Journal of Agricultural and Resource Economics, 20(1):1–21

System Misspecification Testing and Structural Change in the Demand for Meats

Anya McGuirk, Paul Driscoll, Jeffrey Alwang, and Huilin Huang

A misspecification testing strategy designed to ensure that the statistical assumptions underlying a system of equations are appropriate is outlined. The system tests take into account information in, and interactions between, all equations in the system and can be used in a wide variety of applications where systems of equations are estimated. The system testing approach is demonstrated by modeling U.S. consumer demand for meats. The example illustrates how the approach can be used to disentangle issues regarding structural change and other forms of model misspecification.

Key words: econometric modeling, misspecification testing, regression diagnostics, systems of equations

Introduction

The statistical adequacy principle first proposed by R. A. Fisher asserts that to evaluate any theory using econometrics, the theory must be viewed in the context of a valid statistical model.[1] A valid statistical model is one whose underlying assumptions are appropriate for the data being analyzed. In econometrics these assumptions usually relate to the (conditional) distribution and moments of the observable random variables (Spanos 1989). Although published research does not always reflect full appreciation of this principle, most researchers are aware that tests of economic theory have no statistical validity unless model assumptions are valid. Test statistics will not have their expected distributions when underlying model assumptions are inappropriate.

Verifying that model assumptions are adequate for the data being analyzed is not necessarily trivial. For example, Alston and Chalfant (1991a) illustrate that researchers can be misled by isolated tests of model assumptions as often as 100% of the time. McGuirk, Driscoll, and Alwang (MDA) recently proposed a practical misspecification strategy designed to help researchers confirm that a single-equation linear regression model is statistically valid. Their strategy helps minimize the erroneous conclusions often reached when, as is typical in applied econometric studies, only one or two assumptions are checked in isolation. They advocate checking all testable statistical assumptions underlying a model using a battery of individual- and joint-misspecification tests and interpreting these tests as a whole, rather than separately.

Statistical adequacy is as important for systems of equations as it is for single-equation models. Even so, misspecification tests for systems are conducted even less frequently than

The authors are, respectively, an associate professor in the Departments of Agricultural and Applied Economics and Statistics, Virginia Tech; associate professors in the Department of Agricultural and Applied Economics, Virginia Tech; and an economist at the Tainan District Agricultural Improvement Station, Taiwan.

The authors thank Aris Spanos and two anonymous referees for their helpful comments on an earlier draft of this manuscript. The authors are also grateful to Jim Eales and Laurian Unnevehr for providing some of the data used in the analysis.

[1] This principle is also referred to as Fisher's Axiom of Correct Specification (Spanos 1989).

single-equation tests. When "system" tests are conducted, they are usually single-equation tests performed equation by equation.[2] While such tests can be useful in detecting misspecifications, they do not take into account information in and interactions between other equations in the system. For example, (single equation) Durbin-Watson tests will not detect correlations between residuals from different equations. Similarly, single-equation homoskedasticity tests do not examine heteroskedasticity in the contemporaneous covariance between residuals of different equations. These misspecifications can only be detected using system tests.

The purpose of this article is to propose and demonstrate a misspecification strategy designed to ensure that the statistical assumptions underlying a system of equations are appropriate. The testing strategy incorporates system tests suggested by Spanos (1986, 1995a) and is a natural extension of the single-equation testing regime advocated by MDA. To demonstrate the proposed strategy, consumer demand for U.S. meat products is modeled. This example illustrates how the proposed approach can disentangle issues regarding the existence of structural change and other forms of model misspecification. Numerous authors have examined the issue of structural change in U.S. meat demand, but most have tested for structural change without considering the wide variety of other potential model misspecifications.[3] The statistical adequacy principle challenges inferences drawn from models that may be misspecified; in fact, it is likely that the different conclusions regarding whether and to what extent structural change has occurred are due to inappropriate model specifications (Alston and Chalfant 1991b).

Statistical Assumptions Underlying Systems of Equations

The purpose of this section is to discuss briefly the statistical assumptions underlying systems of equations. A system of equations can be nonsimultaneous or simultaneous. Nonsimultaneous systems of equations, henceforth called multiequation linear regression models (MLRM), are considered first, as they represent the most straightforward extension of the single-equation linear regression model (LRM).

The Multiequation Linear Regression Model

Define y_t as a q-vector of endogenous variables and X_t as a $(K+1)$-vector including a constant and K exogenous or predetermined variables, all at time t. The multiequation linear regression model can be formulated as $y_t = \mathrm{B}'X_t + u_t$, where B is a $(K+1) \times q$ matrix of unknown parameters, and u_t is a q-vector of random disturbances. In this formulation all regressors are assumed to appear in all equations. Restricted versions of this most general model can be tested once the assumptions underlying the MLRM are verified for the data being examined. By first establishing that the MLRM is statistically appropriate, valid inferences regarding which regressors appear in which equation are ensured.

[2]For example, both Heien and Durham and Moschini and Meilke test for autocorrelation and only use single-equation tests despite the fact they estimate systems of equations. Theil and Shonkwiler, on the other hand, describe and examine a "system" Monte Carlo test of autocorrelation. However, their test is based on a simple average of the equation-by-equatoin Durbin-Watson tests and, thus, is not a full-system test in the spirit described below. We are only aware of three empirical studies which have examined the assumptions underlying a system of equations using full-system tests: Alston and Chalfant (1991a) employ a system Chow test, and Spanos (1990) and Assarsson et al. conduct various system tests of several model assumptions.

[3]See Capps and Schmitz and Smallwood, Haidacher, and Blaylock for a review of recent meat demand studies.

The MLRM assumptions can be summarized as follows:

a. Normality: $f(y_t|X_t;\,\theta) \sim$ multivariate normal—the distribution of y_t conditional on X_t is normal, where $\theta = (\mathrm{B}, \Omega)$ and Ω is the $(q \times q)$ conditional variance-covariance matrix;

b. Functional Form: $E(y_t|X_t = x_t) = \mathrm{B}'x_t$ or the functional form of the conditional mean is linear;

c. Homoskedasticity:

  i. Static: the conditional variance, $\mathrm{var}(y_t|X_t = x_t)$, does not depend on x_t;

  ii. Dynamic: $\mathrm{var}(y_t|X_t = x_t)$ does not depend on the past history of u_t, y_t, or x_t;

d. Parameter Stability: $\theta = (\mathrm{B}, \Omega)$ is stable. That is, the parameters of the conditional mean and variance do not vary with t;

e. Independence: $Y \equiv (y_1, y_2, \ldots, y_T)'$ represents an independent sample sequentially drawn from $f(y_t|X_t;\theta), t = 1, 2, \ldots, T$.

f. Weak Exogeneity: The marginal distribution of X_t does not contain relevant information for the estimation of θ. Thus, it can be ignored; and

g. No Perfect Collinearity: $\mathrm{Rank}(X) = K + 1$.

If these assumptions hold, then ordinary least squares (OLS) estimation yields minimum variance, linear unbiased estimators of B and Ω; the OLS estimator of $\mathrm{B}(\hat{\beta})$ is normally distributed; and $T\hat{\Omega}$ (where $\hat{\Omega}$ is the OLS estimator of Ω) follows a Wishart distribution (Spanos 1986).[4] Just as in the LRM, only (*a*)–(*e*) are directly testable and any violation of these assumptions invalidates *all* finite and most asymptotic tests of B and Ω (MDA, p. 1045). Consequently, each of these assumptions should be tested and verified before conducting any specification tests such as standard *t*-tests of parameter significance or tests of theoretical restrictions (e.g., symmetry and homogeneity). We refer to a model satisfying (*a*)–(*e*) as "statistically adequate."[5]

Assumptions (*a*)–(*e*) relate to a *system* of equations, and thus, system misspecification tests should be conducted (Spanos 1986, 1990). As noted earlier, these tests differ from single-equation tests conducted equation by equation as they assess the appropriateness of assumptions relevant to each individual equation as well as cross-equation assumptions. In the MLRM, the relevant cross-equation assumptions pertain to the cross-equation error covariances; these covariances are assumed to be homoskedastic, independent, and nonvarying over the index t. Before describing the MLRM misspecification tests, we consider how the statistical assumptions underlying simultaneous systems of equations differ from those of the MLRM.

[4]For this system of equations, OLS estimators of B and Ω (obtained equation by equation) are equivalent to generalized least-squares estimators. Note also that these MLRM assumptions (and the tests outlined below) are applicable to models incorporating time series and or cross-sectional data. That is, the index t does not necessarily refer to time. It can represent any dimension over which it makes sense to order observations. For example, when using cross-sectional data, the relevant orderings over which independence, parameter stability, and dynamic homoskedasticity should hold may be "space" and/or "size of unit" (Spanos 1995a, Chapter 13).

[5]All of the assumptions underlying the MLRM are statistical assumptions. The first step in any modeling exercise is to find a statistical model that adequately summarizes the information in the data. Issues regarding errors in variables, latent or omitted variables, and simultaneity are considered theoretical rather than statistical. They are important when initially defining the data whose relevant features the statistical model is supposed to model and again when identifying the theoretical model from within the statistical model. For a more in-depth discussion of this distinction see Spanos (1986, 1989, 1990, 1995b).

Simultaneous Systems of Equations

Traditionally, estimation and testing with simultaneous systems of equations are different from single-equation or nonsimultaneous system models. In the case of simultaneous systems, a structural model is usually specified and then estimated using one of the many simultaneous systems estimators (for instance, three-stage least squares). These estimators are designed to purge the right-hand-side endogenous variables of their "stochastic" components, assumed to be correlated with the structural equation error. Occasionally, though infrequently, researchers check for correlation among the errors of the structural equation (e.g., report a Durbin-Watson test). However, more often than not, researchers immediately begin interpreting parameter estimates, conducting tests of theory, and perhaps, investigating simultaneity issues.

The current approach to estimating simultaneous systems of equations will not, generally, lead to valid statistical inferences (Spanos 1990). This follows because few, if any, steps are taken to ensure that the structural model is statistically valid. To ensure statistical validity, and thus safeguard against erroneous conclusions, Spanos (1986, pp. 608–21; 1990) suggests first estimating the reduced form implied by the structural model of interest and verifying that its underlying statistical assumptions are met. A statistically adequate reduced form is critical since it is, by construction, the statistical model from which the structural model is derived. Consequently, if the statistical assumptions underlying the reduced form are invalid, statistical inferences drawn from the structural model will be invalid as well.

To see the motivation behind this alternative approach, notice first that the reduced form is simply a MLRM, whose underlying statistical assumptions are straightforward and well known [see (*a*)–(*g*) above]. Further, once the assumptions underlying this model are verified, the distributions of the reduced-form parameter estimators are known. If every equation in a simultaneous system is just identified, the parameters of the structural model can be easily recovered as unique functions of the reduced-form parameters. Once the statistical properties of the reduced-form parameter estimators are verified, the properties of the structural parameter estimators have also been implicitly verified. In this case, testing for misspecification in the reduced form is equivalent to testing for misspecification in the structural model.

Usually, however, the equations constituting simultaneous systems are overidentified. That is, there is more than one combination of reduced-form parameters that yields the same structural parameter. Thus, estimation of an overidentified structural equation or system is equivalent to estimation of a restricted reduced form—a reduced-form system where different combinations of the reduced-form parameters are forced to yields the same structural parameter (Spanos 1986, p. 630). This correspondence presents an obvious way to assess the relevance and statistical adequacy of overidentified structural models. Initially, an unrestricted reduced form should be estimated and its statistical validity assessed. Once the MLRM is statistically adequate, the appropriateness of the theoretical (structural) model can be assessed by testing the implied (overidentifying) restrictions imposed on the reduced form. Because the statistical assumptions underlying the unrestricted reduced form have been verified, valid inferences about the overidentifying restrictions are ensured. Finally, if the structural model is not rejected (i.e., the overidentifying restrictions are not rejected), and if further inferences are to be drawn from this model, Spanos advocates checking to verify that the MLRM is still valid once the restrictions are imposed.

As a consequence of the correspondence between the reduced form and the structural model, assessing the statistical adequacy of a system of equations, whether simultaneous or not, boils down to initially assessing the appropriateness of the statistical assumptions underlying the relevant MLRM. In the next section, we outline a practical set of misspecification tests similar to those advocated by MDA but designed to verify that a MLRM is appropriate for the data being analyzed. Following this, we demonstrate their practical application by modeling the demand for meat.

System Misspecification Tests

MDA demonstrate that a practical set of single-equation tests should check the relevance of all testable model assumptions using various versions of individual and joint tests. The individual tests are designed to check a single assumption, while the joint tests simultaneously assess the relevance of multiple assumptions. Although both types of tests should be conducted, MDA illustrate that joint tests, which require fewer maintained hypotheses, are often instrumental in finding the misspecification's source. System versions of the single-equation individual and joint tests examined by MDA are advocated here. As will become evident, most of the system tests differ from their single-equation counterparts only by the inclusion of extra regressors and/or extra auxiliary regressions incorporated to assess the relevance of cross-equation assumptions. The otherwise similar structure of these tests suggests that the proposed system testing regime will be effective in detecting sources of misspecification. The system version of the LRM tests advocated by MDA are described below. For more details see Spanos (1986, 1990, and 1995a).[6]

Normality

The skewness and kurtosis coefficients of a random $q \times 1$ vector u_t with mean 0 and covariance, Ω, are defined by: $\alpha_3 = \left[E(u_t' \Omega^{-1} u_t)^3\right]^{1/2}$ and $\alpha_4 = E(u_t' \Omega^{-1} u_t)^2$, respectively. These coefficients can be estimated by $\hat{\alpha}_3^2 = (1/T^2)\Sigma_t\Sigma_s g_{ts}^3$ and $\hat{\alpha}_4 = (1/T)\Sigma_t g_{tt}^2$, where $g_{ts} = \hat{u}_t'\Omega^{-1}\hat{u}_s, t, s = 1, 2, \ldots, T$. The null hypothesis of multivariate normality is $\alpha_3 = 0$ and $\alpha_4 = q(q+2)$, which can be tested jointly using the C_w^2 (small sample approximation) omnibus test proposed by Mardia and Foster.

In addition to this multivariate test, the univariate D'Agostino-Pearson K^2 omnibus test described in MDA can be used to examine the normality assumption in each separate equation. If multivariate normality is rejected, the univariate tests provide insight into which equations are likely to be responsible for system nonnormality.

Functional Form

System functional form tests are multivariate versions of the RESET and Kolmogorov-Gabor functional form tests examined by MDA. These system tests are based on the significance of Γ_F, a $p \times q$ matrix of unknown parameters in the q-equation auxiliary regression system,

$$\hat{u}_t = \mathrm{B}_o' x_t + \Gamma_F' \Psi_{Ft} + v_t, \quad t = 1, 2, \ldots, T.$$

[6]With the exception of the normality test, most of the system tests proposed in Spanos (1986) and employed here use what Godfrey refers to as "locally equivalent alternatives" (Assarsson et al.).

In the multivariate RESET2 test, $\Psi_{Ft} \equiv (\hat{y}_{1t}^2, \hat{y}_{2t}^2, \ldots, \hat{y}_{qt}^2, \hat{y}_{1t}\hat{y}_{2t}, \ldots, \hat{y}_{q-1t}\hat{y}_{qt})'$, where $\hat{y}_{it}$ is the fitted value from the ith equation. Thus, the number of regressors (p) in Ψ_{Ft} is $(½)q(q + 1)$.[7] The significance of Γ_F is examined using the small-sample-adjusted likelihood-ratio test proposed by Rao (p. 556).[8] The Rao test statistic, $F = [(1 - \Lambda^{1/t})/(\Lambda^{1/t})][(rt - 2z/pq]$ is distributed approximately $F(pq, rt - 2z)$, where $\Lambda = |\hat{\Omega}|/|\tilde{\Omega}|$, the ratio of the determinants of the unrestricted and restricted variance-covariance matrices, $r = v - (p - q + 1)/2$, v is the degrees of freedom for the error, $z = (pq - 2)/4$, and $t = [(p^2 q^2 - 4)/p^2 + q^2 - 5)]^{0.5}$ if $(p^2 + q^2 - 5) > 0$ or $t = 1$ otherwise.

In addition to the Rao test which simultaneously assesses the appropriateness of all pq coefficients in Γ_F, the significance of the relevant p coefficients in each equation can be examined separately using F-tests. As in the equation-by-equation normality tests, these separate F-tests often provide insight into the source of problems if the "full-system" Rao test is rejected. For example, it may be that a particular equation is the sole cause of system misspecification. The equation-by-equation *system* tests will be useful in identifying the culprit in this case. Notice that these equation-by-equation system tests differ from their single-equation counterparts described in MDA, in that, the latter only include transformations of the fitted values from the particular equation being examined.

A second-order Kolmogorov-Gabor polynomial test (KG2) is conducted in addition to the RESET2 test. In the KG2 test, Ψ_{Ft} includes $x_{it}x_{jt}$, $i \geq j$, $i, j = 2, \ldots, K + 1$, where x_{it} is the ith element of X_t, the $(K + 1)$-vector of regressors in the MLRM. As above, a Rao test is used to examine the significance of Γ_F, and separate F-tests can be conducted to examine the significance of the relevant coefficients equation by equation.

Homoskedasticity

The system homoskedasticity tests are based on the significance Γ_H in the auxiliary regression system:

$$w_t = c_0 + \Gamma_H' \Psi_{Ht} + v_t, \qquad t = 1, 2, \ldots, T,$$

where $w_t \equiv (\hat{u}_{1t}^2, \hat{u}_{2t}^2, \ldots, \hat{u}_{qt}^2, \hat{u}_{1t}\,\hat{u}_{2t}, \ldots, \hat{u}_{q-1t}\,\hat{u}_{qt})'$ and $\hat{u}_{it}$ is the residual of the ith equation. This system of auxiliary equations includes $(½)q(q + 1)$ equations and is used to investigate whether the variances *and* covariances of residuals from the different equations are homoskedastic.

We propose two static homoskedasticity tests. These include a multivariate RESET2-type test where $\Psi_{Ht} \equiv (\hat{y}_{1t}^2, \hat{y}_{2t}^2, \ldots, \hat{y}_{qt}^2, \hat{y}_{1t}\,\hat{y}_{2t}, \ldots, \hat{y}_{q-1t}\hat{y}_{qt})'$ and a multivariate WHITE test where Ψ_{Ht} (only) includes the terms $x_{it}x_{jt}$, $i \geq j$, $i, j = 2, \ldots, K + 1$.[9] The dynamic homoskedasticity test used is a system autoregressive conditional heteroskedasticity (ARCH) test (Engle). It

[7]If q is large relative to the sample size, a degrees of freedom problem can occur. In this case, the tests proposed (here and throughout the paper) can be easily varied. For example, in the RESET2 functional form test, the crossproducts can be left out. Several such simplifications had to be made in the illustrative example given below (e.g., see table 2, footnote b).

[8]Many other small-sample adjustments to the likelihood ratio test have been proposed. We use Rao' s test as Monte Carlo studies indicate it is a consistent performer (see Seber, pp. 414–15; and Rao, p. 556).

[9]The RESET2-type homoskedasticity test described here is sometimes referred to as a Breusch-Pagan test (see, for example, Assarsson et al.).

differs from the static tests in that $\Psi_{Ht} \equiv w_{t-1}$. As in the full-system functional form test, the significance of Γ_H is assessed using a (full-system) Rao test and separate F-tests can be conducted to examine the relevance of the homoskedasticity assumption equation by equation. Notice that conducting these equation-by-equation system tests is different from conducting LRM tests equation by equation in that the system tests check whether the relevant covariances are homoskedastic, and the system RESET2-type test includes extra terms in each auxiliary equation.

Parameter Stability

The stability of B, the MLRM conditional mean parameters, can be assessed using a multivariate version of the single-equation Chow test.[10] Because this mean-stability test assumes that the conditional variance-covariance matrices from the two prespecified sample periods (composed of the first T_1 and last T_2 observations, respectively) are equal, $(\Omega_1 = \Omega_2 = \Omega)$, this latter hypothesis should also be tested. A Rao test can be conducted in both instances, $\Lambda = (|\hat{\Omega}_1|^{T_1/T} |\hat{\Omega}_2|^{T_2/T})/|\hat{\Omega}|$ and $\Lambda = (|\hat{u}_1'\hat{u}_1 + \hat{u}_2'\hat{u}_2|)/|\hat{u}'\hat{u}|$ for the variance and mean tests, respectively. To gain insight into the relevance of the parameter-stability assumption for each equation in the system separately, the single-equation Chow and variance-stability test used in MDA are also conducted.

Independence

Independence can be examined using a Rao test to assess the significance of in the q-equation auxiliary regression system

$$\hat{u}_t = \mathrm{B}_o'x_t + \Gamma_I'\Psi_{It} + v_t, \qquad t = 1, 2, \ldots, T,$$

where $\Psi_{It} \equiv \hat{u}_{t-1}$. This test is a system equivalent of the Breusch-Godfrey test. As above, equation-by-equation F-tests can be conducted to help identify specific problems if the full-system test is rejected. These equation-by-equation system tests differ from the single-equation independence test in MDA as each auxiliary equation includes lagged residuals from all equations in the system.

Joint Conditional Mean Test

The system joint-mean test simultaneously checks the appropriateness of functional form, independence, and the stability of B as each of these assumptions refer to aspects of the conditional mean. The test is based on the auxiliary regression system:

$$\hat{u}_t = \mathrm{B}_o'x_t + \Gamma_F'\Psi_{Ft} + \Gamma_I'\Psi_{It} + \Gamma_S'\Psi_{St} + v_t, \qquad t = 1, 2, \ldots, T,$$

[10]Obviously, there are many possible tests of parameter stability including much more general tests which treat the "breakpoint" as endogenous (see, for example, Zivot and Andrews). We use a multivariate extension of the commonly used Chow test simply to illustrate the possibilities of converting single-equation tests to system tests.

where $\Psi_{Ft} \equiv (\hat{y}_{1t}^2, \hat{y}_{2t}^2, \ldots, \hat{y}_{qt}^2, \hat{y}_{1t}\hat{y}_{2t}, \ldots, \hat{y}_{q-1t}\hat{y}_{qt})'$, allowing for alternative functional forms, $\Psi_{It} \equiv \hat{u}_{t-1}$, $\alpha_3 = \left[E(u_t'\Omega^{-1}u_t)^3\right]^{1/2}$ to capture nonindependence, and Ψ_{St} includes a trend and trend squared to model possible instability in B. Other specifications of Ψ_{St} are possible (e.g., the inclusion of slope and intercept shifters) but are not illustrated here due to limited degrees of freedom (see example). The null hypothesis that the conditional mean is properly specified is $\Gamma_F = \Gamma_I = \Gamma_S = 0$, which can be examined using a Rao test as well as equation-by-equation F-tests. The most likely cause of a rejection, if it results, can be investigated by assessing the significance of Γ_F, Γ_I, and Γ_S separately.

Joint Conditional Variance Test

The full-system joint variance test simultaneously checks for static and dynamic heteroskedasticity as well as stability of Ω. It is based on the auxiliary regression system:

$$w_t = \Gamma_{HS}'\Psi_{HSt} + \Gamma_{HD}'\Psi_{HDt} + \Gamma_S'\Psi_{St} + v_t, \qquad t = 1, 2, \ldots, T,$$

where w_t and Ψ_{St} are defined as above, $\Psi_{HDt} \equiv w_{t-1}$ capturing potential dynamic heteroskedasticity, and $\Psi_{HSt} \equiv (\hat{y}_{1t}^2, \hat{y}_{2t}^2, \ldots, \hat{y}_{qt}^2, \hat{y}_{1t}\hat{y}_{2t}, \ldots, \hat{y}_{q-1t}\hat{y}_{qt})'$ allowing for possible static heteroskedasticity. The null hypothesis that the conditional variance is properly specified is $\Gamma_{HS} = \Gamma_{HD} = \Gamma_S = 0$. As in the mean test, this hypothesis can be examined using a Rao test and equation-by-equation F-tests. The most likely cause of a rejection can also be investigated by assessing the significance of Γ_{HS}, Γ_{HD}, and Γ_S individually.

The System Testing Approach

Following the recommendations of MDA, we conduct each of the individual and joint tests described above; the resulting p-values are interpreted as the weight of evidence against the assumption(s) being tested—the smaller the p-value the more evidence against the assumption(s) holding. As discussed in MDA, exactly what constitutes a "small" value should depend on the *overall* level of significance desired (a subjective decision) as well as the number of tests conducted. In the following example, p-values lower than 0.05 are interpreted as "weak" evidence against the null, and values less than 0.01 as "strong" evidence.[11] However, by reporting all p-values, readers can freely interpret the test evidence.

After conducting all tests, the p-values from the full-system tests are examined.[12] If an overall test p-value indicates a problem, the equation-by-equation p-values are examined to help determine the problem source. Following the recommendations of MDA, no single test is interpreted in isolation. That is, before any conclusions are drawn regarding possible misspecification source, all test results are examined. Furthermore, once a conclusion is

[11] Although the choice is subjective, MDA argue that an overall test size of 20–25% may be necessary to achieve decent power when conducting multiple misspecification tests. Given that we are conducting 11 separate system misspecification tests, the corresponding (approximate) overall level of significance (ω) associated with individual p-values 0.01 and 0.05 are 10% and 43%, respectively (where ω is calculated using the Sidak approximation described in MDA). For useful discussions and additional references on the choice of appropriate significance levels see Maddala or Assarsson et al.

[12] All tests were conducting using programs written in GAUSS. Most were performed using SAM, *An Interactive Regression Program*, written in GAUSS, by Robertson, McGuirk, and Spanos.

drawn and the model respecified, the same battery of tests is conducted to ensure that the misspecification problems have been remedied. To clarify and demonstrate the recommended approach we now introduce the example.

Model Specification, Data, and Assumptions

Variants of the Almost Ideal Demand System (ALIDS) of Deaton and Muellbauer have been employed in many recent meat demand studies (e.g., Moschini and Meilke; Eales and Unnevehr 1988, 1993). We use the static linear approximation ALIDS as our point of departure:

$$w_{it} = \alpha_i + \Sigma_j \gamma_{ij} \log p_{jt} + \beta_i \log(x_t / P_t), \qquad i = 1, \ldots, q+1; \quad t = 1, 2, \ldots, T, \tag{1}$$

where w_{it} and p_{it} represent budget share and price for the ith good, respectively; x_t is total expenditures on meats; P_t is the ALIDS price index, approximated with the Stone index; and α, γ, and β are the model parameters.

We examine the annual demand for beef, pork, and chicken from 1960 to 1988. Retail price and per capita disappearance data (quantity) for beef, pork, and chicken are obtained from the U.S. Department of Agriculture's *Livestock and Poultry Situation and Outlook Report.* These data have been used in many studies of U.S. meat demand (e.g., Capps and Schmitz; Chavas; Dahlgran; Moschini and Meilke).

The static ALIDS model described in (1), with the usual adding up, symmetry, and homogeneity restrictions imposed, is the theoretical model of interest. Model (1) with no theoretical restrictions imposed is the relevant statistical model—a MLRM. Thus, in this example, "solving" for the underlying statistical model is trivial.

As indicated, we begin by estimating and testing the unrestricted static ALIDS model (MLRM) describing the demands for beef, pork, and chicken. All of the system tests outlined above are conducted. For the multivariate Chow test, we define T_1 = 1960–76 and T_2 = 1977–88. This breakpoint was selected based on the finding of many researchers that structural change occurred in the late 1970s.[13] As usual, one share equation is omitted to avoid a singular variance-covariance matrix; thus, $q = 2$. All of the full-system tests described above are invariant to the equation dropped, except the RESET-type heteroskedasticity tests. To obtain an invariant test in these cases, the fitted values from all equations, including the omitted one, compose the RESET terms. Despite these invariant system tests, the system is estimated three times (omitting a different equation each time) to obtain all the relevant equation-by-equation test results.

[13]By conducting a multivariate Chow-like test we assume we know when the structural change occurred. By basing our decision to test this breakpoint on results of studies using the same or very similar data, it is possible that the p-value of this test is biased in favor of rejection (Christiano). However, by reporting the p-value for this (and all other) tests, the readers are free to interpret the p-value evidence themselves.

Table 1. U.S. Meat Demand Models: The p-values for Full-System Misspecification Tests

Item	Model A	Model B	Model C	Model D
Individual Tests				
Normality	0.022	0.161	0.475	0.011
Functional Form:				
RESET2	0.000	0.119	0.016	0.026
KG2	0.001	0.175[a]	0.110[a]	0.199[a]
Heteroskedasticity:				
Static: RESET2	0.710	0.997	0.997	0.982
WHITE	0.947	0.982[a]	0.991[a]	0.878[a]
Dynamic	0.005	0.074	0.061	0.046
Autocorrelation	0.000	0.648	0.326	0.046
Parameter Stability:				
Variance	0.000	0.015	0.000	0.000
Mean	0.000	0.709	0.070	0.019
Joint Tests				
Overall Mean Test	0.000	0.088	0.024	0.013
Parameter Stability	0.003	0.073	0.056	0.096
Functional Form	0.454	0.207	0.081	0.552
Autocorrelation	0.019	0.550	0.562	0.150
Overall Variance Test[b]	0.146	0.538	0.931	0.276
Parameter Stability	0.351	0.736	0.797	0.300
Static Hetero.	0.865	0.900	0.954	0.527
Dynamic Hetero.	0.068	0.119	0.656	0.078

Note: Model A is a static ALIDS model; Model B is a static ALIDS model augmented with *CHOL* and *WWOM*; Model C is a static ALIDS model augmented with *B&S* and *WWOM*; Model D is an interrelated partial-adjustment model.
[a]KG2 and WHITE tests only include squares of independent variables since degrees of freedom are limited.
[b]In the joint variance test, the RESET terms incorporate the fitted values from the beef and pork equations only (regardless of which equation is omitted), due to limited degrees of freedom.

Results

Static ALIDS

The results from the static ALIDS misspecification tests are reported in tables 1 and 2 (Model A). The p-values from the full-system tests are reported in table 1 and equation-by-equation p-values in table 2. The full-system p-values indicate possible violation of all assumptions, except perhaps homoskedasticity. The equation-by-equation results confirm that the model is severely misspecified. It is impressive that multiple misspecifications are detected, given that we are using what can be considered a relatively small sample size ($T = 29$).

Specifically, the system tests of individual misspecification indicate possible problems with functional form, misspecified dynamics (autocorrelation), parameter instability, and dynamic heteroskedasticity (all p-values ≤ 0.005). The equation-by-equation tests indicate that all equations exhibit these symptoms, the severest in beef and chicken equations.

Table 2. U.S. Meat Demand Models: The *p*-values for Equation-by-Equation System Misspecification Tests [a]

Item		Model A Beef	Model A Pork	Model A Chicken	Model B Beef	Model B Pork	Model B Chicken
Individual Tests							
Normality		0.164	0.448	0.516	0.658	0.151	0.162
Functional Form:							
RESET2		0.011	0.047	0.001	0.998	0.740	0.025
KG2		0.001	0.014	0.000	0.637[a]	0.391	0.100
Heteroskedasticity:[b]							
Static RESET2	Beef	0.234	0.286	0.191	0.853	0.878	0.852
	Pork		0.335	0.263		0.905	0.923
	Chicken			0.157			0.940
Static WHITE	Beef	0.305	0.381	0.250	0.838[a]	0.884	0.817
	Pork		0.492	0.322		0.910	0.965
	Chicken			0.231			0.643
Dynamic	Beef	0.000	0.000	0.000	0.786	0.814	0.152
	Pork		0.001	0.000		0.813	0.058
	Chicken			0.000			0.053
Autocorrelation		0.000	0.000	0.000	0.643	0.862	0.330
Parameter Stability:							
Variance		0.220	0.417	0.111	0.319	0.306	0.483
Mean		0.000	0.000	0.000	0.697	0.750	0.455
Joint Tests							
Overall Mean Test		0.000	0.000	0.000	0.497	0.205	0.121
Parameter Stability		0.015	0.018	0.016	0.093	0.035	0.475
Functional Form		0.886	0.378	0.360	0.497	0.270	0.267
Autocorrelation		0.005	0.029	0.040	0.740	0.718	0.338
Overall Variance Test:[b]							
	Beef	0.003	0.004	0.003	0.956	0.964	0.278
	Pork		0.009	0.004		0.964	0.270
	Chicken			0.006			0.164
Parameter Stability:							
	Beef	0.144	0.096	0.249	0.521	0.492	0.819
	Pork		0.071	0.184		0.544	0.735
	Chicken			0.402			0.384
Static Heteroskedasticity:							
	Beef	0.564	0.410	0.769	0.804	0.833	0.511
	Pork		0.289	0.669		0.869	0.855
	Chicken			0.881			0.391
Dynamic Heteroskedasticity:							
	Beef	0.002	0.002	0.003	0.757	0.850	0.112
	Pork		0.004	0.003		0.869	0.064
	Chicken			0.008			0.072

Note: Model A is a static ALIDS model; Model B is similar but also includes *CHOL* and *WWOM* to model possible changes in structure.

[a]KG2 and WHITE tests only include squares of independent variables since degrees of freedom are limited.

[b]The results from all $0.5q(q+1)$ equations are reported. The column and row titles identify the variance or covariance being tested. For example, the test result in the beef row under the pork column is for the equation assessing assumptions regarding the covariance between residuals from the pork and beef equations. The RESET terms in this test incorporate the fitted values from the beef and pork equations only (regardless of which equation is omitted), due to limited degrees of freedom.

Table 2. Continued

Item		Model C Beef	Model C Pork	Model C Chicken	Model D Beef	Model D Pork	Model D Chicken
Individual Tests							
Normality		0.946	0.376	0.091	0.165	0.368	0.003
Functional Form:							
RESET2		0.773	0.633	0.003	0.800	0.465	0.032
KG2		0.356[a]	0.360	0.058	0.481[a]	0.504	0.121
Heteroskedasticity:[b]							
Static RESET2	Beef	0.986	0.975	0.712	0.898	0.853	0.895
	Pork		0.957	0.854		0.648	0.906
	Chicken			0.815			0.946
Static WHITE	Beef	0.727[a]	0.824	0.877	0.484[a]	0.291	0.730
	Pork		0.921	0.935		0.221	0.885
	Chicken			0.818			0.863
Dynamic	Beef	0.298	0.906	0.010	0.071	0.002	0.507
	Pork		0.901	0.030		0.022	0.115
	Chicken			0.013			0.516
Autocorrelation		0.466	0.806	0.113	0.015	0.048	0.125
Parameter Stability:							
Variance		0.161	0.823	0.377	0.403	0.488	0.410
Mean		0.422	0.313	0.008	0.476	0.965	0.001
Joint Tests							
Overall Mean Test		0.336	0.178	0.033	0.087	0.173	0.013
Parameter Stability		0.069	0.026	0.481	0.129	0.680	0.096
Functional Form		0.244	0.222	0.087	0.724	0.550	0.552
Autocorrelation		0.895	0.615	0.517	0.031	0.055	0.150
Overall Variance Test:[b]							
	Beef	0.952	0.968	0.510	0.098	0.010	0.453
	Pork		0.974	0.456		0.137	0.354
	Chicken			0.652			0.520
Parameter Stability:							
	Beef	0.472	0.479	0.909	0.047	0.165	0.071
	Pork		0.550	0.575		0.708	0.221
	Chicken			0.598			0.117
Static Heteroskedasticity:							
	Beef	0.837	0.855	0.602	0.153	0.169	0.225
	Pork		0.867	0.901		0.495	0.521
	Chicken			0.477			0.329
Dynamic Heteroskedasticity:							
	Beef	0.699	0.782	0.386	0.120	0.004	0.606
	Pork		0.842	0.218		0.049	0.203
	Chicken			0.589			0.442

Note: Model C is a static ALIDS model but also includes *B&S* and *WWOM* to capture possible changes in structure. Model D is an interrelated partial-adjustment model.

[a]KG2 and WHITE tests only include squares of independent variables since degrees of freedom are limited.

[b]The results from all $0.5q(q+1)$ equations are reported. The column and row titles identify the variance or covariance being tested. For example, the test result in the beef row under the pork column is for the equation assessing assumptions regarding the covariance between residuals from the pork and beef equations. The RESET terms in this test incorporate the fitted values from the beef and pork equations only (regardless of which equation is omitted), due to limited degrees of freedom.

The joint tests seem to help pinpoint the possible problem(s) in the static ALIDS model (Model A). The full-system joint conditional-mean test (table 1) suggests that the misspecification problems may stem from trending parameters (p-value = 0.003) and/or misspecified dynamics (p-value = 0.019), rather than inappropriate function form. The equation-by-equation joint-mean tests support this finding. They indicate that all equations suffer from these (likely) problems.

The relative magnitude of the p-values on separate test components in the joint-mean tests provides evidence that parameter instability may be the main problem: the parameter instability p-values are lower than those for the autocorrelation tests in two of the three equations (Model A, table 2). However, the evidence is not conclusive and any conclusions regarding the existence or nonexistence of structural change is premature. Dynamic misspecification in all three equations could plausibly result, for example, from unmodeled habit persistence as discussed by Heien and Durham. In fact, many meat demand studies have specified dynamic models (e.g., Moschini and Meilke; Eales and Unnevehr 1993; Chavas). At this point, we can only conclude that the statistical assumptions underlying the static ALIDS model are not satisfied for these data, and thus, any hypothesis tests based on this model, including those regarding simultaneity, will be of questionable validity.

Two approaches to dealing with the misspecifications found in the static ALIDS model seem reasonable. The first approach, "modeling" structural change, considers the possibility that problems stem from structural change in the demand for meat, resulting in unstable parameters. Structural change may have caused the other misspecification tests to indicate problems, although more evidence is needed to draw this conclusion. The only way to assess the plausibility of this explanation is to "model" the structural change such that a model with stable parameters is obtained. If in doing so, the other misspecification tests are "fixed," it can tentatively be concluded that structural change has occurred.[14]

The second approach, enhanced dynamics, assumes that at least some, perhaps all, of the statistical problems arise because the dynamics of adjustment in meat demand are not adequately modeled with a static ALIDS. One possible means of modeling these dynamics is to employ a first-difference ALIDS model. This model is commonly used to "solve" autocorrelation problems in the context of the ALIDS model (e.g., Eales and Unnevehr 1988, 1993; Moschini and Meilke). A less restrictive specification, which is still relatively parsimonious, is employed here. This alternative is a variant of the interrelated partial-adjustment model of Anderson and Blundell. This variant, proposed by Alessie and Kapteyn, is

$$(2) \qquad w_{it} = \alpha_i + \Sigma_j \Theta_{ij} w_{j,t-1} + \Sigma_j \gamma_{ij} \log p_{jt} + \beta_i \log(x_t / P_t), \qquad i = 1, \ldots, q+1,$$

where $\Sigma_j \Theta_{ij} = 0$ for identification (Assarsson et al.). [Note that the possibility of "true" autocorrelation is not even considered as a possible source of the misspecification problems. Spanos (1987) illustrates why "true" autocorrelation is unlikely for most time-series models.]

The second approach (enhanced dynamics) can be compared to the first (modeling structural change) by assessing the extent to which each approach "cures" the problems indicated by the individual- and joint-system misspecification tests.

[14]The conclusion is only "tentative" because there may be another model which is also statistically adequate but leads to other conclusions (see below). This possibility is particularly plausible here since our sample is relatively small, and thus the misspecification tests may not have much power.

"Modeling" Structural Change

Structural change can be modeled in many different ways. It is commonly modeled by incorporating some function of time as a regressor and/or allowing the intercept and/or slope coefficients to change over the sample period by including binary shifters (dummy variables). Incorporating some measure of the phenomena responsible for the change is a better means of modeling change. Researchers who argue that structural change has occurred in the demand for meat often contend that information about saturated fats and cholesterol is a major determinant of the change. Some authors cite increased participation of women in the labor force and changes in income distribution, age distribution, and racial composition of the population as possible sources of change (Capps and Schmitz). Here we incorporate a measure of the concern for cholesterol and the increased participation of women in the labor force to explain the potential structural change.

The index of cholesterol awareness used is new and differs from the popular Brown and Schrader (*B&S*) index, which has been used to explain changing patterns of meat demand, apparently with some success (Capps and Schmitz). Our index was created by counting

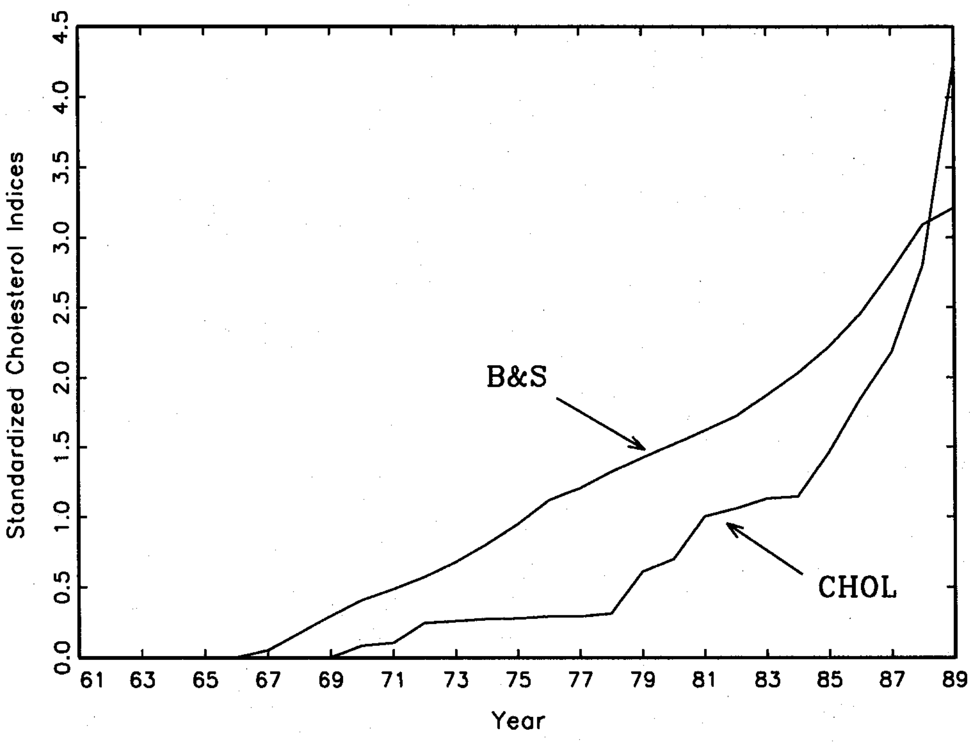

Figure 1. A comparison of the Cholesterol Index, *CHOL*, with that of Brown and Schrader (*B&S*)

articles (weighted by readership) cited in the *Reader's Guide to Periodic Literature* addressing health problems associated with dietary cholesterol. A cumulative count of subscription-weighted articles was created.[15] This variable (*CHOL*) is contrasted in figure 1 with the *B&S* index, which measures the number of articles in medical journals linking cholesterol to health (minus those refuting the link) lagged six months.[16] The pattern demonstrated by the *CHOL* index suggests that the general public may not have been aware of the cholesterol issue until much later than the *B&S* index suggests. Furthermore, the *B&S* index is close to a linear time trend, while the *CHOL* index is less smooth. We contend (and later "test") that the subscription-weighted count of popular press articles is a better representation of consumer cholesterol awareness than the count of articles in professional journals.

The increased participation of women in the work force is captured by using the percentage of women with children under 18 years of age who have been or are married and working. This variable (*WWOM*) was obtained from *Handbook of Labor Statistics*.

To "model" structural change in the context of the ALIDS model, the equation intercepts are allowed to change as *CHOL* and *WWOM* vary.[17] Thus, the relevant statistical model becomes

$$(3)\quad w_{it} = \alpha_i^* + \Sigma_j \gamma_{ij}^* \log p_{jt} + \beta_i^* \log(x_t / P_t) + \delta_{ic}^* CHOL_t + \delta_{ip}^* WWOM_t, \qquad i = 1, \ldots, q+1.$$

The p-values from the misspecification tests of the static ALIDS model augmented with the structural change variables (*CHOL* and *WWOM*) are reported in tables 1 and 2 (Model B). The full-system p-values indicate that all assumptions, except perhaps stable variance-covariance matrix (p-value = 0.015), are reasonable for this model. The equation-by-equation tests shown in table 2 suggest that *all* tested assumptions are appropriate as the variance-stability tests indicate no problems. There are two possible explanations for the difference in results for the variance tests. First, the cross-equation residual covariances may not be stable. Second, the results may simply reflect the fact that the full-system test often gives an inflated measure of departures from covariance homogeneity (Seber, p. 449). Since the suspect variance test is the *only* full-system test to indicate possible problems, this model is judged to be statistically adequate. Consequently, unstable parameters due to changing preferences may have been responsible for the poor misspecification test results of the static ALIDS model. The proxies for concerns over cholesterol (*CHOL*) and increased participation of women in the work force (*WWOM*) adequately model these changing parameters.

Before examining the extent to which misspecified dynamics can also account for the problems in the static ALIDS model, we compare our new cholesterol index with the *B&S* index by reestimating Model B, replacing our index with that of *B&S*. The misspecification results for this modified model are also reported in tables 1 and 2 (Model C). The results suggest that Model C is misspecified. Not only is the p-value on the system variance-stability test low, but the small p-values for the RESET2 functional form and joint conditional-mean tests indicate further problems. The equation-by-equation test results suggest that the chicken equation may be the sole source of these problems.

---

[15] A quarterly circulation variable (measured in millions) was initially constructed. The annual index is obtained by averaging the quarterly index each year.

[16] Both indices have been standardized so that they can be viewed on the same graph.

[17] This means of incorporating the structural change variables is consistent with the entering of demographics in an expenditure system found in other studies. It is akin to the demographic translation technique discussed by Pollack and Wales (1981). This is a parsimonious approach that allows the price and expenditure elasticities to vary with the variables *CHOL* and *WWOM*.

These findings do not necessarily mean that the *B&S* index does not adequately reflect concerns for cholesterol, since the use of the *B&S* index with another functional form, or with a different dynamic specification, could lead to another statistically adequate model. However, use of the *B&S* index, in contrast to the *CHOL* index, within the augmented static ALIDS framework, cannot be justified with these data. No claims can be made regarding the existence of structural change using results from Model C. Furthermore, any other inferences drawn from this misspecified system will be invalid.

Modeling the Dynamics of Meat Demand

The possibility of finding another model that passes all misspecification tests is not precluded by the fact that a statistically adequate model was obtained by "modeling" structural change within an augmented static ALIDS framework (Spanos 1986).[18] Thus, we estimate the interrelated partial-adjustment model (without variables modeling structural change) in order to establish the extent to which misspecified dynamics could account for the problems of the static ALIDS model. The results are reported in tables 1 and 2 (Model D). The full-system tests indicate that this model is not statistically adequate. Although autocorrelation no longer appears to be a real problem, the full-system tests indicate parameter instability in the mean and variance, and possible violations of normality and of functional form. The equation-by-equation tests confirm these findings. They indicate that the symptoms are most severe in the chicken equation and that parameter instability is the most likely problem. If we had not already estimated the structural change model (Model B) above, these results may have led us to try to fix these problems by "modeling" the structural change at this point.

If tests of autocorrelation alone had been conducted, the modeler would have concluded that the dynamic misspecifications identified in the static ALIDS model are largely "cured" by the partial-adjustment specification. In contrast, when all model assumptions are examined this dynamic specification is found to have serious misspecifications.[19] These very different conclusions illustrate the pitfalls associated with interpreting a single misspecification test in isolation and highlight the benefits of conducting a full battery of tests.

Testing Theoretical Restrictions

Because Model B is statistically adequate, valid tests of restrictions implied by consumer theory can be conducted.[20] A test (Rao) of homogeneity leads to a rejection ($F_{2,21}$ = 13.04; p-value = 0.0002). Next, homogeneity and symmetry were imposed and jointly rejected ($F_{3,42}$ = 7.02; p-value = 0.0006). Thus, neither of these theoretical restrictions is supported by the data.

The rejection of the restrictions from demand theory is not surprising. The data are constructed using total disappearance of meat from national stocks and average retail prices. There is good reason to doubt that such data are consistent with the theory of a representative utility-maximizing consumer. Problems such as aggregation over consumers, together with the measurement problems inherent in disappearance data, would invalidate theoretical restrictions.

[18] It is for this reason that Spanos uses the term "a statistically *adequate* model" rather than "*the statistical model.*"
[19] Although not reported here, a first-difference ALIDS model was estimated and found to exhibit similar problems.
[20] Parametric tests of theory can always be conducted, however, unless a valid statistical model is used as a basis of such tests, the underlying distributions of the test statistics are unknown.

Simultaneity problems, leading to biased and inconsistent estimates of the theoretical parameters of interest, could also potentially explain why the theoretical restrictions are rejected. LaFrance, for example, shows that (meat) expenditures will not be exogenous under reasonable assumptions, and Eales and Unnevehr (1993) question the exogeneity of prices. To examine these issues carefully and in a statistically coherent manner, one should first specify a system of supply and demand equations, solve for the reduced form, ensure that the reduced form is statistically adequate, and then conduct a Durbin-Wu-Hausman (DWH) test. Such an approach would ensure the validity of the DWH test (*Simultaneous Systems of Equations* above; Spanos 1986, pp. 653–54). Although desirable, this exercise is beyond the scope of this article. As a second-best solution, we examine the two simultaneity issues raised above by conducting two separate DWH tests using Model B.[21] The endogeneity of expenditures is examined using $\log(CDI_t/P_t)$ as an instrument for $\log(x_t/P_t)$, where CDI_t denotes per capita disposable income (Edgerton), and the endogeneity of all prices and expenditures is examined using the instruments employed by Eales and Unnevehr (1993).[22] The F-statistics (Rao approximation) for these two tests are $F_{2,20} = 0.644$ (p-value = 0.536) and $F_{8,34} = 2.201$ (p-value = 0.052), indicating little evidence of simultaneity.[23]

Implications for Meat Demand

Even though the theoretical restrictions of homogeneity and symmetry are rejected, the implied elasticities of Model B, the statistically adequate model, can be interpreted as percentage changes in aggregate disappearance resulting from changes in the independent variables. The unrestricted coefficient estimates, their t-statistics, and the implied elasticities at the data means are presented in table 3. The point estimates of the price and expenditure elasticities are within the range of estimates reported elsewhere (Smallwood, Haidacher, and Blaylock, pp. 108–11). All own-price elasticities (conditional on meat expenditures) are negative, and the beef own-price and expenditure elasticities are most elastic.

The concern for cholesterol (*CHOL*) effect on consumption (disappearance) is striking. Although the cholesterol elasticities at the mean of the data (table 3) are low, the same elasticities evaluated at the 1988 cholesterol awareness level (and the means of the other variables) are much higher. The 1988 cholesterol elasticities for beef, pork, and chicken are –0.08, 0.08, and 0.12, respectively. Increases in cholesterol awareness negatively affect the consumption of beef and positively affect pork and chicken consumption, the largest being for chicken consumption.

Increased cholesterol awareness has a substantially greater impact on predicted shares of meat expenditures. If prices and income are held at their mean values, and *CHOL* varies from its low in 1960 to its high in 1988, the predicted share of beef decreases 8.5% (0.575

[21]System extensions of the artificial regression DWH test outlined in Davidson and MacKinon (p. 239, eq. 7.62) are used. Note that by following this second-best approach, we cannot be sure that our DWH tests are statistically valid (Spanos 1986, pp. 653–54).

[22]The instruments used by Eales and Unnevehr include (logs of) price of corn (calendar year), average beef carcass dressed weight, pork carcass fat removed per 100 pounds, broiler-feed conversion ratio, 90-day Treasury Bill yields, an energy price index, meat-packing wages, price of nonfoods, and per capita personal consumption expenditures, as well as, a linear time trend. Data on these instruments, originally obtained from Eales and Unnevehr, were "updated" to include data on these same variables in 1960 and 1961. All instruments are used except the corn price series which was replaced with a (complete) season average price series found in *Agricultural Statistics*.

[23]The theoretical argument that consumer meat demands and consumer expenditures on meat are simultaneously determined is sound. However, using these aggregate disappearance data and an arguably ad hoc set of instruments, we do not find strong evidence of simultaneity.

Table 3. Parameter Estimates and Elasticities from Static ALIDS Model with Structural Change

Variable	Beef Equation		Pork Equation		Chicken Equation	
	Parameter Estimate[a]	Elasticity[b]	Parameter Estimate	Elasticity	Parameter Estimate	Elasticity
Constant	−1.150		1.621		0.529	
	(−4.81)		(7.13)		(5.07)	
log Beef Price	0.082	−1.138	−0.015	−0.318	−0.067	−0.221
	(4.12)		(−0.79)		(−7.71)	
log Pork Price	0.016	−0.121	−0.002	−0.813	−0.014	0.075
	(0.88)		(−0.14)		(−1.70)	
log Chicken Price	−0.010	−0.049	−0.050	−0.114	0.060	−0.420
	(−0.37)		(−1.97)		(5.12)	
log(*X*/*P*)	0.288	1.447	−0.211	0.424	−0.076	0.402
	(6.68)		(−5.17)		(−4.06)	
CHOL[c]	−0.110	−0.012	0.072	0.014	0.038	0.022
	(−2.97)		(2.04)		(2.37)	
WWOM	−0.435	−0.393	0.154	0.361	0.281	1.409
	(−4.05)		(1.51)		(5.99)	
$\bar{R}^2$	0.895		0.863		0.951	

[a]The *t*-statistics are in parentheses.
[b]Because a linear approximation ALIDS model is used, the corrected price elasticity formulae suggested by Green and Alston are used. In the case of the expenditure elasticities and the elasticities with respect to *CHOL* and *WWOM*, cases not covered by Green and Alston, the formulae are derived in a similar fashion. The elasticities are calculated as follows:

(*a*) Price elasticities: $\varepsilon_{ij} = -\delta_{ij} + \frac{\gamma_{ij}}{w_i} - \frac{\beta_i}{w_i}[w_j + \sum_k w_k \log P_k(\varepsilon_{kj} + \delta_{kj})]$;

(*b*) Expenditure elasticities: $\eta_i = 1 + \frac{\beta_i}{w_i} - \frac{\beta_i}{w_i}[\sum_k w_k \log P_k(\eta_k - 1)]$;

(*c*) Cholesterol elasticities: $C_i = \delta_{ic}\frac{CHOL}{w_i} - \frac{\beta_i}{w_i}\sum_k C_k w_k \log P_k$;

(*d*) Proportion of women working elasticities: $\phi_i = \delta_{ip}\frac{WWOM}{w_i} - \frac{\beta_i}{w_i}\sum_k \phi_k w_k \log P_k$,

where δ_{ij} = Kronecker delta; β_i, γ_{ij}, δ_{ic} and δ_{ip} are model parameters. The parameter δ_{ic} (δ_{ip}) is the coefficient of *CHOL* (*WWOM*) for the *i*th equation.
[c]Circulation in millions. Reported coefficient estimate is actual times 1,000.

to 0.526), while the predicted shares of chicken and pork increase 9.8% (0.317 to 0.348) and 15.6% (0.109 to 0.126), respectively. Thus, changes in *CHOL* alone have large predictable impacts on budget shares.

Changing cholesterol awareness also affects the estimated price and expenditure elasticities, although in a minor way. All three goods become slightly more expenditure elastic as cholesterol awareness increases, while holding prices and expenditures constant. The largest changes occur in the chicken and pork elasticities, with the chicken elasticity increasing by 20% (from 0.387 to 0.469) and pork by 14.6% (from 0.414 to 0.468) from the low to high range of information about cholesterol.[24] Own-price elasticities changed by less than the expenditure elasticities. The only sizable change occurs in the own-price elasticity of chicken, which declines from – 0.40 to – 0.48 (down 17.3%) as cholesterol awareness increases.

Increases in married females in the labor force also tend to lower beef demand and raise the consumption of chicken. Before drawing specific conclusions based on this variable, however, it is important to point out that 98% of the variability in *WWOM* (measured around a constant mean) can be modeled by a linear time trend. The significance of this variable may thus reflect the importance of a variety of factors that have trended over time in a similar fashion. The factors captured by this trending variable are important determinants of the demand for meats and have their largest impact on the demand for chicken. The elasticity suggests that a 1% increase in the proportion of married women working leads to a 1.4% increase in chicken consumption.

Conclusions

In this article, we illustrate how the single-equation misspecification testing strategy advocated by McGuirk, Driscoll, and Alwang can be extended to systems of equations. This extended approach is applicable to any system of equations, simultaneous or not. We outline how the correspondence between a simultaneous system and its reduced form can be used to examine the statistical assumptions underlying a structural model. A complete battery of individual- and joint-system misspecification tests for linear models is proposed. These system tests differ from usual single-equation tests in that they take into account information in, and interactions between, all equations in the system.

We then demonstrate the proposed approach to system misspecification testing by modeling the annual demand for beef, pork, and chicken in the United States. In doing so, we show how the proposed misspecification testing approach can be used to disentangle issues regarding structural change and other forms of model misspecification.

The misspecification tests reveal that the static ALIDS model of U.S. meat consumption is misspecified for the data we use and that the problems are likely to be related to unmodeled dynamics and/or changing parameters. The fact that we conclusively find model misspecifications with a small sample indicates that the misspecifications may be serious. Any tests of theory based on this or similarly misspecified models, including those of simultaneity, homogeneity, and/or symmetry, have no statistical validity.

When changes in meat consumption are expressed as a function of concern for cholesterol and changes in the labor force, a statistically adequate model is found. The results suggest that increases in cholesterol awareness and the proportion of married working females have significantly depressed beef consumption and caused people to switch to chicken and pork.

An interrelated partial-adjustment ALIDS model apparently "cures" the autocorrelation problems found in the static model but is still not statistically adequate. Other problems causing model misspecification persist. These results highlight the importance of conducting a full battery of misspecification tests on respecified models.

---

[24]To calculate these changes, prices and expenditures are held constant and *CHOL* is varied from its low to its high value.

The use of misspecification tests illustrated here has wide applicability for empirical modeling. As long as we use statistics to draw conclusions from econometric models, we owe it to ourselves and those who use our work to ensure, as best we can, that our models are appropriate. Without doing so, results will be at best suspicious and at worst completely erroneous. While there is no exact formula on how to use the tests advocated here to diagnose sources of misspecification, we illustrate how a careful modeler can find some very useful information in them.

[*Received May 1994; final version received October 1994.*]

References

Alessie, R., and A. Kapteyn. "Habit Forming and Interdependent Preferences in the Almost Ideal Demand System." *Econ. J.* 101(1991):404–19.

Alston, J. M., and J. A. Chalfant. "Unstable Models and Incorrect Forms." *Amer. J. Agr. Econ.* 73(1991a):1171–181.

———. "Can We Take the Con Out of Meat Demand Studies?" *West. J. Agr. Econ.* 16(1991b):36–48.

Anderson, G., and R. Blundell. "Testing Restrictions in a Flexible Dynamic Demand System: An Application to Consumers' Expenditure in Canada." *Rev. Econ. Stud.* 50(1983):397–410.

Assarsson, B., D. Edgerton, A. Hummelmose, I. Laurila, K. Rickersten, and P. Vale. *The Demand for Food in the Nordic Countries.* Forthcoming, 1995.

Brown, D. J., and L. F. Schrader. "Cholesterol Information and Shell Egg Consumption." *Amer. J. Agr. Econ.* 72(1990):548–55.

Capps, O., Jr., and J. D. Schmitz. "A Recognition of Health and Nutrition Factors in Food Demand Analysis." *West. J. Agr. Econ.* 16(1991):21–35.

Chavas, J. P. "Structural Change in the Demand for Meats." *Amer. J. Agr. Econ.* 65(1983):148–53.

Christiano, L. J. "Searching for a Break in GNP." *J. Bus. and Econ. Statis.* 10 (1992):237–50.

Dahlgran, R. A. "Changing Meat Demand Structure in the United States: Evidence from a Price Flexibility Analysis." *N. Cent. J. Agr. Econ.* 10(1988):165–75.

D'Agostino, R. B., A. Belanger, and R. B. D'Agostino, Jr. "A Suggestion for Using Powerful and Informative Tests of Normality." *Amer. Statistician* 44(1990):316–21.

Davidson, R., and J. G. MacKinnon. *Estimation and Inference in Econometrics.* New York: Oxford University Press, 1993.

Deaton, A., and J. Muellbauer. "An Almost Ideal Demand System." *Amer. Econ. Rev.* 70(1980):312–26.

Eales, J. S., and L. J. Unnevehr. "Demand for Beef and Chicken Products: Separability and Structural Change." *Amer. J. Agr. Econ.* 70(1988):521–32.

———. "Simultaneity and Structural Change in U.S. Meat Demand." *Amer. J. Agr. Econ.* 70(1993):259–68.

Edgerton, D. "On the Estimation of Separable Demand Models." *J. Agr. and Resour. Econ.* 18(1993):141–46.

Engle, R. F. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50(1982):987–1007.

Godfrey, L. G. *Misspecification Tests in Econometrics.* Cambridge: Cambridge University Press, 1988.

Green, R., and J. Alston. "Elasticities in ALIDS Models." *Amer. J. Agr. Econ.* 72(1990):442–45.

Heien, D., and C. Durham. "A Test of the Habit Formation Hypothesis Using Household Data." *Rev. Econ. and Statis.* 73(1991):189–99.

LaFrance, J. T. "When is Expenditure 'Exogenous' in Separable Demand Models?" *West. J. Agr. Econ.* 16(1991):49–62.

Maddala, G. S. *Introduction to Econometrics.* New York: Macmillan Publishing Company, 1992.

Mardia, K. V., and K. Foster. "Omnibus Tests for MultiNormality Based on Skewness and Kurtosis." *Communications in Statistics, Theory and Methods* 12(1983):207–21.

McGuirk, A. M., P. Driscoll, and J. Alwang. "Misspecification Testing: A Comprehensive Approach." *Amer. J. Agr. Econ.* 75(1993):1044–55.

Moschini, G., and K. D. Meilke. "Structural Change in Meat Demand." *Amer. J. Agr. Econ.* 71(1989):253–61.

Pollack, R. A., and T. J. Wales. "Demographic Variables in Demand Analysis." *Econometrica* 49(1981):1533–558.

Rao, C. R. *Linear Statistical Inference and Its Applications,* 2nd ed. New York: John Wiley and Sons, 1973.

Reader's Guide to Periodic Literature. Bronx NY: H.W. Wilson Company. Various issues.

Seber, G. A. F. *Multivariate Observations.* New York: John Wiley and Sons, 1984.
Smallwood, D. M., R. C. Haidacher, and J. R. Blaylock. "A Review of the Research Literature on Meat Demand." In *The Economics of Meat Demand: Proceedings of the Conference on the Economics of Meat Demand,* ed., R. C. Buse, Charleston, 1989.
Spanos, A. *An Introduction of Modern Econometrics.* New York: Cambridge University Press, forthcoming, 1995a.
———. "On Theory Testing in Econometrics: The Case of the Efficient Market Hypothesis." *J. Econometrics* 67(1995b), forthcoming.
———. "The Simultaneous-Equations Model Revisited: Statistical Adequacy and Identification." *J. Econometrics, Annals 1990–92: Contributions to Econometric Methodology in Honor of T. W. Anderson.* 44(1990):87–105.
———. "The Specification Error Argument Revisited." Discus. Pap. No. E89-05-06, Dept. of Econ., Virginia Tech, October, 1989.
———. "Error-Autocorrelation Revisited: The AR(1) Case." *Econometric Rev.* 6(1987):285–94.
———. *Statistical Foundation of Econometric Modelling.* New York: Cambridge University Press, 1986.
Theil, H., and J. S. Shonkwiler. "Monte Carlo Tests of Autocorrelation." *Econ. Letters* 20(1986): 157–60.
U.S. Department of Agriculture. *Agricultural Statistics.* Washington DC: Government Printing Office. Various issues, 1960–88.
———. *Livestock and Poultry Situation and Outlook Report* (formerly *Livestock and Meat Situation*). Washington DC: Economic Research Service. Various issues, 1960–88.
U.S. Department of Labor, Bureau of Labor Statistics. *Handbook of Labor Statistics.* Bull. No. 2340. August 1989.
Zivot, E., and D. W. K. Andrews. "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis." *J. Bus. and Econ. Statis.* 10(1992):251–70.