Can We Take the Con Out of Meat Demand Studies?

Julian M. Alston and James A. Chalfant

Whimsy in specification choices leads to fragility of inference in econometric studies of structural change in meat demand. The literature contains a variety of results, with many contradictions, attributable largely to differences in specifications. This article reviews that literature, uses synthetic data to demonstrate the sensitivity of results to specification choices and to evaluate the power of nonparametric tests, and uses Canadian data to demonstrate a preferred approach to testing the hypothesis of structural change.

Key words: meat demand, parametric and nonparametric tests, structural change.

A fragile inference is not worth taking seriously.
—E. Leamer (1985)

In a well-known article (entitled "Let's Take the Con Out of Econometrics"), Leamer (1983) cautioned economists against drawing inappropriate inferences from their econometric work. In a nutshell, he objected to "whimsy" (in relation to specification choices) and "fragility" (in relation to sensitivity of results to those choices). He concluded (p. 43) that "If we are to make effective use of our scarce data resource, it is therefore important that we study fragility in a much more systematic way."

The literature on demand for food would provide little comfort to Leamer.[1] It abounds with whimsical specification choices and fragile results and, in most cases, scant attention is paid to these issues. That specification choices affect results is fairly obvious and is certainly not new [e.g., Chavas (1989) and the other papers in Buse]. That results are therefore always conditional on specification choices is equally obvious but almost invariably ignored (at best it might be mentioned in passing) in studies of demand. This is noticeably so in the large number of recent studies of structural change in demand for meat.

A typical study runs as follows. First, it is noted that the specification of the functional form can influence results and, in consideration of this, a flexible functional form is used—but usually only one functional form is tried. After estimating the parameters of the system, diagnostic tests are performed. Rejection of the model is interpreted as a rejection of stable preferences (with an appeal to demographic shifts or health concerns). It is rare for such studies to examine whether an alternative demand system would have resulted in different conclusions. In part this may be due to the widespread (and questionable) notion that flexible functional forms do a good job of approximating unknown models, but it is more probably due to a preference for avoiding the appearance of mining the data. A typical concluding comment in such a study is to recommend that the beef industry spend money promoting beef and doing research into product development.

For instance, Moschini and Meilke (1989) estimated an almost ideal demand system and concluded that

> the observed meat consumption patterns of the last two decades cannot be fully explained by the dynamics of

The authors are associate professors in the Department of Agricultural Economics, University of California, Davis, and the Department of Agricultural and Resource Economics, University of California, Berkeley, respectively.

This is Giannini Foundation Paper No. 972. An earlier version of this paper was presented at the Western Agricultural Economics Association Meetings, Vancouver, British Columbia, August 1990.

The authors thank Oscar Burt, Colin Carter, Jean-Paul Chavas, John Constantine, Richard Green, Art Havenner, George Judge, Jeff LaFrance, and *Journal* referees for comments on drafts. Special thanks are due to Linda Robbins and Mike Wohlgenant for providing data sets.

[1] His subsequent article (Leamer 1985) also is pertinent.

> prices and income ... this movement toward an increased importance of white meats further supports the idea that dietary concerns are partly responsible for the perceived changes in meat consumption patterns. The implications of this are particularly relevant for the beef industry, calling possibly for a quality adjustment in production and increased efforts in promotion and marketing. (p. 260)

In our view, conclusions such as these are not adequately supported by the data or the econometric work in the literature. Any such conclusions ought to be qualified much more clearly as being conditional upon specification choices (e.g., as done by Choi and Sosin). In particular, the conclusions are subject to the untested hypothesis that the estimated demand system is of the correct functional form (or, at least, that results are insensitive to that choice). Economic theory is not informative about functional forms. Functional form choices are whimsical, in the sense used by Leamer. Equally clearly (witness the abundant studies of U.S. meat demand that use different models and get different results), the results are likely to be fragile, yet we do not know of any previous study that has systematically analyzed the sensitivity of structural change tests to specification choices.[2]

In an earlier paper we suggested that the use of a nonparametric approach, using revealed preference axioms, could avoid the problem of functional form as a joint hypothesis. An application of Varian's nonparametric approach to meat consumption data from Australia and the United States indicated that

> the data from both countries could have been generated by stable preferences. Therefore, any conclusions from these data sets that tastes have changed must come in the form of restrictions on the nature of these demand systems (e.g., to be of the almost-ideal form). The data alone do not indicate changes in preferences. Relative prices, instead, can account for the observed shifts in consumption patterns. (Chalfant and Alston, p. 406)[3]

Two concerns with the nonparametric approach have been raised. One is that the power of the tests is unknown (which is not to say it is low).[4] This has two aspects. First, data generated by a particular demand system subject to structural change (e.g., a translog) conceivably may be consistent with a stable form of some other system (e.g., a Rotterdam model). Second, the nature of the data may mean that the nonparametric tests are incapable of detecting structural change (e.g., when income growth dominates relative price changes).[5] In either of these cases it must be noted that a parametric approach faces the same challenge as does a nonparametric one. These are cases where the data alone are unlikely to tell us what we want to know. A second drawback with the nonparametric approach is that—in the case where the data are consistent with stable preferences—it doesn't give any clues as to the nature of those stable preferences, how to identify the functional form and parameters of the demand equations, nor whether the results from estimation will be plausible.

In this article we address some of these issues. The broader context of the article is the general area of specification choice, interpretation of results, and inference in demand analysis, but we focus in particular on testing for structural change in demand for meat. We begin with a brief review of literature in which we identify some loose connections between specification choices and results (e.g., choosing almost ideal forms results in structural change findings) and between results and specification choices (e.g., finding autocorrelated errors leads to the adoption of dynamic specifications). Following that review, we report empirical work relating primarily to Canadian meat consumption data.

The empirical work comprises three separate parts. The first illustrates how easily specification errors can lead to erroneous conclusions. We generated data using a stable double-log model, fit linear demand equations to the generated data, and then tested for stability of the model in terms of dynamics (autocorrelation) and discrete structural change (Chow tests). We then reversed the process, fitting double-log equations to data from linear demand equations. The result was surprisingly strong. Apparently minor specification errors led to highly significant tests for structural change.

---

[2] White and others have examined the properties of parameter estimates when incorrect models are estimated.

[3] These results reinforce the findings of several studies using the parametric approach (e.g., Wohlgenant) that found functional forms capable of explaining U.S. meat consumption patterns with a stable set of preferences.

[4] Chalfant and Alston (pp. 403–06) discuss this issue at some length.

[5] As discussed by Chalfant and Alston; Landsburg; Varian (1982); and Thurman, when income growth is large (compared to relative price variation), the consumption bundle in every period will be revealed to be preferred to those in all previous periods and effectively there will be no comparable data points.

The second part concerns nonparametric tests. First, the nonparametric test was applied to Canadian meat consumption data. The result indicates that, as we found previously for Australia and the United States, Canada's meat consumption data could have been generated by a stable system of well-behaved demand equations. Then, to address the issue of the power of nonparametric tests, we conducted some Monte Carlo experiments using data generated by a system of Cobb-Douglas preferences (with actual Canadian meat prices). The results indicate that the nonparametric test is capable of detecting moderate structural changes unless there is relatively large growth in total expenditures. However, in these experiments, the power of the nonparametric test (i.e., the frequency of correctly rejecting the hypothesis of stable preferences) was disappointingly low.[6] We cannot say to what extent the results would carry over to different data sets. Further work is needed to clarify the relation between data characteristics and power of both nonparametric and parametric tests.

Finally, several commonly used parametric specifications were estimated using Canadian meat consumption data to illustrate the difference in results among models and (with less confidence) to attempt to identify the stable set of equations that generated the data. We tested for significant trends, first-order autocorrelation, or both. Two single-equation models (linear and double log) and four demand systems were tried. The demand systems include two versions of the linear approximate almost ideal demand system (in standard form and in first-difference form) and two versions of the Rotterdam model (the absolute price formulation and the relative price formulation). The results reinforce the observation that conclusions about structural change are sensitive to the functional form estimated.

A Mixture of Past Results

There has been considerable investment by the profession in studies of structural change in the demand for meat in the United States and elsewhere, mostly published within the last five years. These studies have varied in that they have used (*a*) a variety of data (different time periods, frequencies of observations, places, commodity aggregations, and commodities), (*b*) a variety of model specifications (different functional forms for demand equations, single-equation or systems estimation, and choices about imposing parametric restrictions—e.g., for homotheticity or separability), (*c*) different criteria (violation of homogeneity or symmetry restrictions, presence of apparent dynamic influences or significant unexplained trends, or unstable model parameters), and (*d*) an impressive variety of statistical tests.

The premise of the studies is mostly the same: there has been a shift of consumption away from red meat (especially beef) and toward white meat (especially chicken) that reflects a change in consumer preferences due to increased health consciousness.[7] The alternative explanation is that changes in consumption patterns are due entirely to changes in relative prices and incomes. Also, the data being studied to test these hypotheses are largely the same (i.e., post-World War II U.S. meat consumption data). The dichotomy in the literature is between the studies that found structural change and those that did not. We attribute that dichotomy primarily to choices about specification of the functional form of the demand equations. Of secondary importance is the choice of testing procedures and criteria.[8]

[6] This is not to say that the nonparametric tests are biased in favor of a finding of stable preferences, but it does reduce our confidence in concluding, based on nonparametric test results alone, that there has been no structural change in Canada's meat demand. Nor does it mean that the nonparametric tests are low powered relative to parametric ones. We generally do not know the power of structural change tests using a particular data set, even under the assumption that the functional form is correct. An important point in this context is the idea that power relates very much to the nature of the data being studied. While we expect the nonparametric tests to have low power relative to the correct (but unknowable) parametric specification, at the same time we expect to find fewer false rejections as would result from imposition of the wrong functional form.

[7] The fact that most of the growth in chicken consumption seems to have been in the relatively unhealthy end of the spectrum of ways to eat meat is rarely mentioned. It is difficult to ascribe an increase in consumption of fast food to the behavior of rational consumers actively pursuing a healthier diet. Convenience of chicken products is a plausible alternative, although apparently not viewed as such by the beef and pork industries given their advertising campaigns for "lean beef" and the "other white meat."

[8] A fairly extensive list of these studies is provided by Chalfant and Alston and, more recently, Moschini and Meilke (1989). The findings of a sample of them are summarized by Dahlgran. Several are contained in the book edited by Buse. As a partial summary, (*a*) Braschler; Chalfant and Alston; Chavas (1983); Choi and Sosin; Dahlgran; Haidacher; Haidacher et al.; Menkhaus, St. Clair, and Hallingbye; Moschini and Meilke (1984, 1989); Thurman; and Wohlgenant studied U.S. data; (*b*) Atkins, Kerr, and McGivern; Chen and Veeman; and Young studied Canadian data; and (*c*) Martin and Porter; and Chalfant and Alston studied Australian data.

Type I Errors in the Parametric Approach

That specification choices affect results is certainly not news. Indeed, it is a basic part of training in econometrics. What is less clear is how sensitive results will be to the types of specification choices that we make. In particular, how likely are we to cause a false finding of structural change by choosing the wrong functional form for demand equations? This is an important question given that we can never know the true functional form and, especially when we try only one, we will almost surely be using the wrong one. The problem is widely recognized. However, little seems to be known about the effect of specification errors on the probability of finding structural change where none has occurred (i.e., Type I error).

To explore and illustrate this question and its answer we use a simple example with synthetic data. The approach is as follows. First we fit a set of demand equations to actual meat consumption data. Then we generated predictions from those estimated equations (using actual prices and expenditures to generate synthetic quantities that are therefore known to be from a stable set of preferences, treating the fitted model as a "true" model). Then we estimated alternative demand equations using the synthetic data and tested for structural change in the estimated equations using tests for autocorrelated residuals, tests for significant trends, and Chow tests for discrete structural change.

For "true" models we used a double-log model and a linear model. Then—as is the usual situation—we proceeded as if we did not know the data-generating mechanism. With data generated from the double-log model we estimated a linear model, and with the data generated from the linear model we fit a double-log model.[9] This experiment was carried out using annual data for both the United States (1947–83) and Canada (1960–88) for four meats (beef, pork, poultry, and fish).[10] The linear model had four equations (one for each meat type) in which the dependent variable was per capita consumption of that type of meat and the explanatory variables were real prices of the four meats and the real value of total expenditure on all four meats. Real values were obtained by deflating nominal values by the Consumer Price Index for all goods. In the double-log model the logarithms of variables replaced the corresponding variables in the linear model.

Results of tests for structural change in this experiment are reported in table 1. The first set of tests are sequential Chow tests in which an F-statistic is computed at each data point to test the hypothesis that the parameters of the model are different before and after that point [using the DIAG option in SHAZAM (White, Haun, and Horsman)]. In table 1 the largest F-statistic for each equation is reported. As can be seen in table 1, the use of the incorrect functional form resulted in significant Chow tests in all four equations for both countries for both cases (i.e., when fitting a double-log model to data from a linear model and vice versa). In many instances, results less compelling than these have led to conclusions that tastes have changed.

It is questionable practice to search over the data for the most significant point at which to split the sample and at the same time to test for whether the data should be split. The search makes some sense if there is doubt about when structural change might have occurred, but clearly a smaller rejection probability should be used when the split point is not specified in advance. Elsewhere (Alston and Chalfant), we report Monte Carlo results indicating that rejection probabilities can be several times as great as the nominal size of the test when we search for the maximum Chow test. However, in the application here our general conclusions about functional forms are unaffected when the sample is split at the midpoint instead, although the strength of the finding is reduced (from eight to four equations indicating significant structural change, three in the U.S. and one in Canada). Considering that these results were obtained with data from stable models, it is little comfort that the number falls only to four.

In the Canadian data the use of the incorrect functional form did not result in autocorrelation problems. However, in the U.S. data the use of the incorrect model resulted in significant autocorrelation in the equations for beef,

[9] A non-nested test or a Box-Cox approach obviously could help us choose between these two functional forms. When we know that the true functional form is either double log or linear, it is routine to test the functional form and eliminate the wrong one. Such a test would be pointless in the present exercise, the purpose of which is to illustrate how wrong we can be by capriciously choosing a particular functional form.

[10] The U.S. data were from Wohlgenant and the Canadian data were supplied by Linda Robbins from Agriculture Canada.

Table 1. Consequences of Specification Error for Test Results

	True Model—U.S. Data		True Model—Canadian Data	
	Linear	Double log	Linear	Double log
F-statistics for Chow Tests				
Beef	52.83*	32.27*	7.09*	7.27*
	*15.05**	*21.56**	*0.92*	*0.65*
Pork	4.19*	4.71*	7.93*	6.16*
	2.44	*1.80*	*0.51*	*0.41*
Poultry	92.38*	68.89*	23.75*	11.86*
	*17.60**	*40.79**	*1.42*	*1.88*
Fish	94.67*	89.42*	4.85*	7.17*
	*30.51**	*42.71**	*2.43*	*2.71**
Autocorrelation Coefficients				
Beef	0.79*	0.81*	0.12	0.10
	(0.10)	(0.10)	(0.18)	(0.18)
Pork	0.11	0.06	−0.04	−0.07
	(0.16)	(0.16)	(0.19)	(0.19)
Poultry	0.97*	0.99*	0.09	0.18
	(0.04)	(0.03)	(0.18)	(0.18)
Fish	0.95*	0.97*	0.22	0.25
	(0.05)	(0.04)	(0.18)	(0.18)

Note: When the true model is linear (double log), the estimates are obtained using a double-log (linear) model. With the U.S. data, the calculated F-statistics should be compared to $F_{6,25}$ (the critical values of which are 2.49 and 3.63 at the 95% and 99% confidence levels, respectively) and with the Canadian data, the calculated F-statistics should be compared to $F_{6,17}$ (the critical values of which are 2.70 and 4.10 at the 95% and 99% confidence levels, respectively). The first figure is the F-statistic for the maximum Chow test while the lower figure (in italics) is for a Chow test at the midpoint of the sample. The figures in the lower half of the table are first-order autocorrelation coefficients computed using "AUTO" in SHAZAM, and the figures in parentheses are the corresponding approximate standard errors.
* Denotes significance at the 95% confidence level.

poultry, and fish (but not for pork) in both the linear and double-log specifications of the true model.

The problem is not unique to the single-equation models. We also found autocorrelation when we fit a linear approximate almost ideal demand system model to the data generated from the linear and double-log models for U.S. consumption. With data generated from the double-log model, the estimated first-order autocorrelation coefficient in the almost ideal demand system was .94 with a t-value of 39.77; with data generated from the linear model, the estimated first-order autocorrelation coefficient was .93 with a t-value of 26.24. Similarly, when trends were used as added regressors, they tended to have very significant coefficients; moreover, the beef equation showed a strong negative trend and the poultry equation showed a positive one, replicating typical results in the literature.

Results like this are typical when the almost ideal system is estimated using U.S. meat consumption data, and a common interpretation involves habit persistence or a gradual response to price changes, if not structural change. These results show why we should be reluctant to reject static utility theory or stable preferences based on evidence of this type alone; a check for specification error would be a better first step.

We have the advantage, in this instance, of knowing the true structure of the model and knowing that specification error is the source of the serially correlated residuals. The following is pertinent:

> There are, however, circumstances in which the assumption of a serially independent disturbance term may not be very plausible. For example, one may make an incorrect specification of the *form* of the relationship between the variables. Suppose we specify a linear relation between Y and X when the true relation is, say, a quadratic. Even though the disturbance term in the true relation may be non-autocorrelated, the quasi-disturbance term associated with the linear relation will contain a term in X^2 [emphasis in original]. (Johnston, pp. 243–44)

The results here go slightly beyond reinforcing Johnston's statement. They indicate that not only might an apparently innocuous specification choice lead to autocorrelation, it might lead to autocorrelation of a very serious magnitude—of the type that in previous studies has led people either to assert the presence of dynamic influences in consumption (e.g., habit persistence) and to incorporate lagged dependent variables in demand equations (e.g., Blanciforti and Green) or to estimate their model in difference form (e.g., Moschini and Meilke 1989). Furthermore, since those specification errors involve prices, they are likely to be correlated with time or other proxies for changes in tastes, as is indicated by our results.

Type II Errors in the Nonparametric Approach

The parametric approach to testing for structural change involves the imposition of functional forms for demand equations (and possibly other restrictions) as joint, maintained hypotheses. Thus, the results from such tests are always conditional. If we could know the true functional forms, the imposition of the truth as a restriction would increase the power of the parametric tests. Since we can never know the true functional form, any such gains in power from the imposition of functional form (or other) restrictions may be illusory—at least it is speculative to presume such gains. It is important to recognize this distinction between increasing rejection probability, as seems to occur when the wrong functional form is chosen, and increasing power which requires that the correct functional form was chosen.

An ideal test for structural change would test only the hypothesis that preferences are stable (rather than that they are stable and of a particular form). This advantage is possessed by the nonparametric test that we used previously to test for structural change in demand for meat in Australia and the United States (Chalfant and Alston). We applied Varian's (1982, 1983) generalized axiom of revealed preference (GARP) to test for consistency of the Canadian meat consumption data with the existence of a stable well-behaved set of preferences among the four meat types (see Varian 1982, 1983 for details of the theory and Chalfant and Alston for a discursive treatment). There were no violations of GARP. These results suggest that there exists a stable well-behaved set of demand equations among the four goods that can "rationalize" the data. Thus we can treat the four meats as comprising a weakly separable group in which changes in per capita consumption may be explained by prices of the meats included in the group, total per capita expenditure on the group, and measurement errors.[11]

Two issues remain. First, how much confidence ought we to put in the nonparametric test results? This is the power question: How likely is the nonparametric approach to find evidence of a structural change when such change has occurred; or, how large must a structural change be in order to be detected by the nonparametric approach? Second, what is the nature of the stable set of demand equations? We approach the power question next and discuss estimation of demand equations later.

As mentioned earlier, the issue of power may relate more to the nature of the data than to the testing method being used. For example, when consumption data are characterized by relatively large trends in total expenditure and little variation in relative prices, it is difficult to identify substitution effects among goods using any procedure—parametric or nonparametric. Engel curves alone explain most of the variation in the data. To illustrate the relationship between the characteristics of data and the power of the nonparametric tests, we conducted a series of Monte Carlo experiments.

The design of the Monte Carlo experiments was as follows. Cobb-Douglas demand equations were defined for three meat types.[12] In these demand equations per capita consumption of meat type i in year $t(q_{it})$ depends on its real price (P_{it}) and real expenditure on the group in that year (Y_t) as follows:

$$(1) \qquad q_{it} = \alpha_i(Y_t/P_{it}) \qquad \text{for all } i, t,$$

where α_i is the constant expenditure share of good i and $\Sigma_i\alpha_i = 1$. This functional form was

[11] As in most studies of structural change in meat demand, it is a maintained hypothesis in this approach that the four meats are weakly separable from all other goods. The fact that the nonparametric results support this assumption is reassuring. The only practicable alternative to making such an assumption seems to be including a "nonmeat" aggregate in the analysis—aggregating fruits, vegetables, clothing, housing, etc. This is simply an alternative separability assumption that introduces additional potential for specification errors in the aggregation of all other goods and may mask what is going on when interest is specifically in changes within the meat group.

[12] We applied the nonparametric test for separability (Varian 1982, 1983; Belongia and Chalfant) and found that the data satisifed the necessary and sufficient conditions for weak separability of fish from the remaining three meat types (beef, poultry, and pork). Thus, to reduce the size of the experiment, we excluded fish from the Monte Carlo work.

chosen because it is very simple and convenient with expenditure shares (α_i) completely characterizing the demand equations. As our base case, the three goods were assigned expenditure shares of $\alpha_1 = .50$, $\alpha_2 = .25$, and $\alpha_3 = .25$ (corresponding roughly to beef, poultry, and pork in the United States or Canada).

Taste changes were represented in the model as a discrete shift of the demand equations for beef and chicken occurring at the middle of the series (i.e., in the fifteenth of 29 annual observations). To do this, the share of beef was decreased by an amount ($\delta = .001, .010, .025, .050$, or $.100$) from its base value (of .50), and the share of poultry was increased by the same amount (from its base value of .25). Thus, in general, $\alpha_1 = .50 - \delta$, $\alpha_2 = .25 + \delta$, and $\alpha_3 = .25$ (i.e., the expenditure share of pork is constant), where δ measures the size of the change in tastes in favor of poultry and away from beef.

To generate consumption data, these demand equations were combined with the actual time series of Canadian meat prices (29 observations from 1960 to 1988) and a variety of series of total expenditures. These total expenditures were generated by choosing a compound annual growth rate ($\gamma = .5\%, 1\%, 2\%$, or 3%—bracketing the historical growth rate of real meat expenditures, about 1.5% in Canada) on a base of 100: $Y_t = 100(1 + \gamma)^t$. Then, the consumption data were augmented with independent and identically distributed random normal measurement errors (ϵ_{it}) to obtain replications of consumption data from the same underlying data-generating mechanism. This makes it possible to discuss probabilities. Two sizes of standard deviations of the measurement errors were tried ($\sigma_\epsilon = 20$ or 40) to show the effect of measurement errors on both power and size of nonparametric tests. To generate the measurement errors, standard normals were drawn and multiplied by the relevant magnitude of σ_ϵ. Thus, the estimated quantities were obtained from:

$$(2) \qquad q^*_{it} = \alpha_i\,(Y_t/P_{it}) + \epsilon_{it}$$

where $\epsilon_{it} \sim N(0, \sigma^2_\epsilon)$ for all i, t.

Combining the consumption data sets with the actual price series yields data sets that reflect the measurement errors in quantities. These actual prices and estimated values for consumption and total expenditures were then tested for structural change using the weak axiom of revealed preference (WARP). Because WARP tests for pairwise inconsistency between observations but does not test for higher order intransitivities, this approach possibly understates the extent to which data are inconsistent with GARP, the more rigorous test.[13] At each of the 40 design points (five sizes of taste change, four rates of income growth, two sizes of measurement error variances), 200 draws of random quantity measurement errors were used to generate 200 data sets. These data then were tested for consistency with WARP, and the number of observations that violated WARP was recorded. Table 2 reports, for each design point, the frequency with which the data violated WARP.

Three features of the results in table 2 are notable. First, although the frequency of violations tends to increase as the size of the taste change increases and as the growth rate of total expenditures decreases, it does not always do so. Thus, the relationship between power and these variables (γ and δ) is only approximately in the directions that we had suggested. Second, increasing the variance of errors of measuring quantities tends to increase the frequency of violations but does not always do so. Third, the frequency of violations is disappointingly small. Only when the rate of expenditure growth is relatively small (say, less than 1%) and the size of the taste change is relatively large (say, greater than 10%) is the probability of finding a violation greater than 20%. Further, it is not clear that such violations necessarily would be interpreted as evidence of taste changes. Even in the most successful case ($\gamma = .5\%$, $\delta = .10$, $\sigma_\epsilon = 20$ with a frequency of violations of 81.5%), the average number of violations per data set was only 1.5 and the maximum was 4.

One of the problems with the experiment described above is that too many things are allowed to vary at once. In particular, in pursuit of realism and to simplify the experimental design, we used actual Canadian prices. The

[13] WARP was chosen because we wanted to use a simple FORTRAN program to check consistency as part of the program used to generate the data sets. Varian's software could be used to check each data set for GARP but only in a very laborious process. We suspect there would be limited returns to doing that. In our experience, violations of GARP are almost invariably also WARP violations. That is, higher order intransitivities in the data detected by GARP are usually associated with pairwise inconsistencies that would be detected by WARP. It seems unlikely that evidence for structural change would ever be very convincing when WARP, but not GARP, could be satisfied.

Table 2. Power of Nonparametric Tests: Probability of Violation of WARP Using Canadian Meat Price Data

	Growth Rate of Total Expenditure on Meat: γ							
	(σ_ϵ = 20)				(σ_ϵ = 40)			
	0.5	1.0	2.0	3.0	0.5	1.0	2.0	3.0
Taste Change: δ (Δ Beef Share)	(%)							
0.001 (.2%)	0.0	0.0	0.0	0.0	2.5	1.5	0.5	0.5
0.010 (2%)	0.0	0.5	0.5	0.0	4.5	3.0	0.5	0.5
0.025 (5%)	4.5	3.0	1.5	0.0	12.5	5.0	2.5	1.0
0.050 (10%)	24.0	25.0	8.0	14.5	26.5	27.0	6.5	7.0
0.100 (20%)	81.5	73.0	23.5	57.0	72.5	69.0	37.0	37.0

Note: The figures in parentheses under "Taste Change" express the decrease in share (δ) as a percentage of the initial beef share of expenditure (0.5). The corresponding percentage increase in poultry's share of expenditure is twice the percentage change in beef share. σ_ϵ denotes the standard deviation of measurement errors.

cost of doing this is that the prices vary in an uncontrolled way across the data. Trends in prices—when confounded with structural change, expenditure growth, and quantity measurement errors—may be (most likely are) responsible for all three notable features of the results in table 2: ambiguous effects of expenditure growth, size of taste change and error variance, and low frequency of WARP violations.[14] Still, the results here show that the nonparametric tests are capable of detecting structural changes of the types that have been suggested for meat consumption data.

Parametric Tests for Structural Change

To explore further the effects of specification choices, we estimated demand equations for meat in Canada and applied parametric tests for structural change.[15] As background for this work, the nonparametric tests with the Canadian meat data indicated that the data were consistent with having been generated by a stable system of well-behaved demand equations.

One response to this finding would be to insist that the parametric tests confirm that result, that is, to insist that the parametric model does not reject the null hypothesis of no structural change. To do this, it would be necessary to continue searching across functional forms for demand until one was found that satisfied all possible tests for structural change. That would be a potentially endless and fruitless exercise. Even if a stable system of well-behaved demand equations were found, it might not be the correct one and it might not have plausible characteristics (e.g., it might be one in which beef and chicken are complements or beef is a Giffen good—these are characteristics which are permissible in theory but difficult to accept, at least for these goods). It also would be a bad thing to do. An unstructured search over functional forms will quickly exhaust the information content of the data.

An alternative response is to be slightly more open on the question of structural change but to make an attempt to reduce the chance of Type I errors (false rejections of a true null hypothesis of stable preferences). One way to do this is to conduct a limited search over functional forms; another way is to use relatively flexible functional forms (e.g., the "semiparametric" forms used by Chalfant; Gallant; and Wohlgenant) and thereby reduce the severity of the constraints of joint hypotheses. A sensible compromise is to do some specification search using relatively flexible forms (e.g., as done by Murray). The other component of this alternative response is to

[14] The Cobb-Douglas preferences were maintained throughout the experiment and this limitation on the nature of demand response is likely to have affected the power of nonparametric tests. Further work is planned to extend the experiment to use more flexible demand systems to generate data.

[15] As with our nonparametric work, it is a maintained hypothesis here that meats comprise a weakly separable group. In the parametric models we use expenditure on the meat group as the income variable and treat it as an exogenous variable, as is customary. However, we are conscious of the potential hazards of doing so raised by LaFrance. An advantage of the nonparametric approach is that matters such as simultaneity are not an issue.

Table 3. Test Results for Single-Equation Models of Meat Demand in Canada

	Model	
	Linear	Double log
Time Trends		
Beef	−.50	−.02*
	(.26)	(.007)
Pork	−.22*	−.01*
	(.09)	(.003)
Poultry	.49*	.025*
	(.15)	(.005)
Fish	.08	.012
	(.07)	(.008)
Autocorrelation Coefficients		
Beef	.96*	.94*
	(.05)	(.06)
Pork	.41*	.25
	(.17)	(.18)
Poultry	.94*	.88*
	(.06)	(.09)
Fish	.93*	.93*
	(.07)	(.07)

Note: Figures in parentheses are approximate standard errors.
* Denotes that the parameter (either the time-trend coefficient or the first-order autocorrelation coefficient) is significantly different from zero at the 95% confidence level.

limit the number of tests for structural change rather than to apply all possible tests. In this context, the advantage of a clearly specified alternative hypothesis is clear (as opposed to the general alternative that some form of structural change was present).[16]

The work that follows does not go beyond locally flexible functional forms for demand. We try two forms of single-equation models (linear and double log) and two versions of each of two forms of demand systems (the linear approximate almost ideal demand system and the Rotterdam model) and test for significant linear trends or first-order autocorrelation coefficients in the demand equations. We interpret the existence of statistically significant autocorrelation or trends as evidence that we can reject the joint hypothesis that (*a*) the functional form (along with all other aspects of the specification of the model) is correct and (*b*) preferences are stable. The results of this work are summarized in tables 3 through 5. Complete details are available from the authors.

Table 3 shows the results of the tests for significant trends and autocorrelation of residuals for the two single-equation demand models. Both models had either statistically significant trends, significant autocorrelation, or both, in at least one equation. Table 4 shows the results of the tests for significant trends and autocorrelation of residuals for the two demand systems. The linear approximate almost ideal demand system had both statistically significant trends and significant autocorrelation. First differencing was highly successful as a way of eliminating the autocorrelation (the first-order autocorrelation coefficient in the undifferenced model was very close to 1.0, as happens surprisingly often with this model), but it did not eliminate the trends (which are reflected as intercepts in the differenced model). A test of the hypothesis that the intercepts were zero yielded a test statistic of 9.97 as compared to the χ^2 (.05, 3) value of 7.81. This is very similar to the results of Moschini and Meilke (1989) using U.S. data.

As in the first-differenced almost ideal demand model, intercepts in the Rotterdam model imply trends in quantities consumed. When the Rotterdam model was estimated, in either absolute price or relative price form, the restriction of zero intercepts was rejected using the likelihood ratio test. Two interesting further results led us to discount these findings. First, there was no statistically significant trend affecting either beef or pork, implying that if the results are to be taken as evidence of structural change, they suggest the peculiar result that it is a taste change in which chicken is being substituted for fish. Second, if there has been a structural change due to health concerns, most observers would say it began in the 1970s or early 1980s. We used a dummy variable instead of an intercept to allow these trend effects to enter only after the midpoint of the sample (1974). They were no longer significant (LR = 4.35 and LR = 6.09 for the absolute and relative price versions, respectively, compared to a critical value of 7.81). This suggests to us that the significant trends across the full sample show evidence of the

[16] Alternatively one could insist on a complete battery of diagnostic tests, as suggested by Beggs and illustrated in an application to Australian demand for meats. While this approach is potentially informative, it becomes increasingly difficult to work out the probability of Type I and Type II errors in tests for structural change when several tests are applied. In short, it seems almost certain that some test procedure will reject the hypothesis of no structural change in any model if a sufficiently large number of tests are tried.

Table 4. Test Results for AIDS and Rotterdam Models of Meat Demand in Canada

	AIDS		Rotterdam	
	Static	Dynamic	Absolute	Relative
Likelihood Ratio for Chow Test			16.85	17.33
Time Trends				
Beef	−.002*	−.001	−.00049	.0006
	(.0006)	(.002)	(.002)	(.0025)
Pork	−.002*	−.003*	−.0022	−.0014
	(.0006)	(.0015)	(.0015)	(.0018)
Poultry	.004*	.004*	.0038*	.0041*
	(.0007)	(.0013)	(.0013)	(.0013)
Fish	.0000	−.0001	−.0011	−.0033
LR	22.60	9.97	8.94	9.53

Note: The Chow test refers to the hypothesis that all parameters in the Rotterdam model are stable before and after the midpoint of the data. The likelihood ratio statistic should be compared to a χ^2. The critical values for the χ^2 with nine degrees of freedom (for the absolute price version) and 10 degrees of freedom (for the relative price version) at the 95% confidence level are 16.92 and 18.31, respectively. The lower part of the table refers to tests for significant time trends. Figures in parentheses are approximate standard errors. *LR* is the test statistic for the hypothesis that the time trends are jointly insignificant and should be compared to the χ^2 critical value of 7.81. The fish equation was left out for estimation. The trend coefficient for the fish equation was computed by applying the adding-up restriction.

* Denotes that the time-trend parameter is significantly different from zero at the 95% confidence level.

model not fitting the data well over the entire sample period (whether due to the functional form or gradual changes in the composition of the commodities).

In addition, we tested for a discrete change in preferences in both versions of the Rotterdam model (without intercepts, i.e., without trends) using a midpoint Chow test. The test failed to reject the hypothesis of constant parameters in both versions of the model ($LR =$ 16.85 for the absolute price version and $LR =$ 17.33 for the relative price version). In both cases the test statistic was slightly smaller than the critical value for the 5% test.

On the whole the results do not reject the null hypothesis of a stable well-behaved demand system of the Rotterdam form. There were some significant trends, but they were not of the type associated with taste changes. The Chow tests do not reject a stable model. Finally, the elasticities mostly conform to expectations (see table 5). This is in contrast to the almost ideal model results. We are not convinced that the Rotterdam model is the best model for these data. Some other test may reject the stable form of the Rotterdam model. But we have shown that an arbitrary choice between two models—both commonly used and theoretically plausible—can lead to different conclusions concerning structural change.

Conclusion

It is difficult to learn much about demand with the aggregate per capita time-series data that are typically available. When the data are characterized more by long-term trends in prices, consumption, and total expenditure (and perhaps preferences) than by year-to-year relative price movements, it is difficult to sort out the causes of changes in consumption and deliver a definitive conclusion. As suggested by Chavas (1989) and Haidacher, it is probably asking too much to seek to identify the existence of structural change in demand using the same data that we use to estimate that structure. It is surely ambitious to propose to measure the direction and size of the structural change.

Previous work has yielded mixed results on the question of structural change in meat demand. Even the studies that found statistically significant structural change typically found that only small changes (albeit economically very important changes) in consumption could be attributed to changes in demand [e.g., Moschini and Meilke (1989) suggested a 6% decline in U.S. beef consumption—i.e., from a budget share of .5050 to .4747 at the means—would be attributable to structural change]. It is difficult indeed to measure changes of that magnitude, with any confidence, with the data and

Table 5. Uncompensated Elasticities of Demand for Meat in Canada

	Model			
	Linear	Double log	AIDS	Rotterdam
η_{bb}	−.96	−.84	−1.04	−.66
η_{bp}	.00	.01	−.23	.01
η_{bc}	−.47	−.50	−.19	−.06
η_{bf}	−.14	.05	−.46	−.12
η_{bY}	1.90	1.63	1.93	.82
η_{pb}	−.03	.02	.03	.01
η_{pp}	−.81	−.79	−.84	−.74
η_{pc}	−.26	−.25	.10	−.02
η_{pf}	−.05	.10	−.21	−.10
η_{pY}	.95	.79	.92	.85
η_{cb}	.16	.13	.17	.03
η_{cp}	.22	.25	.24	.07
η_{cc}	−.48	−.58	−.62	−.74
η_{cf}	−.15	−.39	−.25	.20
η_{cY}	.07	.46	.47	.44
η_{fb}	−.44	−.44	−.12	−.64
η_{fp}	−.53	−.57	−.03	−.42
η_{fc}	.52	.64	−.16	−.13
η_{ff}	−.45	−.57	.29	−.90
η_{fY}	.76	.82	.02	2.09

Note: These elasticities are conditional uncompensated price elasticities. That is, they hold constant expenditure on the meats group (rather than total expenditures on all goods) and use shares of expenditures on meat rather than shares of total expenditures on all goods. Linear, double-log and AIDS elasticities are at sample means based on models that include trends. The absolute price version of the Rotterdam model is used. The subscripts denote: *b*, beef; *p*, pork; *c*, chicken; *f*, fish; and *Y*, the total expenditure on the four meats.

methods that are available. Even if such changes in structure of per capita demands did occur and were measured accurately, the question remains open as to whether the cause was a change in tastes due to greater health consciousness. Equally plausible alternatives include, for instance, demographic changes and increased consumption of meals away from home. Thus, the diagnosis and prescription are not clear even when the symptoms can be measured.

We remain open on the question of whether there has been a structural change in Canadian meat demand. Almost everyone can cite some anecdotal evidence of changes in meat consumption behavior. But this may not extrapolate well to the behavior of the representative consumer that underlies our per capita demand models. What is clear from our results is that the aggregate data do not permit the hypothesis of stable preferences to be rejected without the imposition of additional nonsample evidence concerning the functional form of demand equations. The nonparametric tests indicate that the data could have been generated by a stable, well-behaved system of per capita demand equations. We suspect that, due to the nature of the data, the power of this test is lower than we would like [say, a 25% to 50% chance of finding a structural change, of the type that Moschini and Meilke (1989) found in the U.S. data, if one had occurred]; but it is hard to know what is normally accepted by the profession as sufficient power. In any event, combining that result with our parametric results using the Rotterdam model implies not only that a stable demand system could exist but that a very plausible and commonly used example can be found.

It is also questionable whether parametric models have more power as such or merely more probability of rejecting stable preferences. Perhaps simulations such as ours, applied to parametric tests, could show that a particular parametric model has higher power but does not have a higher probability of incorrectly finding a taste change. We know of no such evidence at present and suspect that the probabilities of rejection may depend as much on the behavior of relative prices as on the choice of testing approach and, with parametric models, the choice of functional form.

We illustrated how easy it is to generate important Type I errors with apparently innocuous model specification errors. Then we found a variety of results among alternative, popular specifications when they were applied to Canada's meat consumption data. We see no a priori reason to prefer the almost ideal demand system over the Rotterdam model or vice versa. The almost ideal demand system results indicated significant structural change while the Rotterdam model results were consistent with our nonparametric results in supporting the existence of stable preferences. In order to maintain the conclusion of structural change based on these data in the face of these results, one would have to show either that the Rotterdam model is biased in favor of stability or that prior expectations give low weight to the probability of it being the correct functional form. While model selection criteria can compare functional forms, their value may be limited when the set of alternatives is large.

In conclusion, it seems that there continues to be widespread support in the profession for the structural change hypothesis. However, better data or better methods are needed to provide statistical confidence, in the form of valid inferences, to support that view. Purcell (pp. 17–18) criticized the profession for the fact that "as late as 1987, journal articles still reflected disagreement on whether a shift in demand had occurred." On the contrary, the conclusion of this article is that there remains too little disagreement. We urge the profession to take the "con" out of demand analysis, to pay more attention to the fragility of the inferences that can be drawn from models that are built on a whim, and to be much more cautious in basing recommendations upon fragile inferences.

[*Received April 1990; final version received January 1991.*]

References

Alston, J. M., and J. A. Chalfant. "Unstable Models from Incorrect Forms." *Amer. J. Agr. Econ.* 73(1991): (forthcoming November).

Atkins, F. J., W. A. Kerr, and D. B. McGivern. "A Note on Structural Change in Canadian Beef Demand." *Can. J. Agr. Econ.* 37(1989):513–24.

Beggs, J. "Diagnostic Testing in Applied Econometrics." *Econ. Record* 64(1988):81–101.

Belongia, M. T., and J. A. Chalfant. "The Changing Empirical Definition of Money: Some Estimates from a Model of the Demand for Money Substitutes." *J. Polit. Econ.* 97(1989):387–97.

Blanciforti, L., and R. Green. "An Almost Ideal Demand System Incorporating Habits: An Analysis of Expenditure on Food and Aggregate Commodity Groups." *Rev. Econ. and Statist.* 65(1983):1–10.

Braschler, C. "The Changing Demand Structure for Pork and Beef in the 1970s: Implications for the 1980s." *S. J. Agr. Econ.* 15(December 1983):105–10.

Buse, R. C., ed. *The Economics of Meat Demand.* Proceedings of a Conference held 20–21 October 1986, Charleston SC. Madison: Department of Agricultural Economics, University of Wisconsin, 1989.

Chalfant, J. A. "A Globally Flexible Almost Ideal Demand System." *J. Bus. and Econ. Statist.* 5(1987): 233–42.

Chalfant, J. A., and J. M. Alston. "Accounting for Changes in Tastes." *J. Polit. Econ.* 96(1988):391–410.

Chavas, J.-P. "On the Structure of Demand for Meat." In *The Economics of Meat Demand,* ed., R. C. Buse, pp. 125–35. Madison: Department of Agricultural Economics, University of Wisconsin, 1989.

———. "Structural Change in the Demand for Meat." *Amer. J. Agr. Econ.* 65(1983):148–53.

Chen, P. Y., and M. M. Veeman. "Demand for Meat in Canada: Model Specification and Structural Change." Staff Paper 89-11, Department of Rural Economy, University of Alberta, 1989.

Choi, S., and K. Sosin. "Testing for Structural Change: The Demand for Meat." *Amer. J. Agr. Econ.* 72(1990): 227–36.

Dahlgran, R. A. "Changing Meat Demand Structure in the United States: Evidence from a Price Flexibility Analysis." *N. Cent. J. Agr. Econ.* 10(1988):165–76.

Gallant, A. R. "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form." *J. Econometrics* 15(1981):211–45.

Haidacher, R. C. "Assessing Structural Change in the Demand for Food Commodities." *S. J. Agr. Econ.* 15(July 1983):31–37.

Haidacher, R. C., J. A. Craven, K. S. Huang, D. M. Smallwood, and J. R. Blaylock. *Consumer Demands for Red Meats, Poultry, and Fish.* Staff Report No. AGES 820818. Washington: U.S. Department of Agriculture–Economic Research Service, 1982.

Johnston, J. *Econometric Methods.* 2nd ed., International Student Edition. Tokyo: McGraw-Hill Kogakusha Ltd., 1972.

LaFrance, J. "When Is Expenditure 'Exogenous' in Separable Demand Models?" *West. J. Agr. Econ.* 16(1991): 49–62.

Landsburg, S. E. "Taste Change in the United Kingdom, 1900–1955." *J. Polit. Econ.* 89(1981):92–104.

Leamer, E. E. "Let's Take the Con Out of Econometrics." *Amer. Econ. Rev.* 73(March 1983):31–43.

———. "Sensitivity Analyses Would Help." *Amer. Econ. Rev.* 75(June 1985):308–13.

Martin, W., and D. Porter. "Testing for Changes in the Structure of the Demand for Meat in Australia." *Aust. J. Agr. Econ.* 29(1985):16–31.

Menkhaus, D. J., J. S. St. Clair, and S. Hallingbye. "A Reexamination of Consumer Buying Behavior for Beef, Pork, and Chicken." *West. J. Agr. Econ.* 10(1985): 116–25.

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

———. "Parameter Stability and the U.S. Demand for Beef." *West. J. Agr. Econ.* 9(1984):271–82.

Murray, J. "Retail Demand for Meat in Australia: A Utility Theory Approach." *Econ. Record* 60(1984):45–56.

Purcell, W. D. "The Case of Beef Demand: A Failure by the Discipline." *Choices* 2(1989):16–19.

Thurman, W. N. "The Poultry Market: Demand Stability and Industry Structure." *Amer. J. Agr. Econ.* 69(1987): 30–37.

Varian, H. "Non-Parametric Tests of Consumer Behavior." *Rev. Econ. Stud.* 50(1983):99–110.

———. "The Nonparametric Approach to Demand Analysis." *Econometrica* 50(1982):945–73.

White, H. "Using Least Squares to Approximate Unknown Regression Functions." *Int. Econ. Rev.* 21(1980):149–70.

White, K. J., S. A. Haun, and N. G. Horsman. *SHAZAM: The Econometrics Computer Program Version 6—User's Reference Manual.* Vancouver: University of British Columbia, Department of Economics, 1987.

Wohlgenant, M. K. "Estimating Cross Elasticities of Demand for Beef." *West. J. Agr. Econ.* 10(1985):322–29.

Young, L. J. *Canadian Meat Demand.* Working Paper 10/87. Ottawa: Agriculture Canada, Policy Branch, Commodity Coordination Directorate, June 1987.