Structural Econometric Models: Past and Future (with Special Reference to Agricultural Economics Applications)

by

William G. Tomek

Structural Econometric Models: Past and Future (With Special Reference to Agricultural Economics Applications)

William G. Tomek*

Introduction

Agricultural economists have often attempted to estimate the parameters of behavioral equations which derive from economic theory, such as demand, supply, and investment functions. These estimates were thought to be useful for a wide range of problems: analysis of proposed policies, evaluation of existing policies, forecasting, and improved understanding of the economy. We now know, however, that estimates of structural equations are not necessary for all of these applications. A time-series model may, for example, do a better job of forecasting than does a structural model. We also know that the introduction of a new policy may cause the structural parameters to change. If so, estimates based on the pre-change period are not applicable to the new policy.

But, for some problems, estimates of structural coefficients are important. To evaluate the effect of reducing food stamp expenditures on the consumption of food, an estimate of the marginal propensity to consume food out of food stamp income is needed. To evaluate the effect of changing the level of advertising expenditures for milk (or other generic commodities) on producer and consumer welfare, estimates of demand and supply relations are needed. Often, only observational (non-experimental) data are available at a reasonable cost. Thus, for some research problems, the question remains, can useful estimates of structural parameters be obtained from observational data?

In this context, this paper has three objectives: to review the historical contributions of the literature on structural models, to outline the problems of obtaining robust estimates of parameters, and to discuss how to improve the quality of results. The paper emphasizes applications in agricultural economics. The outline of the paper follows these objectives. I first discuss the nature of the contributions to the literature, but not the details of results. Then, I characterize the problems that have arisen in estimation. This naturally leads to a discussion of the consequences of the problems for empirical results and the impacts this has had on empirical econometrics. The final section looks to the future: how can we improve the quality of our work?

---

*Tomek is a professor in the Department of Agricultural, Resource, and Managerial Economics, Cornell University, Ithaca, NY 14853-7801. This paper is the written version of the 1997 Havlicek Memorial Lecture in Applied Econometrics, Ohio State University; the lecture attempts to synthesize the wisdom that has accumulated from building and estimating structural models over the 1960-1997 period. H. Hanawa, B. F. Stanton, and L. S. Willett provided helpful comments, but of course, I am responsible for the opinions expressed.

Historical Contributions

Econometrics circa 1960 had been heavily influenced by the work of the Cowles Commission (e.g., see Christ). Time-series observations generated by the economy can be viewed as stochastic, dynamic, and (often) simultaneously determined, and the Cowles Commission used an explicit probabilistic framework to consider the data generating processes for the current endogenous variables. Their contributions start from the distinction between endogenous and exogenous variables. The current endogenous variables are viewed as being generated by a system of (usually) simultaneous equations. The exogenous variables, which help determine the endogenous variables, are themselves determined outside the system of structural equations. That is, the exogenous variables are not modeled.

A system of simultaneous equations can be written in either structural or reduced form. The structural form incorporates all of the (presumably correct) restrictions on the model. In this context, the definition and solutions of the identification problem were developed, as were alternative estimators for the structural parameters. A parallel literature emphasized distributed lag models as an explanation of the dynamics of endogenous variables (for a literature review of that era, see Griliches).

Another strand of the literature dealt with systems of linear regression equations in which each equation is assumed to have only one current endogenous variable, i.e., are "seemingly unrelated." But, the equations are linked via correlations among the error terms of the different equations and/or via restrictions on the parameters (e.g., Zellner). This literature has been important in agricultural economics *inter alia* as a basis for estimating systems of demand or cost-share equations, which impose adding-up and symmetry restrictions on the parameters of the model (Capps and Havlicek; Johnson, Hassan, and Green; Paris, Caputo, and Holloway).

Still another important component of the literature relates to limited dependent variable models (Maddala). Probit and logit models have enjoyed widespread use in agricultural and resource economics. Generalized logit models of demand systems have been shown to work well in comparison with other popular models, like the Almost Ideal Demand System (e.g., Tyrrell and Mount; Weng and Mount). Multinominal logit models are useful in analyzing individuals' choices among a number of qualitatively different alternatives, like modes of transportation (for a discussion of model specification issues, see Kling and Thomson). This literature also includes models of censored and truncated variables, and endogenous variables are sometimes viewed as being generated by markets that are in disequilibrium (e.g., an application to the watermelon market by Goldfeld and Quandt). Thus, the literature collectively provides a rich array of ways of viewing the processes by which the observations on endogenous variables are generated.

The largest area of applications of the various models by agricultural economists was, however, the early adoption of the Cowles framework to the empirical analyses of agricultural product markets. But, Fox hypothesized that many supply-demand models for agricultural commodities could be treated as recursive systems. Current production is determined by prior decisions, and quantity consumed is essentially equal, Fox argued, to the (predetermined) quantity produced. The quantity, in turn, determines the market-clearing price in a recursive sequence.

The predetermined supply is sometimes modeled as being allocated among competing uses, so that the quantities utilized are simultaneously determined with the prices of the respective uses. For example, at harvest-time, the predetermined production of apples can be allocated to use as fresh fruit, to a variety of processing uses, and to storage for future use. Thus, a model for a commodity may have a simultaneous block of equations even if production is determined by decisions made in prior months or years. Many of the contributions of agricultural economists relate to specific ways to model the unique features of agricultural supply and demand, including the modeling of price expectations, risk, and distributed lag effects (for more detail and references, see Tomek and Myers).

The numerous empirical results constitute another contribution of econometrics-related research by agricultural economists. We have estimates of many different elasticities based on single equations (which may implicitly be part of a larger system), simultaneous or recursive supply and demand models, or from systems of demand equations which impose the restrictions suggested by theory. These models were fitted to a wide variety of data sets, sometimes emphasizing retail-level demand for finished goods and sometimes emphasizing farm-level derived demand relations.

There has been relatively little winnowing of results and models. Consequently, a diverse range of elasticity estimates exist for any particular commodity. And competing model specifications appear to fit the historical data about equally well. It has also become increasingly clear that particular results are sensitive to the model specification, to the sample period (data set) used to estimate the model, and to the estimator used in the analysis. (Evidence is discussed below, in the consequences section.) Moreover, the reported results are typically a consequence of much experimentation with the model's specification using a fixed data set, i.e., pretesting. This implies that the type I error for hypothesis tests, using the final results, are much larger than the nominal levels (Wallace).

Another contribution of the literature, therefore, is the increased appreciation of the difficulty of obtaining stable, useful results when estimating structural equations from observational data. Nonetheless, agricultural economists have not, in my opinion, been at the forefront of developments to improve the quality of econometric results. This literature, among other things, provides tests for model misspecification, for stationarity and for exogeneity of variables, and discusses the consequences of pretesting. But, researchers in agricultural economics have not been early adopters of the ideas in this literature. Evaluations of results continue to be rather shallow, and although the consequences of pretesting have been well-known for over 20 years, regression results are still reported as statistically significant at the one, five, or ten percent level, when in fact the true level of type I error is unknown.

Problems in Empirical Econometrics

Framework for Modeling

The Neyman-Pearson framework implicit in the Cowles Commission approach to econometrics turned out to be an issue in doing empirical econometrics. This framework places heavy weight on the correctness of the model--the maintained hypothesis--since the properties of the estimators and hypothesis tests are conditional on the model. The Commission's econometricians assumed that researchers could correctly classify variables as endogenous and exogenous and could correctly specify the restrictions on the model. The latter assumption means that each equation of the model contains the relevant variables (those with non-zero parameters) and that each equation correctly excludes irrelevant variables. The exclusion restrictions in a simultaneous system of equations have been the principal way of determining the degree of identification, and a simultaneous equations estimator will not be a consistent estimator unless the equation(s) being estimated includes all relevant variables. In other words, the traditional approach in econometrics assumed that the true data generating process is known and correctly modeled. Thus, emphasis was placed on obtaining optimal estimators under the assumed data generating process.

This carried over into empirical applications in agricultural economics. Much published research has emphasized the selection of the estimator assuming the model was correctly specified. It has been rather uncommon to discuss model selection procedures or to go beyond a rather superficial evaluation of the quality of results. The classification of variables as endogenous and exogenous has typically gone untested; if "t" ratios are large, signs of coefficients logical, etc., the model has been assumed to be adequate.

In the typical research situation, however, the correct model specification is uncertain. The analyst has a research problem, observational data, and has access to relevant theory and previous research. But, there are often competing theories, and published results can differ substantially. The true data generating process is unknown, and numerous plausible representations of the data exist (Hendry, pp. 16f.). Consequently, there is uncertainty about the appropriate inference methods, and an estimator, which is optimal for one conceptualization of the data generating process, may be seriously biased in practice.

The problem is even worse for forecasting. In this case, the model that was assumed valid for the sample period must be assumed to remain valid for the forecast period. Moreover, ancillary forecasts of the exogenous variables are required in order to forecast the endogenous variables. As already noted, the exogenous variables are not explicitly modeled in the typical structural model, but they are usually stochastic variables with unknown data generating processes. A structural model is potentially most useful for forecasting when a major change occurs in one of the explanatory exogenous factors, such as a large reduction in production. But, this is precisely when estimated models have failed.

There are two problems. One is the quality of the estimates of the parameters. The event being analyzed may have caused a structural change; the parameters do not remain invariant in face of the change. Thus, estimates based on the historical time series are not relevant to the new conditions. Even if the structure has not changed, the historical sample may not have provided adequate variability in the explanatory variables to obtain precise estimates of the relevant parameters. Also, a linear functional form which seemed adequate for the historical sample may turn out to be wrong in light of the larger range of variability in the data.

The second problem is that it is often impossible to forecast the magnitude of the realized change in the explanatory variables. The high grain prices in Spring 1996 were not forecast accurately the prior summer because the relevant levels of supply and demand factors were not accurately forecast. The actual crop size was smaller than expected, and use of this crop turned out to be larger than expected. Thus, ending stocks for the 1995-96 crop year for corn and soybeans were small.

Clearly, forecasting (and analysis of proposed changes in economic policy) make more demands on the model than does obtaining the best estimate of structural coefficients for some historical period. Not only is the true model unknown, but the definition of the correct model depends on the research problem. I expand on these points below, and also see Bessler and Covey for additional discussion.

Modeling Problems

As already noted, early applications in empirical econometrics assumed that variables could be correctly classified as endogenous and exogenous. Subsequent research indicates that such classification is not easy or obvious. If the research problem involves inference about structural parameters, then the explanatory variables need to be only weakly exogenous, i.e., predetermined. For forecasting, the variables must be strongly exogenous, and for analyzing future policy changes, the variables must be super exogenous (Engle, Hendry, and Richard). Statistical tests exist for various types of exogeneity, but they are not commonly used in agricultural economics.

The alternative notions of exogeneity in the literature suggest that the researcher needs to specify, at least implicitly, models of the exogenous variables. For example, to be weakly exogenous, a variable must be generated so that its contemporaneous covariance with the error term of the structural equation is zero.

Two inter-related, difficult problems are constancy of the structure over the sample period and correctness of the model specification (Alston and Chalfant). A correct model is understood to be one that includes all relevant variables (those with non-zero parameters), that has a correct functional form, and that excludes irrelevant variables. The typical model specification also assumes that the parameters of interest are constants over the sample period used for estimation. One can test for parameter constancy, but such tests are conditional on the correctness of the remaining model specifications. Or, one can test the correctness of selected model specifications assuming the parameters are constant over the sample period.

Specification of relevant variables includes how to model possible distributed lag effects. For time series data, lagged effects are likely to be important, but there are few definitive guides for modeling the length and form of the lag structure. If the analyst fits a relatively comprehensive, unrestricted model, the estimates are likely to have small bias, but large variances. If the analyst fits a relatively restricted model (by excluding variables, by restricting the form of the lag, etc.), bias is increased while the variance of the coefficients is reduced. The trade-offs in a mean squared error sense are difficult to evaluate.

Restrictions on the model also affect identification. The Cowles Commission approach assumed that numerous variables in the system would be excluded from individual equations. Hence, each equation would be highly over-identified. While Liu suggested that this was not true in 1960, it wasn't until Sims' 1980 paper that the credibility of over-identifying restrictions was seriously questioned. Modern macro-economics suggest that most structural equations are under-identified, but nonetheless Keynesian-type macro-econometric models are still used for analyses of the economy. Likewise, structural models in agricultural economics continue to be fitted under the assumption that the equations are identified, i.e., that the identification restrictions are correct. It is uncommon to test for the validity of identification restrictions.

Joint tests can be conducted for hypotheses related to model specification (nicely summarized in McGuirk, Driscoll, and Alwang and in McGuirk, et al.). For example, a researcher might test jointly for possible omitted variables and parameter constancy, the joint problem noted earlier in this paper. But, ultimately, the tests are conditioned by some minimal specification, which is assumed to be correct.

Data Issues

Errors in variables is a potentially serious problem in empirical econometrics. Errors may be made in recording, compiling, and manipulating data, but care in data handling can minimize this problem. A more difficult issue is that observed variables are not good measures of the underlying economic concept. For example, expected output of an agricultural commodity is a function of prices expected to prevail at harvest-time, but how do farmers form expectations? Empirical models have used notions of naive expectations, adaptive expectations, quasi-rational expectations, and rational expectations. In empirical analyses, rational expectations must be specifically defined in terms of the model decision-makers are assumed to use in forming expectations. The limited evidence available suggests that empirical results are sensitive to the definitions of expectations. In an analysis of fed cattle supply, for example, current quotes of prices for future delivery gave quite different estimates of the supply elasticity than an alternative definition of a rationally expected price based on a particular model (Antonovitz and Green). Since risk is defined as a function of the difference between expected and realized outcomes, the appropriate empirical definition of risk is unknown.

Researchers must understand how the data series are constructed, so that they may be used appropriately in their research. Defining the data generating process for variables is not only a theoretical, but also an empirical issue. An important pre-condition for developing proxy variables for unobservable concepts is to understand the precise definition of the observed variables. Regrettably, some analysts are not careful, and therefore do not understand how data series are constructed.

For instance, domestic feed use of grains in the United States is derived from balance sheets, using estimates of beginning inventories, production, etc. Since feed use is measured as a residual in this computation, it necessarily includes the aggregation errors contained in the various components of the balance sheet (Gardner). Thus, measured feed use may be seriously in error relative to actual feed use, and if feed use is treated as a right-hand side explanatory variable in an inverse demand function, a serious errors-in-variable problem exists.

The marketing margin for beef is based on particular price series for choice grade beef at the retail and farm levels, and is not a margin for all beef sales. It appears that this margin behaves differently than a more comprehensive measure of margins for beef. The transmission of farm prices to retail prices appears to be faster for the comprehensive measure (Tomek, 1996). These and other examples suggest that understanding the specific definition of data series is important.

Observational data from government sources are often subject to revision. The researcher needs to assure that the data are internally consistent over the length of the sample period being used in the analysis. Also, the data should be carefully documented, so that subsequently an analyst could confirm the research. Perhaps the biggest problem in attempting to confirm the results of earlier published research is the inability to duplicate the data set used in the analysis (Tomek, 1993). Many revisions can be viewed as reducing errors in variables. If, for example, an input into the balance sheet used to estimate feed consumption is changed, then presumably the new estimate of consumption is more nearly like actual consumption (for more on the conceptualization of data revisions, see Tomek, 1993).

Analyses of time series are often plagued by collinearity of the variables. We shouldn't be surprised that, say, the prices of substitutes are correlated or that some consumption series have common seasonal components. In my view, collinearity is one the most misunderstood and difficult problems in empirical econometrics. Data for a particular set of variables have a given level of independent variability. This cannot be changed by transformations. Of course, transformations can change the correlation among regressors; the ultimate transformation is to orthogonalize the variables. But, the inherent independent information content of the sample is unchanged. Collinearity is analogous to having a small sample (Goldberger). A lack of an informative sample is regrettable, but can't be changed.

Thus, solutions to collinearity must involve additional information: either additions to the sample or the use of non-sample information, which may take the form of restrictions on the model. The non-sample information may be in a classical or Bayesian framework. If restrictions on the model are used as a solution to collinearity, then the trade-off, discussed earlier, arises; namely, restrictions can reduce the variance of the coefficients, but likely at the cost of increasing the bias. Whether or not the restricted estimator is better in a mean squared error sense is difficult to tell. Clearly, there are no easy solutions to collinearity.

Another question is, are time-series variables stationarity? Can we assume that the mean, variance, and auto-covariances (for fixed time lags) are constants? The usual large sample theory used to appraise least squares and instrumental variable estimators assumes that the variables are stationary, but some may not be. Numerous tests for stationarity exist. A common set-up is to test the null that the variable is once-integrated, i.e., has a unit root (Dickey and Fuller). The usual tests have low power when the root is near one, but is not in fact one.[1] Thus, the null hypothesis cannot be rejected unless strong evidence exists against it. Consequently, much uncertainty attaches to whether the time series commonly used in agricultural economics are stationary or not.

Consequences of Problems

This section considers two kinds of consequences. One relates to the impact of the problems on empirical results. The other major consequence is the influence of the various difficulties on the development of econometric tools and on how empirical research is conducted.

Empirical Results

Many econometric results are fragile. Estimates of parameters are sensitive to changes in the sample as well as to changes in the model specification. My first experience with this problem occurred almost 30 years ago, when I decided to refit a supply of milk equation to illustrate alternative models in class. The alternative model required an additional lag, and I found that when the first data point was dropped, the coefficient of price changed from positive with a large t-ratio to negative with a small t-ratio. It turned out that the "significant" price elasticity of supply in the original model was dependent on the first data point. Subsequently, I found that the distributed lag effect that Nerlove had obtained for cotton supply disappeared if a plausible change was made in the model specification (Tomek, 1972).

Since that time, much evidence has accumulated about the sensitivity of results to changes in model specification or data. Estimates of the marginal propensity to consume out of food stamp income have varied from 0.17 to 0.86 (Fraker). The estimated price elasticity of supply for fed cattle ranges from -0.84 to 0.71 depending on the definition of expected price (Antonovitz and Green). Alston and Chalfant, as well as Eales, have attributed the diverse estimates of demand parameters for meats (and of whether or not structural change has occurred) to model misspecification.

Diebold and Lamb point out that the estimates of supply response for many agricultural commodities, using a geometric form distributed lag model, have been highly variable. They argue that part of this variability may be attributed to using inappropriate estimators, given the lagged dependent variable in the model specification. Likewise, Baltagi and Griffin found that the estimates of the parameters of the demand for gasoline, in a distributed lag specification, varied widely with the estimator used. The long-run price elasticity of demand ranged from -0.30 to

-1.42. In their case, the estimators were related to alternative models for pooled time-series, cross-section data.

The variability of results can be further illustrated using Nerlove and Waugh's classic study of the effect of advertising on the demand for oranges. Their research used the sample period 1909 to 1958, omitting the years 1941-1945, inclusive. They used a relatively simple specification, in which the explanatory variables are quantity of oranges sold, consumer income, current advertising expenditures, and average advertising expenditures for the previous 10 years. The dependent variable was farm-value of sales. All variables are on a per capita, deflated basis, and are in logarithms. The data are published in the article.

It is possible to exactly confirm the published results. An additional result is that the Durbin-Watson statistic for their analysis is 1.264. (It was uncommon to test for autocorrelated errors at the time they did their research.) The residuals are autocorrelated, and today this would be taken as a sign that the model may be misspecified. I did not, however, explore alternative models; rather, I considered the effects of alternative sample periods on their results. Selected results are summarized in Table 1. The coefficient of quantity is moderately sensitive to the deletion of the first data point, 1909, changing from -0.39 to -0.48. Among the alternative sample periods explored, I found estimates of the quantity coefficient that ranged from -0.27 to -1.22 and estimates of the current advertising coefficient that ranged from 0.38 to -0.04. Some coefficients have small t-ratios. Whether or not advertising was "significant" is dependent on the sample period used in the analysis.

Since evidence of autocorrelation existed, I reestimated the model by a Feasible Generalized Least Squares estimator, using the 1910-58 sample. The estimated autocorrelation coefficient is 0.336 with a t-ratio of 2.199. The coefficient of advertising is 0.287 for OLS and 0.220 for FGLS, and while this is a small absolute change, it is over a 23% change. (Given the standard errors, the confidence intervals overlap.) The coefficient of quantity changed from -0.494 to -0.376, over 25%. In my view, this is a "moderate" change in response to using an alternate estimator.

The Nerlove-Waugh paper was, on the one hand, a path-breaking and hence an award-winning paper. It was a good first attempt to estimate the effects of advertising for a generic commodity. But, on the other hand, it would not be publishable by today's standards. We would not believe that such a simple model is an adequate representation of the data generating process for oranges over a 50 year period. That is, it is highly unlikely that the parameters can be treated as constants for such a long period in the context of the relatively simple model specification.

An extension of the concept of fragility is the notion of spurious results. Not only are the estimated coefficients likely to have serious biases, but so are their estimated variances. Moreover, if both the dependent and an independent variable have unit roots, the distribution of the ratio of the regression coefficient to its standard error is not standard, i.e., does not have the "t" distribution. Thus, erroneous conclusions can be drawn from hypothesis tests. In the extreme case, two variables which have no relationship may be found to have a "significant" relationship, a spurious regression.

Table 1.

Alternative Estimates of Selected Parameters of Model of Advertising Oranges
(Nerlove and Waugh)

Alternative samples	Coefficients of: log Q	log A
1909-58, OLS	-0.390	0.233
(41-45 deleted)	(0.198)	(0.125)
1910-58, OLS	-0.476	0.273
(41-45 deleted)	(0.209)	(0.132)
1910-58, OLS	-0.494	0.287
(42-45 deleted)	(0.202)	(0.126)
FGLS	-0.367	0.220
	(0.204)	(0.142)
1936-58, OLS	-1.215	0.278
(42-45 deleted)	(0.404)	(0.195)
1938-58, OLS	-0.963	0.380
(42-45 deleted)	(0.517)	(0.240)
1946-58, OLS	-0.275	-0.045
	(0.617)	(0.325)

Comments: All variables are per capita, deflated series and are in logarithms. Q is based on sales of oranges; A is based on advertising expenditures. The dependent variable is per capita farm value of sales in logarithms. Estimated price elasticities ranged from -0.45 to -0.78. Standard errors of coefficients in parenthesis.

These problems are compounded by pretesting. Since it is essentially impossible for an economist to specify "the correct" model, it is inevitable that the researcher will use the data to search for a preferred model. Such a search can prevent egregious errors in model specification, but a consequence of pretesting, as noted earlier, is that the probability of type I error is far larger than the nominal levels used in "t" and "F" tests. Naturally, the researcher will be searching for "significant" results and will select models that have large test statistics. The results may simply represent the peculiarities of the data set.

An additional consequence of the problems of model specification and estimation is that research results have not been very cumulative. Researchers emphasize the novel features of their model and results. Typically, new results differ from the old, but the comparisons made (if they are made at all) are superficial. A researcher may argue that her/his results are preferred, but not provide a fundamental analysis of the differences. As I have argued elsewhere (Tomek, 1993), it is essential to confirm the historical research before comparisons can be made. The results of the old model are likely to have changed if applied to new, revised data. Moreover, the earlier work may contain errors. Thus, to make results cumulative, one must fit the old and new model to both the earlier and current data sets. This is the only way one can hope to tell why results differ.

Of course, such a research strategy takes time. In fact, it may be impossible to confirm the earlier work. But, as I shall argue below, making empirical research more cumulative requires greater depth of analysis than has been common in the past.

Impacts on Econometrics

The realization of how fragile empirical results can be grew dramatically in the 1970s, and this had a major impact on econometric methods and methodology. *First,* it stimulated developments in other topics than estimation of structural models. The most notable is the vast expansion of the literature on time-series econometrics. This literature includes developments in specification, estimation and hypothesis testing for autoregressive moving average (ARMA) equations, vector autoregressions (VAR), the "state-space representation" of a dynamic system and Kalman filters, models of heteroscedasticity, particularly autoregressive conditional heteroscedasticity (ARCH) and its extensions, and models of univariate and multivariate time series with unit roots and of variables that are possibly cointegrated (Box and Jenkins; Hamilton).

Vector autoregression models were seen by Sims (1972) and others as an alternative to structural models. VARs have been used for forecasting and to some extent for policy analyses. Granger's causality test can be seen as a two variable VAR. Hence, VAR models can be used to test theories which suggest, say, that one variable in the system should fail to Granger-cause another. Theory may also suggest hypotheses about the time-series properties of a variable, or alternatively theorists may seek to explain empirical results obtained from VARs.

While econometricians have tended to divide into time-series and structural camps, developments in the time-series literature complement structural modeling. As just suggested, time-series models are an alternative approach to some problems, such as forecasting. In addition, the literature on integrated and cointegrated time series and on conditional heteroscedasticity has helped improve structural models. Error correction models, for example, are potentially important specifications for dynamic structural equations.

A second major impact has been on methodological approaches to doing empirical econometrics. The development of time-series econometrics can be viewed as one such development, but I prefer to think of the time-series literature as complementing the structural modeling literature. Within the structural modeling literature, however, several different philosophies have been developed (for an in-depth discussion, see Darnell and Evans). One can be associated with Bayesian and the other with the classical statistics literature.

The Bayesian approach to empirical econometrics is perhaps most associated with the work of Edward Leamer. He argues that because classical statistical methods were developed for use with experimental data, they are inappropriate for estimating structural models. The Bayesian approach does not offer any "Golden Rules" about integrating model uncertainty with data uncertainty, but the hope is that uncertainties about model specification can be formalized in the prior densities (the non-sample information) used to estimate the model. In other words, the researcher is forced to formalize her/his prior knowledge before estimating the model.

Leamer also proposes an "extreme bounds analysis" as a way of reporting results. He argues that researchers should explore the effects of the entire range of plausible prior information that might be used in the model. This is a type of sensitivity analysis. For example, if the research objective focuses on a particular parameter, the analyst should report the range of estimates of that parameter, which have been obtained from different model specifications that incorporate the alternative versions of the prior information. In a sense, Leamer is just urging researchers to recognize the uncertainty about model specification and to provide a more honest report about the process that led to the final, reported equation.

If the bounds found for the coefficient are narrow, the result is pronounced robust, and the researcher can have considerable confidence in the result. If the bounds are wide, the result is fragile, and the researcher must address whether it is possible to obtain a more robust result. At a minimum, Leamer is arguing for a systematic rather than an ad hoc approach to modeling.

A second methodological approach to structural modeling is called "general to specific modeling." David Hendry and his colleagues (e.g., Davidson, et al.) are strong proponents of this approach, which is undertaken in a classical statistical framework Here, the researcher starts with a general autoregressive distributed lag (ADL) specification that (hopefully) contains the simpler, but correct model of the phenomenon being analyzed. Consequently, one can conduct a sequence of tests of hypotheses nested within the general model in order to obtain the "final," simpler specification. A clear example is to start with a specification of lag length of "m" which is larger than the logical lag length "n," and then test down to find "n." By having a logical sequence of tests, it is possible to control the level of type I error in the final test. The ADL specification also helps deal with potential non-stationary variables that may be cointegrated.

In this approach, the "final" model is subjected to a series of tests of model adequacy, which are intended to act as quality control devices. The proposed tests are far more comprehensive than the usual examination of Durbin-Watson statistics, t-ratios, and coefficient signs, that were the historical practice in agricultural economics.

The proponents of the Leamer and Hendry approaches have written some informative and entertaining papers critiquing the alternative methodologies (e.g., see McAleer, Pagan, and Volker versus Cooley and LeRoy). It is not possible for me, in this limited space, to summarize all of the arguments (for one summary, see Darnell and Evans). My main point here is that the failures of empirical econometrics have stimulated much thinking about how to do empirical econometrics. These approaches stand in contrast the Cowles Commision methodology, which could be characterized as "specific" modeling in a classical statistical framework. As noted earlier, the Cowles Commission approach placed great faith in the econometrician's ability to use economic theory as a basis for writing down the correct model. An extensive search for the final model was not contemplated. When models failed, as they often did, then researchers tried to "generalize" their models, usually in an ad hoc fashion.

One way to summarize the extant methodological approaches to structural modeling is in terms of a two by two matrix. On one axis, we can think of Bayesian and classical statistical approaches. On the other axis, we have "general to specific" versus "specific to general" methods. I would classify the historical or Cowles approach as a "specific to general classical" approach, while the Hendry methodology is a "general to specific classical approach." Leamer sees his approach as "specific Bayesian." It is not my purpose to recommend one approach over another, but I do think that it is important that the researcher take a systematic approach to modeling. A mindless, ad hoc specification search is not acceptable. If the researcher can not reconstruct how the final model was obtained and cannot characterize the procedures to readers, then a problem exists.

(This literature contains an important message for referees of journal manuscripts. Namely, it is proper to criticize ad hoc specification searches, inadequate discussion of how the final model was obtained, and incomplete evaluations of results. It is not proper, in my judgment, to recommend rejection of a paper that has used a clear, logical procedure simply because the referee prefers an alternative. That is, those wedded to Leamer's methodology should not prevent the publication of work based on Hendry's methodology. Both, at our current level of understanding, are acceptable. What is not acceptable is the lack of any clear methodology.)

The third major impact arising from the poor performance of structural econometric models is on econometric methods, especially the development of a large number of statistical tests of model misspecification (e.g., Godfrey; MacKinnon; McGuirk, Driscoll, and Alwang). This literature includes tests for omitted variables and incorrect functional forms, tests for autocorrelation, heteroscedasticity and non-normality of the errors, and tests for the constancy of regression parameters. There is also a large literature, already mentioned, on the exogeneity and stationarity of variables. In addition, methods exist to find outliers and observations with large influence on the results. Tests have been developed in the context of linear regression models, simultaneous equations models, and qualitative and limited dependent variable models.

The proliferation of tests relates, in part, to the large number of different data generating processes (models) to which the tests might be applied. Theoretical econometricians try to determine the statistical size and power of tests under varying circumstances. Our understanding of preferred tests evolves slowly however, and in empirical applications, the tests used are influenced by their convenience. It has been long known, for example, that the Durbin-Watson test for first-order autocorrelation is seriously biased when the model contains lagged endogenous variables. But, this test statistic continues to be reported simply because it is routinely computed by econometrics software packages.

Typically, the null hypothesis is that a particular specification is correct. If the null is rejected, the interpretation of the alternative may be ambiguous. For example, if the null hypothesis that the errors have zero first-order autocorrelation is rejected, it does not necessarily follow that the true errors have first-order autocorrelation. The rejection may instead be a sign of omitted variables or perhaps of higher-order autocorrelation. In general, rejection of a null tells the researcher that something is wrong, but usually not what is wrong. Nonetheless, the development of a wide variety of tests of model adequacy represents an important advance in evaluation of empirical results.

Finding that variables are not stationary, that models are inadequate, and so on has influenced the development of new models and estimators. A notable example is the literature on cointegration and error correction models. It is quite plausible that two (non-stationary) variables have a stable long-run relationship, but that short-run departures from the long-run equilibrium exist and that this behavior needs to be captured by the model specification (by an error correction mechanism). Although the error correction model is just one particular specification, it encompasses a number of commonly used models in the literature, such as partial adjustment and finite distributed lag models (Hendry, Chapter 7).

Given the diversity of plausible models, the literature on alternative estimators continues to grow as well. Naturally, this literature attempts to find the preferred estimator for each of the alternative models. If the error term is heteroscedastic but the researcher is uncertain about the form of the heteroscedasticity, what is the preferred estimator (e.g., White)? If one observation seems to have "too large" an influence, how should it be trimmed (e.g., Huber)? If we have errors in variables and potentially "weak instruments," what estimator should be used (e.g., Staiger and Stock)? And so on. Authors also continue to try to integrate the individual estimators into a logical whole (e.g., Bates and White).

My summary impression is, however, that the effect of poor model performance on the development of methods has been greatest for hypothesis testing. This seems true in two senses. One is that a deeper appreciation has developed about the potential problems of using conventional "t" and "F" tests for hypotheses about model parameters. While it is natural to prefer a large "t" value to a small one, specific conclusions from a test may be seriously in error. The second sense is that we have a far larger array of tests for model misspecification. The issue of minimizing model misspecification is relatively more important in the empirical econometrics literature than it was 20 years ago. Thus, greater emphasis is being placed on misspecification testing, though in my view, this transformation has occurred rather slowly within agricultural economics

Looking to the Future

In this section, I cover some of the necessary conditions for obtaining robust estimates of structural parameters, and I try to do this in a forward looking manner. What have we learned from the past that can help improve results in the future? The emphasis is on various econometric topics, but I also comment on the incentives for doing high quality empirical work.

Technical Improvements

First, it is important to define the research problem precisely. What is the focus of the research? Does the focus require estimation of one or more key structural parameters? Problem definition not only helps define the appropriate model, but also the evaluation of results. Some published research has not had a precise focus. For example, many studies have addressed the question, has the demand structure for red meats and chicken changed? Conflicting answers exist, and this is likely because the answer is conditioned by the model used to test the hypothesis (Alston and Chalfant). That is, conditional on model A, the structure is found to have changed, but conditional on model B, the structure is found not to have changed. One interpretation is that both results are correct and that the general question (has the structure changed?) is not stated precisely enough.

Why are we interested in testing for structural change? Is it to find a preferred forecasting equation? Is it to ascertain whether such factors as life style, nutritional education, or changing age distribution have affected demand? Or? If the problem is to test for "health effects" or "age distribution effects," then the model must be specified to capture these effects (McGuirk, et al.). It is not possible to specify the globally correct structural model, and testing for structural model is conditioned by the intended use of the model.[2]

Second, the analyst must ascertain whether the observational data available for analysis are adequate for answering the research question. As already discussed, the data may be inadequate in several senses. The observed variables may not be a good measure of the underlying economic concept, and/or no data exist for the concept. Then, one must decide how to define proxy variables or to determine the consequences of omitted relevant variables. If the resulting estimates of the key parameter to answer the research question are fragile because of such problems, the conclusion may have to be that a precise answer is not possible.

In looking to the future, two opposite trends seem to be occurring relative to data. One regrettable trend is that budget pressures are adversely influencing the quantity and quality of data collected by the federal government. This seems particularly true when one considers that the economy is becoming more complex while relatively fewer resources are being put in the federal statistical system. On the other hand, data from private and experimental sources are becoming increasingly available. Retail stores are generating vast quantities of scanner data, providing observations on transactions for many individual items. Trade associations and other private groups collect data or commission surveys. Economists are also generating data via formal experiments. These types of data have limited, but specific uses. For example, it is sometimes possible to test for the effect of advertising in particular cities, using location-specific data and not be limited to broad regional or national aggregate analyses (e.g., Liu and Forker). It is important that we look for relevant data sets from new sources as ways to solve problems.

Third, given a specific research problem and relevant data, the researcher needs to adopt a logical philosophy of modeling and of model evaluation. This starts with a careful study of the literature. What alternative, plausible theories exist in the literature, and how does the proposed model relate to this literature? Are alternative hypotheses nested (or not nested) within the proposed model? In some cases, it may be important to confirm selected results in the literature. Can they be duplicated? If so, how sensitive are these results to changes in the sample? Having started with a comprehensive specification, is a logical process of model simplification and evaluation followed? Once a tentative model has been obtained, can it pass a battery of tests of model adequacy? If the model is claimed to be better than alternatives in the literature, can this be demonstrated?

A complete evaluation would include using old data sets from previous research, current data, and ultimately new data. The best test of a model is its performance with new data. Thus, model evaluation should include replications, i.e., refitting the model to new data. To make research more cumulative, it is also important to compare alternative models using new data. Of course, if the data are observational time series, one must wait for time to pass for such analyses to occur. But, it would be possible to develop a program of research that went back into the literature to see how existing models perform with new data. At present, it is uncommon to try to duplicate published results and then use new data as a type of replication.[3]

The foregoing suggestions imply a depth of analysis that has been uncommon in agricultural economics. Doing high quality empirical research is time-consuming, and moreover the pay-off from this effort is uncertain. After much work, the researcher may still have fragile results.

I would like to think, however, that there is a place in professional journals for negative or less-than-perfect results. Failures or partial-successes can assist future research. Over 20 years ago, a researcher reported that capital punishment was a statistically significant factor in deterring murders. Subsequently, other researchers could not confirm this result (Passell and Taylor). It may be exceedingly difficult to explain the variability of the murder rate via a structural model, and it is important to know whether estimates of the coefficient of the capital punishment variable are fragile or not. If they are fragile, then the next step is to understand why.

Incentives and Teaching

Unfortunately, the incentive system in academia does not encourage in-depth, evaluative empirical analyses. The highest rewards go to those who publish large numbers of papers and to those that seem to have innovative papers. The emphasis on quantity, of course, encourages a rush to publication that probably will not involve a confirmation of prior work, careful evaluation of data, or thorough evaluations of model results. After all, a comprehensive evaluation may find that the model can not past the tests of adequacy.

With respect to innovation, the incentive is to propose a new model and to test it with a readily available data set. Model revisions are naturally biased toward selecting results that are consistent with researcher's pre-conceived notions. The marginal benefit to the researcher of a careful comparison of models is likely small. The alleged innovation might not be better if an in-depth analysis were undertaken.

Thus, academia needs to revise its criteria for promotion. Greater weight must be given to quality than to quantity. Allegedly innovative models need to be held to a higher standard of evaluation. Imperfect results can contribute to knowledge. As noted earlier, it helps to know whether results are robust or fragile.

The way graduate-level courses in econometrics are taught contributes to how empirical research is done. As noted above, the traditional approach to econometrics was to expect theory and logic to help specify the correct model; then the econometrician is just concerned with the questions of estimation and hypothesis testing. Also, it is natural to want students to have a strong foundation in statistics. Thus, textbooks and courses start with a heavy emphasis on the classical linear regression model, and of course if the data were actually generated in this way, the ordinary least squares estimator is the best unbiased estimator. This is often followed by a discussion of the generalized model in which the true error terms are assumed to be autocorrelated and/or heteroscedastic. If the true variance-covariance matrix is known, then generalized least squares is the best unbiased estimator.

While these topics are necessary background, they emphasize models that are usually unrealistic in economics. Students are perhaps led to false perceptions about the ease of building correct models or about the likelihood that estimators are unbiased. Models are far from perfect. Regressors are mostly stochastic, not non-stochastic. For time-series data, dynamic models with lagged endogenous variables are likely to be important. The true autocorrelation coefficient is unknown, and indeed the most likely reason for autocorrelated residuals is model misspecification, not autocorrelation in the true errors. And so on.

Graduate students must be exposed to the problems associated with model uncertainty and with stochastic regressors. This in turn requires discussion of evaluation of data and models. Courses also need to deal with econometric methodology, such as general to specific modeling. In addition, practical topics, like use of models for forecasting and policy analysis, should be addressed in courses (Granger).

All of this takes time, and I don't have a perfect solution for how to cover all of this material in two or three semesters. A preferred solution would be for students to have obtained better foundations in statistics as undergraduates, so that graduate courses could contain the in-depth material essential for high quality empirical research. Since this is unlikely, part of the teaching must occur via individual mentoring of students by faculty. Faculty have an important responsibility to convey the requirements of in-depth scholarship in applied economics.

Conclusions

My conclusions can be summarized as follows. First, the quality of data remains an important constraint on obtaining high quality empirical results. We must support adequate funding for the data generating agencies of the federal government, and we need to seek new sources of useful data. Second, given the potential limitations of the available data, it is not always possible to estimate the desired structural parameters. The data sometimes will be inadequate to answer to the question asked.

Third, it is nonetheless true that the literature of econometrics is providing a broad range of tools to do higher quality empirical research. At present, these tools are not being fully utilized. One constraint on their use appears to be the lack of incentives for their use. Incentives can be improved by changing the profession's expectations about what constitutes high quality empirical results and by improved graduate teaching.

Thus, I remain moderately optimistic that it is possible to obtain high quality estimates of structural parameters from observational data in some applications. If this is to happen, however, the research must involve greater depth--scholarship--than has been common in the historical agricultural economics literature. I can only hope that this paper can help encourage such depth.

End Notes

[1] A stochastic variable may be generated by an autoregressive process, e.g., as

$$X_t = b_0 + b_1 X_{t-1} + v_{t,}$$

where v_t represents an error term with the usual "classical" properties. If $b_1 = 1$, then the mean and variance of X_t are infinite (for definitions of the mean and variance, see Hamilton, p. 53). To test the null hypothesis $b_1 = 1$, it is common to rewrite the equation as

$$X_t - X_{t-1} = b_0 + (b_1 - 1)X_{t-1} + v_t = b_0 + dX_{t-1} + v_t\,,$$

or a variant of this equation which includes additional regressors. Then, the null hypothesis is $d = 0$. If the null cannot be rejected, this implies the variable has a unit root. The alternate hypothesis is that b_1 is less than one (d is negative), and the problem of low power of a test, mentioned in the text, occurs when b_1 is less than, but close to, one.

[2] Bessler and Covey emphasize that the definition of structure depends on the intended use of the model and that the treatment of omitted variables is important in interpreting results. In practice, model specifications likely do not include all potentially relevant variables, but the omission of particular variables may or may not be important for answering a research question. To illustrate, I assume that the observations on red meat consumption, q_t, are generated by

$$q_t = ay_t + bx_t + e_t \text{ and } x_t = cy_t + v_t,$$

where $c > 0$ for $y_t > k$, $c = 0$ otherwise, and $\mathrm{cov}(e_t v_t) = 0$.

I think of x_t as an index of health awareness such that $b< 0$ and y_t as income such that $a > 0$. Thus, health awareness is positively related to income after income reaches a threshold level, but meat consumption decreases as health awareness increases. The research objective is, say, to estimate the marginal propensity to consume meat out of income. To consider alternative estimates of "a" with x_t included and excluded from the model, rewrite the model as

$$q_t = ay_t + b(cy_t + v_t) + e_t = (a + bc)y_t + bv_t + e_t = dy_t + w_t.$$

If y_t is sufficiently small, $c = 0$ and $d = a$; whether or not x_t is included in the model, least squares gives an unbiased estimate of "a" (ignoring other possible problems). If, on the other hand, $y_t > k$ always, then x_t must be included to obtain an unbiased estimate of "a," but if x_t is omitted, an unbiased estimate of "d" is obtained. This estimate of the combined effect of y_t and x_t on q_t may be all that is required for the research objective; it is an estimate of the collective effects of x_t and y_t, as captured by y_t. Since such variables as age distribution and health indexes are often omitted from consumption and demand functions, the foregoing interpretation is implicit in many fitted models. With respect to measuring structural change, "d" is constant as long as "a," "b," and "c" are constant. In this example, however, "c" is defined to switch from zero to a positive number at a threshold $y_t = k$. Thus, if x_t is omitted and if y_t moves from below to above "k," then "c" and hence "d" change. A variable like health awareness could have a changing effect on consumption as income and educational levels change, and model specification and structural change can be interrelated issues (see also Alston and Chalfant).

[3] An analyst usually compares her/his model, fitted to a current sample period, with published results in the literature. The published results involve older data and a different model. In such comparisons, it is unclear whether the older results can be duplicated. Moreover, we do not know how the previous model would perform with the current data set, nor how the new model would perform with the old data set. Considering two models and three data sets (old, current, and future), fair comparisons of models require that the analyst use precisely the same data for both models. A full comparison implies six sets of results--both models fitted to the old, current, and new data. If it is not possible to wait for new data to become available, a full comparison of models would require fitting both to old and current data--four sets of results.

References

Alston, J. M., and J. C. Chalfant. "Can We Take the Con Out of Meat Demand Studies?" West. J. Agr. Econ. 16(1991): 36-45.

Antonovitz, F., and R. Green. "Alternative Estimates of Fed Beef Supply Response to Risk,' Amer. J. Agr. Econ. 72(1990): 475-487.

Baltagi, B. H., and J. M. Griffin. "Pooled Estimators vs. Their Heterogeneous Counterparts in the Context of Dynamic Demand for Gasoline," J. of Econometrics 77(1997): 303-327.

Bates, C., and H. White. "A Unified Theory of Consistent Estimation for Parametric Models," Econometric Theory 1(1985): 151-178.

Bessler, D. A., and T. Covey. "On the Search for Econometric Structure in Agriculture," Amer. J. Agr. Econ. 75(Oct. 1993, 75th anniversary issue): 41-47.

Box, G. E. P., and G. M. Jenkins. Time Series Analysis: Forecasting and Control, revised edition. Holden-Day, Inc.,1976.

Capps, O., Jr., and J. Havlicek, Jr. "Concepts of Consumer Demand Theory," in Food Demand Analysis, R. Raunikar and C.-L. Huang, eds. Iowa State Univ. Press, 1987. Chapter 1.

Christ, C. F. "The Cowles Commission's Contributions to Econometrics at Chicago, 1939-1955," J. of Econ. Lit. 32(1994): 30-59.

Cooley, T. F., and S. F. LeRoy. "What Will Take the Con Out of Econometrics? A Reply," Amer. Econ. Rev. 76(1986): 504-507.

Darnell, A. C., and J. L. Evans. The Limits of Econometrics. Edward Elgar Publishing Limited, 1990.

Davidson, J. E. H., D. F. Hendry, F. Srba, S. Yeo. "Econometric Modelling of Aggregate Time-Series Relationship between Consumers' Expenditures and Income in the United Kingdom," Econ. J. 88(1978): 661-692.

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

Diebold, F. X., and R. L. Lamb. "Why are Estimates of Agricultural Supply Response so Variable?" J. of Econometrics 76(1997): 357-373.

Eales, J. "A Symmetric Approach to Canadian Meat Demand Estimation," J. Agr. & Res. Econ. 21(1996): 368-380.

Engle, R. F., D. F. Hendry, and J.-F. Richard. "Exogeneity," Econometrica 51(1983): 277-304.

Fox, K. A. Econometric Analysis for Public Policy. Iowa State College (now Univ. Press, 1958. Chapter 2.

Fraker, T. M. The Effects of Food Stamps on Food Consumption: A Review of Literature. A report prepared for the Food and Nutrition Service, USDA, Oct. 1990. Table IV.1, p. 62.

Gardner, B. L. "How the Data We Make Can Unmake Us: Annals of Factology," Amer. J. Agr. Econ. 74(1992): 1066-1075.

Godfrey. L. G. Misspecification Tests in Econometrics. Cambridge Univ. Press, 1988.

Goldberger, A. S. A Course in Econometrics. Harvard Univ. Press, 1991. Chapter 23.

Goldfeld, S. M., and R. E. Quandt. "Estimation in a Disequilibrium Model and the Value of Information," J. of Econometrics 3(1975): 325-348.

Granger, C. W. J. "A Review of Some Recent Textbooks of Econometrics," J. of Econ. Lit. 32(1994): 115-122.

Griliches, Z. "Distributed Lags: A Survey," Econometrica 35(1967): 16-49.

Hamilton, J. D. Time Series Analysis. Princeton Univ. Press, 1994.

Hendry, D. F. Dynamic Econometrics. Oxford Univ. Press, 1995.

Huber, P. J. Robust Statistics. John Wiley & Sons, 1981

Johnson, S. R., Z. A. Hassan, and R. D. Green. Demand Systems Estimation: Methods and Applications. Iowa State Univ. Press, 1984.

Kling, C. L., and C. J. Thomson. "The Implications of Model Specification for Welfare Estimation in Nested Logit Models," Amer. J. Agr. Econ. 78(1996): 103-114.

Leamer, E. E. Specification Searches: Ad Hoc Inference with Nonexperimental Data. John Wiley & Sons, 1978.

Liu, D. J., and O. D. Forker. "Optimal Control of Generic Fluid Milk Advertising," Amer. J. Agr. Econ. 72(1990): 1047-1055.

Liu, T. C. "Underidentification, Structural Estimation, and Forecasting," Econometrica 28(1960): 855-865.

MacKinnon, J. G. "Model Specification Tests and Artificial Regressions," J. of Econ. Lit. 30(1992): 102-146.

Maddala, G. S. Limited-Dependent and Qualitative Variables in Econometrics. Econometric Soc. Mono. No. 3. Cambridge Univ. Press, 1983.

McAleer, M., A. R. Pagan, and P. A. Volker. "What Will Take the Con Out of Econometrics?" Amer. Econ. Rev. 75(1985): 293-307.

McGuirk, A. M., P. Driscoll, and J. Alwang. "Misspecification Testing: A Comprehensive Approach," Amer. J. Agr. Econ. 75(1993): 1044-1055.

__________, P. Driscoll, J. Alwang, and H. Huang. "System Misspecification Testing and Structural Change in the Demand for Meat," J. Agr. & Res. Econ. 20(1995): 1-21.

Nerlove, M., and F. V. Waugh. "Advertising Without Supply Control: Some Implications of a Study of the Advertising of Oranges," J. Farm Econ. 43(1961): 813-837.

Paris, Q., M. R. Caputo, and G. J. Holloway, "Keeping the Dream of Rigorous Hypothesis Testing Alive," Amer. J. Agr. Econ. 75(Oct. 1993, 75th anniversary issue): 25-40.

Passell, P., and J. B. Taylor. "The Deterrent Effect of Capital Punishment: Another View," Amer. Econ. Rev. 76(1977): 445-451.

Sims, C. A. "Money, Income, and Causality," Amer. Econ. Rev. 62(1972): 540-552.

"Macroeconomics and Reality," Econometrica 48(1980): -48.

Staiger, D., and J. H. Stock. "Instrumental Variable Regression with Weak Instruments," Econometrica 65(1997): 557-586.

Tomek, W. G. "Distributed Lag Models of Cotton Acreage Response: A Further Result," Amer. J. Agr. Econ. 54(1972): 108-110.

__________. "Confirmation and Replication in Empirical Econometrics: A Step Toward Improved Scholarship," Amer. J. Agr. Econ. 75(Oct. 1993, 75th anniversary issue): 6-14.

__________. Farm-Retail Price Spreads in a Changing Economy. Cornell Univ. Agr., Res., & Mgr. Econ. Staff Paper SP96-07, Dec. 1996.

__________, and R. J. Myers. "Empirical Analysis of Agricultural Commodity Prices: A Viewpoint," Rev. of Agr. Econ. 15(1993): 181-202.

Tyrrell, T., and T. Mount. "A Nonlinear Expenditure System Using a Linear Logit Specification,' Amer. J. Agr. Econ. 64(1982): 539-546.

Wallace, T. D. "Pretest Estimation in Regression: A Survey," Amer. J. Agr. Econ. 59(1977): 431-443.

Weng, W., an T. D. Mount. Demand Systems for Energy Forecasting: Practical Considerations for Estimating a Generalized Logit Model. Cornell Univ. Agr., Res., & Mgr. Econ. Working Paper WP97-02, Feb. 1997.

White, H. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity," Econometrica 48(1980): 817-838.

Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," J. Amer. Stat. Assoc. 57(1962): 348-368.