Journal of Agricultural and Resource Economics 27(2):335–347

AIDS versus the Rotterdam Demand System: A Cox Test with Parametric Bootstrap

Alix Dameus, Francisca G.-C. Richter, B. Wade Brorsen, and Kullapapruk Piewthongngam Sukhdial

A Cox test with parametric bootstrap is developed to select between the linearized version of the First-Difference Almost Ideal Demand System (FDAIDS) and the Rotterdam model. A Cox test with parametric bootstrap has been shown to be more powerful than encompassing tests like those used in past research. The bootstrap approach is used with U.S. meat demand (beef, pork, chicken, fish) and compared to results obtained with an encompassing test. The Cox test with parametric bootstrap consistently indicates the Rotterdam model is preferred to the FDAIDS, while the encompassing test sometimes fails to reject FDAIDS.

Key words: First-Difference AIDS model, parametric bootstrap, Rotterdam model

Introduction

Functional form is an important issue in empirical production and consumption studies. Different functional forms often result in very different elasticity estimates. The two most commonly used models in demand analysis are the Almost Ideal Demand System (AIDS) and the Rotterdam model. Most researchers arbitrarily pick one model or the other, but recent interest has focused on developing proper nonnested tests of the two demand systems.

Two prominent studies have presented techniques to select between the AIDS and the Rotterdam demand systems (Alston and Chalfant; and LaFrance). Alston and Chalfant (AC) used a compound-model approach to select between the First-Difference Almost Ideal Demand System (FDAIDS) and the Rotterdam models, using U.S. meat demand data (beef, pork, chicken, and fish). They found support for the Rotterdam model. However, LaFrance pointed out that Alston and Chalfant's approach was biased and inconsistent because they had not considered endogeneity of budget shares, and their prices in the Stone index were not mean scaled. Using the same data, LaFrance conducted both a Lagrange multiplier and a likelihood-ratio test, and failed to reject either demand system. Compound-model approaches typically have correct asymptotic size, but low power (Pesaran). Thus, the failure to reject either null hypothesis may simply be the

---

Alix Dameus is a professor and head of the Agricultural Economics Department at l' Univerite' d' Etat d' Haiti; Francisca G.-C. Richter is a postdoctoral research associate at Oklahoma State University; B. Wade Brorsen is a regents professor and Jean and Patsy Neustadt Chair in the Department of Agricultural Economics at Oklahoma State University; and Kullapapruk Piewthongngam Sukhdial is a graduate student in management information systems at Oklahoma State University. Senior authorship is shared by the first two authors. The authors wish to thank Jeffrey LaFrance and Derrell Peel for supplying the data used in this study. LaFrance also provided numerous helpful comments.

Review coordinated by Gary D. Thompson.

result of using a test with low power.[1] Most of the previous nonnested tests have been developed for models having the same dependent variables (e.g., Pesaran). Coulibaly and Brorsen show that a Cox test based on the parametric bootstrap has high power, is relatively easy to use, and is applicable to any model which can be simulated. The approach appears promising as a method for selecting among nonnested functional forms in demand systems.

In this study, a Cox nonnested test with parametric bootstrap is developed to test FDAIDS versus Rotterdam demand systems. The test is then used to determine whether the Rotterdam or the FDAIDS is preferred for U.S. meat demand. A difficulty in using the parametric bootstrap is in simulating quantities from the Rotterdam model. The approach adopted is based on a Taylor's series expansion similar to the approach of Kastens and Brester.

Tomek's recommendations on how to make research more cumulative are followed here. Tomek suggests using both the data and methods from past research. In this way, it is possible to determine whether differences in results are due to different data or different methods. LaFrance uses Moschini and Meilke's (MM's) 1967–1988 data[2] on U.S. meat demand that includes four commodities: beef, pork, chicken, and fish. The updated data have a 1970–1997 time span, come from a different source, and do not include fish.[3] Our tests are performed using MM's data set with and without fish, as well as with the updated data set.

Nonnested Hypothesis Tests

Nonnested hypothesis tests select between two regression models where one model cannot be written as a special case of the other. In such a case, the models are said to be nonnested. Suppose we have two nonnested models, A and B, with the same set of explanatory variables to choose from using the same set of data. To test that model A is the true model, the nonnested hypotheses for the two models can be written in the following general form:

$$\text{(1)} \qquad H_0: f_i(y_{it}) = \mathbf{X}_t'\beta_{0i} + u_{0it}, \quad i = 1, \ldots, n \text{ (model A)},$$

$$\text{(2)} \qquad H_1: g_i(y_{it}) = \mathbf{X}_t'\beta_{1i} + u_{1it}, \quad i = 1, \ldots, n \text{ (model B)},$$

where there are n goods, and thus n equations. Observations are indexed with $t = 1,\ldots,T$. The variable y_{it} is quantity of the ith good for period t, $\mathbf{X}_t$ is a vector of explanatory variables, β_{0i} and β_{1i} are parameter vectors under the null and alternative hypotheses, and u_{0it} and u_{1it} are the error terms under the null and alternative hypotheses. The two approaches considered to select between nonnested models are the encompassing test and the Cox test, discussed below.

[1] Note that the titles of the papers by Alston and Chalfant and by LaFrance are misleading—i.e., the lambdas in Alston and Chalfant are not "silent" since they "speak" in favor of the Rotterdam model, and the lambdas in LaFrance do not "bleat" since the low power of the LaFrance approach leads to no significant difference between the two approaches and thus it is silent about which approach is preferred.

[2] In fact, the data used by LaFrance are the same data used by Alston and Chalfant.

[3] The fish data are an aggregate of a wide variety of high-value and low-value products, and thus the fish data are widely considered unreliable.

Encompassing Test

The encompassing test is based on a composite model obtained by forming a linear combination of the two models in the null and alternative hypotheses. For models A and B in equations (1) and (2), the composite model can be specified in the following way:

$$(3) \qquad (1-\lambda)f_i(y_{it}) + \lambda g_i(y_{it}) = \mathbf{X}_t'\beta_i + u_{it},$$
$$\beta_i = (1-\lambda)\beta_{0i} + \lambda\beta_{1i},$$
$$u_{it} = (1-\lambda)u_{0it} + \lambda u_{1it},$$

where $i = 1, ..., n$, and $t = 1, ..., T$. The parameter λ linearly combines the two models. Note that in the empirical section, additional restrictions are imposed upon the coefficients.

Testing that model A is the true model is equivalent to testing that the parameter λ is equal to zero. Similarly, testing that model B is the true model corresponds to a test of λ equal to 1. Because the model is nonlinear in the parameters, a likelihood-ratio test is used to test the null hypotheses. Greene (p. 301) argues that the encompassing test does not really distinguish between the null and the alternative hypotheses, but rather distinguishes between the alternative and a hybrid model.

Cox Test and Parametric Bootstrap

The Cox test in its generic version proposed by D. R. Cox is based on the log-likelihood ratio of two models under consideration. In our example of the two models A and B, the log-likelihood ratio statistic under the null hypothesis can be computed as the difference between the log-likelihood values of models A and B. In general, the Cox test statistic has the following representation in testing the null hypothesis H_0 against H_1:

$$(4) \qquad T_0 = L_{01} - E_0(L_{01}),$$

where $L_{01} = L_0(\hat{\theta}_0) - L_1(\hat{\theta}_1)$ is the difference in estimated maximum log likelihoods (i.e., the log of the likelihood ratio) under H_0 and H_1; $E_0(L_{01})$ is the expected value of L_{01} under H_0; and $\hat{\theta}_0$ and $\hat{\theta}_1$ are the maximum-likelihood parameter estimates of the null and the alternative models, respectively. T_0 is asymptotically distributed with mean zero and variance v_0^2 under H_0 (Cox). Similarly, the test statistic for testing H_1 against H_0 would be $T_1 = L_{10} - E_1(L_{10})$.

The difficulty in implementing the Cox test resides in obtaining analytical formulas for $E_0(L_{01})$ and v_0^2. Pesaran derived analytical results for the linear regression models with the same dependent variable. Both Pesaran and Deaton, and Pesaran and Pesaran developed a version of the Cox test with transformed dependent variables such as needed for testing linear versus log-linear models. However, their test statistics exhibit incorrect sizes in small samples.

More recently, Coulibaly and Brorsen showed that a Cox test associated with a parametric bootstrap approach gives a test statistic with correct size and high power, even in small samples. The test statistic is the likelihood ratio of the two models, and the parametric bootstrap is used to estimate the distribution of this test statistic under the null.[4] With the parametric bootstrap, Monte Carlo samples are generated using the parameters estimated under the null hypothesis. Samples are generated with the same number of observations as the original data.

The hypothesis test is performed by computing a p-value as the percentage of simulated likelihood-ratio statistics that are less than the likelihood ratio computed from the actual data. This p-value is calculated using the actual and the generated data in the following way (Coulibaly and Brorsen):

$$(5)\qquad p\text{-}value = \frac{\left(numb\left[L_0(\hat{\theta}_{0j}, \mathbf{y}_j) - L_1(\hat{\theta}_{1j}, \mathbf{y}_j) \le L_{01}\ \forall j = 1, ..., N\right] + 1\right)}{N + 1},$$

where $numb[\cdot]$ stands for the number of realizations for which the specified relationship is true; N is the number of samples of size T generated under each model; L_{01} is the actual value of the log-likelihood ratio; and $L_0(\cdot)$ and $L_1(\cdot)$ are the values of the log-likelihood function with the generated data under the null and the alternative hypotheses, respectively. The value of one is added to the numerator and denominator as a small-sample correction. This p-value estimates the area to the left of the Cox test statistic L_{01}. A small area indicates that the statistic is far from the mean according to H_0, so we can reject the null hypothesis. In other words, a small p-value indicates rejection of the null hypothesis.

Selecting Between the AIDS and the Rotterdam Models for U.S. Meat Demand

The Selected Models

Previous studies by Alston and Chalfant (AC) and LaFrance used encompassing tests to select between the AIDS and the Rotterdam models for U.S. meat demand. For the Rotterdam, AC present two alternative models with seasonal dummy variables. One uses the Divisia volume index as real income, and the other uses deflated expenditures with the Stone index. Based on their findings, these two specifications give nearly the same parameter estimates. For the AIDS model, AC use four alternative specifications of the first-difference model with seasonal dummy variables. For the purpose of this study, only the standard specifications which AC denote Rotterdam II and AIDS VI are considered.[5]

The first-difference linearized version of the AIDS model with quarterly seasonal dummies and real expenditure variables (using the Stone index), presented as AC's model VI with the time subscripts suppressed, is designated by:

$$(6)\qquad \Delta s_i = \tau_i + \sum_{k=1}^{4} \theta_{ik} D_k + \sum_{j=1}^{4} \gamma_{ij} \Delta\ln(p_j) + \beta_i [\Delta\ln(x) - \Delta\ln(P)], \qquad i = 1, ..., 4.$$

In this model, s denotes budget share; D_k's are quarterly seasonal dummy variables; p_j is price of good j; x is the total expenditure on the n goods; τ, θ, γ, and β are parameters; Δ is a first-difference operator; and P is the Stone index.

---

[4] The likelihood ratio is not an asymptotically pivotal statistic, yet Coulibaly and Brorsen find it provides tests at least as good as a parametric bootstrap based on Pesaran and Pesaran's asymptotically pivotal statistic. As argued by Maasoumi, "asymptotically pivotal statistics are, in general, neither necessary nor sufficient for good bootstrap performance" (p. 86).

[5] AC were unable to test these exact specifications. Thus, their approach is to test the Rotterdam versus "almost AIDS," and to test the AIDS versus "almost Rotterdam."

The Rotterdam model II has the following specification in AC's paper:

(7) $$\bar{s}_i \Delta\ln(y_i) = \tau_i + \sum_{k=1}^{4} \theta_{ik} D_k + \sum_{j=1}^{4} \gamma_{ij} \Delta\ln(p_j) + \beta_i \left[\Delta\ln(x) - \sum_{j=1}^{4} \bar{s}_j \Delta\ln(p_j) \right],$$

where $\bar{s}_j$ is the average budget share of good j (four goods are considered), y_i denotes quantity of good i, and all the other variables are as defined above. The term in brackets is real expenditure.

Encompassing Tests and Selection Between the AIDS and Rotterdam Models

The studies by AC and LaFrance are based on encompassing tests, with a difference in estimation methods and in the representation of the compound model equation. AC do not account for endogeneity of budget shares. To correct for endogeneity, LaFrance restates the models and shares as explicit functions of quantity, and estimates the parameters with full-information maximum likelihood. Also, LaFrance recognizes the lack of invariance of Stone's price index to units of measurement, and thus scales the prices in that index by the means. Note rescaling is used only in the calculation of Stone's index.

AC present two compound models: one to test the Rotterdam model in equation (7) against an approximate FDAIDS, while the other is used to test the linearized version of the FDAIDS in equation (6) against an approximate Rotterdam. AC's compound models are:

(8) $$(1-\lambda)\bar{s}_i \Delta\ln(y_i) + \lambda \Delta s_i = \tau_i + \sum_{k=1}^{4} \theta_{ik} D_k + \sum_{j=1}^{4} \gamma_{ij} \Delta\ln(p_j) + \beta_i \left[\Delta\ln(x) - \sum_{j=1}^{4} \bar{s}_j \Delta\ln(p_j) \right],$$

(9) $$(1-\lambda')\Delta s_i + \lambda' \bar{s}_i \Delta\ln(y_i) = \tau_i + \sum_{k=1}^{4} \theta_{ik} D_k + \sum_{j=1}^{4} \gamma_{ij} \Delta\ln(p_j) + \beta_i \left[\Delta\ln(x) - \Delta\ln(P) \right],$$

where λ' linearly combines the two models under H_1. Equation (8) compounds AC's Rotterdam II with their FDAIDS IV, which is an approximation to the FDAIDS. This approximation leads to both models having a common right-hand side, and thus the linear combination is only applied with the left-hand-side variables. In this compound model, testing $\lambda = 0$ is equivalent to testing that the Rotterdam model is the true model. Equation (9) compounds AC's FDAIDS VI with their approximate Rotterdam; again, this allows combining only the left-hand side of both models. Testing $\lambda' = 0$ corresponds to testing that FDAIDS is the true model.

LaFrance conducted an encompassing test based on a likelihood-ratio statistic using a compound model like the one presented in (3) which combines all aspects of the two models. LaFrance's compound model is as follows:

$$(10)\qquad (1-\lambda)\bar{s}_i(y_i)\Delta\ln(y_i) + \lambda\Delta s_i(y_i) = \tau_i + \sum_{j=1}^{4}\theta_{ij}D_j + \sum_{j=1}^{4}\gamma_{ij}\Delta\ln(p_j) + (1-\lambda)\beta_{i0}\left(\Delta\ln(x) - \sum_{j=1}^{4}\bar{s}_j(y_i)\Delta\ln(p_j)\right) + \lambda\beta_{i1}\big(\Delta\ln(x) - \Delta\ln(P)\big),$$

where all the elements are as previously defined.

As LaFrance points out, AC do not account for the endogeneity of budget shares. For instance, in equation (8), share s_i, which is a function of quantity y_i, appears on the right-hand side of the equation. Also, LaFrance's compound model takes into account both the AIDS and the Rotterdam model's expenditure terms, whereas AC's models approximate these variables. The estimable version of equation (10) for meat demand is specified as:

$$(11)\qquad u_{it} = (1-\lambda)\frac{1}{2}\left[\left(\frac{p_{it}}{x_t}\right)y_{it} + s_{it-1}\right]\log\left(\frac{y_{it}}{y_{it-1}}\right) + \lambda\left[\left(\frac{p_{it}}{x_t}\right)y_{it} - s_{it-1}\right] - \tau_i - \sum_{j=1}^{3}\theta_{ij}\left(D_{jt} - D_{4t}\right) - \sum_{j=1}^{3}\gamma_{ij}\log\left(\frac{p_{jt}p_{4t-1}}{p_{jt-1}p_{4t}}\right) - \beta_{i0}(1-\lambda)\left\{\log\left(\frac{x_t}{x_{t-1}}\right) - \frac{1}{2}\sum_{j=1}^{3}\left[\left(\frac{p_{jt}}{x_t}\right)y_{jt} + s_{jt-1}\right]\log\left(\frac{p_{jt}p_{4t-1}}{p_{jt-1}p_{4t}}\right) - \log\left(\frac{p_{4t}}{p_{4t-1}}\right)\right\} - \beta_{i1}\lambda\left\{\log\left(\frac{x_t}{x_{t-1}}\right) - \sum_{j=1}^{3}\left(\frac{p_{jt}}{x_t}\right)y_{jt}\log\left(\frac{p_{jt}}{p_{4t}}\right) + \sum_{j=1}^{3}s_{jt-1}\log\left(\frac{p_{jt-1}}{p_{4t-1}}\right) - \log\left(\frac{p_{4t}}{p_{4t-1}}\right)\right\}$$

for $i = 1, 2, 3$ meat commodities, and $t = 1, \ldots, T$ observations. Here, $\mathbf{u}_t = [u_{1t}, u_{2t}, u_{3t}]$ is assumed to be *i.i.d.* $N(\mathbf{0}, \Sigma)$, and symmetry requires $\gamma_{ij} = \gamma_{ji}\ \forall i \neq j$. Homogeneity and adding-up restrictions are embedded in the system of equations. Equation (11) is equivalent to LaFrance's (p. 229) equation (23), except for a typo in the latter, which shows the term $-\log(p_{4t}/p_{4t-1})$ multiplied by $(1+s_{4t-1})/2$ when $\lambda = 0$.

LaFrance's reported estimation results, however, are correct and consistent with (11). Note how the budget shares are now explicit functions of quantity. The parameters in this equation can be determined by maximum-likelihood estimation. From AC's perspective, a test of one model against the other could be conducted based on the estimated value of the parameter λ. In LaFrance's view, "a likelihood ratio test should be used to discriminate between the two competing models, rather than simply examining the t-ratio for the estimated lambda" (p. 229). Because the model is nonlinear in the parameters, a t-test, which is a Wald test, would be sensitive to the units of measurement. Therefore, we agree with LaFrance that a t-test would be a poor choice.

LaFrance's suggested approach, however, has possible weaknesses in both power and size. When λ is restricted to zero, the AIDS expenditure coefficients (β_{i1}'s) are undefined. The β_{i1}'s are called nuisance parameters because they appear under the alternative, but are undefined under the null of $\lambda = 0$. In the presence of nuisance parameters, likelihood-ratio statistics may no longer have the usual χ^2 distributions (Andrews; Davies). The only restriction imposed is on λ, because the β_{i1}'s can be any value. Yet, LaFrance assumes a $\chi^2_{[4]}$ distribution for the likelihood-ratio statistic. Thus, LaFrance considers the β_{i1}'s as also being restricted.

With linear models and in the presence of more than one nonoverlapping parameter (a parameter is nonoverlapping if it is included under one hypothesis, but not the other), encompassing tests are known to have low power (Pesaran). Pesaran's results do not apply directly here because the compound model is nonlinear and the nonoverlapping parameters are associated with nuisance parameters, but LaFrance's findings that neither model can be rejected are hardly surprising. Alternatively, a Cox test based on a parametric bootstrap can be expected to have good size and power properties (Coulibaly and Brorsen; Hall and Titterington).

A Monte Carlo study was conducted to determine if there were serious problems with the size of the test from LaFrance's compound model. Ten thousand data sets were generated from the estimated AIDS model. The AIDS and compound models were then estimated with each data set. To avoid convergence problems, a grid search over values of λ was used to select starting values. The results of the Monte Carlo study are presented in figure 1. As figure 1 shows, the distribution of the likelihood-ratio statistic closely approximates a $\chi^2_{[4]}$. Thus, the size of the test with the compound model seems fine. In a recent study, Andrews found that statistics sometimes have standard distributions even in the presence of nuisance parameters. The presence of a nuisance parameter, however, does cause the Rotterdam expenditure coefficients to be inestimable whenever the estimated λ is close to one.

Cox Test and Parametric Bootstrap with the FDAIDS and Rotterdam Models

Using the Cox test with the parametric bootstrap for selecting between the AIDS and the Rotterdam models requires the following steps: (*a*) estimate the two models under consideration using the actual data set; (*b*) based on the likelihood values of the two estimated models, compute the actual likelihood ratio of the two models; (*c*) assuming the null hypothesis model, estimate a distribution function for the original data and, based on this estimate, generate a large number of data sets of the same size; (*d*) re-estimate the two models for each of the generated samples; (*e*) compute the simulated log-likelihood ratio for each simulated data set; and (*f*) compare the true and simulated log-likelihood ratios to compute the *p*-value presented in equation (5). Steps (*c*)–(*f*) must be performed twice—once by letting one of the two models (say, FDAIDS) be the null hypothesis, and the second time, under the assumption the other model (say, the Rotterdam) is the null hypothesis.

Parametric Bootstrap and Difficulties in Data Generation

The data that must be generated in the context of the FDAIDS and Rotterdam models are the quantities. However, quantity is not explicit in the left-hand side of either the AIDS or the Rotterdam.

The approach used requires predicted quantities. First, estimates for the dependent or left-hand-side terms are generated under a normal process. Normality of residuals cannot be rejected for any of the models using the Henze-Zirkler test provided as an option in SAS's PROC MODEL (SAS Institute, Inc.). Also, the Mardia tests from the same SAS procedure fail to detect skewness or kurtosis in the system. If normality had

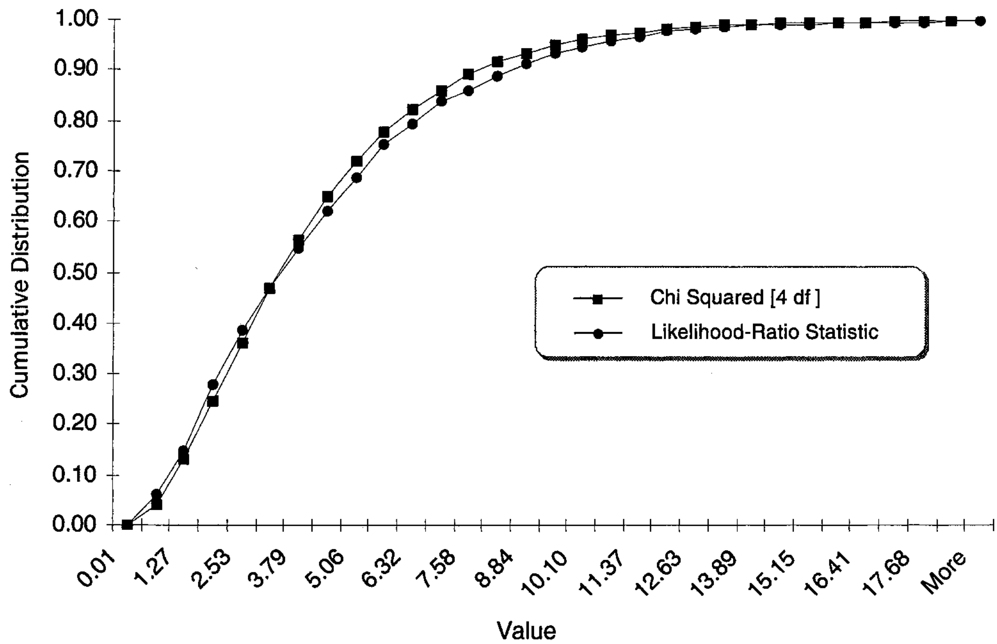

Figure 1. Cumulative distributions for a likelihood-ratio statistic of AIDS vs. compound model when data are from an AIDS model and the cumulative distribution of a chi-squared statistic with four degrees of freedom

been rejected, a nonparametric bootstrap could have been used. While the asymptotic properties of parametric and nonparametric bootstraps are the same if the parametric distribution is correctly specified, Monte Carlo experiments have shown that the numerical accuracy of the parametric bootstrap is much higher (Horowitz, p. 34).

Once left-hand-side terms are generated, quantities need to be approximated. However, it is difficult to simulate data from the Rotterdam model. "Since the Rotterdam involves a nonlinear transformation of quantity on the left-hand side, predicted or expected quantities are not immediately derived by taking the inverse functional transformation of the model-predicted left-hand side" (Kastens and Brester, p. 303). Kastens and Brester proposed a method for obtaining the expected quantities from the Rotterdam model using the predicted left-hand side (*predLHS*) and a second-order Taylor-series expansion of the dependent variable. While the derivation and final equation used here differs from that of Kastens and Brester, the principle remains the same.

We start with the predicted equation of the Rotterdam model:

$$(12)\qquad E\left[\frac{1}{2}\left(s + s_{t-1}\right)\left(\ln(y) - \ln(y_{t-1})\right)\right] = \mathbf{X}_t'\hat{\beta}_i = predLHS,$$

where the variables s and y without a subscript are current budget shares and current quantities, respectively. The dependent variable or term within the expectation operator can be approximated by a second-order Taylor series expansion around y_0, the expected value of y. Then, the expected value of this approximation can be used to approximate (12) as follows:

(13)
$$f(y) = \tfrac{1}{2}\left(py/x + s_{t-1}\right)\left(\ln(y) - \ln(y_{t-1})\right)$$
$$f(y) \approx f(y_0) + f'(y_0)(y - y_0) + \left(\tfrac{1}{2}\right)f''(y_0)(y - y_0)^2$$
$$predLHS = E(f(y)) \approx f(y_0) + f''(y_0)E(y - y_0)^2,$$

where the sample variance of y, denoted by v, is used to estimate $E(y - y_0)^2$. Thus, we can solve for y in:

(14)
$$\left(py + s_{t-1}x\right)\left(\ln(y) - \ln(y_{t-1})\right) + \left(v/2y^2\right)\left(py - s_{t-1}x\right) - 2x\,predLHS = 0.$$

Solving for y in equation (14) gives an approximation for the predicted quantity of the Rotterdam model. Predicted quantities for the AIDS model are obtained without the need of a Taylor-series approximation because the left-hand side is already a linear function of quantity. Following Kastens and Brester, predicted quantities are specified as:

(15)
$$y_0 = \frac{\left[predLHS + y_{t-1}\right]}{p}x.$$

The simulation procedures using both (14) and (15) produced quantities with moments similar to those of the actual data (table 1). The moments are not an exact match because the moments of the actual data depend on the order of the random realizations. Ten thousand random samples of meat quantities were generated for each model to construct the p-value of each test. The estimation methods incorporate the homogeneity, symmetry, and adding-up restrictions.

AIDS and Rotterdam Likelihood Functions

To use the Cox statistics, the likelihood functions of both the AIDS and the Rotterdam models must be converted to the same units. The dependent variables in the FDAIDS model are budget share differences. In the Rotterdam model, the dependent variables are the differences in the natural logarithms of quantities multiplied by average expenditure shares. The log-likelihood functions for the dependent variables in both models are transformed to log likelihoods of quantity by adding a Jacobian term (see LaFrance). Then the transformed values are compared.

Meat Demand Data

AC and LaFrance used Moschini and Meilke's (MM's) data on U.S. quantities and prices of beef, pork, chicken, and fish to select between the AIDS and the Rotterdam. The data used in their studies are quarterly per capita disappearance and retail prices of beef, chicken, pork, and fish in the United States, for the years 1967–1988.

We use the same data used by AC and LaFrance, plus an additional set of updated quarterly data on beef, pork, and chicken. The latter data set does not include fish because of the poor quality of the U.S. fish data. For comparison purposes, we also run both the encompassing and the Cox tests with parametric bootstrap on MM's data set without fish. Such an approach allows identification of the effect on the model choice of difference in method, difference in data, and differences in both data and method, as recommended by Tomek.

Table 1. Statistics on the Quantities Generated for the Bootstrap Procedure with Updated Data, 1970–1997

Statistic	Actual Data	Simulated Quantities: Rotterdam	Simulated Quantities: FDAIDS
Means:			
Beef	19.646	18.964	18.715
Pork	12.853	13.712	13.690
Chicken	16.746	17.776	17.665
Standard Deviations:			
Beef	2.373	2.218	2.244
Pork	1.095	1.755	1.726
Chicken	3.998	4.069	4.125
Sample Size	112	112	112
No. of Samples Generated		10,000	10,000

Note: Values for approximations are averages over all 10,000 samples.

Estimation Methods

The Model Procedure (PROC MODEL) in SAS with the full-information maximum likelihood (FIML) option was used to conduct the encompassing test on the data. The Interactive Matrix Language Procedure (PROC IML) in SAS and PROC MODEL were used to implement the Cox test with parametric bootstrap, and SAS Macros[6] were used to provide the necessary loops.

Results

With the 1967–1988 data including fish, we are able to replicate AC's results using iterative seemingly unrelated regression. For instance, using AC's compound model to test AIDS VI versus the (almost) Rotterdam II, we obtain 0.3579 as an estimate for λ, as compared to LaFrance's 0.36 and AC's 0.35997. LaFrance's results are also replicated in full.

The likelihood-ratio test results using LaFrance's method with three data sets are presented in table 2. Results for the MM data with fish confirm LaFrance's findings that neither the Rotterdam nor FDAIDS can be rejected. Results for the other two data sets favor the Rotterdam model, but each model has some problems in the estimation. For the MM data without fish, the estimates of λ and one of the expenditure parameters (one of the nuisance parameters) are perfectly correlated (table 2). Because the estimate of λ is close to zero, the expenditure parameter is not estimable. For the updated data set, the estimation process fails to converge for some starting values, and when it does converge, it produces a "maximized" log-likelihood value for the compound model that is lower than the corresponding value for the Rotterdam. This difficulty in obtaining convergence is another drawback of the compound model. Like AC, with the new data, we are only able to report results when the expenditure terms are equal, although we realize this is not a true test of FDAIDS versus Rotterdam. The Cox test with parametric

[6] The program is available online at http://go.okstate.edu/~brorsen/WP/cox1.SAS.

Table 2. Results of Likelihood-Ratio Test to Select Between FDAIDS and Rotterdam Models for U.S. Meat Demand, as Proposed by LaFrance

Data Set	Model	λ Estimate	Log Likelihood	p-Value[a]	Test Result
1. Moschini & Meilke's	▸ Compound Model	0.056	68.603		
(with fish)	▸ Rotterdam	(set to 0)	68.544	0.998	Fail to reject
1967–1988	▸ FDAIDS	(set to 1)	64.377	0.076	Fail to reject
2. Moschini & Meilke's	▸ Compound Model	1.8 E-05	67.580		
(w/o fish)	▸ Rotterdam	(set to 0)	66.919	0.724	Fail to reject
1967–1988	▸ FDAIDS	(set to 1)	63.226	0.033	Reject
3. Updated Data	▸ Compound Model	0.159	25.027		
(w/o fish)	▸ Rotterdam	(set to 0)	24.559	0.817	Fail to reject
1970–1997	▸ FDAIDS	(set to 1)	18.146	0.003	Reject

Notes: Estimators for λ and a nuisance parameter (expenditure term parameter) in the compound model were biased for the Moschini and Meilke data set without fish, showing that, as λ approaches zero, estimation problems arise due to the nuisance parameters. Similarly, estimators for λ and nuisance parameters in the compound model failed to converge for the updated data. Thus, we restricted the compound model by imposing equality of the coefficients on the expenditure term.

[a] The p-values are based on a χ^2 distribution.

Table 3. Statistics from the Cox Test with Parametric Bootstrap for U.S. Meat Demand

			Moschini & Meilke's 1967–1988 Data		Updated Data (w/o Fish)
Statistic	Data Set	Estimated Model	with Fish	w/o Fish	1970–1997
Log Likelihood	Actual	FDAIDS	64.377	63.222	19.015
Log Likelihood	Actual	Rotterdam	68.544	66.919	25.370
Difference			-4.167	-3.697	-6.355
Average LLV[a]	H_0: FDAIDS[b]	FDAIDS	67.441	66.700	22.576
Average LLV	H_0: FDAIDS	Rotterdam	61.557	62.831	14.598
Difference			5.884	3.869	7.978
Average LLV	H_0: Rotterdam[b]	FDAIDS	68.462	62.469	18.294
Average LLV	H_0: Rotterdam	Rotterdam	73.829	66.820	22.762
Difference			-5.367	-4.351	-4.468
p-Value[c]		H_0: FDAIDS	0.003	0.009	0.002
p-Value		H_0: Rotterdam	0.406	0.438	0.694
Test Result:			Reject FDAIDS	Reject FDAIDS	Reject FDAIDS

[a] LLV is log-likelihood value.
[b] The data sets are simulated using the model estimated with the actual data.
[c] A small p-value indicates a rejection of the null hypothesis and a large p-value indicates a failure to reject the null hypothesis.

bootstrap avoids these estimation problems because there is no need to estimate the compound model, and thus, there are no nuisance parameters present.

As reported in table 3, the Cox test rejects the FDAIDS and fails to reject the Rotterdam model for all data sets. Thus, this study gives additional evidence of the high power of the Cox test, as compared to an encompassing test. The Rotterdam model is clearly favored for use in U.S. meat demand. Tomek's approach was also advantageous because it separates effects due to differences in method from those that are due to differences in data.

Conclusions

A Cox test based on a parametric bootstrap is developed for use in testing the first-difference AIDS model versus the Rotterdam model. Parametric bootstrap tests are known to have good size and power properties, whereas encompassing tests like those used by Alston and Chalfant or LaFrance have low power.

The new approach and an approach like that suggested by LaFrance were used to select between the FDAIDS and Rotterdam models for U.S. meat demand. With the parametric bootstrap, the FDAIDS was consistently rejected in favor of the Rotterdam model. Thus, the results support using the Rotterdam model for U.S. meat demand. In addition to low power, another disadvantage of the encompassing test is that the compound model in one instance converged to a local rather than a global optimum.

The Cox test with parametric bootstrap can be used to test nonnested hypotheses involving most functional forms—for instance, a double-log demand model, the AIDS in levels, the Rotterdam, and the AIDS with different expenditure deflators. The Cox test requires more computational work than encompassing tests, but the additional effort is rewarded with a test of higher power. We recommend using the Cox test with parametric bootstrap to conduct nonnested tests among alternative demand systems.

[*Received July 2001; final received June 2002.*]

References

Alston, J. M., and J. A. Chalfant. "The Silence of the Lambdas: A Test of the Almost Ideal and Rotterdam Models." *Amer. J. Agr. Econ.* 75(May 1993):304–13.

Andrews, D. W. K. "Testing When a Parameter Is on the Boundary of the Maintained Hypothesis." *Econometrica* 69(May 2001):683–734.

Coulibaly, N., and B. W. Brorsen. "Monte Carlo Sampling Approach to Testing Nonnested Hypotheses: Monte Carlo Results." *Econometric Rev.* 18,2(1999):195–209.

Cox, D. R. "Further Results on Tests of Separate Families of Hypotheses." *J. Royal Statis. Society*, Series B 24(1962):406–24.

Davies, R. B. "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative." *Biometrika* 64(1977):247–54.

Greene, W. H. *Econometric Analysis*, 3rd ed. Upper Saddle River NJ: Prentice-Hall, 1997.

Hall, P., and D. M. Titterington. "The Effect of Simulation Order on Level Accuracy and Power of Monte Carlo Tests." *J. Royal Statis. Society*, Series B 31(1989):459–67.

Horowitz, J. L. "The Bootstrap." Work. paper, Department of Economics, University of Iowa, November 2000. Online. Available at www.biz.uiowa.edu/faculty/horowitz/papers/Bootstr_Rev2.pdf. [Accessed June 2002.]

Kastens, T. L., and G. W. Brester. "Model Selection and Forecasting Ability of Theory-Constrained Food Demand Systems." *Amer. J. Agr. Econ.* 78(May 1996):301–12.

LaFrance, J. T. "The ~~Silence~~ Bleating! of the Lambdas: Comment." *Amer. J. Agr. Econ.* 80(February 1998):221–30.

Maasoumi, E. "On the Relevance of First-Order Asymptotic Theory to Economics." *J. Econometrics* 100(2001):83–86.

Moschini, G., and K. D. Meilke. "Modeling the Pattern of Structural Change in U.S. Meat Demand." *Amer. J. Agr. Econ.* 71(May 1989):253–61.

Pesaran, M. H. "On the General Problem of Model Selection." *Rev. Econ. Stud.* 41(1974):153–71.

Pesaran, M. H., and A. Deaton. "Testing Nonnested Nonlinear Regression Models." *Econometrica* 46(1978):677–94.

Pesaran, M. H., and B. Pesaran. "A Simulation Approach to the Problem of Computing Cox's Statistic for Testing Nonnested Models." *J. Econometrics* 57(1993):377–92.

SAS Institute, Inc. *SAS/ETS User's Guide*, Version 8. Cary NC: SAS Institute, Inc., 1999.

Tomek, W. G. "Structural Econometric Models: Past and Future (with Special Reference to Agricultural Economics Applications)." In *NCR-134 Conference Proceedings: Applied Commodity Price Analysis, Forecasting, and Market Risk Management*, ed., T. C. Schroeder, pp. 1–23. Manhattan KS: Kansas State University, 1998.