

Health Knowledge and Nutritional Adequacy
of Female Heads of Households
in the United States*

CYRUS RAMEZANI

and

CLAUDIA ROEDER

Department of Consumer Science

University of Wisconsin-Madison

February 19, 1995

*Cyrus Ramezani is an assistant professor and Claudia Roeder was a masters student in Consumer Science. We thank Helen Jensen, Robin Douthitt, Karen Holden, Wen Chern, three journal referees, and seminar participants at the University of Wisconsin-Madison for their valuable comments. Any errors remaining are the authors'. Please direct comements to the senior author at the Dept. of Consumer Science, University of Wisconsin, 1300 Linden Dr., Madison, WI 53706.

**Health Knowledge and Nutritional Adequacy of Female Heads of Households
in the United States**

ABSTRACT

Female-headed households – consumer units with a female reference person, with or without children and with no spouse – are a growing segment of the U.S. population. Data from a recent consumption survey (CSFII-89) are used to identify the socioeconomic determinants of nutritional adequacy for this group. Nonparametric procedures are used to select a parametric model that best describes the data. The results indicate that education, health (nutrition) knowledge, income, and degree of urbanization are important predictors of the overall nutritional status of this population. Among these, health knowledge is the most significant, suggesting that improving such knowledge could lead to more informed decisions and an enhancement of diet quality. Perceptions, demographic variables, time use, and spatial indicators are found to be less important predictors of intake.

Health Knowledge and Nutritional Adequacy of Female Heads of Households in the United States

Over the last decade, a large number of studies have attempted to identify the influence of socioeconomic factors on individuals' nutrient intake. This literature has provided insightful explanations of what determines nutrient consumption patterns, and from a public policy perspective, what steps may be taken to change consumer behavior and enhance the nutritional status of the population as a whole (Senauer, Asp & Kinsey 1991, Morgan 1986). A criticism of the existing literature is that individuals from various social strata have been lumped together in broad studies of the population as a whole. Consequently, certain segments of the population have received less attention. The purpose of this study is to address this criticism by focusing on a specific group.

The target population is female-headed households – consumer units with a female reference person, with or without children and with no spouse present. This group is a growing segment of the U.S. population (U.S. Bureau of the Census 1990) and its growth is generally attributed to changing patterns of marriage, fertility, divorce, and the tendency to live alone, and an increased earnings capacity for women (Wojkiewicz, McLanahan & Garfinkel 1990). Poverty is the predominant characteristics of this segment of the population and as a consequence a large portion are dependent on welfare (McLanahan & Booth 1989). Despite a growing concern for the welfare of this segment of the population, no comprehensive analysis of factors influencing their nutritional adequacy currently exists. The present analysis aims to fill this gap. In general, a better understanding of the link between nutrient intake and socioeconomic characteristics of individuals is fundamental to the design of public policies that target specific groups – e.g., the Food Stamp Program (FSP) – or the public at large – e.g., food product labeling.

The literature on socioeconomic determinants of nutritional adequacy is extensive.¹ No previous work, however, has considered the nutritional adequacy of female-headed

families using the types of variables considered in this study. A number of studies provide descriptive analyses of the economic characteristics of this population, particularly single mothers. Recently, Lino & Guthrie (1994) compared food expenditures, shopping behavior, and diet quality of single mothers with married mothers and concluded that the former have diets of lower quality. This conclusion was based on the finding that the mean value of three composite measures of diet quality were statistically different in the two samples. Lino and Guthrie (1994, Page 20) suggest that “nutrition education programs that consider the income constraints and food shopping behavior of single-parent families are needed.” Without multivariate analysis, however, it is unclear how they arrive at such recommendations. The present study removes such ambiguities by using multivariate analysis and new estimation techniques that reduce the errors associated with model specification – an important consideration in the context of modelling nutrient demand (Behrman & Deolalikar 1988, Ramezani 1993, Ramezani 1994).

Section 2 outlines the proposed modelling approach. Section 3 describes the data and the construction of *new* explanatory variables related to health knowledge, dietary perception, and educational attainment. Section 4 reports the results of the analysis. The key finding of the study is that nutrition knowledge and education are highly significant predictors of intake, suggesting that marginal food choices are a likely consequence of informed decisions. Policies that aim to enhance nutritional knowledge are therefore likely to improve diets, independent of other household socioeconomic characteristics. Furthermore, as Mothersbaugh, Herrmann & Warland (1993) have found, higher levels of nutritional knowledge could mitigate the negative effects of time constraints, which are more severe for this segment of the population. These findings suggest that the returns to public investment in expanding nutritional knowledge could be significant, particularly in the long run. The final section provides a summary of the results and suggests directions for further research.

Modelling Nutrient Intake

The “household production” models discussed in Deaton & Muellbauer (1980), Behrman & Deolalikar (1988), and Eastwood (1985) provide the theoretical basis for linking nutrient demand to a host of socioeconomic variables and income. The estimated relationships are the “reduced form” from an optimizing model of consumer choice in which individuals select goods so as to maximize the utility derived from the goods’ attributes, in this case, nutrients. Under such formalism, it can be shown that the demand for any nutrient (N) is dependent upon a set of socioeconomic variables (S) which include prices and income, $N = f(S)$. Economic theory offers little guidance on the appropriate functional specification for $f(S)$, and in empirical work reliance on Ordinary Least Square (OLS) estimation implies a linearized approximation such as:

$$N = f(S) \cong \alpha + \beta_1 S_1 + \beta_2 S_2 + \dots + \beta_n S_n + \epsilon \quad (1)$$

where α is the intercept term, β_i 's are the regression coefficients associated with variables S_i (its square or logarithm), and \cong signifies the fact that this dependence is presumed to be approximately linear. Estimations via OLS or instrumental variable techniques maintain the linearity assumption but differ with respect to assumptions about the residuals of the fitted model. A critical limitation of the linear specifications is that the estimated parameters (β_i 's) measuring the marginal response of the dependent variable remain constant over the range of the explanatory variable(s). In reality, the dependent variable's response to changes in any explanatory variable(s), say income, is likely to vary considerably with its level. Forcing a linear form on data that reflect such variability will lead to significant misspecification error and biased estimates (Judge, Griffiths, Hill & Lee 1980, pp. 616-619).

In modelling nutrient demand, there are intuitive priors that intake and certain explanatory variables – income, education, age, and physical measures such as weight and height, etc. – are likely to be related in a nonlinear manner. For example, intake of most nutrients is likely to initially rise with income and then become invariant to changes in income (or fall) beyond a certain threshold (Behrman & Deolalikar 1988, Ramezani 1993, Ramezani 1994). To allow for this type of interaction, previous studies have relied on *ad hoc* specifications such as the quadratic or the logarithmic transformation, a choice that could lead to erroneous inferences.

This analysis relies on semiparametric functional form estimation procedures to avoid errors arising from inappropriate specification. The Locally Weighted Regression (LWR) technique of Cleveland & Devlin (1988), which aims to ascertain the best functional relationship between the dependent and the explanatory variables is used to assess the nutrient intake relationship. LWR is primarily a graphic tool that is particularly useful for exploratory data analysis. The “smoothed” plots generated provide a basis for parametric specification and estimation. In this study the LWR technique is described in the context of modelling nutrient intake. It is hypothesized that the intake of specific nutrient N is related to a vector of demographic and economic variables S via the following additive structure:

$$N = \alpha + \sum_i^n \beta_i \phi_i (S_i) + \epsilon \quad (2)$$

where the subscript i refers to the i -th explanatory variable, including interaction terms that may be defined as $S_k = S_l * S_m$, and ϵ is an error term with zero expectation. This is a relatively general specification, providing good approximation to arbitrary functions $f(S)$. Equation (2) is a generalization that replaces level representation of the explanatory variables in (1) with functions, $\phi_i (.)$, whose value can change over the range of the explanatory variable(s) – e.g., $\log(S_i)$. The advantage in using nonparametric methods

is that the data determine the appropriate transformation, $\phi_i(\cdot)$. As in the OLS setting, Analysis of Variance can be used to identify potential interaction terms ($S_k = S_l * S_m$).

LWR assumes the $\phi_i(S_i)$'s are smooth functions and ϵ 's are independent and identically distributed with mean zero and finite variance. The latter assumption is similar to the standard regression and is more likely to be met under LWR, as this specification permits variation over the range of the data. To obtain a best estimate of the pairwise functional relation of N and any S_i , the LWR algorithm operates in the same manner as a moving average, where a prespecified portion of the data nearest to each observation is used to estimate a locally linear or quadratic weighted regression line. The LWR algorithm assigns weight to each observation. The weights sum to unity and closer observations receive higher weight.²

To provide a more concrete example, consider the pair-wise relationship between calorie intake and income: $\phi_i(\text{income})$. Suppose we select the neighborhood (window) size to be 50% of the data. Then at any data point – say income equals \$1,000 – the LWR techniques would fit a weighted regression line to 50% of the total sample nearest to this point, where the weights assigned to each observation decrease with distance. In this manner, a curve is fitted at each data point. The loci of all fitted curves provide an estimate, $\hat{\phi}_i(\text{income})$, of the overall relationship between intake and income, $\phi_i(\text{income})$. To capture the curvature in the data, the weighted regression at each point – the local approximation – can be selected to be linear or quadratic. Plots of the smoothed regression surface are used as an aid in determining the appropriate parametric specification.

In general, the greater the fraction of data used to fit the regression, the smoother the estimated surface. At one extreme the window size may be single observations ($\frac{1}{n}$), resulting in a line connecting each observation as the estimate of the surface (high variability but low biased estimate). At the other extreme the window size may be selected to be the total sample ($\frac{n}{n}$), resulting in a smooth fit (low variability but high bias). Hence, the

window size has important implications for the relationship that is uncovered. There are numerous methods for selecting the appropriate window size (Altman 1992). The results reported below, however, are robust to the choice of the window size. To support this assertion, the plots reported include estimates of the functional relationship using Kernel estimators that adaptively select the appropriate window size (Friedman 1984).

The LWR and other nonparametric curve fitting procedures are now widely available as part of various statistical packages and can be used for data analysis with little programming effort.³ Their usefulness has been discussed in many recent surveys (Delgado & Robinson 1992, Altman 1992, are among the most lucid presentations). Interested readers are also referred to manuscripts by Delgado & Robinson (1992), Cleveland and Devlin (1988), and Härdle (1989) for discussions of the properties of these estimators and issues related to inference and forecasting. A detailed discussion is not presented here since the primary goal in using these techniques is to develop a reasonable parametric model that lends itself to least-square estimation, whose statistical properties are well understood.

The Data

The data used in this analysis were obtained from the 1989 Continuing Survey of Food Intake by Individuals (CSFII) in conjunction with the 1989 Diet and Health Knowledge Survey (DHKS). These are among the latest food consumption surveys conducted by the the U.S. Department of Agriculture (USDA). The CSFII contains information for 5,204 individuals of all ages from 2,164 randomly chosen households. The DHKS is one of the first surveys to obtain information about knowledge and attitudes towards “Dietary Guidelines for Americans.” Lino & Guthrie (1994) provide a comprehensive description of this survey data. CSFII/DHKS includes information on self-perceptions of diet quality, health- and nutrition-related knowledge, food labels, and food safety issues. Findings in the previous literature are used to select the appropriate socioeconomic and demographic variables. This literature is advanced by considering variables that are new to CSFII/DHKS and therefore have not been analyzed before.

The CSFII respondents were asked to provide three consecutive days of dietary data. The estimates of individuals’ nutrient intake were constructed by the recall (1st day) and diary (2nd and 3rd days) methods. For this study, respondents with two or three full days of food records were chosen. Nutrient intake measures were constructed by averaging the actual intake over the number of days. The data set contained 513 non-pregnant, non-lactating female household heads, 19 years and older, who participated in both the CSFII and the DHKS. Outliers with incomes over \$70,000 (four observations) and an individual with a caloric intake below 10% of RDA were excluded from the sample (final $N = 508$). Nonparametric procedures are sensitive to extreme outliers and these exclusions insure that a few observations do not overtly influence the results.

The DHKS survey gathered much useful and never-before-collected information on households’ food consumption behavior, health and nutrition knowledge, and attitudes

and habits. These data provide a wealth of information that will be analyzed for many years and are likely to result in new understandings and controversies. The present study utilizes a number of new variables, primarily selected to obtain the largest sample size (fewest missing observations). While the variables related to education and perceptions are used in their “raw” form, variables measuring nutrition knowledge are combined to create an index. The inclusion of an index that measures individuals’ awareness and knowledge of diet-health relationships is the most innovative aspect of the present analysis. The index permits us to separate the influence of education (number of years of schooling) from health (nutrition) knowledge. The latter has a more immediate impact on food choice.

Table 2 provides definitions and basic statistics for all variables for all female heads of households and single mothers. It is important to note that the distributional characteristics reported in Table 2 are very similar to that of female-headed households in the total U.S. population (Frazao 1992).

The dependent variables considered measure specific and overall nutrient intake of the female head of household. The first indicator is caloric intake as a percentage of recently revised (National Research Council 1989) Recommended Dietary Allowances (RDA). This is a good measure of overall nutritional quality since most intakes show high correlations with total caloric intake. The second measure – fat consumption as a percentage of total calorie intake – controls for quality of diets and overconsumption.⁴ The third measure is calcium as a percent of its RDA. The danger of calcium deficiency in this population (Rolls 1992) provides the rationale for creating this measure. The last indicator is a score measuring the number of nutrients (maximum 15) for which the individual’s intake falls below 67% of RDA.⁵ These indices are useful for assessing the overall diet quality (Murphy et al. 1992).

The exogenous variables considered (Table 2) are grouped into logical categories. The

first group – previous year’s pre-tax income, homeownership, and expenditures on foods at home and away (reported as “usual” expenditure for the three months prior to the survey) – measure a household’s purchasing ability. Using current income and expenditures in intake regressions could lead to econometric problems, as these variables are highly correlated. Moreover, intake and current income (expenditures) are likely to be “simultaneously” determined. In the following analysis, explanatory variables are selected so as to reduce the problems associated with collinearity. As in other studies, previous year’s income is used to reduce the simultaneity problem. Following Horton & Campbell (1991), homeownership is used as a proxy for the existence of food-processing facilities. Household size, presence of children, race, and age are the key demographic variables. Regional location and degree of urbanization measure spatial effects.

Employment status and time spent in food preparation measure time-use constraints. Hours of employment is expected to be negatively related to nutrient intake because of the reduced time available for at-home food production (Horton & Campbell 1991). Yet as Mothersbaugh et al. (1993) have shown, higher incomes and education can mitigate the effect of reduced time availability by permitting consumers to purchase better quality foods or food-preparation services.

Highest school grade completed, use of nutrition labels on food products, and an index of health and nutrition knowledge (described below) measure the level of knowledge of health determinants. Three variables – self-assessed healthfulness of diets, whether vitamin supplements are taken, and self-assessed level of health – are intended to measure individuals’ perceptions of their diets. The key policy variable is whether a household is an FSP participant. Food-assistance programs are believed to have a positive influence on nutrient intake (Senauer et al. 1991, Horton & Campbell 1991). Finally, to control for an individual’s physical requirements, the Body Mass Index (BMI) – a composite of height and weight – is included.

The index of health and nutrition knowledge was constructed in the following manner: Respondents were given a set of questions asking them to relate consumption of nine food components (fat, fiber, salt, calcium, etc.) to health problems caused by over- and under-consumption of each component. For example, the typical question asked: “Have you heard about any health problems that might be related to: how much fat a person eats?” Whenever a respondent showed awareness of health problems caused by consumption of a food component, one unit was added to the index. The more points a respondent was able to accumulate, the higher her awareness of diet-health relationships (range 0-9). Of course, more sophisticated knowledge indices can be constructed but experimentation showed that such indices, are more difficult to interpret and offer no additional predictive power.

In order to strengthen the analysis and in an effort to build a parsimonious model with minimal collinearity, some explanatory variables in Table 1 were combined to form new variables. Following bivariate correlation analysis, income as percent of the federal poverty measure was retained instead of income and household size. Food expenditures at home and away were considered jointly (added) and separately. The study uses a composite measure of food expenditures – weekly expenditures away divided by at-home expenditures. This variable is least correlated with other explanatory variables and does not hide the distinction between the two types of expenditure. All remaining variables are used in the multiple regression analysis. The correlation between these variables, though sometimes significant, is generally very small in magnitude (in all cases below .26). Interaction between the explanatory variables, particularly race, income, and education, is of particular interest for the present sample. Multivariate Analysis of Variance indicated no interaction among these variables. Hence, no interaction terms are included in the analysis.

Nonparametric Analysis

The first step in the analysis is to consider whether the intake measures are nonlinearly related to the continuous variables. Based on this analysis parametric models are then estimated. Figure 1 contain the results of the LWR analysis for calorie and calcium intake. The findings for number of low nutrients and fat are very similar and are excluded to save space. The plots in these figures depict the relationship between intake and income (relative to poverty level), level of education, and the BMI. The relationship between intake and level of income, food expenditures at home and away, as well as the interaction between the intake measures were also investigated. These were generally linear without surprising features. The plots showing these results are excluded to save space. Instead, the discussion is focused on the variables used in the regression model.

In Figure 1, the left column is for calories and the right column is for calcium. Actual data are represented by “*”, the thick solid line is the shape suggested by LWR, and the thin solid line is obtained from a Kernel nonparametric curve fitting procedure with automatic window size selection.⁶ The latter contrasts well with the LWR plots where we subjectively selected the window size to be 50% of the data. As is apparent, however, both procedures suggest similar nonlinearities, independent of the window size. The plots show the relevant range of data where most observations occur and nonlinearity is most apparent. Plotting the total range would include extreme values and make the nonlinearities seem less severe.

Beginning with the top row, calorie and calcium intake smoothly increase, albeit at a very low rate, with relative income. Indeed, we find that for this sample a linear specification in relative income will be adequate for all intake measures. The intake-food expenditure (at home) relationship was found to be nonlinear, rising at first and then becoming flat suggesting a logarithmic transformation could improve the fit. However,

at home food expenditures and its logarithm were highly correlated with other explanatory variables. To avoid this problem we use the ratio of outside the home to at home food expenditures. Nonparametric plots indicate that the relationship between intake measures and the latter expenditure measure is linear though essentially flat over the range of data.

The second row of plots show that the intake-education relationship is generally linear, except for calcium, for which the benefits of increasing education are most distinct for those with at least a high school education. This suggests increasing benefits to education, which is quite interesting in itself. It is possible to control for this type of nonlinearity using the method of switching regression (Judge et al. 1980). However, even without such correction, education is a highly significant determinant of calcium intake. To ease interpretation, linear specification rather than switching regression results are reported even though the latter improves the fit. The last row of Figure 1 shows a puzzling relationship between intake measures and the BMI. A quadratic specification may be appropriate and hence, BMI squared is included in the regressions.

In concluding this analysis, it should be noted that the curves represent the *mean* response of the dependent variable to changes in the explanatory variables and that actual data are in fact evenly spread across the surface.⁷ Indeed, it appears that a large number of intake values occur, vertically, at around \$5,000 income. This pattern – common to other data – is perhaps responsible for the weak intake-income relationship (low but statistically significant regression coefficients) often reported in the literature. The high variability in intake for individuals with similar incomes cannot be adequately explained by other socioeconomic differences. Possible explanations for this phenomena include omitted variables bias and error associated with measuring individual's intake. The CS-FII sampling procedure is fairly standardized, attempting to control for these type of problems.

Regression Analysis

The LWR plots can now be used for specifying a parametric model; the relationship between intake and relative income appears to be linear for calorie and perhaps logarithmic for calcium intake. Intake is linearly related to education and quadratic with respect to BMI. The regression models reported below are linear in all explanatory variables except BMI. Experiments with a switching regression that permits a kinked relationships with respect to education improves the fit. The improvement in the fit, however, is not worth the added difficulty in interpretation. One therefore finds limited evidence of nonlinearity *in this sample*. The evidence from large cross-sections of the population, however, suggests extensive nonlinearities (Ramezani 1994), which may be reflective of the heterogeneous nature of that sample.

Table 2 contains the regression results for the total sample. The subsample of single mothers is distinguished with the dichotomous variable accounting for the presence of children. Of course, it is possible to run separate regressions for each group. However, this leaves a much smaller number of observations and weakens the results. A statistical test for whether these groups should be treated separately (the Chow test) suggests the contrary; the structural intake relationship for single mothers is not statistically different from the rest of the sample. This has important implications for design and implementation of assistance programs that target these populations.

Before discussing specific results, we note that the regressions have marginally higher R^2 than previous studies, primarily because of the inclusion of additional variables, particularly the nutrition knowledge index. For all regressions, the F-statistics are highly significant. Using ‘canned’ procedures in SAS (SAS Institute Inc. 1985), sensitivity analyses to assess the influence of outliers were also conducted. One observation with calcium intake more than 5 standard deviations above the mean appeared overly influential and

was deleted. Tests for heteroscedasticity were also rejected. Overall, the signs of the regression coefficients are consistent with *a priori* expectations.

The various categories of the independent variables provide a natural order for the discussion of specific findings. The measures of purchasing power – income, home ownership, and food expenditure – are discussed first. Income (previous year, before tax, and excluding any welfare payment) relative to the federal poverty measure, which is based on household size, is a highly significant predictor of all intake measures. Keeping household size constant, increases in income lead to unambiguous improvement in all intake measures. Similarly, keeping income constant, increases in household size worsen intake.

The food expenditure variable, which measures food expenditures away relative to at-home food expenditures, is statistically insignificant in all regressions. This is consistent with results from the nonparametric analysis. Homeownership, which proxies for the effects of wealth and the existence of kitchen facilities, is significant for calorie intake and the intake of all nutrients but not calcium and fat.

Turning to the demographic variables, the presence of children is statistically insignificant in all regressions, confirming the results of the Chow test discussed above. Race is clearly an important predictor of calcium intake – whites have higher intake relative to African Americans and other minority groups – but is insignificant for other measures. Respondents' age is an insignificant factor in determining intake except for fat, where percentage of calories from fat declines with age.

The variables reflecting spatial location – degree of urbanization and regional variables – are generally insignificant. Among the significant variables, city dwellers have lower calorie intake, Western residents have higher calcium intake, and individuals in the Northeast and Midwest seem to obtain a higher percentage of their calories from fat. These results may be reflective of differences in habits and the availability of particular food items in each region (e.g., dairy products and fresh fruits).

Overall, time-constraint variables – employment status and actual meal-preparation time – are not important predictors of intake. Employment status significantly worsens the overall intake and calcium intake. Prolonged food preparation time appears to be associated with a higher percentage of calories obtained from fats.

The next group of variables measures the influence of educational attainment, and health and nutrition knowledge on intake. This group is by far the most significant, both statistically and in magnitude. Higher educational attainment reduces the number of low nutrients in one’s diet and significantly increases calcium intake. The two-part kinked relationship with respect to education presented in Figure 2 is borne out by switching regression results (not reported), which suggests that the marginal impact of rising education on intake, though positive overall, is higher at higher education levels. Educational attainment, however, does not capture the effect of other important impetus to consumer behavior such as health knowledge. The nutrition knowledge index, which partly captures the latter effect, is significant for all quantitative measures of intake except fat. Calorie and calcium intake significantly increase with the index, while the number of low nutrients declines with the index. This suggests that dissemination of nutrition information could lead to informed food decisions that improve overall nutrient intake. Therefore, policies that aim to enhance nutritional knowledge could result in improved diets independent of other socioeconomic characteristics of individuals.

The last group of variables measures the influence of perceptions, policy variables, and physical requirements. The individual’s perceptions regarding personal diet positively influences the number of low nutrients in the data. Interestingly, those who consider their health to be good receive a significantly higher percentage of their calories from fat. One possible interpretation of this finding is that the “causality” is from health to fat intake. That would explain the insignificance of nutrition knowledge index in the estimated equation for fat intake. Taking vitamin and mineral supplements is insignificant

in all regressions.

The regression coefficient for the Food Stamp Program measures current participation status (dichotomous). Controlling for income and other socioeconomic variables, participation in FSP is statistically insignificant for all measures except calcium intake, where the regression coefficient has the largest relative magnitude. This is an important finding as individuals in this particular sample are disproportionately poor and at greater risk of calcium deficiency. Despite statistical insignificance, which may be due to small sample size, these regressions suggest that FSP is an important tool for improving nutrient intake.

Finally, the composite variable measuring individuals' physical requirements is highly significant for calorie intake and percent of calorie from fats. The regression results confirm the nonlinearities suggested by nonparametric analysis; BMI has a negative sign and its squared term has a positive sign. The puzzling finding related to BMI – that intake falls with BMI – is also indicated by the regressions. As a referee has noted, the anomalous BMI-Calorie findings may be due to misspecification: “That is, individuals who have recently gained some weight eat (and report) less in an attempt to maintain their weight.”

Summary and Conclusions

Nutrient intake is affected by many factors, including inadequate knowledge of proper nutrition, poor food preparation and eating habits, and insufficient purchasing power (Senauer et al. 1991). This study considered the influence of these and other socioeconomic variables on four measures of intake using a sample of female household heads. The key contributions of the study were the focus on this important segment of the population, the introduction of new variables that measure the influence of health (nutrition) knowledge, education, and measures that capture perceptions of personal health status and diet. Nonparametric estimation techniques were shown to be a useful tool for model-building and reducing specification error.

The results indicate that for this population nutrient intake is strongly influenced by a variety of socioeconomic variables. Overall, factors related to health knowledge, educational attainment, and purchasing power were found to be more significant relative to other explanatory variables. Based on the sign and the significance of the regression coefficients it appears that increases in health knowledge leads to improvements in intake. This result, however, may be a consequence of self selection in that those concerned with their diets have better intake and an incentive to learn about the relationship between diet and disease. Such self selection may mean that the estimated coefficients are upwardly biased. The results should therefore be interpreted with care. Clearly improvements in modelling this relationship are needed to correct for this and other biases. In addition to enhancing the modelling effort, future research should make simultaneous use of the three-year sample of CSFII/DHKS, which is now available.

References

- Adrian, J. & Daniel, R. (1976), 'Impact of socioeconomic factors on consumption of selected food nutrients in the United States', *American Journal of Agricultural Economics* **58**(1), 31–38.
- Akin, J. (1984), 'The school lunch program and nutrient intake: a switching regression analysis: reply', *American Journal of Agricultural Economics* **66**(4), 513–514.
- Altman, N. (1992), 'An introduction to Kernel and nearest-neighbor nonparametric regression', *The American Statistician* **46**, 175–186.
- Basiotis, P., Brown, M., Johnson, S. & Morgan, K. (1983), 'Nutrient availability, food costs, and food stamps', *American Journal of Agricultural Economics* **65**(4), 685–693.
- Basiotis, P., Johnson, S., Morgan, K. & Chen, J. (1987), 'Food stamps, food costs, nutrient availability, and nutrient intake', *Journal Policy Modeling* **9**(3), 383–392.
- Behrman, J. & Deolalikar, A. (1988), Health and nutrition, in Chenery & Srinivasan, eds, 'Handbook of Development Economics', North Holland.
- Chavas, J. & Keplinger, K. (1983), 'Impact of domestic food programs on nutrient intake of low-income persons in the United States', *Southern Journal of Agricultural Economics* **15**, 155–163.
- Cleveland, W. & Devlin, S. (1988), 'Locally weighted regressions: an approach to regression analysis by local fitting', *Journal of American Statistical Association* **83**, 596–610.
- Deaton, A. & Muellbauer, J. (1980), *Economics and Consumer Behavior*, Cambridge University Press, New York.
- Delgado, M. & Robinson, P. (1992), 'Nonparametric and semiparametric methods for economic research', *Journal of Economic Surveys* **6**(3), 201–249.
- Eastwood, D. B. (1985), *The Economics of Consumer Behavior*, Allyn and Bacon, Newton, MA.
- Frazao, E. (1992), 'Food spending by female-headed households', *USDA-ERS Technical Bulletin No. 1806*.
- Friedman, J. (1984), *A variable span smoother*, Tech. Rep. No. 5, Dept. of Statistics, Stanford University.
- Härdle, W. (1989), *Applied Nonparametric Regression*, Cambridge University Press, New York.
- Horton, S. & Campbell, C. (1991), 'Wife's employment, food expenditures, and apparent nutrient intake: Evidence from Canada', *American Journal of Agricultural Economics* **73**, 784–794.
- Jensen, H. H., Nusser, S. H., Riddick, H. & Sands, L. (1992), 'A critique of two methods for assessing the nutrient adequacy of diets', *Journal of Nutrition Education* **24**, 123–129.
- Judge, G. G., Griffiths, W. E., Hill, R. C. & Lee, T. C. (1980), *The Theory and Practice of Econometrics*, John Wiley and Sons, Inc.
- Lino, M. & Guthrie, J. (1994), 'Food situation of families maintained by single mothers: expenditures, shopping behavior, and diet quality', *Family Economic Review* **7**(1), 9–21.

- McLanahan, S. & Booth, K. (1989), 'Mother-only families: Problems, prospects, and politics', *The Journal of Marriage and the Family* **12**, 557–580.
- Morgan, K. J. (1986), 'Socioeconomic factors affecting dietary status: An appraisal', *American Journal of Agricultural Economics* **68(5)**, 1240–1246.
- Morgan, K. J. & Price, D. W. (1987), Food consumption and nutrient intake patterns of school-age children, in R. Raunikaar & C.-L. Huang, eds, 'Food Demand Analysis: Problems, Issues, and Empirical Evidence', Iowa State University Press.
- Mothersbaugh, D., Herrmann, R. & Warland, R. (1993), 'Perceived time pressure and recommended dietary practices: The moderating effect of knowledge of nutrition', *The Journal of Consumer Affairs* **27**, 106–126.
- Murphy, S. P., Rose, D., Hudes, M. & Viteri, F. E. (1992), 'Demographic and economic factors associated with dietary quality for adults in the 1987-88 Nationwide Food Consumption Survey', *Journal of the American Dietetic Association* **92(11)**, 1352–1357.
- National Research Council (1989), *Recommended Dietary Allowances*, 10th Edition, National Academy Press, Washington, D.C.
- Ramezani, C. A. (1993), 'The calorie-income link in the united states: A nonparametric analysis', Working Paper, *Dept. of Consumer Science, Univ. of Wisconsin-Madison*.
- Ramezani, C. A. (1994), 'Determinants of nutrient demand: A nonparametric analysis', Working Paper, *Dept. of Consumer Science, Univ. of Wisconsin-Madison*.
- Rolls, B. A. (1992), Calcium nutrition, in A. F. Walker & B. A. Rolls, eds, 'Nutrition and the Consumer: Issues in Nutrition and Toxicology 1', Elsevier Applied Science.
- SAS Institute Inc. (1985), *SAS User's Guide: Statistics, Version 5 Edition*, SAS Institute Inc., Cary, NC.
- Scarce, W. & Jensen, R. (1979), 'Food stamp program effects on availability of food nutrients for low income families in the Southern region of the United States', *Southern Journal of Agricultural Economics* **11(2)**, 113–120.
- Senauer, B., Asp, K. & Kinsey, J. (1991), *Food Trends and the Changing Consumer*, Eagan Press, St. Paul, Minnesota.
- U.S. Bureau of the Census (1990), 'Current Population Survey', *Government Printing Office*.
- Wojkiewicz, R. A., McLanahan, S. S. & Garfinkel, I. (1990), 'The growth of families headed by women: 1950-1980', *Demography* **10**, 19–30.

Table 1

Table 2

Figure 1: Locally Weighted Regressions (thick line) and the Kernel Estimator of the Intake Relationships

Notes

¹ A partial list of recent studies includes Adrian & Daniel (1976), Horton & Campbell (1991), Murphy, Rose, Hudes & Viteri (1992), Ramezani (1994), Ramezani (1993), Murphy et al. (1992), Basiotis, Brown, Johnson & Morgan (1983), Basiotis, Johnson, Morgan & Chen (1987), Akin (1984), Chavas & Keplinger (1983), Morgan & Price (1987), Searce & Jensen (1979). Senauer et al. (1991) and Morgan (1986) provide a complete survey of this literature.

² The tri-cube function assigns a weight w_i to observation x_i according to $w_i = (1 - (x_i/t)^3)^3$ where t is a fixed constant which depends upon the window size.

³ The LWR and Kernel analysis reported below were conducted using the S-Plus statistical package. The programs are available from the authors upon request.

⁴ This variable is constructed by multiplying the absolute fat intake (measured in grams) by its caloric equivalent (9 kcal/grams of fat) and then dividing by total calorie intake.

⁵ The nutrients are: Protein, Vitamins A (Ret), E, and C, Thiamin, Riboflavin, Niacin, Vitamin B_6 , Folate, Vitamin B_{12} , Calcium, Phosphorus, Magnesium, Iron, and Zinc. These measures are discussed in detail by Murphy et al. (1992) and Lino & Guthrie (1994). Jensen et al. (1992) present a lucid discussion of other methods for constructing population-specific indices of nutritional adequacy.

⁶ Kernel estimators are discussed in Härdle (1989) and Delgado & Robinson (1992). The Kernel used here is known as the “super-smoother” (Friedman 1984). The name refers to the fact that the algorithm determines the optimal window size – the bandwidth – for any given data. The Kernel estimators were not made the focus of this analysis because these would require more detailed descriptions (Altman 1992, provides a recent survey).

⁷ Note that in all cases the fit suggested by the Kernel estimator is more jagged, reflecting the variable window size used. However, the suggested patterns are essentially identical.