AGRICULTURAL ECONOMICS AND RURAL SOCIOLOGY

CLEMSON UNIVERSITY

working paper

WP 072584 July

THE IMPLICATIONS OF SPATIAL AGGREGA
TO FOOD EXPENDITURE PROJECTIONS

by

James C. O. Nyankori*

clemson university
south carolina agricultural experiment station

department of agricultural economics
and rural sociology
clemson university
clemson, south carolina 29631

WP 072584 July 1984

THE IMPLICATIONS OF SPATIAL AGGREGATION TO FOOD EXPENDITURE PROJECTIONS

by

James C. O. Nyankori*

Working Papers are not subject to internal review within the Department of Agricultural Economics and Rural Sociology, Clemson University.

*Associate Professor, Department of Agricultural Economics and Rural Sociology, Clemson University, Clemson, South Carolina.

THE IMPLICATIONS OF SPATIAL AGGREGATION TO FOOD EXPENDITURE PROJECTIONS

Introduction

Empirical studies of food consumption using Engel analysis show that household income has a significant influence on household total food expenditures (Blanchiforti et al., 1981; Sexauer, 1979; Smallwood and Blaylock, 1981). The influence of income on food expenditures has also been shown to vary across income groups. This suggests that changes in income distribution of households have considerable impact on total food expenditures and, consequently, bear important implications for producers, marketing and public food policy decision makers.

In practice, strategic food marketing and policy adjustments to changing household income distribution require nationwide projections of household total food expenditures. However, given the regional differences in the structures of the food expenditure relationship there arises a question of the relative merit of regional versus national projections.

In this paper, the results from projections of food expenditures using aggregate (national) and subaggregate (regional) data are compared. In the former, a model of U.S. household food expenditures is estimated and used to make national projections and in the latter, the U.S. projections are obtained by aggregating projections from regional models.

Aggregation problems have received considerable treatment in theoretical and empirical studies. In a formal way, Theil developed conditions under which bias can be introduced in aggregating from microeconomic to macroeconomc relations. Green, in a survey of earlier works on aggregation reiterates that the problem of the appropriate type and degree of aggregation belongs to the field of statistical decision theory. In practice, the purpose of the investigator is paramount but ought to be weighed against likely errors from a high degree of aggregation and the cost of using disaggregated model.

Yet the resolution of the problem is compounded by the specification argument (Grunfeld and Griliches, 1960); the loss of information argument (Edwards and Orcutt, 1969; Orcutt, Watts and Edwards, 1968), and the measurement argument (Aigner and Goldfeld, 1973).

Changes in the U.S. Household Income Distribution

Since 1970, there have been noticeable annual variations in the income distributions of U.S. households. Generally, there have been increases in the percentage of households in the lowest two and the top income groups and decreases in the third and fourth income groups between 1970 and 1980.

At the same time, the number of households grew at a steady annual rate. However, the national and regional income distribution patterns and the annual rates of growth in the number of households differ markedly. Furthermore, there are interregional differences in the magnitude and direction of changes in the income distributions and number of households.

Given that the food expenditure-income response structures are specific to income groups, clearly the national total food bill will vary with the number and income distributions of households.

But the intranational and interregional differences in the expenditure response structure, income distributions and the total number of households raise the still unanswered question regarding aggregation and projections and suggests, for practical purposes, a comparison of total food expenditure projections from national and regional models.

Accurate projections are important to a number of food marketing and public policy decisions. Public decisions regarding the allocation of funds to food programs at the national and regional levels are not independent. Projections at the national level are needed to determine the gross program costs, yet regional level information is needed for a more equitable interregional allocation of the program budget. Similarly, in the private sector, national and regional projections are needed to determine gross capital requirements and the allocations to regional processing and distribution capacity development, respectively.

Data and the Models

Information on the income, total food expenditures and demographic characteristics of U.S. households are from the Consumer Expenditures Survey, 1980-81 of the Bureau of Labor Statistics This is a nationwide survey of a sample of approximately 20,000 U.S. households and contains information used to update the Consumer Price Index.

Suppose the household food expenditures relationship is characterized by the Engel function of the form

$$E_i = a + bY_i \qquad (1)$$

where E_i $(i = 1,2,\ldots,n)$ is the total food expenditures for the ith household, Y_i is the the annual household income with the constant term, a, and income coefficient, b.

To explicitly introduce food expenditure-income relationships at different income ranges in relationship (1), household income, Y, is tarnsformed as

$$
\begin{aligned}
y_{1i} &= Y_i \\
y_{ki} &= Y_i - Y_k \text{ if } Y_i > \dot{Y}_k \\
&= 0 \qquad \text{if } Y_i \leqq \dot{Y}_k \\
& \quad k = 1,2,\ldots,K .
\end{aligned}
\tag{2}
$$

The specific values of $\dot{Y}_k$ are exogenously given and correspond to the upper limit income values of the kth fifth (k = 1,2,3,4) of the percentage share of aggregate income received by households in 1980.

The transformed variables in (2) are used to reformulate the Engel relationship (1) as

$$E_i = \alpha + \beta_1 y_{1i} + \beta_2 y_{2i} + \beta_3 y_{3i} + \beta_4 y_{4i} + \beta_5 y_{5i} + u_i \tag{3}$$

where α is the constant term and β_1 is the marginal propensity to spend. The change in the marginal propensity to spend is β_k $(k > 1)$ at household income $Y_i = \dot{Y}_k$ $(k = 2,3,\ldots,5)$, and u_i is the error term. Equation (3) is estimated with the logarithm of E_i as the dependent variable.

The set of income values Y_k: $\dot{Y}_1 < \dot{Y}_2 < \dot{Y}_3 < \dot{Y}_4$ defines five income groups and for each income group so defined, the mean total food expenditure, $\bar{E}_g$, is computed as

$$\bar{E}_g = \alpha + \bar{Y}_g \sum_{k=1}^{r} \beta_k, \quad (r,g = 1,2,\ldots,K) \tag{4}$$

where $\bar{Y}_g$ is the sample mean household income for the gth income group and the term in parentheses is the expenditure-income coefficient for the gth income group.

The total household food expenditure for the gth income group is, therefore,

$$E_g = nw_g\bar{E}_g \qquad g = 1,2,\ldots,k$$

$$\sum_{g=1}^{k} w_g = 1$$

where w_g is the proportion of households in the gth income group and n is the total number of households. The total household food expenditure for all income groups is, consequently, the sum of all income group expenditures

$$E = \sum_{g=1}^{k} E_g \qquad g = 1,2,\ldots,k.$$

Suppose the original (1980) vector of household income distribution is characterized by the vector, W_o,

$$W_o = \{w_{o1}\ w_{o2}\ w_{o3}\ w_{o4}\ w_{o5}\}\ . \tag{5}$$

Because of the numerous possibilities of income distribution outcomes, alternative income distribution vectors are generated from (4), for the purposes of projections, using a cumulative distribution factor, $0 < \tau < 100$, such that the resulting percentage households in the gth income group is

$$\begin{aligned} w_g &= (w_{og}/\sum_{g=1}^{h} w_{og})\tau \qquad 1 \leqq g < h \\ &= (w_{og}/\sum_{g=h}^{G} w_{og}) * (100-\tau) \qquad h \leqq g < G\ . \end{aligned} \tag{6}$$

By varying the values of τ, and h, several household income distribution vectors were generated for the U.S. and each of the four regions: Northeast, North Central, South and West.

The total national food expenditure projections, E_n, were computed directly from the estimated parameters of the U.S. model, the generated income distribution vectors and the data on U.S. mean expenditures for each income group. The aggregated projections, E_a, were similarly obtained but using the estimated parameters of of the regional models and the regional group mean expenditures and income distribution vectors, and finally summing across regions

$$E_a = \sum_{i=1}^{4} E_i \qquad (i = 1,2,3,4) .$$

Since the effect of aggregation is not known a priori, the two sets of projections were compared for equality on the basis of a significance test. If it could be established that there were no significant differences between national projections, E_n, and the aggregated projections, E_a, then the problem of spatial aggregation would be considered to be of no empirical consequences in this case. If, however, there were significant differences between the two sets of projections some conclusions that have important implications to the analysis of household food expenditure for policy and marketing decisions could be drawn.

Therefore, the national and the aggregated expenditure projections were compared using the test that the two vectors of projections, E_n and E_a, are simultaneously equal

$$H_0: (E_a - E_n) = 0, \qquad (7a)$$

$$H_1: (E_a - E_n) \neq 0. \qquad (7b)$$

Model Results

The estimated model parameters are presented in table 1. The main effect of income on household total food expenditures (β_1) is significant and positive in all the regional and U.S. models. There are, however, differences in the magnitudes of the estimated main income effect ranging from 0.0000027 in the West to 0.0000466 in the Northeast.

Table 1. Estimated Income Coefficients of Household Food Expenditures in the U.S. and in the Northeast, North Central, South and West Regions of the United States

Coefficient Estimates	United States	Northeast	North Central	South	West
α	1.2818890	1.2081450	1.3027020	1.3496930	1.3592910
	(89.1)	(35.9)	(39.9)	(50.6)	(32.2)
β_1	0.0000204	0.0000466	0.0000081	0.0000119	0.0000027
	(7.2)	(7.3)	(1.3)	(2.1)	(1.34)
β_2	0.0000101	-0.0000307	0.0000236	0.0000193	0.0000414
	(1.8)	(-2.3)	(1.8)	(1.6)	(2.7)
β_3	-0.0000186	-0.0000065	-0.0000060	-0.0000198	-0.0000362
	(-3.4)	(-0.5)	(-0.5)	(-1.6)	(-2.7)
β_4	-0.0000028	0.0000037	-0.0000195	-0.0000022	0.0000051
	(-0.8)	(0.5)	(-2.7)	(-0.3)	(0.6)
β_5	-0.0000077	-0.0000144	-0.0000017	-0.0000077	-0.0000108
	(-3.6)	(-3.1)	(-0.4)	(-1.6)	(-2.5)
R^2	0.19	0.22	0.28	0.17	0.20

Notes: β_1 is the main income coefficient; β_2 is the change in the income effect at $7,999; β_3 is the change in income effect at $12,4999; β_4 is the change in income effect at $20,000; β_5 is the change in income effect at $34,999. The dollar values refer to the annual household income levels and the numbers in parentheses are the t-values.

The changes in the effect of income on total household food expenditures at household income \$7,9999 (β_2), were significant and positive in all except the Northeast and varied in magnitude from -0.0000307 (Northeast) to 0.0000414 (West). At household income \$12,499, there were no significant changes in the effect of income on total food expenditure expenditures (β_3) in the Northeast and North Central but there were negative changes in the South, West and U.S.

In all but the North Central, there are no significant changes in the effect of income, (β_4), on total food expenditures at household income \$19,999. Finally, at household income \$34,999 there was a decrease in the effect of income (β_5) on total food expenditures in all but North Central.

The estimated income coefficients shown in table 1 were used to generate two sets of U.S. food expenditure projections: the national model projections and the aggregated projections. The ratios of the national projections expressed as percentage of the aggregated regional projections are shown in figure 1. Clearly the national projections are lower than aggregated projections in all income distribution scenarios.

On the basis of the test of significance, the hypothesis that there is no difference between the national and aggregated projections is rejected. Projections using the national model of U.S. households were invariably lower than the those computed by aggregating projections from regional models of households.

Summary and Conclusions

In this paper, the problem of aggregation was discussed within the specific framework of total food expenditure projections using household food expenditures data aggregated at national and regional levels.

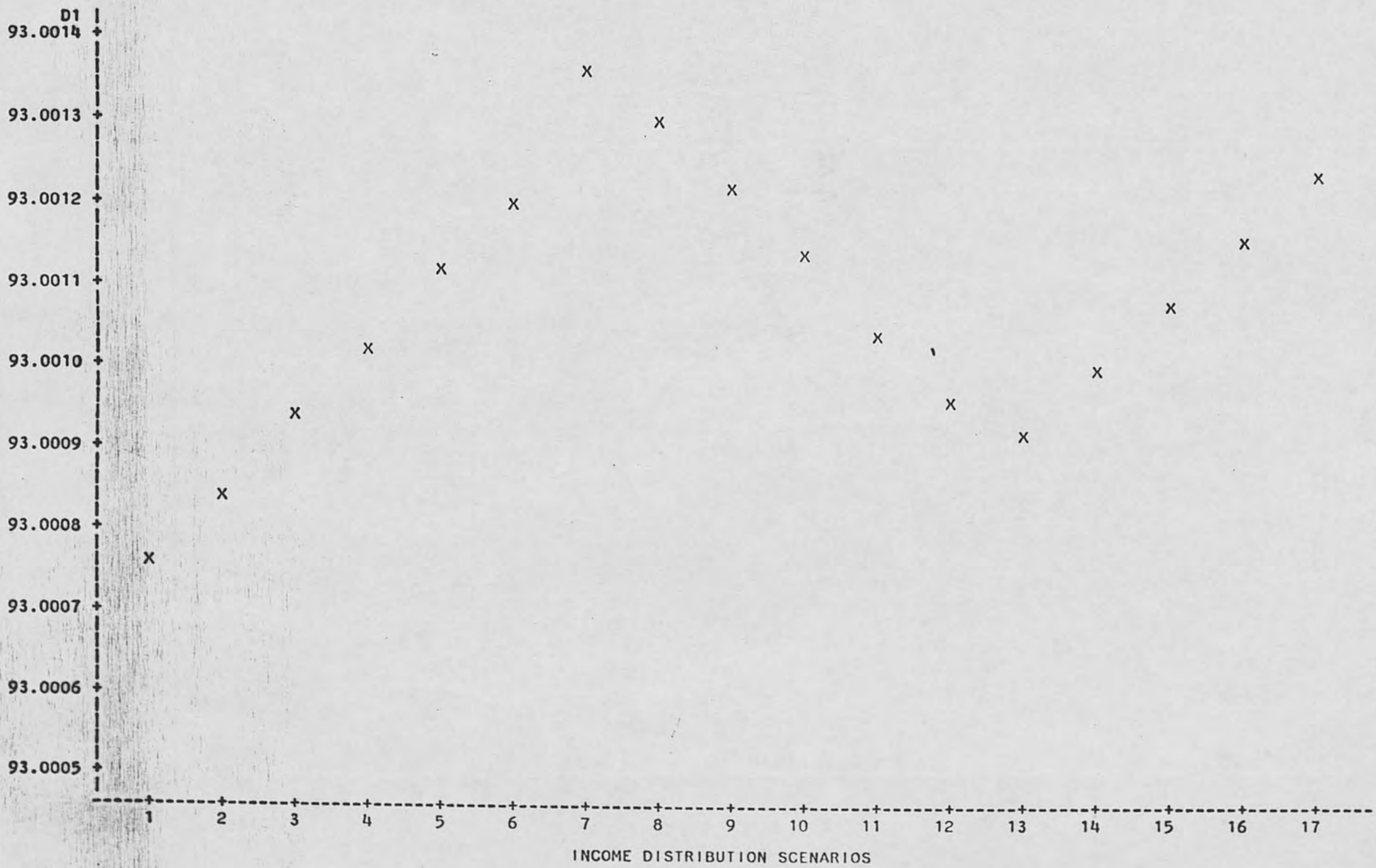

Figure 1. The rations of national to aggregated regional food expenditure projections (Percentages)

Regression estimates of the food expenditure responses across prespecified income groups were generated for U.S. households and for households in the Northeast, North Central, South and West. These, together with national and regional income data were used to make household total food expenditure projections under alternative income distribution scenarios.

A comparison of the national projections using the estimated national expenditure parameters and aggregated regional projections showed that the former were invariably lower than the national projections. Furthermore, a significance test led to the rejection of the null hypothesis that the national and aggregated projections were equal.

The general conclusion is that there are significant differences in food expenditure projections from aggregated and subaggregated data. Since subaggregate data contains additional information inevitably lost in aggregation, aggregated projections are preferred, especially in situations where underestimation is more costly than overestimation.

References

Aigner, D. J. and S. M. Goldfeld, "Simulation and Aggregation: A Reconsideration." Rev. Econ. Stat. 55(1973):114-18.

Blanchiforti, L., R. Green, and S. lane. "Income and Expenditure for Relatively More Nutritious Versus Relatively Less Nutritious Foods Over the Life Cycle.' Amer. J. Agr. Econ. 63(1981):255-60.

Dunn, D. M., W. H. Williams, and T. L. DeChaine, "Aggregate Versus Subaggregate Models in Local Area Forecasting". J. Amer. Stat. Assoc. 71(1976):68-71.

Edwards, John B. and Guy H. Orcutt, "Should Aggregation Prior to Estimation be the Rule?" Rev. Econ. Stat. 51(1969):409-20.

Green, John H. A., Aggregation in Economic Analysis, An Introductory Survey, Princeton, New Jersey: Princeton University Press, 1970.

Grunfeld, Yehuda and Zvi Grilliches, "Is Aggregation Necessarily Bad?" The Review of Economics and Statistics, 42(1960):1-13.

Orcutt, Guy, H., Harold W. Watts and John B. Edwards, "Data Aggregation and Information Loss." Amer. Econ. Rev. 58(1968):773-87.

Sexauer, B. "The Efect of Demographic Shifts and Changes in the Income Distribution on Food-Away-From-Home Expenditures." Amer. J. Agr. Econ. 61(1979):1046-57.

Smallwood, D., and J. Blaylock. Impact of Household Size and Income on Food Spending Pattern. ESCS Tech. Bull. 1650. 1981.

Theil, Henri, Principles of Econometrics, New York: John Wiley Sons, Inc., 1971.

SOUTH CAROLINA AGRICULTURAL EXPERIMENT STATION,
CLEMSON UNIVERSITY, CLEMSON, SOUTH CAROLINA
W. C. GODLEY, ASSOCIATE DEAN AND DIRECTOR
S. C. EXPERIMENT STATION
LUTHER P. ANDERSON, DEAN, COLLEGE OF AGRICULTURAL SCIENCES