providing most of the labor and management
- Have total family incomes from farm and nonfarm sources below the median nonmetropolitan family incomes in their States
- Depend on farming for a significant portion, though not necessarily a majority, of their incomes.

About 52 percent of all U.S. farm families had incomes below the median nonmetropolitan family level in 1975 (table). Most of these families lived in the South and North Central regions.

These guidelines depict a group whose members share a problem that distinguishes them from other farm families—moderate to low income. The criteria imply two conclusions about the Department's small-farm policy and research that have been noted before.[2] First, the appropriate unit of concern is the family, not the farm. Second, small-farm programs would develop human resources and both farm and nonfarm income opportunities would receive attention. Thus, the guidelines define a target group, a policy problem, and a framework within which to address that problem.

As a policy tool, the guidelines serve well. As an analytical concept, however, they are deficient regarding the requirement that small-farm families should depend on farming for a "significant" share of their incomes. Being a subjective notion, "significant" is difficult to quantify. Leaving its precise meaning up to individual program managers could cause uncertainty about the families encompassed by the small farm guidelines. Some managers might interpret "significant" to mean 40 to 50 percent, whereas others might choose a much lower figure. Such disparity would cloud any profile that might be drawn of the group served and would likely hinder analysis of programs based on the recommended criteria.

We propose that, at least for statistical purposes, families covered by the guidelines should derive a designated portion of their income from farming operations. While the choice of a percentage is arbitrary, we suggest that it be set at 10 percent—a figure roughly equal to the total nonmetropolitan income accounted for by farm earnings since 1970. In other words, to meet the Assistant Secretaries' criteria, families would be required to have incomes from farming that, at a minimum, are as significant to them as farm earnings are generally to the greater nonmetropolitan population.

For many of the farm families with incomes below the median nonmetropolitan family level in 1975 farm income is a minor part of total family income. Interpreting "significant" to mean that at least 10 percent of family income comes from farming would reduce the number of farm families included in the new policy effort to about 38 percent of all farm families. (This percentage includes low-income families whose farm expenses exceed their farm sales.) The resulting set of small-farm guidelines would be suitable for analysis as well as policy.

The Department's new definition

In Earlier Issues

. . . it is desirable for the creative minds to be protected from the buffeting that comes from the overuse of red-tape but they are being buffeted by administrative problems of their own creation.

William T. Wolfrey
AER, Vol. IV, No. 1,
Jan. 1952, p. 24

*David Brewster is an historian with the National Economics Division and Thomas A. Carlin is an economist with the Economic Development Division, ESCS.

[1] Secretary Memorandum No. 1967. "Assistance to Small Farm Operators," Jan. 3, 1979.

[2] Brewster, David. "Perspectives on the Small Farm." *Small Farm Issues: Proceedings of the ESCS Small Farm Workshop, May 3-4, 1978* (forthcoming); and Lewis, James A., and Peter Emerson. "A Note on Small Farms." *Agr. Econ. Res.* Vol. 30, No. 4, Oct. 1978, pp. 42-43.

We propose that, at least for statistical purposes, families covered by the guidelines should derive a designated portion of their income from farming operations

differs substantially from the older definition of a small farm as any operation with less than $20,000 in annual agricultural sales. The older definition emphasizes the farm business as the primary policy concern. It is so broad that it fails to describe unique problems. The farms, so defined, account for nearly three quarters of the Nation's total. Their operators tend to be slightly older and less well educated than other farmers. But as Lewis, Larson, and Emerson have shown, farmers in the under $20,000 sales class share little else in common.[3] Some live in poverty yet most do not. Nor do most depend primarily on farming for their incomes. They engage in all types of crop and livestock production. They include in their ranks both white and minority operators. Except for their low agricultural sales, they exhibit much the same characteristics as the general farm population.

Because of its convenience, the $20,000 sales definition has been widely cited and has sometimes been misconstrued as officially established by the Department or the Census Bureau. But use of the definition is required by statute only in connection with certain research and extension programs authorized by the Rural Development Act of 1972, as amended. Apart from those programs, the Department is free to develop alternative guidelines to direct services to small-scale operators and their families.

[3] Larson, Donald K., and James A. Lewis. "Small Farm Profile." *Small Farm Issues: Proceedings of the ESCS Small-Farm Workshop, May 3-4, 1978* (forthcoming).

Farm families with income below nonmetropolitan median family income, by region, 1975*

Region[1]	Farm families		Proportion with incomes below median non-metropolitan family income
	Total	Incomes below median non-metropolitan family income	
	Thousands		*Percent*
Northeast	131	71	54
New England	38	25	66
Middle Atlantic	93	46	49
North Central	976	483	49
East North Central	409	190	46
West North Central	567	293	52
West	249	106	43
Mountain	112	55	49
Pacific	138	51	37
South	844	482	57
South Atlantic	241	144	60
East South Central	301	167	55
West South Central	302	171	57
United States	2,200	1,142	52

*Estimated.

[1] New England: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. Middle Atlantic: New Jersey, New York, and Pennsylvania. East North Central: Illinois, Indiana, Michigan, Ohio, and Wisconsin. West North Central: Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota. South Atlantic: Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia. East South Central: Alabama, Kentucky, Mississippi, and Tennessee. West South Central: Arkansas, Louisiana, Oklahoma, and Texas. Mountain: Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming. Pacific: Alaska, California, Hawaii, Oregon, and Washington.

Source: U.S. Bureau of the Census, *Current Population Reports:*

(1) Series P-60, No. 105, "Money Income in 1975 of Families and Persons in the United States," 1977;

(2) Series P-60, Nos. 110, 111, 112, 113, "Money Income in 1975 of Families and Persons in the United States (Spring 1976 Survey of Income and Education," 1978.

THE FOOD STAMP PROGRAM AND THE PRICE OF FOOD

By Mike Belongia and William T. Boehm*

The Food Stamp Program (FSP) provides additional "food buying income" to low-income households. It allows participating households to purchase nutritionally adequate diets through regular market channels. Prior to the passage of the Food and Agriculture Act of 1977, all FSP participant households of a given size received the same total value of coupons. The purchase requirement, the amount paid to receive the stamps, depended mainly on the household's earned income. The difference between what a participating household *paid* for the coupons and the value *received* in food buying income represented the addition to purchasing power, or "bonus." In fiscal year 1977, the value of the bonus exceeded $4.3 billion. The 1977 act eliminated the purchase requirement.

Questions have arisen about the FSP's effectiveness in improving the nutritional status of recipients' diets.[1] Questions also have been asked about the effect, if any, of the transfer of food-buying resources on prices.

Some have argued that the program represents a "hidden tax"—higher income households not only pay for the resource transfer directly, but pay indirectly through higher food prices. In this note, we report research on the FSP's *time related* influence on the overall price of food and on prices of selected commodity groups. The FSP, results indicate, has had a statistically significant, positive, but small influence on food prices generally.

*Mike Belongia and William T. Boehm are agricultural economists with the National Economics Division, ESCS.

[1] Sullivan, Dennis H. "A Note on Food Stamp Reform," *Am. J. Agr. Econ.* Vol. 58, No. 3, Aug. 1976.

RESEARCH PROCEDURES

The bonus is most appropriately treated as a segmented increase in participants' income, rather than as a decrease in the prices participants pay for food.[2] If the supply of food is assumed to be inelastic, increases in the quantity taken from the market tend to put upward pressure on prices. The extent of the price or quantity impact will depend upon the size of the shift in demand, the income and price elasticities of demand, and the price elasticity of supply.

A set of single-equation models—one for each major food commodity group—was developed to isolate the effect of the bonus on the price of food. Factors thought to be responsible for shifting either demand for or supply of the foods were included. Parameter estimates were obtained using ordinary least squares regression. The Consumer Price Index (CPI) for each of the groups was the dependent variable.

The models estimated were:

1. $CPI\ (\text{all food})_t = f(Y_t, B_t, M_{t-1}, FP_{t-1}, Z)$,
2. $CPI\ (\text{meats})_t = f(Y_t, B_t, L_{t-1}, G_{t-1}, M_{t-1}, Z)$,
3. $CPI\ (\text{cereals and bakery products})_t = f(Y_t, B_t, G_{t-1}, M_{t-1}, Z, R)$,
4. $CPI\ (\text{dairy})_t = f(Y_t, B_t, D_t, DS_t, M_{t-1}, Z)$.

[2] Southworth, Herman M. "The Economics of Public Measures to Subsidize Food Consumption." *J. Farm Econ.* Feb. 1945; and Sullivan, Dennis H. "A Note on Food Stamp Reform," *Am. J. Agr. Econ.* Vol. 58, No. 3, Aug. 1976.

where:

CPI_t = Consumer Price Index for the relevant food commodity in quarter t as reported by the U.S. Department of Labor;

Y_t = Real disposable per capita personal income, 1972 dollars;

B_t = Real per capita food stamp bonus, 1972 dollars;

FP_{t-1} = Index of prices received by farmers;

L_{t-1} = U.S. total meat production, in millions of pounds;

G_{t-1} = Quantity of all grain stocks, in millions of short tons;

M_{t-1} = Index of food processing and marketing costs;

D_t = Milk production, in millions of pounds;

DS_t = Dairy stocks, in millions of pounds;

Z = Zero-one variable indicating the quarters that national price controls were in effect beginning in August 1971;

R = Zero-one variable indicating unusual foreign grain sales in 1973-74.

Quarterly data were obtained from secondary sources. The period studied extended from July 1970 to June 1977, during which the CPI for all food increased 64.1 percent from 115.8 to 190.0 (1967=100). Price levels for each of the commodity groups increased about the same amount but varied substantially more quarter to quarter. The total value of bonus stamps issued per quarter increased from $314 million in 1970 to a high of $1,392 million in the first quarter of 1976. The issuance in spring 1977 was $1,311 million. FSP participation increased from 6.9

The FSP has not noticeably increased food prices. . . . Significant increases in the percentage of the total market influenced by the FSP could exert more noticeable upward pressure on food prices.

million persons per month in 1970 to more than 19 million in 1976. Participation in 1978 averaged about 16 million persons per month. Mean values for all variables appear in table 1.

RESULTS

The models explained a statistically significant proportion of the variability in food prices (table 2). No evidence was found of any major problem with intercorrelation among the independent variables. Production variables were inversely related to price. Increases in real disposable per capita income were associated with increases in price.

The bonus food stamp variable is significantly different from zero at the 99-percent level in all models but dairy. Results of the all-food model indicate that a 1-percent increase in real per capita bonus would increase the CPI for food by less than 0.07 percent. This change amounts to only 0.14 points on the price index.

The coefficient estimates for bonus in the meats and in the cereal and bakery equations are larger than in the all-foods model. This implies that, given an increase in real bonus, other things being equal, there will be a proportionally greater impact on prices for these components than on food prices in general.

The bonus coefficient was smaller for dairy than for grain and meat. First, low-income households tend to allocate a greater percentage of food expenditures to the cereal and bakery products group than do higher income households. Second, a relatively large proportion of the food stamp households live in the South where dairy product purchases are relatively low.[3] Finally, prices of dairy products are influenced by Government price support and marketing order provisions. In such cases, production—rather than price—adjusts to changes in market forces.

[3] Boehm, William T. "The Household Demand for Major Dairy Products in the Southern Region," *So. J. Agr. Econ.* Vol. 7, No. 2, Dec. 1975.

CONCLUSION

The FSP has not noticeably increased food prices. Currently, food purchases with bonus food stamps account for little more than 2 percent of all food purchases. USDA estimates that the present program *adds* about $2 billion (less than 1 percent) to food sales. Significant increases in the percentage of the total market influenced by the FSP could exert more noticeable upward pressure on food prices.

Table 1—Means and standard deviations of variables included in the models, July 1970 to June 1977

Item	Unit	Mean	Standard deviation
CPI: all food	1967=100	148.13	27.74
CPI: meat, poultry, and fish	1967=100	150.78	27.07
CPI: cereals and bakery	1967=100	144.96	32.48
CPI: dairy	1967=100	138.13	23.26
Y	1972 dollars	995.21	43.31
B	1972 dollars	3.29	1.10
L	Million pounds	9,366.30	530.79
G	Million short tons	94.91	45.52
D	Million pounds	29,449.71	1,858.62
DS	Million pounds	14,547.00	3,446.00
FP	1967=100	1567.33	33.30
M	1967=100	148.55	30.68

Table 2—Results of the parameter estimation

Dependent variable	Number of observations	Constant	Independent variables					
			Y_{dt}	B_t	M_{t-1}	Z	FP_{t-1}	L_{t-1}
CPI_{food}	27	1.53	0.027 [1] (1.04)	3.046 (2.13)	0.489 [3] (8.82)	−2.834 [2] (−2.12)	0.250 [3] (7.46)	—
CPI_{meat}	27	−113.69	.351 (6.00)	14.695 [3] (4.09)	−.077 (−.48)	−4.728 (−1.37)	—	−0.013 [3] (−4.89)
CPI_{grain}	27	92.75	−.084 [2] (−2.10)	14.825 [3] (6.00)	.565 [3] (5.25)	−4.422 (−1.79)	—	—
CPI_{dairy}	27	−4.237	.074 (1.76)	4.647 (1.71)	.408 [2] (3.36)	−3.920 (−1.46)	—	—

Dependent variable	Number of observations	Constant	Independent variables						
			G_{t-1}	R	D_t	DS_t	R^2	D-W	F
CPI_{food}	27	1.53	—	—	—	—	0.99	1.63	751.34
CPI_{meat}	27	−113.69	−0.024 (−.86)	—	—	—	.96	2.12	78.64
CPI_{grain}	27	92.75	.029 (1.49)	14.822 [3] (7.13)	—	—	.99	2.22	253.78
CPI_{dairy}	27	−4.237	—	—	−0.0007 (−1.21)	0.001 [2] (2.61)	.96	1.05	95.77

[1] Values in parentheses are *t* ratios. [2] Significant at the .05 level. [3] Significant at the .01 level.

DISTORTIONS OF AGRICULTURAL INCENTIVES

Theodore W. Schultz, editor, Indiana University Press, Bloomington, Ind., and London, England, 1978, 353 pages, $12.95.

*Reviewed by Lyle P. Schertz**

The agricultural economics profession has a long tradition of analyzing conflicts among political and economic interests. *Distortions of Agricultural Incentives* states that the outcome of these conflicts in low-income countries has distorted economic incentives in ways that suppress farmers' economic opportunities. The most intriguing of the many secondary themes to this reviewer is that although equity is often used to justify the political decisions causing the distortions, the distortions decrease equity.

People who judge that producer prices in many countries are lower than they should be, as I do, will find comfort and support in this book. Yet those who judge, as I do, that agricultural economists should give more attention to distribution issues will find that this book reinforces their earlier conclusions.

The book includes 22 papers presented at a 1977 workshop. The 22 authors have considerable experience with development and trade problems of low-income countries. Only one author, however, comes from such a country. Half the authors are associated with universities in North America. Four are members of U.S. institutions, such as foundations and the Federal Government. Three work in international research institutes. The others are associated with the World Bank, the British Government, or the Australian National University.

The book reflects the wisdom and influence of Theodore W. Schultz, organizer of the conference and editor of the publication. There are six parts:

1. Constraints on Agricultural Production
2. Resources and Environment
3. Distortions in Incentives
4. International Markets
5. Agricultural Research, Education, and New Institutions
6. Quest for Equity

Authors of part 3, "Distortions of Incentives," the largest, devote major attention to price distortions. In addition, David Hopper reviews many government nonprice decisions which have affected production and marketing efficiency. Keith Finlay's comments illustrate issues related to the transfer of technology among scientists in different countries. Randolph Barker points up the importance of public decisions on public investments. Martin Abel reminds us that the process of reducing and eliminating distortions is not continuous. He emphasizes the importance of developing analytical capability to respond to situations amenable to policy change. Other parts of the book, such as the one focused on research, provide important perspectives of forces that influence decisions distorting agricultural incentives.

Implicitly, "distortions" are taken by most of the authors to mean deviations from international trade prices. Schultz, however, avoids saying that international prices should be the standard for the measurement of distortions. Instead, he points to conditions where "an optimum economic incentive provides the information that leads producers to allocate resources in ways that result in a maximum production that will clear the market at the price that maximizes the utility of consumers."

Two aspects of the book merit special mention: the meaning of distortions and the concept that these have been inconsistent with the realization of greater equity. There are limits to how far even a task master as superb as Ted Schultz can push contributors. They do not make his concept of distortions operational. Sir John Crawford appropriately asks: "distortions from what?" He concludes that most of the authors use world prices. Yet, as he points out, international prices are obviously affected by all sorts of activities directly applicable to trade—including tariffs, levies, and export subsidies. Domestic policies in developed and developing countries that affect production, consumption, and prices also influence international trade and prices.

Apart from Sir John, D. Gale Johnson is the only author who deals with the appropriateness of international prices as a standard. He recognizes that international prices are not necessarily fair nor equitable nor do they always reflect underlying conditions of supply and demand. What he emphasizes, too many of us overlook. These prices are reality! And, therefore, such prices reflect alternatives and opportunities realizable by countries to the extent that they are involved in world trade.

The possible inconsistency of distortions and equity is an important issue for our profession. Schultz argues ". . . that economic theory and evidence tell a consistent story about the adverse effects of distortions in incentives on agricultural production and on welfare." Gilbert Brown similarly states: "My hypothesis is that agricultural production, income distribution and economic growth would all benefit from reduction or elimination of distortions that reduce agriculture's domestic terms of trade."

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

Two aspects of the book merit special mention: the meaning of distortions and the concept that these have been inconsistent with the realization of greater equity.

The authors confirm the postulated effects of distortions on production, but the postulated effects of distortions on welfare and income distribution are not treated in depth empirically. I would like to have seen a closer scrutiny of the wording of this proposition and a closer analysis of the supporting evidence. If supported, the political decisions leading to the distortions will ultimately be exposed as a rip-off of the poor. Such exposure might do much to bring about different decisions. On the other hand, if the empirical evidence indicates a mixed record—as I suspect it might—the evidence will, nevertheless, make political decisions wiser.

I wish that the authors' disagreements had been highlighted. We all know that no one thinks very much when everyone thinks alike. Casual reading of this book may lead one to think the disagreements are minor. Perhaps. But some disagreements in the book appear fundamental, such as (1) the differences between Sir John and others on the use of international prices and (2) the differences between Schuh's advocacy of uncoupling price and welfare policies and Willett's argument that prices do many things—only one of these being to guide allocation of resources.

Further, the text does not indicate if the participants accepted, rejected, or simply ignored Schultz' concept of optimum economic incentives that is set in the neoclassical paradigm of maximizing utility of consumers. One has a hunch that Emery Castle is not completely satisfied when he points out "... a different efficiency solution will result if the income distribution is varied."

The book makes several important contributions. The authors point up the pervasiveness of distortions in prices, but they also examine distortions in such areas as administrative decisions and public investments. For me, the book has a major lesson. Researchers should be doing much more about measuring income distribution effects of price policies. They need to examine seriously programs justified on equity grounds. Do the programs undercut the realization of equity as well as efficiency? Surely, policies which sacrifice both efficiency and equity should be identified and exposed. Our responsibility here is a serious one.

In Earlier Issues

Ever since the Physiocrats—1756-76—declared that there were natural laws which governed the operation of the economic system, economists have sought to discover what these laws were. We have seen a procession of theoretical systems—classical, Austrian, neoclassical, imperfect competition, Keynesian—each having its day and then yielding to another. Some of these schools had hoped to create permanent systems of economic thought that would have universal validity.

The historical school denied the existence of economic laws; the relativists declared it was impossible to construct a body of theory universally true in both time and space. Each system of thought, they held, was valid only for a particular and definite economic order. As one result of all this, today, economic theory is in a chaotic state—the despair of the professional economist and the laughing stock of the market place.

Max J. Wasserman
AER, Vol. IV, No. 1, Jan. 1952, p. 30

MODELS IN THE POLICY PROCESS: PUBLIC DECISION MAKING IN THE COMPUTER ERA

Martin Greenberger, Matthew A. Crenson, and Brian L. Crissey. Russell Sage Foundation, New York, 1976, 355 pages, $15.

*Reviewed by William E. Kost**

As society becomes more complex and interdependent, its social and economic problems become increasingly complicated. Decisionmakers need to comprehend the structure and behavior of the socioeconomic system to diagnose problems and to propose and evaluate actions that could lead to solutions. They have increasingly turned for help to methodologies based on mathematical formulations of system interactions. Identifying a socioeconomic system and its underlying structure, and then representing it mathematically, is called policy modeling by the authors of *Models in the Policy Process: Public Decision Making in the Computer Era.*

This book focuses on the role of policy modeling and policy modelers in the public decisionmaking process. The authors' basic approach is historical, sometimes anecdotal. They concentrate on case studies that show the development of particular methodologies and describe how models enter the public policy decisionmaking arena. They concentrate on the political process of using models, rather than on their technical nature. Thus, this is one of the few books on modeling that does not require of the reader a strong technical or quantitative background.

The book's first section provides several brief, descriptive cases of the interaction of policy models, modelers, and decisionmakers: (1) the Club of Rome *Limits to Growth* model, (2) gross national product forecasting and the Laffer-Ranson model, (3) the program-planning-budgeting system (PPBS) movement, and (4) a Jamaica Bay tidal estuary simulation.

The second section contains one of the clearer expositions I have read of the relationships between the "real world," theory, methodology, data, and the formal model. The authors emphasize the idea that a major advantage of policy modeling is that it forces one to think logically about a complex situation. They also emphasize the important role of the modeler in interpreting and promoting the model. This section includes a chapter outlining the historical development of several methodologies for policy modeling: linear economics, operations research, statistical economics, urban and regional development, and engineering. The authors look at the reasons underlying each approach and how each was designed to answer certain questions as well as highlight the implicit biases of each.

Probably other authors would have come up with other categories of modeling methodologies. For example, I would have included the time series modeling of Box, Jenkins, Granger, and others. The categories, however, are not the important issue. What is important is that methodologies are shown to have been developed to fill certain needs and to cope with specific problems. Once this is recognized, a good policy modeler should be able to tailor methodologies to particular issues. Unfortunately, as the authors point out, this is not always the case. Individual researchers tend to have their own favorite methodology which they attempt to apply, with varying success, to all situations.

The third section of the book details three case studies involving the interaction of modeling with policy. The first case involves the rise of systems dynamics—and the controversy surrounding it—through its applications documented in Forrester's *Urban Dynamics* and *World Dynamics and in Limits to Growth.* The authors center on objections raised concerning the validity of the model structure, the conclusions reached, and the absence of supportive data. Particular attention is paid to the system dynamics-econometrics methodology controversy.

The second case study involves the parallel rise of large-scale computers and econometric modeling, as embodied in large-scale macroeconomic models. This case study emphasizes the methodology's reliance on formal economic theory, data, and estimation procedures. The approach is contrasted with systems dynamics. The authors follow this general discussion with a rather complete genealogy of modern macroeconomic models of the U.S. economy. They start with Tinbergen, follow through "the Klein line," and culminate in the institutionalization of models in large research-modeling consulting firms such as Data Resources, Inc., Wharton Econometric Forecasting Associates, and Chase Econometric Associates, Inc.

Finally, the authors consider the life cycle of the Rand Institute and its relationship with New York City, particularly with its fire department and health services administration. This case study emphasizes not so much the methodologies used but the political and institutional environment in which policy modelers work.

Decisionmakers have high expectations of policy modelers. However, authors feel that:

> . . . the growth in the useful application of policy models to the problems facing gov-

---

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

The authors feel that a considerable gulf exists between the decisionmaker and the modeler regarding expectations of realistic roles for models.

> ernment decision makers is not keeping up with the increase in either the number or complexity of these problems. . . . At the same time, the use of models for political purposes is expanding noticeably. . . . The use of models to dramatize or publicize particular points of view is overshadowing their use for the enlightenment of policy makers (p. 337).

The authors feel that a considerable gulf exists between the decisionmaker and the modeler regarding expectations of realistic roles for models. The political setting and institutional environment in which models are developed and applied becomes a key factor in bridging this gap. Because model construction and use educates the modeler more than the decisionmaker, modelers—especially those who speak for a model—often prove more important than the model in determining how or if it is used. Therefore, recommendations for improving the potential of policy models should include emphasis on education, communication, and institution building. Decisionmakers must better understand the capabilities and limits of models and expend more effort communicating their problems and objectives to the modeler.

The modeler needs to develop models for problem solving rather than as an academic exercise. Results must be presented in a timely manner, be clearly communicated, and keyed to the client, if they are to have any impact on decision processes.

The authors see two other serious (and related) deficiencies in policy modeling. Professional standards for modeling are virtually nonexistent. Documentation standards for data, assumption specifications, model structure, programs, validation, and results do not exist. This severely hampers effective use. The authors also find that modelers build or work on their own models and pay little attention to others' models. This phenomenon is reinforced by the modelers' belief that specific models have to be built for each specific problem, and by the professional reward system recognizing individual innovative work as opposed to group efforts or to a series of marginal extensions of others' models. This situation not only hinders the development of modeling standards, but it also limits a model's useful life.

What are needed, the authors strongly believe, are "living" models with long life spans. These models, as they are improved and expanded, are more likely to aid in the analysis of the complex and interdependent socioeconomic problems that face society. To give models a life of their own requires developing organizational frameworks in which these models can exist. It also requires changing the professional reward system so that model maintainers and analyzers, as opposed to model builders and users, are rewarded for doing the day-to-day work necessary to keep models viable. Modeling efforts must be additive.

Anyone interested in modeling, either builders or users, will find this book fascinating. It does not teach how to model. It does, however, give an excellent perspective on the environment in which models are being used and on some of the problems faced by policy modelers. For economists, the chapter on macroeconomic modeling alone almost justifies picking up this book.

In Earlier Issues

Zoning originated and developed in crowded cities. Its basic regulations were designed for urban ends. Pioneers in the field of rural zoning took these basic raw materials and shaped them to serve the rural community. In the early twenties, in Wisconsin and in a few other States, rural communities began to adopt zoning techniques for protecting rural values on that day's more restricted urban fringe. During the following decade, they established forestry and recreational districts to help bring order out of land-use chaos in the cut-over areas of the North Central States. More recently, in the open country and on an expanded urban fringe, which sometimes extends 30 to 50 miles beyond city limits and often overlaps fringe areas of neighboring cities, many rural communities are exercising zoning powers to guide residential growth, to assign commercial and industrial activities to designated areas, and to preserve the safety and carrying capacity of their highways.

Rural people in increasing numbers are recognizing the value of rural zoning as one available regulatory measure for protecting their community and for guiding its growth.

Erling D. Solberg
AER, Vol. III, No. 4,
Oct. 1951, p. 135

ECONOMETRIC MODELING OF WORLD COMMODITY POLICY

F. Gerard Adams and Jere R. Behrman, editors, Lexington Books, Lexington, Mass., 1978, 223 pages, $17.95.

*Reviewed by Jack Rower**

Adams and Behrman have produced a book that provides something for everyone interested in international agricultural policy and econometric modeling of commodities. The chapters are research papers that present preliminary results of a broad-based study evaluating policy strategies for stabilization of world commodity markets. The major objective is to discover and develop general principles about commodity markets that can be used to formulate optimal policy strategies.

The specific policy issues examined are the strategies for implementing the Integrated Commodity Program proposed by the United Nations Commission on Trade and Development (UNCTAD). UNCTAD's policy objective is to minimize price fluctuations of agricultural and other nonpetroleum commodities through multinational government market intervention. The purpose of modeling international commodity policy is to aid policymakers in choosing "best" or "least detrimental" policy instruments to achieve stabilization goals. The goal of policy modeling is to create, as Adams states, a tool that will generate logically consistent scenarios simulating outcomes of commodity policy applications.

The chapters trace the policy modeling process chronologically. This process falls into three general categories: methodology, implementation, and simulation analysis.

Methodology defines the logical framework in which policy modeling occurs. Model building begins with theoretical specifications of a commodity market structure. Economic theory provides the basis for abstracting the complexity of commodity markets and reducing, as much as possible, webs of interactions into a set of relationships that describe the economic underpinnings of the market.

The empirical part of model specification is critical to the development of a useful simulation model. In this stage, the theoretical model is translated into an empirical model. No matter how well a commodity model is specified in terms of economic theory, there is no guarantee that an empirical model based on it will yield useful results. Accuracy relies on such factors as the availability and quality of data required, as well as proxies; that is, substitute variables that can be used where appropriate data are unavailable. Adams notes, for example, the lack of reliable stock data for developing countries. Detailed modeling of crops such as coffee or food and feed grains is difficult because of the lack of data for countries which are major market participants.

Specification of the empirical model must not only predict and describe the commodity economy, but must also handle the decision-making process. Mariana, an associate author, addresses this issue in chapter 4 with a lucid discussion of control theory and its value as a linkage between the empirical predictive model and the decision process of policymakers. Control theory acts as the mathematical medium in which the commodity market structure is handled as a state variable (that is, the way things are), whereas the policy constraints are contained in a control variable. This technique is a recent innovation in applied economics, one that should prove interesting to agricultural economists. Using control theory, the econometric modeler can create many policy scenarios quickly, with a high degree of internal consistency.

The world coffee model exemplifies an international stochastic commodity policy model that is dynamic and that incorporates both the production and consumption market structures.

Chapter 8, an exhaustive bibliography of current commodity models and related literature, is in itself a reason to read the book. The bibliography, compiled by commodity, represents a handy research tool.

Persons seriously interested in international agricultural commodity policy or modeling will find the book helpful. It presents clearly the policy problems involved in the UNCTAD-initiated Integrated Commodity Program, as well as the problems and techniques with which the applied economist must deal.

In Earlier Issues

Diminishing returns is present in the feed-egg relationship. But the maintenance part of the ration is large and hens are fed in flocks. This leads to the conclusion that a total curve of diminishing physical return can not be experimentally distinguished from a straight line. Full feeding is usually most profitable for egg production. Practical feeding is a matter of dead reckoning.

Peter L. Hansen
and Ronald L. Mighell
AER, Vol. IV, No. 1,
Jan. 1952, pp. 2 and 7

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

ECONOMIC REALISM

Chastain, E. D., and Raymond Ritland, editors, Craftmaster Printers, Inc., Opelika, Alabama, 1978, 245 pages, $6.

*Reviewed by Gerald Schluter**

Messrs. E. D. Chastain and Raymond Ritland deserve an award for courage. Perceiving a void of realistic views on economic phenomena other than those provided by "the helpful textbook, the usual readings book, or the formal and informal oral communication," they commissioned a series of articles to provide these "realistic" views. As if this were not courageous enough, they go on to state that: "the present-day world seems to have lost some of its objectivity earlier associated with academic life and the procedures of science." Therefore, they have tried to solicit "realistic, objective views on several economic topics of great importance in everyday life." As one would expect, *Economic Realism* does not meet these high standards.

Nonetheless, this interesting set of papers covers topics that exceed the range of most nonsurvey undergraduate courses in economics. These topics include capitalism, Adam Smith, economic leadership and communication, economic forecasting, the Great Depression, the persistence of inflation, business cycles, the competitive structure of the American economy, the challenge to the free market economy, John Maynard Keynes, the Davis-Bacon Act, farm price supports, international economics, economic development, systems analysis, and the potential for economics research.

In a book of this type, the authors can achieve objectivity either by imposing the desired standards of objectivity upon the individual authors or by balancing the number of subjective articles. Although many articles in this book have a free market tilt, editor Chastain coauthored an article on Keynes and Keynesian views that fairly presents the government intervention in the market system view of economic affairs.

To achieve objectivity, the editors, therefore, chose a third alternative. Readers may agree or disagree philosophically with individual articles, but most would concur that individual articles are not balanced and the overall mix leans heavily towards the free market view. Yet, by the end, the reader has been exposed to alternative views of the economic system.

Only one article is an outright advocacy piece. One wonders why the editors included the near propaganda by Shirley Smith and Ritland, who advocate repeal of the Davis-Bacon Act, in a set of readings chosen to provide economics students with an objective view of economic events.

Some articles oversimplify economic problems. In his highly critical study of Government farm programs, Clifton Luttrell rather offhandedly dismisses the problem of factor immobility. Ritland in his article, "Capitalism, Past, Present, and Future," chastises labor union members for "economic shortsightedness" in resisting innovations. Yet he simultaneously cites the effectiveness of the buyers' strike against meat products in 1974 as an example of consumers' self-restraint that had the desired effect of lowering prices. One could argue that the 1974-75 recession also reduced demand for

*Gerald Schluter is an agricultural economist with the National Economics Division, ESCS.

In Earlier Issues

In addition to his reputation as an original penetrating economic analyst, Professor Abba P. Lerner is widely known as an indefatigable and persuasive expositor of the economics of John Maynard Keynes. Although his analysis is penetrating, it will hardly be convincing to the skeptics, but he finds it "necessary to remind almost everyone that the human race is not noted for the promptness with which it acts collectively to adopt reasonable methods of dealing with its problems."

James P. Cavin
AER, Vol. IV, No. 1, Jan. 1952, p. 22

▲ ▲ ▲

Direct Federal aid to agriculture has been a part of our national policy since the passage of the Agricultural Marketing Act of 1929 which, among other things, brought the Federal Farm Board into being. The great depression brought further efforts to assist farmers as a part of a general program of economic recovery. Since then, the people of this country through their Government have been accepting an increasing responsibility for the protection of all the people from the harsh winds of economic adversity.

C. Kyle Randall
AER, Vol. IV, No. 1, Jan. 1952, p. 29

beef. But if one grants the effectiveness of the buyers' strike and observes the current higher beef prices resulting from the herd reduction following the 1974 drop in cattle prices, one wonders if Ritland was not inadvertently guilty of economic shortsightedness.

At times the free market rhetoric borders on a callousness and paternalism more characteristic of economic writers in sympathy with active Government involvement in the economic system. For example, from an otherwise informative article on the Brazilian developmental experience, Marshall Martin and Julian Atkinson write:

> But progress came at the cost of individual and political freedom for many. On the other hand, one must assume that those who make up the lower socio-economic classes had little to lose and they are, in general, better off in terms of jobs, wages, and social services, including education.

Inclusion of this book on a supplemental reading list for an undergraduate economics class would help to balance those economic readings which advocate a more active Government role in the economic system. Some of the articles are quite good. Arthur Burns' treatment of the 1961-75 period as one long business cycle offers a contrasting interpretation to Dorothy Sherlings's and E. D. Chastain's view of the first half of this period as one in which Keynesian economics was emphasized and regarded as effective.

Thomas Humphrey provides a thoughtful treatment of the persistence of inflation and leads into Burns' paper. Hugh Macaulay and Bruce Yandles in "The New Competitive Structure of the American Economy," after struggling to prove the existence of competition in our current economy considering the product or brand level rather than the industry level, give an excellent discussion of the role of Government regulation in influencing the level of competition. John Lee provides a useful overview of economic forecasts and forecasting. These papers warrant thoughtful reading by both professional economists and students.

The paper I judge the best was curiously placed at the end of the book. In writing "The Potential For Economics Research," Clark Edwards has provided an excellent framework in which to strive for an objective, realistic view of economic topics, the goal which Messrs. Chastain and Ritland had set forth. The book is worth reading for this article alone.

In Earlier Issues

The importance of cooperation between statisticians and economists is emphasized. An economist should formulate his theories so as to permit measurement, and the statistician should develop "a desire to measure relevant things."

R. J. Foote
AER, Vol. IV, No. 1,
Jan. 1952, p. 23

FOOD AND SOCIAL POLICY, I: PROCEEDINGS OF THE 1976 MIDWESTERN FOOD AND SOCIAL POLICY CONFERENCE

Gary H. Koerselman and Kay E. Dull, editors, Iowa State University Press, Ames, 400 pages, 1978, $8.95.

*Reviewed by Thomas A. Stucker**

The Midwestern Food and Social Policy Conference held October 21-23, 1976, in Sioux City, Iowa, brought together public and private agribusiness leaders, academicians, farmers, and consumers. Many of the papers published in these proceedings were presented by agribusiness leaders.

Food policy and the context in which it functions are the major themes of the conference proceedings. They begin with articles on food and the future of civilization, and move through such issues as land tenure, domestic and international distribution of food, regulations and agreements affecting food distribution, and the priorities involved in maintaining a productive agricultural base.

Given this range of issues, the book represents an admirable overview. Short summaries of each chapter provide a linkage between the separate papers. The overview nature of the articles means that single issues are not treated in depth. Those who want more detail about specific components of the food system and factors affecting it should refer to sources with a more narrow, detailed focus.

The late Currier Holman, delivering the keynote address, expressed his belief that world hunger problems cannot be solved until food distribution is improved. He views the basic challenge to policy workers as that of understanding "interaction of the parts, which in general are the food industries, the governments, and the publics at large (*both domestic and foreign*). . . ."

The growing imbalance between supply and demand raises moral and ethical questions about food use and population growth. Holman views both specialists and generalists (integrators) as necessary in resolving food problems: "a solution to the food problem must encompass a multitude of technical disciplines, both broad and narrow, that together are as comprehensive as civilization itself." Politicians are generalists and synthesizers, and Holman urges that they be trained as professionals in these skills. Although this would certainly be desirable, reality dictates that other criteria enter into the political selection process.

Holman concludes on what he calls a pessimistic note, but he urges positive action. In his view, food is related directly and inextricably to petroleum. The current petroleum use rate cannot be sustained, which in turn implies future reductions in food availability unless the energy problem is solved. Holman admonishes that we should act now to resolve the worldwide problems of shortages of food and natural resources.

The proceedings, although not providing detailed solutions, do set forth an interesting perspective on food and social policy.

*The reviewer is an agricultural economist in the National Economics Division, ESCS.

In Earlier Issues

In a substantial number of fields in economics both theory and empirical research have reached, if not a dead end, at least the area of sharply diminishing returns. An attempt at radical departure from traditional forms of analysis should be welcome. . . . Before the findings of one discipline can become the tools of another, it often proves necessary to overcome a certain disciplinary chauvinism. . . .

Gustav F. Papanek
AER, Vol. III, No. 4,
Oct. 1951, p. 150

Attitudes of owners influence the use and management of land. All kinds of individuals and agencies own forest land. They keep it for different reasons, only one of which may be growing timber. Studies furnish knowledge about the people who own the land. This knowledge helps those who are responsible for administering policies of forest land to do a better job.

Adon Poli
AER, Vol. IV, No. 1,
Jan. 1952, p. 8

RESEARCH BROKERAGE: THE WEAK LINK

The process by which social science knowledge gets from the producer to the consumer is worth examining. The producer of social science knowledge rarely deals directly with the consumer. Two broad categories of intermediaries can be identified—academic and research broker. Academic intermediaries have a flair for interpreting the technical findings of their colleagues. After the findings are in the public domain, it is up to research brokers to present them for the use of policy makers. The research broker is invariably the crucial point of transmission. It is the most likely breaking point. It is also the point of leverage for getting repairs when the system breaks down.

In the Executive Office of the President is the model for all aspiring research brokers—the Council of Economic Advisors. In the executive departments, most of the older departments have research bureaus—the Department of Labor's Bureau of Labor Statistics, the Department of Agriculture's Economic Research Service,* the Department of Commerce's Bureau of Economic Analysis and Bureau of the Census, to name a few. These bureaus are some distance removed from the department heads so there is in most cases an individual who with a small staff serves as a link between the department's research organizations and its policy makers.

An example is the director of agricultural economics in the Department of Agriculture. The director is the research broker who is in the innermost circle in the Department's decision-making process. Backing up the director is the Economic Research Service,* which not only has a thousand researchers on its own payroll but also has links with a network of researchers in the land-grant colleges. Every department and bureau administering social or economic programs needs a research broker to serve in a capacity like that of the Department of Agriculture's director of agricultural economics.

---

*Editor's note—Predecessor agency of the Economics, Statistics, and Cooperatives Service.

Excerpted from *Knowledge and Policy: The Uncertain Connection* (1978), p. 126, with the permission of the National Academy of Sciences, Washington, D.C.

In Earlier Issues

ESCS is a unique institution as a body of social scientists in this town. It has a great opportunity to help broaden Government's role in food and agricultural policymaking. Not only can ESCS service USDA, but I can see a certain advantage in its keeping some distance between itself and the Department. Obviously, to remain part of the Department, ESCS must be responsive to the policymaking apparatus in USDA. Yet it cannot take as an excuse to build in isolation from reality in addressing relevant questions. It makes some sense for ESCS to have credibility as an institution, and that might call for a little bit of distance.

Lynn Daft, "Implications for Domestic Policy," *Agricultural-Food Policy Review: Proceedings of Five Food Policy Seminars*, Econ., Statis., and Coop. Serv., U.S. Dept. Agr., AFPR-2, 1978, p. 15.

Suggestions for Submitting Manuscripts for *Agricultural Economics Research*

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

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

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

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

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

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

U.S. MAIL

IMPORTANT NOTICE TO OUR READERS

Recently we sent each person on our complimentary mailing list a postcard or memorandum asking if you want to continue to receive the Journal. We asked that you indicate this by returning your mailing label, on which you were also to show any corrections.

Unfortunately, we are receiving back many postcards with no labels so that we have no idea which of you sent these postcards. Thus, we will be dropping from the list every name for which no label was returned.

If you recall that you sent back a postcard or memo *without* your label, write us immediately. Include your mailing label from *this* issue with any corrections.