New Economic Approaches to Consumer Welfare and Nutrition
A Food & Agricultural Marketing Consortium Conference
Sponsored by the Economic Research Service

January 14-15, 1999

New Economic Approaches to Consumer Welfare and Nutrition

Conference Proceedings

Food and Agricultural Marketing Consortium
Alexandria, Virginia
January 14-15, 1999

Income, Program Participation, and the Choice of Dietary Pattern

First Draft of a Paper For Presentation at the
Food and Agricultural Marketing Consortium Meeting
Alexandria, Virginia
January 1999

Abstract
This paper employs new Food Guide Pyramid servings data to measure how food intake is affected by income, food stamps, and WIC. The analysis uses a maximum likelihood estimator that combines the seemingly unrelated regression and one-way error component models, to account for two types of intra-household correlations. The empirical results show that participation in the Food Stamp Program leads to higher intakes of meats, dairy group servings, added sugars, and total fats. Participation in the WIC program leads to diets with decreased intakes of added sugars. Additionally, the estimation of intake correlations across household members and across food groups, provides a tool that can help guide nutrition education efforts.

Authors

Paul McNamara[1]

Parke Wilde

Christine Ranney

[1]McNamara is a research associate in the Department of Agricultural, Resource and Managerial Economics (ARME) at Cornell University. Wilde is an economist at the USDA's Economic Research Service. Ranney is an associate professor in ARME and associate dean of the Graduate School at Cornell.

Introduction

An important recent advance in the nutrition literature has been a new methodology for using a national food intake survey to measure people's diets in the same common-sense terms that are employed by the federal government in its dietary recommendations and the Food Guide Pyramid. For example, we can now measure how many servings of vegetables are consumed by the average 35-year-old American male, correctly counting all vegetables that are ingredients in "mixed" foods, and compare that measurement to the recommendation. However, much of the recent nutrition literature has focused on how diets differ across characteristics that cannot be changed by public policy (such as age and sex in the preceding example). Differences across income levels have been given some attention, but only in a univariate context that economists will immediately perceive as unsatisfactory, with no explicit model of the implied behavior and no control for obvious confounding variables.

This paper reports, in the same common-sense language of the dietary recommendations, how nutritional quality for low-income Americans is affected by three real policy instruments: income levels, food stamp participation and benefit levels, and WIC participation. Intake observations may be correlated in two ways – across foods and across individuals within the same household This paper therefore estimates a system of food and nutrient intake equations using a maximum likelihood approach that joins the seemingly unrelated regression model with the one-way error components model, in order to account for these correlations. This research strategy combines the strengths of recent work in both nutrition and applied economics to shed new light on food intake patterns.

Background and Theory

In the recent nutrition literature, researchers have developed the ability to monitor food intakes in the same categories employed in the U.S. Government's Food Guide Pyramid dietary recommendations (Cleveland, et al. 1997). Monitoring intakes with these categories allows the straightforward assessment of population eating patterns relative to public health nutrition goals (see, for instance, Krebs-Smith, et al. 1995). This development in nutrition research is significant, because important nutrition education efforts communicate nutrition messages in terms of food categories, while previous nutrition monitoring focussed on the intakes of micro-nutrients, such as vitamins.

In the applied economic literature, empirical nutrient intake functions have been motivated using two general theoretical approaches. In the "indirect" approach, one typically defines preferences over foods and other goods, estimates a theoretically-consistent complete demand system for these goods, and employs a fixed-coefficient Leontief matrix to describe nutrient intakes indirectly as a simple linear function of the demand for foods (for example, Huang 1996, Ramezani, Rose and Murphy 1995). In the "direct" approach, one typically defines preferences over nutrients themselves and estimates reduced-form nutrient intake equations (for example, Akin, et al. 1985, Butler and Raymond 1996). Sometimes, preferences are defined as in the "indirect approach," but reduced-form nutrient intake equations are nonetheless estimated as in the "direct" approach (for example, Devaney and Moffitt 1991). Such intake functions, however motivated, are intuitive and useful for predicting the effects of changes in important policy variables.

Empirical Model and Data

The mathematical details of the model are described in Wilde, McNamara, and Ranney, but the basic structure is easy to explain. The random variation in the nutrient intake equations is assumed to be independent across households, but we are interested in two types of correlations for the variation within households. First, a given individual's intake of different foods and nutrients may be correlated, as in the seemingly unrelated regression (SUR) model. Second, individuals in the same household may be more alike than individuals in different households, as in the one-way error-components model from the literature on panel data.

In particular, suppose the residual for intake of food *m* by person *t* in household *n* is written u_{ntm}. Let the intake equations for household *n* be stacked first by food $m = 1,...,M$ and then by household member $t = 1,...,T_n$. The "tall" column vector u_n of residuals for household *n* will have the distribution:

(1) $u_n \sim N(0, \Omega_n)$,

where $\Omega_n = (\Sigma_\mu \quad J_{T_n}) + (\Sigma_v \quad I_{T_n})$ is an (MT_n by MT_n) matrix, J_{T_n} is a (T_n by T_n) matrix of ones, I_{T_n} is an identity matrix, Σ_μ is the (*M* by *M*) "Between" cross-equation variance-covariance matrix for the household error component, and Σ_v is the (*M* by *M*) "Within" cross-equation variance-covariance matrix of the individual error component.

The maximum likelihood estimator used here is a combination of the SUR estimator and a one-way error-component estimator. It accounts for correlations across intake equations and across individuals within households.[1] This general model has three advantages over ordinary least squares (OLS) estimation of each nutrient intake equation separately. First, maximum likelihood estimation is expected to be more efficient than OLS if the correlations across equations and the household effects are not negligible. Second, the usual estimates of the standard errors of the OLS estimator are inconsistent in this framework. Third, the cross-equation correlations and the correlations among individuals in the same household are themselves interesting and worthwhile to estimate.

The data, from the 1994-1996 Continuing Survey of Food Intake by Individuals (CSFII), contain observations on individuals in households. This study employs observations on 3,642 low-income persons (in 1901 households) with income less than or equal to 130 percent of the poverty line. There are seven dependent variables, which together suffice to capture the most important aspects of the federal government's dietary recommendations: the number of daily servings for each of the five main food groups in the Food Guide Pyramid, the daily amount of sweets and added sugars, and total intake of fats. The values for the dependent variables are two-day average intakes.

Independent variables include age, sex, household income, food stamp program variables, WIC participation, and binary variables for ethnicity. Variable names, definitions, and mean values are displayed in Table 1. The age effect is modeled as a linear spline with kinks at ages 12, 20, and 60. This non-linear functional form captures key features of the life-cycle profile in food intake. The income effect is modeled using a quadratic term, to capture the potential curvature of the Engel relationship. The effect of food stamps is modeled with both a binary variable for participation and a continuous variable for monthly food stamp benefits, to capture both the “intercept” and the “slope” effects of food stamps on food intake. Thus, while the specification of each equation is linear in the parameters, it captures the potential nonlinearity in the effects of several key variables under study.[2]

Results

The model yields two types of interesting results. First, it gives the estimated effects of the independent variables on the dependent food and nutrient variables. Second, it gives two cross-equation variance-covariance matrices for the residuals: the “Between” matrix indicates the extent to which a household that generally consumes more of one food is also likely to generally consume more of another food; the “Within” matrix indicates the extent to which an individual who consumes more of one food (relative to the household average) is also more likely to consume more of another food.

Effects of the Independent Variables

Space constraints require that this sub-section focus on the key income and program effects. We note in passing only that the effects of age and sex are significant and reasonable, and that the effects of ethnicity vary from one food category to another (and there is no presumption *a priori* about the direction these effects should take).

Because of the non-linear functional form for key variables, the maximum likelihood parameter estimates in Table 2 are most easily interpreted by considering a representative person. The column of “baseline” servings in Table 3 are the predicted intakes of each dependent variable for a person whose household’s monthly income is equal to the first quartile of the low-income sample (approximately $162 per person), who receives no food stamps or WIC, and who has mean values of all other variables. The subsequent columns show the change in predicted intake in response to changes in income, food stamp participation, and WIC participation, by comparison to the baseline case.

The “intercept” effect of participating in the Food Stamp Program is large and statistically significant for meats, dairy, added sugars, and total fats. The “slope” or marginal effect of further benefits is either not significantly different from zero or (for dairy and meats) slightly negative. The combined food stamp effect, as a proportion of the baseline case, is greatest for added sugars, meats, dairy, and total fats (in that order). This effect is smaller and not statistically significant for vegetables, grains, and fruits.

The WIC program, which includes a nutrition education component and only subsidizes certain foods, appears to cause modest increases in fruit, dairy, and vegetable intake and a decrease in intake of added sugar. Only the added sugar effect is statistically significant. One may hypothesize that this effect is due to substitution of WIC-supplied fruit juices for colas and WIC-supplied low-sugar cereals for higher-sugar cereals.

Cross-Equation Variance-Covariance Matrices

The estimated total variance in the residual for each equation is the sum of the appropriate diagonal element of the “Between” matrix in Table 4 and the corresponding diagonal element of the “Within” matrix. These estimated variances show that variation within and between households are both important. The gain in efficiency over OLS is expected to be larger if there is a substantial “Between” component, as there is.

The off-diagonal elements of these two matrices are almost all positive, with the unsurprising implication that people and households likely to consume more of one food or nutrient are also likely to consume more of the others. The corresponding correlation matrices are omitted here, but according to both matrices, the residual for total fat intake is most correlated with the residuals for grains and meats and least correlated with fruit. The gain in efficiency over OLS is expected to be larger if these matrices have relatively large off-diagonal elements, as they do.

Discussion

This paper’s analytic approach may inform nutrition education efforts in several ways. From the parameter estimates for the food intake equations, it is useful to know that food stamp program participation appears to raise intake of fats and added sugars, for example, while WIC participation appears to lower intake of these nutrients.

From the diagonal elements of the variance-covariance matrices, it is useful to know the decomposition of the residual variance between and within households. For example, the “Within” variance for fat intake is slightly higher than the “Between” variance, so individuals with high fat intake frequently have high fat intake relative to their household average. By contrast, the “Within” variance for fruit intake is slightly lower than the “Between” variance, so individuals with low fruit intake are frequently found in “low-fruit” households. These insights may help to target nutrition education efforts.

From the off-diagonal elements of these same two matrices, one may identify correlations across food and nutrient groups. Higher fruit intake, not surprisingly, appears to be uncorrelated with higher fat intake. By contrast, meat and (perhaps surprisingly) grain products are highly correlated with fat intake. These insights may help in designing nutrition education strategies that achieve underlying goals, such as reduction in fat intake, through more palatable and upbeat messages, such as encouraging fruit intake.

This paper illustrates the advantages of combining the increasingly intuitive and useful data sources developed by nutritionists with recent analytic and statistical techniques from the applied economic literature. Future research along these lines may further illuminate the multivariate relationships between income, program variables, and food intake variables. Such research may also yield new ideas for nutrition education policy.

Notes

[1] This type of error component model for SUR systems of equations is discussed in Avery and in chapter 6 of the textbook on panel data by Baltagi. Magnus discusses maximum likelihood estimation for such models. Beierlein, Dunn and McConnon estimate such a model for a system of energy demand equations.

[2] We have not yet estimated a variant of this model that accounts for the potential endogeneity of food stamp participation. In previous studies of nutrient intake, no significant selection bias was found by Devaney and Moffitt or Akin et al., and selection bias was found in just one of two samples used by Butler and Raymond.

Sources

Akin, J. S., D. K. Guilkey, B. M. Popkin, and K. M. Smith. "The Impact of Federal Transfer Programs on the Nutrient Intake of Elderly Individuals." *Journal of Human Resources. Summer* 20(1985): 383-404.

Avery, R. B. "Error Components and Seemingly Unrelated Regression." *Econometrica* 45(1977): 199-209.

Baltagi, B. H. *Econometric Analysis of Panel Data*. New York: John Wiley and Sons, 1995.

Beierlein, J. G., J. W. Dunn, and J. C. McConnon. "The Demand for Electricity and Natural Gas in the Northeastern United States." *Rev. Econ. and Statist.* 63(1981): 403-408.

Butler, J. S., and J. E. Raymond. "The Effect of the Food Stamp Program on Nutrient Intake." *Economic Inquiry* 34(1996): 781-798.

Cleveland, L. E., D. A. Cook, S. M. Krebs-Smith, and J. Friday. "Method for Assessing Food Intakes in Terms of Servings Based on Food Guidance." *American Journal of Clinical Nutrition* 65(1997): 1254S-63S.

Devaney, B., and R. Moffitt. "Dietary Effects of the Food Stamp Program." *Amer. J. Agr. Econ.* 73(1991): 202-211.

Huang, K. "Food Nutrient Demand Elasticities." *Amer. J. Agr. Econ.* 78(1996): 21-29.

Krebs-Smith, S. M., A. Cook, A. Subar, L. Cleveland, and J. Friday. "U.S. Adults' Fruit and Vegetable Intakes, 1989-1991: A Revised Baseline for the Healthy People 2000 Objective." *American Journal of Public Health* 85(1995): 1623-1629.

Magnus, J. R. "Multivariate Error Components Analysis of Linear and Nonlinear Regression Models by Maximum Likelihood." *J. Econometrics* 19(1982): 239-285.

McNamara, P., C. Ranney, L. S. Kantor, and L. Krebs-Smith. “The Gap Between Food Intakes and the Pyramid Recommendations: Measurement and Food System Ramifications.” Ithaca NY: Cornell University, November 21, 1997.

Ramezani, C. A., D. Rose, and S. Murphy. “Aggregation, Flexible Forms, and Estimation of Food Consumption Parameters.” *Amer. J. Agr. Econ.* 77(1995): 525-532.

Wilde, P., P. McNamara, and C. Ranney. “A Seemingly Unrelated Regression Error Components Model for a System of Food Intake Equations Using 'Pyramid' Servings Data.” Ithaca NY: Cornell University, Forthcoming.