

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY  
DIVISION OF AGRICULTURAL AND NATURAL RESOURCES  
UNIVERSITY OF CALIFORNIA AT BERKELEY

---

**Working Paper No. 863**

**INFERRING THE NUTRIENT CONTENT OF  
FOOD WITH PRIOR INFORMATION**

by

Jeffrey T. LaFrance

---

Copyright © 1998 by Jeffrey T. LaFrance. All rights reserved. Readers may make verbatim copies of this document for noncommercial purposes by any means, provided that this copyright notice appears on all such copies.

---

---

**California Agricultural Experiment Station  
Giannini Foundation of Agricultural Economics  
December, 1998**

## **Inferring the Nutrient Content of Food with Prior Information**

Jeffrey T. LaFrance<sup>\*</sup>

**Key Words:** Agricultural Farm and Food Policy, Bayesian Method of Moments, Generalized Maximum Entropy, Kullback-Leibler Cross Entropy, Nutrient Content

### **Abstract**

Given measurements on the nutrient content of the U.S. food supply and a coherent reduced form empirical model of the demand for foods, we can analyze the effect of agricultural farm and food policy on nutrition. Using unpublished documents from the HNIS, estimates of the percentages of seventeen nutrients supplied by twenty-one foods were compiled for the period 1952-1983. The Bayesian Method of Moments is applied to this data set to obtain a proper prior for the purpose of drawing year-to-year inferences about the nutrient content of the U.S. food supply for the period 1909-1994. Information theory and the Kullback-Leibler cross entropy criterion are used to formalize the inference problem.

<sup>\*</sup>The author is a professor in the Department of Agricultural and Resource Economics, and a member of the Giannini Foundation of Agricultural Economics, at the University of California at Berkeley. Contact information: 207 Giannini Hall, University of California, Berkeley, CA, 94720. E-mail: [jeff\\_lafrance@are.berkeley.edu](mailto:jeff_lafrance@are.berkeley.edu).

### **Inferring the Nutrient Content of Food with Prior Information**

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 (Heien; Ippolito and Masson), 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.<sup>1</sup> 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.

I have been interested in the interplay of farm and food policy on consumer choice and nutrition for several years. A central focus of this research has been an effort to establish a direct economic link between food consumption choices and nutrition. Therefore, before getting to the main subject of this paper, I would like to briefly motivate its undertaking. Suppose that we have a stable, theoretically consistent reduced form empirical model of the demand for foods, which might be written in the form  $E(x|\mathbf{p}_x, \mathbf{p}_y, m, s) = \mathbf{h}^x(\mathbf{p}_x, \mathbf{p}_y, m, s)$ , where  $x$  is the  $n_x$ -vector of

foods,  $p_x$  is the corresponding vector of market prices,  $p_y$  is the vector of market prices for all other goods,  $m$  is disposable income, and  $s$  is a vector of demographic variables and other demand shifters.<sup>2</sup> It is well-known that *weak integrability* of the subsystem of demands is necessary and sufficient for virtually all economically relevant analyses in such a model, including, *inter alia*, exact welfare measurement of the price and income effects of farm and food policies (LaFrance and Hanemann).

Moreover, given measurements or estimates on the nutrient content matrix transforming the available supply of foods into the available supply of nutrients, say  $z = Ax$ , where  $z$  is the  $K$ -vector of nutrients consumed by the household and  $A$  is the  $K \times n_x$  matrix of nutrient contents per unit of foods, we also can analyze policy effects on nutritional intakes. This follows from the simple fact that the conditional mean for nutrients, given prices, income, demographics and other demand shifters and  $A$  satisfies  $E(z|p_x, p_y, m, s, A) = Ah^x(p_x, p_y, m, s)$ . For example, nutrient 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  $i^{th}$  nutrient with respect to the  $k^{th}$  price,  $\epsilon_{p_k}^{x_j} \equiv (p_k/x_j) \cdot \partial x_j / \partial p_k$  is the price elasticity of the  $j^{th}$  food with respect to the  $k^{th}$  price, and  $w_{ij} \equiv a_{ij}x_j/z_i$  is the share of the  $i^{th}$  nutrient supplied by the  $j^{th}$  food.

The matrix  $A$  is the rub, however. My primary data set consists of annual time series for the period 1909-1995 on retail prices for and per capita U.S. consumption of twenty-one foods (fresh milk and cream; butter; cheese; ice cream and frozen yogurt; canned and powdered milk; beef and veal; pork; other red meat; fish; poultry; fresh citrus fruit; other fresh fruit; fresh vegetables; potatoes and sweetpotatoes; processed fruit; processed vegetables; fats and oils excluding butter; eggs; cereal and bakery products; sugar; and coffee, tee and cocoa) and the total availabil-

ity of seventeen nutrients (food energy; protein; carbohydrates; fat; cholesterol; calcium; magnesium; phosphorous; iron; zinc; vitamins A, B<sub>6</sub>, B<sub>12</sub>, C and E; niacin; riboflavin; and thiamin), plus a set of age distribution and ethnicity variables, per capita disposable personal income, and the consumer price index for all items excluding food. But year-to-year measures of the nutritional content of food items at this level of disaggregation are neither published by the USDA nor readily available from other sources for this sample period.

Nearly ten years ago, with the generous assistance from Nancy Raper of the Human Nutrition Information Service, using unpublished handwritten documents I was able to compile annual estimates of the percentages of the seventeen nutrients supplied by the twenty-one foods for the period 1952-1983. 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 - e.g., grams of protein per pound of beef. These original 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 1952-1983. Figure 1, which depicts the resulting estimated time paths for the energy content of U.S. foods, is fairly representative of the types of fluctuations that occur for all seventeen of the nutrients over this period. A third problem has been 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 entire 32-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 im-

proved diets, not to mention the simple fact that a hog in 1909 was a very different creature than the typical barrow or gilt of today.

This leads us 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 foods from the U.S. food supply and the total availability of nutrients from those foods for the entire period from 1909-1994. I also have a sample of estimates for the individual nutritional content of each of these food items for the period 1952-1983. However, the food quantity and nutrient availability data has been updated several times by the USDA since the sample of 32 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. The shorter 32-year data set can be used to draw inferences about the likely behavior of the joint distribution of the elements of the nutrient content matrix. 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 forecasts outside of the sample data and to draw inferences on the reasonably likely ranges 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 1909-1994.

### **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* recently developed by Golan (1994), Golan, Judge, and Miller (1996), and Golan, Judge, and Perloff (1996). 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  $z_t \in \mathbb{R}_+^K$  be the  $K$ -vector of nutrients available for consumption per capita in the food supply in year  $t$ , let  $x_t \in \mathbb{R}_+^{n_x}$  be the  $n_x$ -vector of food quantities consumed per capita, and write the linear relationship between food and nutrients as

$$(1) \quad z_t = A_t x_t, \quad t = 1, \dots, T,$$

where  $A_t$  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  $\bar{A}^\circ$ , obtained independently of the current inference problem. But we do not have data on the nutrient content matrices on an individual 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,  $\mathbf{a} \geq \mathbf{0}$  satisfying  $z = \mathbf{a}'x$ , given a prior estimate of the nutrient content vector,  $\mathbf{a}^\circ$ , and observations on  $z$  and  $x$ . We first specify a compact interval of support for each  $\alpha_i$  containing the prior estimate,  $\alpha_i^\circ \in [\underline{\alpha}_i, \bar{\alpha}_i]$ ,  $i = 1, \dots, K$ , say, divide each interval into  $N$  subintervals,

$$\left[ \left( \frac{N-n+1}{N} \right) \underline{\alpha}_i + \left( \frac{n-1}{N} \right) \bar{\alpha}_i, \left( \frac{N-n}{N} \right) \underline{\alpha}_i + \left( \frac{n}{N} \right) \bar{\alpha}_i \right], \quad n = 1, \dots, N,$$

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

$$(2) \quad \alpha_i = \underline{\alpha}_i p_{i0} + \left[ \left( \frac{N-1}{N} \right) \underline{\alpha}_i + \left( \frac{1}{N} \right) \bar{\alpha}_i \right] p_{i1} + \cdots + \left[ \left( \frac{1}{N} \right) \underline{\alpha}_i + \left( \frac{N-1}{N} \right) \bar{\alpha}_i \right] p_{iN-1} + \bar{\alpha}_i p_{iN}$$

$$= \sum_{j=0}^N [\underline{\alpha}_i + (j/K) \delta_i] p_{ij}, \quad i = 1, \dots, K,$$

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  $\alpha$  solves

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

$$p_{ij} \geq 0 \quad \forall i, j,$$

$$\sum_{j=0}^N p_{ij} = 1 \quad \forall i,$$

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

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 non-negativity constraints are slack at the optimal solution. The GME solution can be written in the form

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

with the normalizing condition

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

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



$$(6) \quad \alpha_i = \underline{\alpha}_i + \sum_{j=0}^N \left( \frac{j}{N} \right) \delta_i \frac{\exp\{-\lambda \delta_i x_i (j/N)\}}{\sum_{j=0}^N \exp\{-\lambda \delta_i x_i (j/N)\}}, \quad \forall i = 1, \dots, n_x,$$

while the Lagrange multiplier for the mean constraint is defined by

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

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 seems to me largely subjective and not necessarily the result of truly prior information. However, with regard to the nutritional content of foods, we do know (with probability one) that any given food item can not account for less than zero nor more than 100 percent of a given nutrient's total availability. This gives us a natural choice for the support of the elements of  $\mathbf{a}$ . 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  $\mathbf{a}$  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  $\alpha$ . One way to overcome this subjectivity is to let  $N \rightarrow \infty$  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  $-\sum_{i=0}^N p_i^* \ln(p_i^*) + \ln(N+1)$ , where

$p_i^*$  is the optimal choice for the  $i^{th}$  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  $1/(N+1)^2$ . Thus, an asymptotic approximation will become accurate quite 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  $i = 0, 1, \dots, N$ , define  $p_i(s) \equiv p_{i[sN]}$ . For given  $i$ ,  $N$ , and  $0 \leq j \leq N$ , let  $s$  satisfy  $j/N \leq s < (j+1)/N$ . Then, uniformly in  $s \in [0,1]$ , the  $i^{th}$  cumulative probability distribution function satisfies

$$\begin{aligned}
 (8) \quad F_i(s) &\equiv \frac{\sum_{k=0}^{[sN]} (1/N) \exp\{-\lambda \delta_i x_i (k/N)\}}{\sum_{k=0}^N (1/N) \exp\{-\lambda \delta_i x_i (k/N)\}} \\
 &= \frac{\int_0^{[sN]/N} \exp\{-\lambda \delta_i x_i ([uN]/N)\} du}{\int_0^1 \exp\{-\lambda \delta_i x_i ([uN]/N)\} du + (1/N) \exp\{-\lambda \delta_i x_i\}} \\
 &\xrightarrow{N \rightarrow \infty} \frac{\int_0^s \exp\{-\lambda \delta_i x_i u\} du}{\int_0^1 \exp\{-\lambda \delta_i x_i s\} ds} = \frac{1 - e^{-\lambda \delta_i x_i s}}{1 - e^{-\lambda \delta_i x_i}},
 \end{aligned}$$

which is a truncated exponential cumulative distribution function, with the Lagrange multiplier  $\lambda$  now defined by the mean condition

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

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-forecast prior* distribution, and/or a loss function?

I began this process quite 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 1952-1983 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 it is well known that 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 the GME approach. However, the Kullback-Leibler criterion can be applied to any prior. I also am comfortable with uniform priors on compact intervals before undertaking *any* data analysis.

Thus, the one thing I remain reticent to impose is a specific assumption about the likelihood function for the shorter, but more 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 the use of 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 generate the sample estimates for the mean vector, say  $\hat{\mathbf{a}}$ , and variance-covariance matrix, say  $\hat{\mathbf{S}}$ , by simply applying the method of moments to the 32-year data set. In this instance, this produces a *post-data density* in the form

$$(10) \quad f(\mathbf{a}|D) \sim N\left(\hat{\mathbf{a}}, \frac{1}{n}\hat{\mathbf{S}}\right),$$

where  $\hat{\mathbf{a}}$  is the  $n_x$ -vector of sample means and  $\hat{\mathbf{S}}$  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 1953-1982.

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,<sup>3</sup>

$$(11) \quad f_0(\mathbf{a}) = (2\pi)^{-n_x/2} \left| \frac{1}{n} \hat{\mathbf{S}} \right|^{-1/2} \exp \left\{ -\frac{n}{2} (\mathbf{a} - \hat{\mathbf{a}})' \hat{\mathbf{S}}^{-1} (\mathbf{a} - \hat{\mathbf{a}}) \right\}.$$

For the Kullback-Leibler criterion, the objective is to

$$(12) \quad \text{minimize } \int \cdots \int f_1(\mathbf{y}) \log[f_1(\mathbf{y})/f_0(\mathbf{y})] d\mathbf{y}_0 \cdots d\mathbf{y}_N, \text{ subject to}$$

$$\int \cdots \int f_1(\mathbf{y}) d\mathbf{y}_0 \cdots d\mathbf{y}_N = 1,$$

$$\sum_{i=1}^N x_i \left[ \int \cdots \int y_i f_1(\mathbf{y}) d\mathbf{y}_0 \cdots d\mathbf{y}_N \right] = z.$$

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

$$(13) \quad f_1(\mathbf{a}) = (2\pi)^{-n_x/2} \left| \frac{1}{n} \hat{\mathbf{S}} \right|^{-1/2} \exp \left\{ -\frac{n}{2} (\mathbf{a} - \mathbf{a}_1)' \hat{\mathbf{S}}^{-1} (\mathbf{a} - \mathbf{a}_1) \right\},$$

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

Figure 2 displays the results of the mean calculations for the energy content of each food in year in the period 1909-1994. 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!

## Conclusions

The BMOM and GME solution to the nutrient inference problem produces the same algebraic result as the following classical approach. First we would use a simple least squares procedure to

estimate the sample means and variance-covariance terms. We then would 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  $\hat{\mathbf{a}}$  and  $\mathbf{a}_1$  relative to the quadratic norm  $\frac{1}{n} \hat{\mathbf{S}}$ . But there is a significant difference in both the interpretation and the logic behind these approaches. In finite samples, the Bayesian classical solutions only coincide when the likelihood function is multivariate normal, and this is known *a priori* to be the case. Other likelihood functions generate very different results for the two approaches. In addition, the Bayesian information methods applied here basis for inference on the likely values of the nutrient content elements on a year-to-year basis. As an illustration, figure 3 depicts 99 percent confidence intervals for the four nutrient content elements that are the most volatile in the 32-year complete information data set – butter, ice cream and frozen yogurt, pork, and sugar and other sweeteners. The distributions for these estimates could easily be incorporated into calculations of standard errors or confidence intervals for such entities as price elasticities of nutrient consumption as a way to obtain reasonable bounds on the likely response of nutritional intake levels to changes in farm and food policy. This is a set of very powerful tools that should become common in the economic analysis of agricultural problems and issues.

## References

- Csiszár, I. "Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems." *The Annals of Statistics* 19 (1991): 2032-2066.
- 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* 10 (1994): 1-46.
- 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." *Journal of the American Statistical Association* 91 (1996): 841-853.
- Jaynes, E. "Information Theory and Statistical Mechanics." *Physics Review* 106 (1957a): 620-30.
- \_\_\_\_\_. "Information Theory and Statistical Mechanics II." *Physics Review* 108 (1957b): 171-90.
- \_\_\_\_\_. "Prior Information and Ambiguity in Inverse Problems." in D. McLaughlin, ed., *Inverse Problems*, Providence RI: American Mathematical Society, 1984: 151-166.
- Kullback, J. *Information Theory and Statistics*, New York: John Wiley and Sons, 1959.
- Kullback, J. and R. Leibler. "On Information and Sufficiency." *Annals of Mathematical Statistics* 4 (1951): 99-111.
- LaFrance, J. "The Structure of U.S. Food Demand." unpublished manuscript, University of California, Berkeley, 1998.
- LaFrance, J. and M. Hanemann. "The Dual Structure of Incomplete Demand Systems." *American Journal of Agricultural Economics* 71 (1989): 262-274.

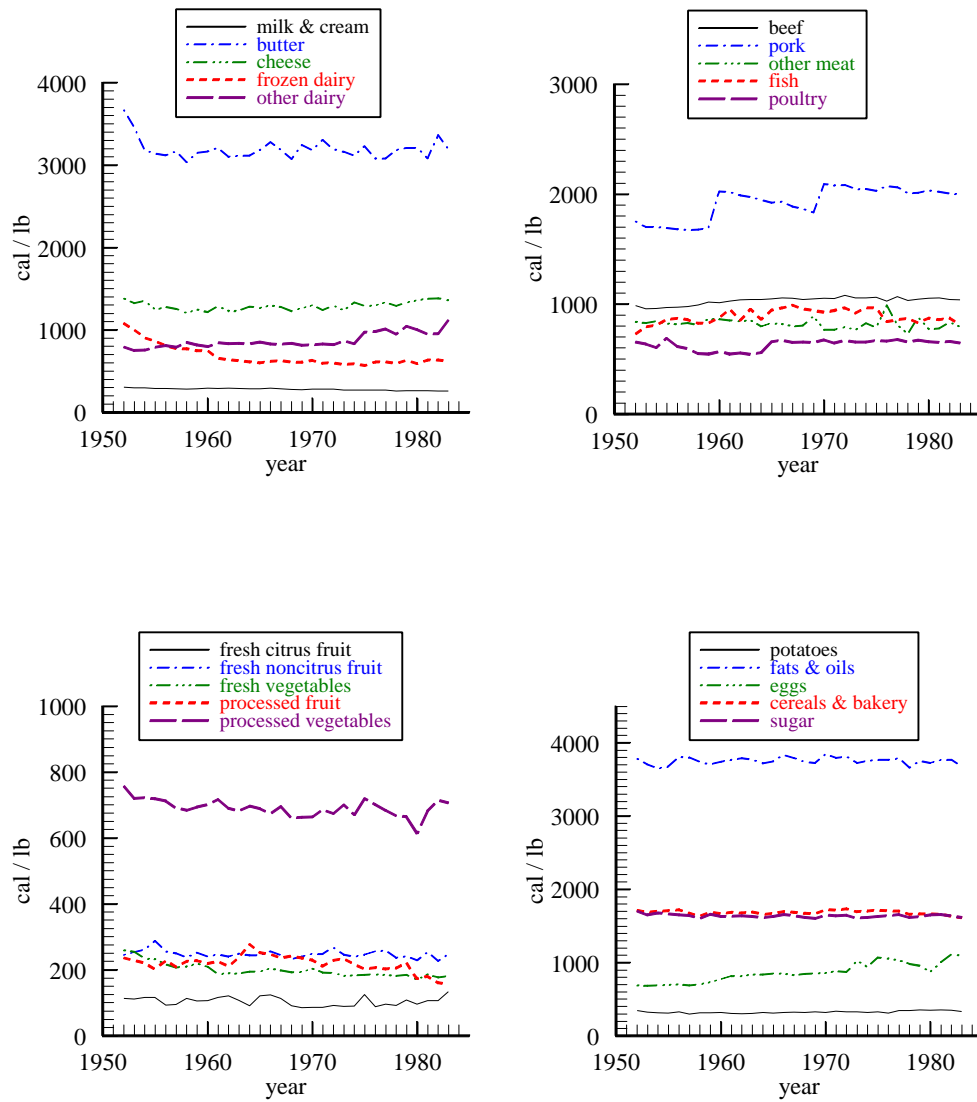
- Shannon, C. "A Mathematical theory of Communication." *Bell System Technical Journal* 27 (1948): 379-423.
- Tobias, J. and A. Zellner. "Further Results on the Bayesian Method of Moments Analysis of the Multiple Regression Model." H.G.B. Alexander Research Foundation, University of Chicago, 1997. Presented at the Econometric Society meeting, June 1997.
- Zellner, A. "The Bayesian Method of Moments (BMOM): Theory and Applications." In T. Fomby and R. Hill, eds., *Advances in Econometrics* 12 (1997): 85-105.
- 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

**Table 1. Sample Means and Standard Errors, Energy Content of U.S. Foods, 1952-1983.**

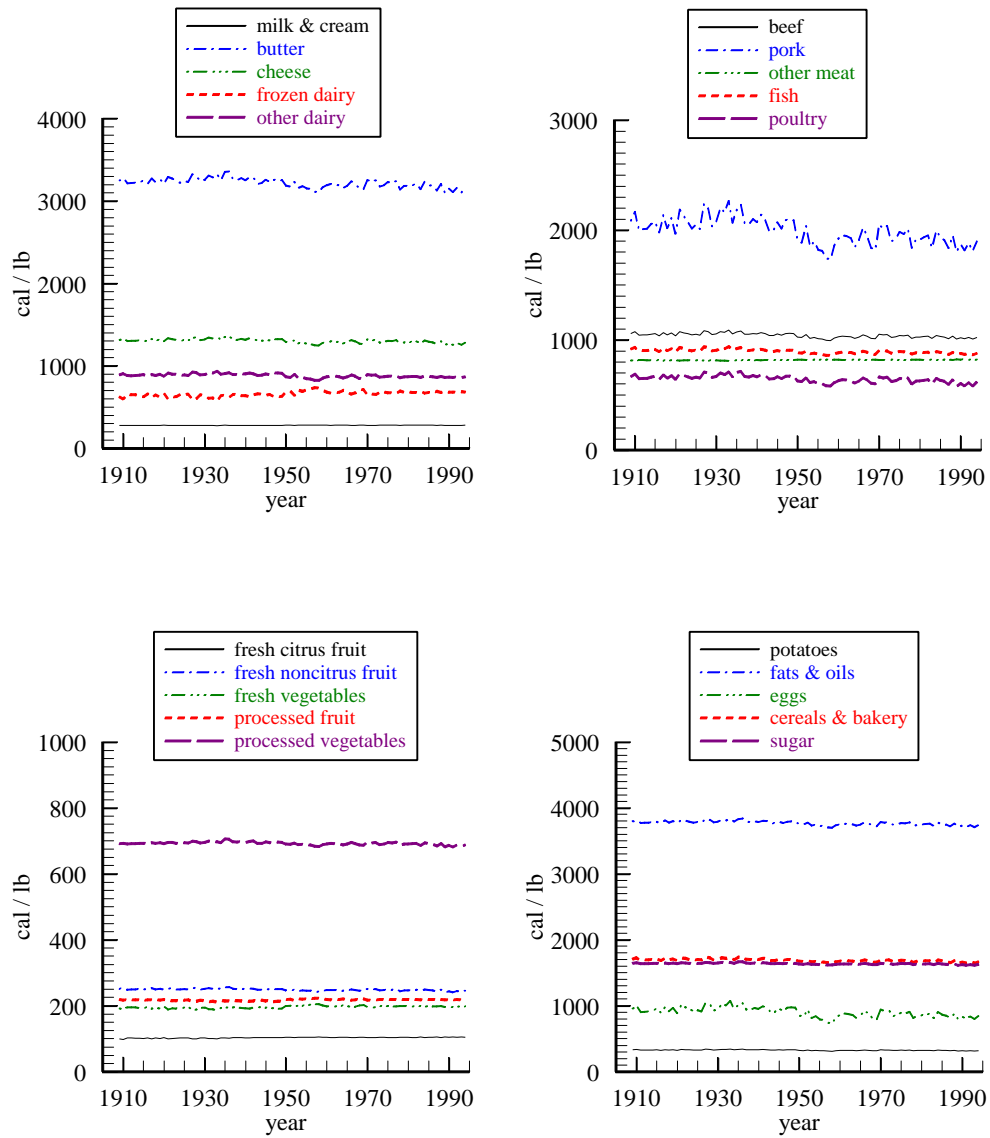
<b>Food Item</b>	<b>Sample Mean</b>	<b>Standard Error</b>
Fresh Milk and Cream	279.853	2.27135
Butter	3193.76	21.3308
Cheese	1290.95	8.86004
Ice Cream and Frozen Yogurt	683.073	22.0997
Canned and Powdered Milk	869.336	15.7867
Beef and Veal	1029.37	5.88297
Pork	1925.22	25.4956
Other Red Meat	822.742	8.22912
Fish	887.037	10.9005
Poultry	630.447	8.17579
Fresh Citrus Fruit	104.220	2.39075
Fresh Non-Citrus Fruit	247.969	2.03088
Fresh Vegetables	199.849	3.69462
Potatoes and Sweetpotatoes	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



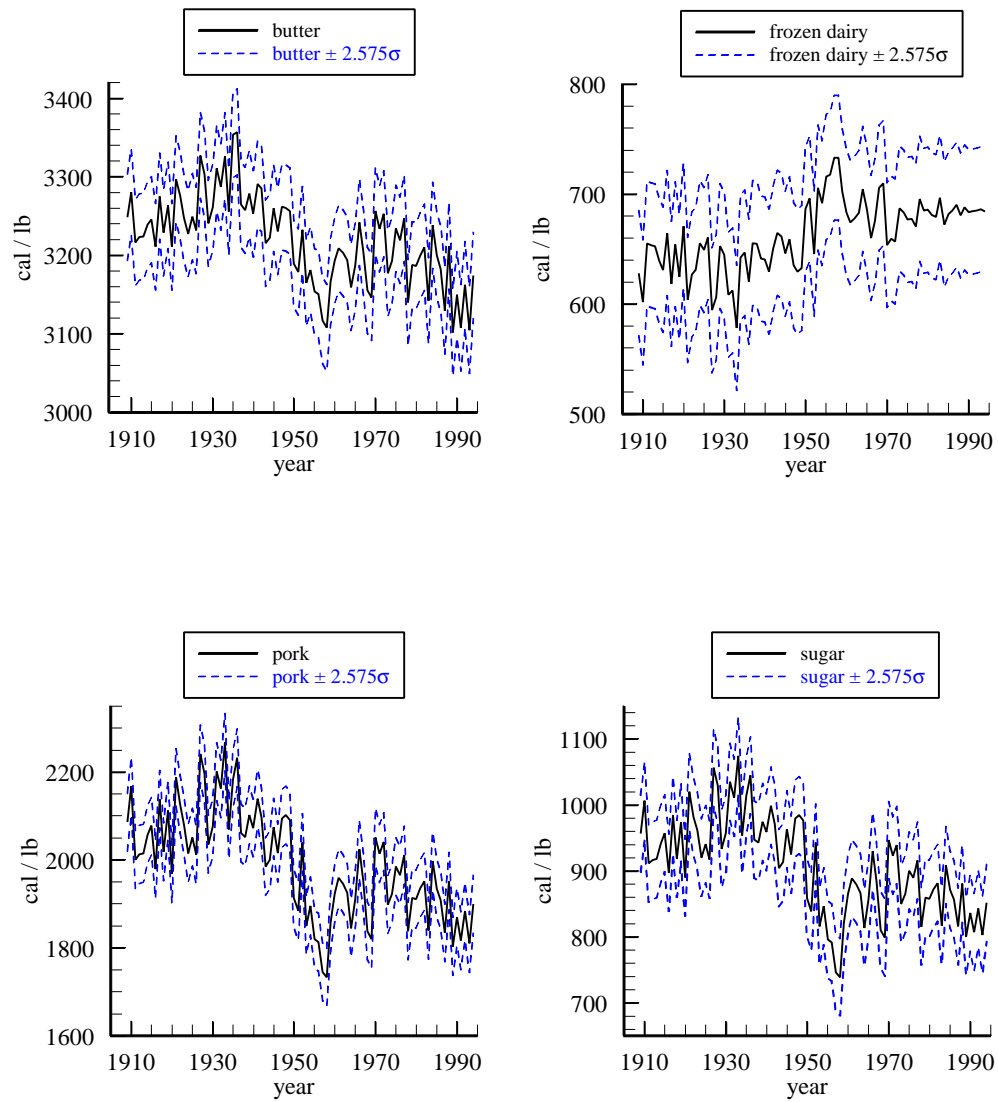
**Figure 1. Energy Content of U.S. Food, 1952-1983.**



**Figure 2. Predicted Energy Content of U.S. Food, 1909-1994.**



**Figure 3. Energy Content of Select Foods, 99% Confidence Intervals.**



### Endnotes

<sup>1</sup> 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).

<sup>2</sup> See, e.g., LaFrance for one example of this type of empirical model.

<sup>3</sup> 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.