INFERRING THE NUTRIENT CONTENT OF FOOD WITH PRIOR INFORMATION

JEFFREY T. LAFRANCE

U.S. farm and food policy is being transformed. Direct cash payments and a movement toward a more open market is replacing many farm-level price and income support programs. Welfare; food stamps; Women, Infants and Children; Aid to Families with Dependent Children; and school lunch programs are being reduced in scope at the federal level and replaced by block grants to states. All of these changes will influence prices and quantities consumed of foods, and therefore the nutritional intake of U.S. consumers. But it is unclear what the overall nutritional effects of these policy changes might be. Food stamps provide direct in-kind subsidies for food consumption, with the goal of increasing the nutritional status of the poor. In contrast, federal milk marketing orders increase the price of fresh milk and lower the prices of manufactured dairy products, creating incentives to substitute away from fresh milk and toward butter and cheese. Other farm-level policies also create consumer incentives at odds with those created by food subsidy programs.[1] Though food aid recipients spend more on food, they eat less healthy foods due to price distortions. Other consumers, who pay the taxes needed to finance farm and food programs, have lower disposable incomes, food expenditures, and nutritional intakes. For this group, policy-induced price distortions also create incentives for less healthy diets.

A central focus of much research on farm and food policy on consumer choice and nutrition has been an effort to establish the economic links between food consumption choices and nutrition. Suppose we have a stable, theoretically consistent reduced-form model of the demand for foods, which might be written in the form $E(\mathbf{x} \mid \mathbf{p}_x, \mathbf{p}_y, m, \mathbf{s}) = \mathbf{h}^x(\mathbf{p}_x, \mathbf{p}_y, m, \mathbf{s})$, where $\mathbf{x}$ is the n_x-vector of foods, $\mathbf{p}_x$ is the corresponding vector of market prices, $\mathbf{p}_y$ is the vector of market prices for all other goods, m is disposable income, and $\mathbf{s}$ is a vector of demographic variables and other demand shifters.[2] It is known that weak integrability of the subset of demands is necessary and sufficient for virtually all economically relevant analyses, including exact welfare measurement of the effects of farm and food policies (LaFrance and Hanemann).

Given measurements on the nutrient content matrix transforming foods into nutrients, say $\mathbf{z} = \mathbf{A}x$, where $\mathbf{z}$ is the K-vector of nutrients consumed and $\mathbf{A}$ is the $K \times n_x$ matrix of nutrient contents per unit of foods, we also can analyze policy effects on nutritional intakes with the above demand model. This follows because the conditional mean for nutrients, given prices, income, demographics and other demand shifters, and $\mathbf{A}$ satisfies $E(\mathbf{z} \mid \mathbf{p}_x, \mathbf{p}_y, m, \mathbf{s}, \mathbf{A}) = \mathbf{A}\mathbf{h}^x(\mathbf{p}_x, \mathbf{p}_y, m, \mathbf{s})$. Nutrient demand price elasticities satisfy $\epsilon_{p_k}^{z_i} = \sum_{j=1}^{n_x} w_{ij}\epsilon_{p_k}^{x_j}$, where $\epsilon_{p_k}^{z_i} \equiv (p_k/z_i) \cdot \partial z_i/\partial p_k$ is the price elasticity of the ith nutrient with respect to the kth price, $\epsilon_{p_k}^{x_j} \equiv (p_k/x_j) \cdot \partial x^j/\partial p_k$ is the price elasticity of the jth food with respect to the kth price, and $w_{ij} \equiv a_{ij}x_j/z_i$ is the share of the ith nutrient supplied by the jth food as one example. The matrix $\mathbf{A}$ is the rub, however. Year-to-year measures of the nutritional content of disaggregated food items are neither published by the USDA nor readily available from other sources.

About ten years ago, with assistance from Nancy Raper of the Human Nutrition Information Service (HNIS) and by using unpublished handwritten documents, I compiled an-

The author is professor in the Department of Agricultural and Resource Economics, and a member of the Giannini Foundation of Agricultural Economics, at the University of California, Berkeley.

[1] Target prices, deficiency payments, and nonrecourse loans increase supplies of feed grains, lower market prices of feed, and increase supplies and lower retail prices of red meat, which is high in cholesterol. Marketing orders and agreements for many fruits, nuts, and vegetables contain regulations that lead to higher prices for fresh products and lower prices for manufactured products, which are less nutritious and contain relatively large amounts of salt (Jamison).

[2] See, e.g., LaFrance for one example of this type of empirical model.

Amer. J. Agr. Econ. 81 (August 1999): 728–734
Copyright 1999 American Agricultural Economics Association

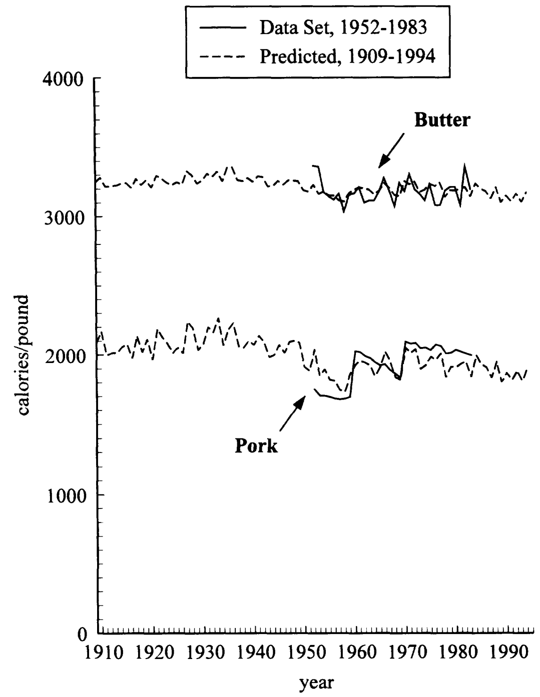

Figure 1. Energy per pound of food

nual estimates of the percentages of seventeen nutrients (energy, protein, fat, carbohydrates, and cholesterol; calcium, iron, magnesium, phosphorous, and zinc; vitamins A, B_6, B_{12}, C, and E; and niacin, riboflavin, and thiamin) supplied by twenty-one foods (fresh milk and cream, butter, cheese, ice cream and frozen yogurt, and canned and powdered milk; beef, pork, other red meat, fish, and poultry; fresh citrus fruits, fresh noncitrus fruits, fresh vegetables, potatoes and sweet potatoes, processed fruits, and processed vegetables; and fats and oils excluding butter, eggs, cereals and bakery products, sugar and sweeteners, and coffee, tea, and cocoa) for the period 1952–83. Each of these percentages was multiplied by the total supply of the corresponding nutrient and divided by the per capita consumption of the corresponding food to generate year-to-year estimates of the nutrient content per pound of each food—for example, grams of protein per pound of beef. These percentage contribution estimates were recorded with only two or three significant digits, suggesting a fair amount of measurement error. Even so, only small changes in the elements of the nutrient content matrices occurred between 1952–83. The two solid lines in figure 1 illustrate the time paths over this period for the energy content of butter and pork. These two food items had far and away the greatest variation in their estimated energy contents over the sample period.

Another problem in estimating the nutrient content of aggregate food items is that, at least until the present, obtaining updated or back dated disaggregated nutrient content estimates has proven to be untenable. As a consequence, in previous work I calculated the average nutrient content matrix over the thirty-two-year period as a first guess for the nutritional content of the U.S. food supply. This has obvious problems, especially in light of recent policy and research emphases on nutritional reporting, health education, and improved diets, not to mention the simple fact that a hog seventy-five years ago was a very different creature than the typical barrow or gilt of today.

This leads to the main focus of this paper. Suppose we wished to make inferences about the likely values of a number of unknown quantities based on a single data point. This is an impossible task using classical statistical methods, unless one is willing to live with infinite uncertainty about the precision of the estimates obtained. But, if a source of reasonable prior information exists, such inference problems can be addressed readily with Bayesian methods. This situation describes quite precisely the nutrient content question I want to address. I have annual observations on the total disappearance of twenty-one foods from the U.S. food supply and the total availability of seventeen nutrients from those foods for the period 1909–94. I also have a sample of estimates for the individual nutritional content of each of these food items for the period 1952–83. However, the food quantity and nutrient availability data has been updated several times by the USDA since the sample of thirty-two observations was originally constructed. Hence, the nutrient content estimates obtained from the extraneous sample are not entirely consistent with the available data on total annual food and nutrient consumption. But it is reasonable to think that the shorter thirty-two-year data set can be used to draw inferences about the joint behavior of the elements of the nutrient content matrix over time. Given this "post data" information, which we will assume has the form of a prior distribution, the longer, incomplete, data set can be used to make year-to-year forecasts to draw inferences on the reasonably likely range of values for the elements of the nutrient matrix. The primary question, then, is how "best" to proceed. In this paper, I outline one possible strategy and apply it to estimating the year-to-year energy content of food commodities in the United States food supply over the period 1909–94.

Inferring the Nutrient Content of the U.S. Food Supply

My initial point of departure is an ingenious approach to ill-posed inference problems known as *generalized maximum entropy* developed recently by Golan; Golan, Judge, and Miller; and Golan, Judge, and Perloff. Although I ultimately pursue a somewhat different strategy for reasons that should become clear below, it is useful to briefly summarize this approach as it relates to the present problem.

Consider the problem of estimating the nutritional content of food items in a given year from aggregate per capita disappearance data and estimates of the total nutrients available in the food supply. Let $\mathbf{z}_t \in \mathbf{R}_+^K$ be the K-vector of nutrients available for consumption per capita in the food supply in year t, let $\mathbf{x}_t \in \mathbf{R}_+^{n_x}$ be the n_x-vector of food quantities consumed per capita, and write the linear relationship between food and nutrients as

$$(1) \quad \mathbf{z}_t = \mathbf{A}_t\mathbf{x}_t, \qquad t = 1, \ldots, T$$

where $\mathbf{A}_t$ is a $K \times n_x$ matrix of positive parameters to be estimated in each year. Suppose that we have an average estimate of the nutrient content matrix, say $\bar{\mathbf{A}}^{\mathrm{O}}$, obtained independently of the current inference problem. But we do not have data on the nutrient content matrices on a year-to-year basis. Let's focus on the case of a single nutrient to simplify the discussion, specifically, the energy content of foods, and omit the time subscripts whenever this is not confusing. The inference problem is to find a vector, $\alpha \geq \mathbf{0}$ satisfying $z = \alpha'\mathbf{x}$, given a prior estimate of the nutrient content vector, $\boldsymbol{\alpha}^{\mathrm{o}}$, and observations on z and $\mathbf{x}$. We first specify a compact interval of support for each α_i containing the prior estimate, $\alpha_i^{\mathrm{o}} \in [\underline{\alpha}_i, \overline{\alpha}_i]$, $i = 1, \ldots, K$, divide each interval into N subintervals, each having the form

$$\left[\left(\frac{N-n+1}{N}\right)\underline{\alpha}_i + \left(\frac{n-1}{N}\right)\overline{\alpha}_i, \right.$$
$$\left.\left(\frac{N-n}{N}\right)\underline{\alpha}_i + \left(\frac{n}{N}\right)\overline{\alpha}_i\right], \qquad n = 1, \ldots, N$$

and write the α_i's as weighted averages of the $N + 1$ endpoints,

$$(2) \quad \alpha_i = \underline{\alpha}_i p_{i0} + \left[\left(\frac{N-1}{N}\right)\underline{\alpha}_i + \left(\frac{1}{N}\right)\overline{\alpha}_i\right]p_{i1} + \cdots$$
$$+ \left[\left(\frac{1}{N}\right)\underline{\alpha}_i + \left(\frac{N-1}{N}\right)\overline{\alpha}_i\right]p_{iN-1}$$
$$+ \overline{\alpha}_i p_{iN} = \sum_{j=0}^{N} [\underline{\alpha}_i + (j/K)\delta_i]p_{ij},$$
$$i = 1, \ldots, K$$

where $\delta i \equiv \bar{\alpha}_i - \underline{\alpha}_i \; \forall i$, $p_{ij} \geq 0 \; \forall \; i, j$ and $\sum_{j=0}^{N} p_{ij} = 1$. The GME choice for α solves

$$(3) \quad \max - \sum_{i=1}^{n_x} \sum_{j=0}^{N} p_{ij}\log(p_{ij})$$

subject to

$$p_{ij} \geq 0 \quad \forall \; i, j,$$
$$\sum_{j=0}^{N} p_{ij} = 1 \quad \forall \; i,$$
$$\sum_{i=1}^{n_x} \sum_{j=0}^{N} [\underline{\alpha}_i + (j/K)\delta_i]p_{ij}x_i = z.$$

This is a straightforward constrained optimization problem with a strictly concave objective function and linear constraints, and a unique solution is guaranteed to exist. Moreover, the logarithmic transformation strictly bounds the solution away from zero, so the nonnegativity constraints are slack at the optimal solution. The GME solution can be written in the form

$$(4) \quad p_{ij} = p_{i0}\exp\{-\lambda\delta_i x_i(j/N)\},$$
$$\forall \; j = 0, \ldots, N,$$
$$\forall \; i = 1, \ldots, n_x$$

with the normalizing condition

$$(5) \quad p_{i0} = 1 \Big/ \sum_{j=0}^{N} \exp\{-\lambda\delta_i x_i(j/N)\}$$

which ensures that the probabilities add up to one for each i. Finally, the optimal posterior choices for the α_i's are the means of the posterior discrete probability distributions,

$$\text{(6)}\quad \alpha_i = \underline{\alpha}_i + \sum_{j=0}^{N}\left(\frac{j}{N}\right)\delta_i \frac{\exp\{-\lambda\delta_i x_i(j/N)\}}{\sum_{j=0}^{N}\exp\{-\lambda\delta_i x_i(j/N)\}},$$

$$\forall\ i = 1, \ldots, n_x$$

while the Lagrange multiplier for the mean constraint is defined by

$$\text{(7)}\quad \sum_{i=1}^{n_x} x_i\left[\underline{\alpha}_i + \sum_{j=0}^{N}\left(\frac{j}{N}\right)\delta_i \frac{\exp\{-\lambda\delta_i x_i(j/N)\}}{\sum_{k=0}^{N}\exp\{-\lambda\delta_i x_i(k/N)\}}\right] = z.$$

This approach always produces a well-defined, unique answer to even highly ill-posed inference problems, including the present one. However, the GME algorithm raises some issues, at least for this application. First, what form does the prior information really take? In the standard GME solution, the choice for the compact support for the coefficients often seems to be a subjective judgment that is not necessarily truly prior information. However, with regard to the nutritional content of foods, we do know (with probability one) that any given food item cannot account for less than zero nor more than 100% of a given nutrient's total availability. This gives us a natural choice for the support of the elements of α. But without looking at any data, I have no other prior knowledge about the percentage of the total energy available in the U.S. food supply that comes from beef, for example. While this level of ignorance is not inconsistent with the standard GME assumption of a discrete equally likely (i.e., uniform) prior, the proper post-data distribution for the elements of α may not, and in most cases will not, be uniform.

A second issue is that each choice for the discrete number of subintervals, N, generates a different solution for the optimal probability weights and therefore for the elements of α. One way to overcome this subjectivity is to let $N \to \infty$ and use a continuous density function for both the prior and the posterior. This is useful for another reason. If we consider the GME solution formally as minimizing the Kullback-Leibler cross entropy criterion function relative to a uniform prior, then the indirect objective function has the form $-\Sigma_{i=0}^{N} p_i^*\ln(p_i^*) + \ln(N + 1)$, where p_i^* is the optimal choice for the ith probability weight. Artificially considering N as a continuous variable, the envelope theorem implies that the optimal entropy level is strictly increasing in N, and that the slope (i.e., the rate of increase) decreases at the rate $1/(N + 1)^2$, and an asymptotic approximation should become accurate rapidly.

To derive the GME approach's limiting distribution, for $s \in [0, 1]$ let $[sN]$ be the largest integer no larger than sN and for each $i = 0, 1, \ldots, N$, define $p_i(s) \equiv p_{i[sN]}$. For given i, N, and $0 \le j \le N$, let s satisfy $j/N \le s < (j + 1)/N$. Then, uniformly in $s \in [0, 1]$, the ith cumulative probability distribution function satisfies

$$\text{(8)}\quad F_i(s) \equiv \frac{\sum_{k=0}^{[sN]}(1/N)\exp\{-\lambda\delta_i x_i(k/N)\}}{\sum_{k=0}^{N}(1/N)\exp\{-\lambda\delta_i x_i(k/N)\}}$$

$$= \left(\int_0^{[sN]/N}\exp\{-\lambda\delta_i x_i([uN]/N)\}\,du\right) \div \left(\int_0^{1}\exp\{-\lambda\delta_i x_i([uN]/N)\}\,du + (1/N)\exp\{-\lambda\delta_i x_i\}\right)$$

$$\underset{N\to\infty}{\longrightarrow} \frac{\int_0^{s}\exp\{-\lambda\delta_i x_i u\}\,du}{\int_0^{1}\exp\{-\lambda\delta_i x_i s\}\,ds} = \frac{1 - e^{\lambda\delta_i x_i s}}{1 - e^{-\lambda\delta_i x_i}}$$

which is a truncated exponential cumulative distribution function, with the limiting value of the Lagrange multiplier defined by the mean condition

$$\text{(9)}\quad \lambda\sum_{i=1}^{n_x}\delta_i x_i\int_0^1 s e^{-\lambda\delta_i x_i s}\,ds/(1 - e^{-\lambda\delta_i x_i}) = z.$$

It is straightforward to verify, using methods from optimal control theory, that this distribution is the continuous GME solution (e.g., Golan, Judge, and Miller, p. 40). Finding the continuous GME posterior leads naturally to the question, what are appropriate choices for a pre-data prior distribution, a post-data posterior distribution, which becomes the pre-

Table 1. Sample Moments for the Energy Content of U.S. Foods, 1952–83

Food Item	Mean (Calories/Pound)	Standard Error
Fresh milk and cream	279.853	2.27135
Butter	3193.76	21.3308
Cheese	1290.95	8.86004
Ice cream and frozen yogurt	683.073	22.0997
Canned and powdered milk	869.336	15.7867
Beef and veal	1029.37	5.88297
Pork	1925.22	25.4956
Other red meat	822.742	8.22912
Fish	887.037	10.9005
Poultry	630.447	8.17579
Fresh citrus fruit	104.220	2.39075
Fresh noncitrus fruit	247.969	2.03088
Fresh vegetables	199.849	3.69462
Potatoes and sweet potatoes	326.975	2.71337
Processed fruit	218.810	4.38049
Processed vegetables	691.288	4.52930
Fats and oils, excluding butter	3754.19	8.39000
Eggs	866.313	22.7001
Cereals and bakery products	1682.45	4.87875
Sugar and sweeteners	1636.86	3.90332

forecast prior distribution, and a loss function?

I began this process ignorant of all of these matters, except perhaps for a small amount of introspection regarding the logical support for the unknown nutrient content quantities. In addition, visual inspection of figure 1 (and similar plots for the other nutrients) suggests that the short but complete data set for the years 1952–83 does not contain very much information about a systematic structure beyond perhaps the first and second moments of the underlying distribution. However, based on the work of Csiszár, Gokhale and Kullback, Jaynes (1957a, 1957b, and 1984), Kullback, and Shannon, the Kullback-Leibler cross-entropy function seems a logical choice for the criterion function. Since the GME solution is equivalent to minimizing the Kullback-Leibler cross-entropy function relative to a uniform prior (Golan, Judge, and Perloff; Gokhale and Kullback), this choice remains logically consistent with GME. However, although I am comfortable with uniform priors on compact intervals before undertaking any data analysis, the Kullback-Leibler criterion can be applied to any post-data prior.

The one thing I remain reticent to impose is a specific assumption about the likelihood function for the shorter, complete data set. Given this, a particularly attractive method is the Bayesian Method of Moments (BMOM), which yields post-data densities for model parameters without an assumed likelihood function (Tobias and Zellner; Zellner; and Zellner, Tobias, and Ryu). In particular, it is known that the proper maximum entropy density given first and second moments is a multivariate normal density (see, e.g., Zellner, Tobias, and Ryu). We obtain sample estimates for the mean vector and variance-covariance matrix by applying the method of moments to the thirty-two-year data set. In this instance, this gives a post-data density of the form

$$(10) \quad f(\alpha \mid D) \sim N\left(\hat{\alpha}, \frac{1}{n}\hat{\Sigma}\right)$$

where $\hat{\alpha}$ is the n_x-vector of sample means and $\hat{\Sigma}$ is the $n_x \times n_x$ matrix of sample variance-covariance terms. To illustrate, table 1 presents the sample means and estimated standard errors of the means for the energy content of foods for the sample period 1953–82.

Thus, for the post-data inference problem, we assume a multivariate normal density function as the prior distribution for each year's observations on total food and nutrient quantities available in the food supply,[3]

[3] In actuality, the compact support for the elements of the nutrient content vector implies a truncated multivariate normal distribution. However, given the sample estimates, the probability of being on or outside the boundary was always on the order of 10^{-9} or smaller, so I ignored it.

$$(11)\quad f_0(\boldsymbol{\alpha}) = (2\pi)^{-n_x/2}\left|\frac{1}{n}\hat{\boldsymbol{\Sigma}}\right|^{-1/2} \times \exp\left\{-\frac{n}{2}(\boldsymbol{\alpha}-\hat{\boldsymbol{\alpha}})'\hat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{\alpha}-\hat{\boldsymbol{\alpha}})\right\}.$$

For the Kullback-Leibler criterion, the objective is to

$$(12)\quad \text{minimize} \int\cdots\int f_1(\mathbf{y})\log[f_1(\mathbf{y})/f_0(\mathbf{y})] \times dy_1 \cdots dy_{N_x}$$

subject to

$$\int\cdots\int f_1(\mathbf{y})\, dy_1 \cdots dy_{N_x} = 1,$$

$$\sum_{i=1}^{N_x} x_i\left[\int\cdots\int y_i f_1(\mathbf{y})\, dy_1 \cdots dy_{N_x}\right] = z.$$

Using techniques from optimal control theory, it can be shown that the optimal choice for $f_1(\boldsymbol{\alpha})$ also is multivariate normal with an updated mean and the same covariance matrix,

$$(13)\quad f_1(\boldsymbol{\alpha}) = (2\pi)^{-n_x/2}\left|\frac{1}{n}\hat{\boldsymbol{\Sigma}}\right|^{-1/2} \times \exp\left\{-\frac{n}{2}(\boldsymbol{\alpha}-\boldsymbol{\alpha}_1)'\hat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{\alpha}-\boldsymbol{\alpha}_1)\right\}$$

$$(14)\quad \boldsymbol{\alpha}_1 = \hat{\boldsymbol{\alpha}} + (\mathbf{x}'\hat{\boldsymbol{\Sigma}}\mathbf{x})^{-1}\hat{\boldsymbol{\Sigma}}\mathbf{x}(z - \mathbf{x}'\hat{\boldsymbol{\alpha}}).$$

We end up with a very simple least squares rule as the solution to what started out as a difficult and highly ill-posed inference problem. I find this quite delightful! The dashed lines in figure 1 display the mean calculations for the energy content of butter and pork in each year in the period 1909–94. This procedure indeed appears to generate reasonable predictions.

Conclusions

The BMOM and GME solution to the nutrient inference problem produces the same algebraic result as the following classical approach. We first use least squares to estimate the sample means and variance-covariance terms. We then take these sample estimates to be the "true" parameter values and calculate a single generalized least squares step in each year to minimize the distance between $\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{\alpha}_1$ relative to the quadratic norm $\frac{1}{n}\hat{\boldsymbol{\Sigma}}$. But there is a significant difference in both the interpretation and the logic behind these approaches. In finite samples, the Bayesian and classical solutions only coincide when the likelihood function is multivariate normal, and this is known a priori. Other likelihood functions generate different results. In addition, the Bayesian method provides a logical basis for inference on the year-to-year distributions of the nutrient content elements. These distributions could be used to calculate standard errors or confidence intervals for such things as the price elasticities of nutrient consumption to obtain reasonable bounds for the changes in nutritional intakes to changes in farm and food policy.

References

Csiszár, I. "Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems." *Ann. Statist.* 19(1991):2032–66.

Gokhale, D.V., and S. Kullback. *The Information in Contingency Tables.* New York: Marcel Dekker, 1978.

Golan, A. "A Multivariable Stochastic Theory of Size Distribution of Firms with Empirical Evidence." *Advances in Econometrics,* vol. 10, pp. 1–46. Greenwich CT: JAI Press, 1994.

Golan, A., G. Judge, and D. Miller *Maximum Entropy Econometrics: Robust Estimation with Limited Data.* New York: John Wiley and Sons, 1996.

Golan, A., G. Judge, and J. Perloff "A Maximum Entropy Approach to Recovering Information from Multinomial Response Data." *J. Amer. Statist. Assoc.* 91(June 1996):841–53.

Jaynes, E. "Information Theory and Statistical Mechanics." *Physics Rev.* 106(1957a):620–30.

———. "Information Theory and Statistical Mechanics II." *Physics Rev.* 108(1957b):171–90.

———. "Prior Information and Ambiguity in Inverse Problems." *Inverse Problems.* D. McLaughlin, ed., 151–66. Providence RI: American Mathematical Society, 1984.

Kullback, J. *Information Theory and Statistics.* New York: John Wiley and Sons, 1959.

Kullback, J. and R. Leibler. "On Information and Sufficiency." *Ann. Math. Statist.* 4(1951): 99–111.

LaFrance, J. "The Structure of U.S. Food Demand." Working paper 862, Department of Agricultural and Resource Economics, University of California, Berkeley, 1998.

LaFrance, J., and M. Hanemann. "The Dual Structure of Incomplete Demand Systems." *Amer. J. Agr. Econ.* 71(May 1989):262–74.

Shannon, C. "A Mathematical Theory of Communication." *Bell System Tech. J.* 27(1948): 379–423.

Tobias, J., and A. Zellner. "Further Results on the Bayesian Method of Moments Analysis of the Multiple Regression Model." H.G.B. Alexander Research Foundation, University of Chicago, 1997. Presented at the Econometric Society meeting, June 1997.

Zellner, A. "The Bayesian Method of Moments (BMOM): Theory and Applications." *Advances in Econometrics,* vol. 12. T. Fomby and R. Hill, eds., pp. 85–105. Greenwich CT: JAI Press, 1997.

Zellner, A., J. Tobias, and K. Ryu. "Bayesian Method of Moments (BMOM) Analysis of Parametric and Semiparametric Regression Models." Alexander Research Foundation, University of Chicago, 1997