Factors Affecting the Macronutrient Intake of U.S. Adults

Jayachandran N. Variyam

Abstract

The purpose of this study is to better characterize factors associated with the likelihood of macronutrient excess or inadequacy among U.S adults by modeling parts of the conditional distribution of dietary intakes other than the conditional mean. The risk of dietary inadequacy or excess faced by an individual tends to increase as his or her intake moves from the mean of a nutrient intake distribution toward its tails. Therefore, marginal effects of explanatory variables estimated at the conditional mean using ordinary least squares may be of limited value in characterizing these distributions. Quantile regression is effective in this situation since it can estimate conditional functions at any part of the distribution. Quantile regressions based on data from USDA's 1994-96 Continuing Survey of Food Intakes by Individuals indicate that differences in mean macronutrient intakes based on sociodemographic characteristics can be quite different from intake differences at other parts of the distributions. Therefore, judging dietary disparities between subpopulations by comparing mean intakes only, and not by comparing intakes at other parts of the distributions, may lead to misleading or incomplete conclusions.

Keywords: Diet quality, health risk, heteroskedasticity, nutrition, quantile regression.

About the Author

Jayachandran N. Variyam is an economist, Food and Rural Economics Division, Economic Research Service, USDA.

Acknowledgments

Contents

Summary

Diet-related diseases and health conditions account for a major share of morbidity and mortality in the United States. Just seven diet-related health conditions cost the United States $71 billion (in 1995 dollars) annually in medical costs and productivity losses, according to the latest ERS estimates. Healthier diets can help reduce these costs. Various segments of the U.S. population bear this burden differently. Differences related to gender, race, ethnicity, income, and educational attainment are major sources of health disparities in the United States. Reducing such health disparities, and disparities in related risk factors, such as obesity and poor diets, is a goal of the Federal Government's Healthy People 2010 initiative.

With better knowledge of the dietary differences and potential excessive energy intakes among population subgroups, public health professionals can devise more effective strategies for improving the diets and correcting the caloric imbalance among vulnerable subgroups. A number of past studies have assessed dietary differences among various subpopulations, but most have drawn conclusions based on a comparison of average nutrient intakes, either unadjusted or adjusted for sociodemographic differences. This approach overlooks an important feature of nutrient intake distributions. For many nutrients, the risk of dietary excess or inadequacy, and therefore, the risk of adverse health effects, is greater at the upper or lower parts of the nutrient intake distributions rather than at the mean. Therefore, judging intake difference between subpopulations by comparing their mean intakes only, and not differences at other parts of the intake distribution, could lead to incomplete or potentially misleading conclusions.

The purpose of this study is to more accurately characterize differences in nutrient intake among U.S. adults by focusing on the tails of the distribution of intakes instead of the mean. The study specifically examines the intakes of five macronutrients—energy, total fat, saturated fat, cholesterol, and fiber. First, the study examines the differences in nutrient intakes between selected sociodemographic groups by comparing their percentile curves. Second, the study uses quantile regression to estimate intake differences at various percentiles attributable to specific characteristics, while controlling for other characteristics. By estimating a family of quantile regressions for each nutrient, the study assesses intake differences between sociodemographic groups—not just at the mean, but along the entire distribution of nutrient intake.

The results of this study indicate that, for many sociodemographic subpopulations, differences in mean macronutrient intakes can be quite different from intake differences at other parts of the distributions. For example, based on estimates of the mean and the quantiles adjusted for other characteristics, both Black men and Black women consume significantly larger amounts of cholesterol than White men and White women, respectively. However, quantile estimates suggest a narrower disparity at the lower quantiles and a much wider disparity at the upper quantiles between these groups, compared with estimates at the mean. Since the risk of excess intake is greater at the upper quantiles, this suggests a more serious nutritional problem than if the conditional differences had been uniform across the distribution of cholesterol intakes.

In economic models of health, educational attainment plays a crucial role as a determinant of health outcomes by influencing health behaviors and choices. Results of this study confirm previous findings that education is positively correlated with better diets. However, compared with previous findings, the results here, especially for men, also show something new. The beneficial effects of education are larger at parts of the conditional distribution that matter most—at the upper quantiles of fat and cholesterol intakes where the risk of excess are higher. Together with similar findings regarding the effects of income and age, this result suggests that, compared with younger, less educated, and lower income men, older, more educated, and higher income men may have benefited more from health and nutrition information initiatives such as the Nutrition Labeling and Education Act.

The results from this study have important implications for future studies evaluating the dietary impact of many nutrition-related policy interventions such as food assistance programs and food labeling regulations. For such studies to fully uncover the extent and nature of the behavioral impact of interventions, it may be essential to look beyond the conditional mean to parts of the dietary intake distributions where the risks of inadequacy or excess are higher.

Introduction

A man was searching under a streetlight. A passerby stopped and asked, "What are you looking for?" "My keys" the man replied. "Where did you lose your keys?" the passerby asked. "Over there" the man said, pointing to the shadows. "Then why are you looking over here?" the passerby inquired. "Because the light is better here" the man replied. Anonymous

The flood of scientific evidence on the health effects of foods, nutrients, and other dietary components has heightened interest in the composition of U.S. food demand and supply, as well as the quality of American diets and their determinants (Adelaja, Nayga, and Lauderbach, 1997; Bowman et al., 1998; Frazao, 1999; Gould, 1996; Kantor, 1998; Lin, Guthrie, and Frazao, 1999; Nayga, 1994). It is well established that dietary excesses and inadequacies are associated with several chronic health conditions that can reduce productivity and hasten mortality (National Research Council, 1989). Just seven diet-related health conditions cost the United States $71 billion (in 1995 dollars) annually in medical costs and productivity losses, according to the latest ERS estimates. Healthier diets can help reduce these costs.

This study focuses on macronutrient intake in the American diet. For significant proportions of the U.S. adult population, excess intakes of total fat, saturated fat, and cholesterol and inadequate intake of fiber constitute a persistent nutritional problem.[1] Further, the rapid rise in the prevalence of obesity among adults suggests an imbalance between total energy intakes and energy use through physical activity (Koplan and Dietz, 1999). Excess dietary fats may contribute to the energy imbalance that leads to obesity. High-fiber diets may protect against obesity by lowering insulin levels (Lichtenstein et al., 1998; Ludwig et al., 1999).
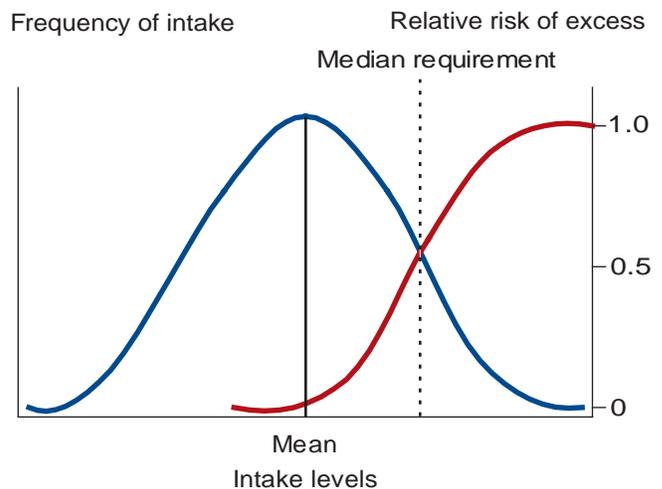
Coexisting with these nutritional excesses and inadequacies for the overall population is the evidence that intake levels differ substantially among population subgroups (U.S. Department of Agriculture, 2000). An accurate understanding of these dietary intake differ-

ences is crucial, given the goal, set by national initiatives such as the Healthy People 2010, of reducing health disparities. Besides the obvious use for monitoring relative nutritional status and guiding nutrition policy, a clear assessment of dietary differences is helpful for other uses such as targeting nutrition education efforts and forecasting dietary trends due to changing demographics.

Understandably, explaining dietary intake differences in the population by identifiable characteristics such as gender, age, income, and racial and ethnic identity has been an active research area. Dietary disparities due to these characteristics have been examined in many studies (e.g., Adelaja, Nayga, and Lauderbach, 1997; Chavas and Keplinger, 1983; Murphy et al., 1992; Nayga, 1994). Such studies have typically used the least squares multiple regression method to estimate the regression-adjusted differences (or marginal effects) in the conditional mean of intakes due to the characteristics. However, when studying distributions such as those of nutrient intakes, marginal effects at the conditional mean provide only a very limited characterization of distributional differences among subpopulations. This is because, for many nutrients, the risk of dietary excess or inadequacy, and therefore, the risk of adverse health effects, is greater at the upper or lower parts of the intake distributions than at the mean. Therefore, for one to infer intake differences between subpopulations by looking at the differences in their conditional mean intakes only, and not differences at other parts of the intake distribution, could lead to incomplete or potentially misleading conclusions.

Figure 1

Risk of excessive intake

---

[1] Dietary fiber is not strictly a macronutrient but a dietary component. For ease of exposition, we include fiber under the term macronutrient in this report.

Consider figure 1 where the bell-shaped curve on the left shows the distribution of the usual intakes of a hypothetical nutrient whose excess intake is of concern (Institute of Medicine, 2000). The "S" shaped curve on the right shows the cumulative distribution of the required intakes for the same nutrient. This risk curve gives the probability that any given intake exceeds the requirement for the individual consuming that intake. For example, for an individual with observed intake at the median requirement level (denoted by the dotted line), there is a 50-percent probability that his or her intake is excessive. For this hypothetical nutrient, the mean intake is well to the left of the risk curve so that the risk of excess at the mean is close to zero. However, a significant portion of the intake curve to the right of the mean overlaps with the risk curve. For individuals consuming the nutrient at these levels, the probability of their exceeding the requirement rises rapidly with intake.

Table 1 reports the mean and selected percentiles of the intakes of five macronutrients by U.S. adults, excluding pregnant or lactating women. These statistics were estimated from USDA's 1994-96 Continuing Survey of the Food Intakes of Individuals (CSFII). For comparison, table 1 also reports the daily intake levels for these macronutrients recommended by the health authorities for the relevant gender and age groups.

Comparing the intake distributions with the recommended intake levels, it is clear that the mean and median intakes of fats and cholesterol are at or close to adequate levels. For example, the recommended total fat intake for men between 21 and 50 years of age is less than or equal to 96.7 grams of fat per day based on a 2,900-calorie diet. The mean and median intakes of total fat for this group of men during 1994-96 were about 98 and 90 grams, respectively, which are close to the recommended level. However, total fat intake at the 90th percentile is 153 grams, considerably above the healthful level. For women aged 21-50, the cholesterol intake even at the 75th percentile, 277 milligrams (mg), is below the recommended daily intake of 300 mg, whereas at the 90th percentile, the intake is 402 mg, well above the healthful level.

Clearly, from a public health and nutrition policy perspective, regression-adjusted intake difference between population subgroups at the 75th and 90th percentiles for total fat, saturated fat, and cholesterol are of greater interest than those at the mean. Similarly, for fiber, although the mean intake is substantially below the recommended level, one would be more interested in the intake difference at the 10th or 25th percentiles.

Using marginal effects estimated at the conditional mean to characterize intake differences over the whole distribution would be meaningful if these effects were equal to the marginal effects estimated at all other parts of the conditional intake distribution. This equality will be realized only under the strict assumption that the intake distributions are identically and symmetrically distributed with respect to the subpopulations of interest. If the intake distribution differs in variance and shape among population subgroups, then the marginal effect of the explanatory variable will not be equal across different parts of the intake distribution and instead will differ from point to point along the distribution. In this case, knowing the marginal effect at the riskier parts of the distribution would be more useful for policy purposes than knowing the marginal effects at the conditional mean.

The purpose of this study is to better characterize the macronutrient excess or inadequacy among U.S. adults by modeling parts of the conditional distribution of dietary intakes other than the conditional mean. This is achieved by employing the method of quantile regression proposed by Koenker and Bassett (1978). While the classical least squares regression estimates the conditional mean of a dependent variable as a linear function of explanatory variables, the quantile regression enables the estimation of any conditional quantile of the dependent variable as a linear function of explanatory variables. Therefore, the estimation of quantile regressions allows us to obtain a more complete characterization of the dependence of macronutrient intakes on population characteristic of interest. More specifically, quantile regression enables us to look at the marginal differences in macronutrient intakes among subpopulations at specific points of interest along the conditional distribution, such as the 90th percentile for total fat, saturated fat, and cholesterol, and the 10th percentile for fiber.

The rest of this report is organized as follows:

- The next section discusses a theoretical framework in economics for examining nutrient intake behaviors.

- The empirical approach for estimating and testing marginal effects of key sociodemographic variables on nutrient intakes are presented next.

Table 1—Recommended intakes of macronutrients and their observed distributions among U.S. adults

Nutrient	Units	Recommended daily intake	Observed daily intake Percentile 10	25	50	75	90	Mean
Men, 21-50 years:								
Energy	Calories	2900	1463.0	1900.2	2396.4	3083.0	3758.1	2576.9
Total fat	Grams	≤96.7	48.0	65.7	90.9	121.8	153.3	98.1
Saturated fat	Grams	<32.2	14.9	21.3	30.0	41.4	53.2	33.4
Cholesterol	Milligrams	≤300	127.6	189.1	289.3	437.2	607.2	340.9
Fiber	Grams	33.4	8.1	11.7	16.3	23.2	29.8	18.2
Men, >50 years:								
Energy	Calories	2300	1172.7	1535.8	1972.2	2450.8	2932.1	2035.1
Total fat	Grams	≤76.7	35.6	51.4	72.3	96.4	122.8	76.7
Saturated fat	Grams	<25.6	11.1	15.9	22.7	32.0	41.6	25.2
Cholesterol	Milligrams	≤300	110.0	163.2	258.7	391.6	542.5	300.3
Fiber	Grams	26.5	7.9	11.7	16.4	22.5	29.8	18.0
Women, 21-50 years:								
Energy	Calories	2200	990.9	1292.7	1618.9	2036.9	2460.6	1693.8
Total fat	Grams	≤66.7	29.0	42.1	58.6	78.3	100.8	62.5
Saturated fat	Grams	<22.2	9.0	13.1	19.4	26.4	35.0	20.9
Cholesterol	Milligrams	≤300	72.9	116.3	179.9	277.2	402.2	215.1
Fiber	Grams	23	5.8	8.4	12.3	16.5	22.0	13.3
Women, >50 years:								
Energy	Calories	1900	879.4	1115.2	1444.8	1754.5	2081.8	1464.7
Total fat	Grams	≤63.3	24.9	36.5	50.5	67.2	83.6	53.2
Saturated fat	Grams	<21.1	7.6	10.9	16.1	22.5	28.6	17.3
Cholesterol	Milligrams	≤300	72.4	112.7	171.7	264.4	371.6	201.8
Fiber	Gram	21.9	6.5	9.2	12.9	17.6	22.7	14.0

Note: Estimates based on two nonconsecutive days of intakes from the 1994-96 CSFII, using sampling weights. The recommended intakes are from Lin, Guthrie, and Frazao (1999), table 4, p.6.

- Details of the data and specific variables used in the analysis are described.

- The results are discussed next.

- The final section presents the conclusions of the study and the implications for future research.

Economics of Nutrient Intake

Why do some consumers have healthful diets while others do not? If healthful diets are a critical component of healthful living and avoidance of chronic disease, why do some people choose poor diets? To economists, at the simplest level, the answer is that individuals choose foods not only to meet nutritional requirements but also for other reasons such as taste, preferences, and pleasure. Further, these choices are constrained by the income at their disposal, food prices and prices of other consumption goods, time available for cooking and food preparation, and the ability to combine foods and other resources to produce a nutritious diet. The observed dietary differences among individuals are the outcomes of a complex interplay of tradeoffs among desires for health, tasteful foods, other goods and services, and the constraints of limited resources. On a more formal level, economists have brought together these preferences and constraints, and the resulting choices, within a framework called the theory of household production.

The theory of household production grew out of Becker's (1965) study of the allocation of time in households and Lancaster's (1966) development of the characteristics model of consumer demand, which views purchased goods in terms of their product attributes. The theory integrates a variety of biological, sociodemographic, and economic factors, all of which interact and influence household consumption decisions. Household production models developed from the theory have been used to analyze many types of consumer and household behaviors (Strauss and Thomas, 1996). Such models are powerful tools for analyzing choices involving intrafamily interaction, such as maternal influence on children's health. Where intrafamily effects are not of interest, the models can be simplified to focus on individuals' choices.

Households seek to maximize satisfaction (or utility) through consumption of commodities. While some of these commodities, such as cars, clothing, or food are purchased in the market, others are "produced" by the households for their members' consumption. For

example, a household may purchase a variety of food items and combine it with cooking skills, nutrition knowledge, preparation time, and kitchen appliances to produce healthful meals. The objects of desire, including good nutrition, children's health, as well as the health of other family members, are not market goods, but are produced with inputs of market goods and time.

The household utility maximization is subject to an array of income, time, and technology constraints. Technology constraints are represented by production functions which capture the notion that all households are not able to produce the same amount of commodities, such as household members' health, from a given quantity of inputs. Households vary in their efficiency of producing commodities, depending on household members' sociodemographic and other characteristics. For example, each family member's health is determined by a unique production function, which depends on time, health inputs, the sociodemographic characteristics of the household, community characteristics, and the genetic endowment of the individual. These production functions could be interrelated because some of the commodities are intermediate goods produced by households as inputs into the production of a final commodity. For example, nutrients produced by combining foods with cooking time and nutrition knowledge appear as inputs in the health production function.

Income constraint ensures that the household expenditure on purchased goods and services does not exceed money income, which is equal to the sum of earnings from wages and any non-labor income. The time constraint ensures that the sum of all time inputs into the production of commodities plus leisure time and time spent at work does not exceed the total time available. Since labor earnings depend on time spent at work, income and time constraints can be combined into one "full income" constraint.

The solution to the household maximization problem, subject to technology and full income constraints, gives the demand functions for commodities produced by the household (such as nutrients and health) and commodities purchased in the markets (such as foods and medical services). These functions depend on prices of all purchased consumption goods and inputs, wage rates, and household income, as well as the sociodemographic characteristics of the households.

Grossman (1972) pioneered the use of the household production model to study the determinants of health, health behaviors, and health inputs. Grossman's approach grew out of the recognition that many consumer choices, such as those relating to the amount of exercise, the nutritional quality of diets, and the purchase of medical services, are not made because consumers gain utility from these choices directly, but rather because these choices influence health. Health, in turn, is demanded because it is a source of utility and because it determines income and wealth. The dependence of income as well as mortality (and thus future time available for work and leisure) on health implies that a fully specified Grossman-type model would be a dynamic programming problem in which households maximize the current value of its stream of future utilities (Sickles and Taubman, 1997). Because of the complexity and very specific data needs to solve these types of problems, fully specified dynamic health models have been used only in very specialized cases (Sickles and Yazbeck, 1998).

In the absence of detailed longitudinal data on prices, income, wages, and consumption, most of the empirical work using Grossman-type models have focused on estimating static (or one-period) reduced-form input demand and production functions (Strauss and Thomas, 1996). Even here the estimation is often not trivial since the production technologies and input choices are simultaneously determined. Nevertheless, the static approach can still answer important questions. For example, estimating health production functions along with health input demands can be useful for examining pathways through which choice of inputs factors affect the health outcomes of interest (Rosenzweig and Schultz, 1983).

In this study, we estimate a set of nutrient intake functions for U.S. adults that express the quantity of nutrients consumed as a function of income and sociodemographic and anthropometric characteristics. These functions are not structural equations derived from a fully specified household utility maximization problem. Rather, they may be viewed as linear approximations of reduced-form health input demand functions derived from the household production model, or alternatively, as Engle functions derived from the traditional consumer theory (Adelaja, Nayga, and Lauderbach 1997; Chavas and Keplinger, 1983). Viewing the estimated functions as reduced-form health input demand functions has the advantage that the effects of key

sociodemographic variables such as educational attainment and age can be interpreted in the context of predictions from the household production theory.

Under the household production theory, sociodemographic factors such as education and age enter input demand functions because they influence production efficiency. Educational attainment affects health production by raising technical and allocative efficiencies of input use (Grossman and Kaestner, 1997). Technical efficiency causes the more educated to produce a larger health output from a given level of health inputs. Allocative efficiency causes the more educated to acquire and use information about the true effects of inputs on health. Similarly, if the demand for health is inelastic, and if health stock depreciates at an increasing rate with age, then health investment will increase with age (Grossman, 1972). In terms of nutrient intake, this implies that older individuals will be more likely to have better diets than younger individuals.

Income enters the nutrient intake functions because it represents the budget constraint facing the consumer. However, the household production theory does not offer a clear prediction of the effect of income on the intake of nutrients. This is because consumers choose foods for multiple attributes that often have opposing effects on health. For example, higher income may increase consumption of fat-rich foods with taste-enhancing attributes. Higher income may also provide better access to health information that tends to reduce consumption of these fat-rich foods. The net effect of income will depend on which of these effects is dominant. If the informational effect is dominant, then income will have an effect similar to the effect of education.

Other sociodemographic factors such as age, gender, race, ethnicity, and a person's height and weight may also influence nutrient intakes, both by affecting the relative amounts of food consumed and by influencing the health production efficiency. Most of the effects of age, gender, and anthropometric characteristics on macronutrient intakes will be due to differences in the amounts of foods consumed. Thus, lower intakes are expected for older adults, women, and those with lower weights and shorter stature. Both cultural differences in food choices and differences in health production efficiencies may be reflected in the influence of race and ethnicity on macronutrient intakes. As with income, the net effect of racial and ethnic status is

uncertain, and left to be resolved through empirical estimation.

Empirical Approach

To estimate the impact of income, educational attainment, age, and other sociodemographic and anthropometric characteristics on macronutrient intakes of U.S. adults, we specified nutrient intake functions of the form:

$$(1) \quad y_i = \beta_0 + \beta_1 \text{Income}_i + \beta_2 \text{Education}_i + \beta_3 \text{Age}_i$$

$$+ \beta_4 x_{4i} + ... + \beta_K x_{Ki} + \varepsilon_i,$$

where y_i denotes the intake of a nutrient by the ith individual, $i=1, ..., N$, $x_{4i}, ..., x_{Ki}$ represent additional explanatory variables influencing y_i, $\beta_1, ..., \beta_K$ represent the coefficients of the explanatory variables, and ε_i represents a random error term that accounts for unobserved factors influencing y_i.

With intake data and the observed characteristics for N individuals obtained from food consumption surveys, equation (1) is typically estimated using Ordinary Least Squares (OLS) regression. In this case, the β coefficients represent the marginal change in nutrient intake for a unit change in the explanatory variable at the conditional mean of intake. Estimating the effect of explanatory variables on intake at the conditional mean is a convenient choice, often dictated by the ease of applying and interpreting OLS regression. However, this focus on change at the conditional mean is not dictated by theory. Neither the theory of consumer demand nor the household production theory gives any guidance as to the parts of distribution of intakes where the effects of income, age, education, or other explanatory variables are likely to occur. Therefore, the question as to which part of the distribution to study will be answered by (a) the nature of the observed distribution of the choice variable in the relevant population and (b) the potential implications of the fact that underlying behavior is different at different points of the distribution. In fact, as noted earlier, for the macronutrient intakes of the U.S. adult population, finding the effects of explanatory variables at the tails of the conditional intake distributions is likely to be more interesting and useful than finding the effects of explanatory variables at the conditional means.

This goal can be accomplished by estimating equation (1) for various quantiles of y_i by quantile regression.

Quantile Regression

Koenker and Bassett (1978) introduced quantile regression as a generalization of the sample quantile to the conditional quantile, where the conditional quantile is expressed as a linear function of explanatory variables. This is analogous to the OLS regression where the conditional mean of a random variable is expressed as a linear function of explanatory variables. By enabling the estimation of *any* conditional quantile, quantile regression allows one to describe the entire conditional distribution of a dependent variable given a set of regressors. The Least Absolute Deviation (LAD) estimator is a special case of quantile regression that expresses the conditional median as linear function of covariates.

The key factor that makes quantile regression's ability to characterize the entire conditional distribution so useful is the presence of heteroskedasticity in the data (Koenker and Bassett, 1982; Deaton, 1997). When the data are homoskedastic, the set of slope parameters of conditional quantile functions at each point of the dependent variable's distribution will be identical with each other and with the slope parameters of the conditional mean function. In this case, the quantile regression at any point along the distribution of the dependent variable reproduces the OLS slope coefficients, and only the intercepts will differ.

However, when the data are heteroskedastic, the set of slope parameters of the conditional quantile functions will differ from each other as well as from the OLS slope parameters. Therefore, estimating conditional quantiles at various points of the distribution of the dependent variable will allow us to trace out different marginal responses of the dependent variable to changes in the explanatory variables at these points. Figure 2 illustrates this point. It shows the scatter diagram of total fat intake against age for a 10-percent random sample of men from the 1994-96 CSFII. A visual examination of the scatter shows that the data are clearly heteroskedastic with a declining variation in intake by age. Regression lines estimated by OLS and by quantile regression at the 0.1 and 0.9 quantiles are superimposed on the scatter diagram. Clearly, the line at the 0.9q has a greater negative slope than the line at the conditional mean. If the data were

Figure 2
Regression of total fat intake on age for a 10-percent random sample of men, CSFII 1994-96

Total fat (grams)
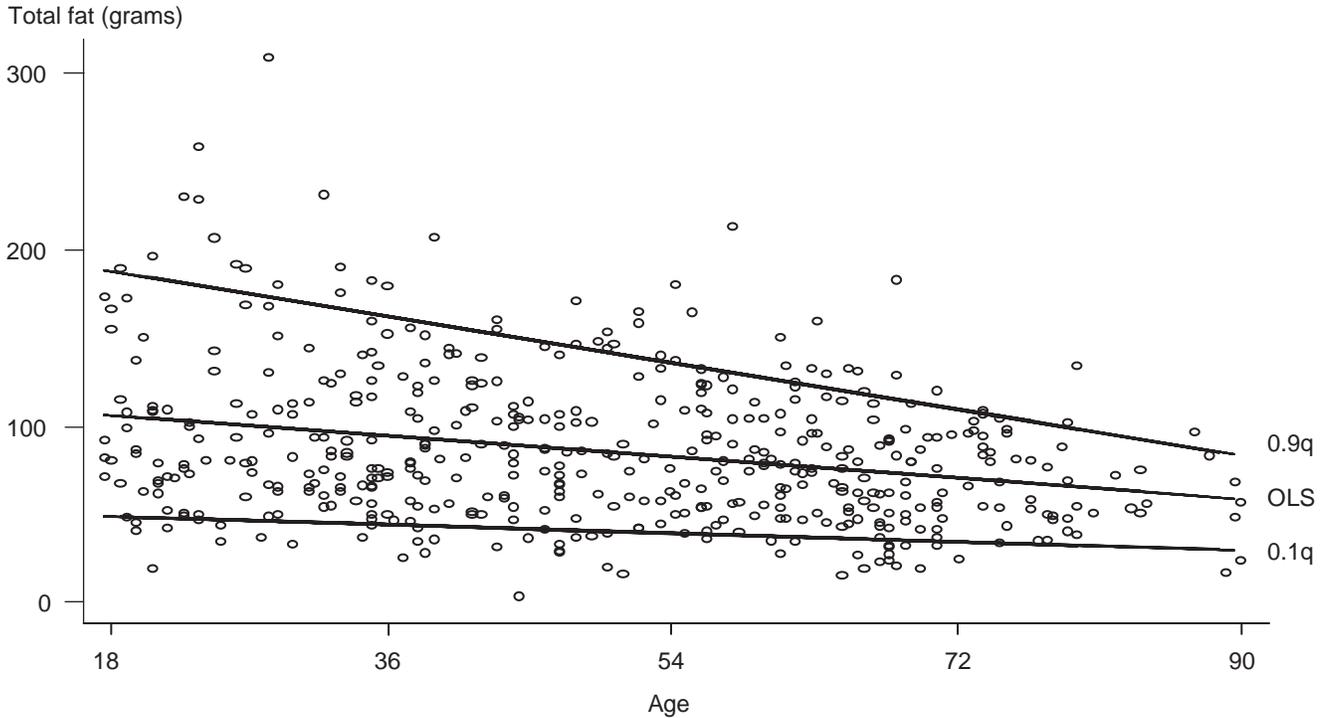
homoskedastic, the three lines would have identical slope and thus would be parallel to each other.

Two additional features of the quantile regression model are relevant to our application (Buchinsky, 1998). First, the classical properties of efficiency and minimum variance of the least squares estimator are obtained under the restrictive assumption of independently, identically and normally distributed (i.i.d.) errors. When the distribution of errors is non-normal, the quantile regression estimator may be more efficient than the least squares estimator. Second, the quantile regression estimator is "robust" when the dependent variable has outliers or the error distribution is "long-tailed." Since the objective function from which the quantile regression estimator is derived is a weighted sum of absolute deviations, the parameter estimates are less sensitive to a few large or small observations at the tails of the distribution.[2] The distributional statistics reported in table 1 show that the mean intakes of macronutrients are consistently higher than the median (50th percentile) intakes. This suggests that the intakes are asymmetrically distributed with some influential observations at the upper tails (a feature visible in fig. 2). Quantile regression ensures that the

parameter estimates are less sensitive to such observations compared with OLS.

Estimation and Testing

The quantile regression model for equation (1) can be written as

$$(2) \quad y_i = \mathbf{x}_i'\beta_\theta + \varepsilon_{\theta i}, \quad Q_\theta(y_i|\mathbf{x}_i) = \mathbf{x}_i'\beta_\theta \quad i = 1,\cdots,N,$$

where $\mathbf{x}_i$ denotes a $(K+1)\times 1$ vector of the explanatory variables, β_θ is the corresponding vector of coefficients, and $Q_\theta(y_i|\mathbf{x}_i)$ denotes the θth conditional quantile of y_i, $(0 < \theta < 1)$. From equation (2), the quantile regression estimator of β_θ is obtained by solving

_____

[2] This can be seen by comparing the objective function for the conditional quantile (equation 3) with the objective function for OLS, which is $\min \sum_{i=1}^{N}(y_i - \mathbf{x}_i'\beta_\theta)^2$. The OLS estimator is sensitive to outliers because the objective function is squared; the farther an outlier, the greater its influence on β. In the quantile objective function, β_θ is chosen so that $y_i - \mathbf{x}_i'\beta_\theta$ is positive for $(\theta *100)\%$ of the observations and negative for the remainder. Therefore, an increase in the value of any observation above or below a conditional quantile has no influence on it.

(3)

$$\min_{\beta_\theta} \frac{1}{N} \left\{ \begin{array}{c} \sum_{i:y_i \geq \mathbf{x}_i'\beta_\theta} \theta |y_i - \mathbf{x}_i'\beta_\theta| \quad + \\ \\ \sum_{i:y_i < \mathbf{x}_i'\beta_\theta} (1-\theta)|y_i - \mathbf{x}_i'\beta_\theta| \end{array} \right\}.$$

When K = 0 and $\mathbf{x}_i$ is a (1×1) vector that includes only the intercept for all i, this minimization problem reduces to an estimator of the sample θ-quantile. The minimization problem in equation (3) has a linear programming representation, which is guaranteed to have a solution in a finite number of simplex iterations (Buchinsky, 1998). Several estimators for the asymptotic covariance matrix for $\hat{\beta}_\theta$ obtained from the above minimization are available, but for obvious reasons, those that rely on the assumption of i.i.d. errors are of limited value (Deaton, 1997). Buchinsky (1995) has shown that the design matrix bootstrap estimator provides a consistent estimator for the covariance matrix under very general conditions. In the design matrix bootstrap, quantile regression is estimated on a sample of N observations $(y_i^*, \mathbf{x}_i^*), i = 1, \ldots, N,$ drawn randomly with replacement from the original sample. The process is repeated B times to obtain bootstrap estimates, $\hat{\beta}_{\theta b}^*$, b = 1, …, B. The covariance matrix of $\hat{\beta}_\theta$ is then obtained as the covariance of $\hat{\beta}_\theta^*$ computed from the B bootstrap estimates with $\hat{\beta}_\theta$ as the pivotal value.

The minimum-distance method can be used to test for the equality of slope coefficients of a given dependent variable across all estimated quantiles (Buchinsky, 1998). Let $\hat{\beta}_P = (\hat{\beta}_{\theta_1}', \ldots, \hat{\beta}_{\theta_P})'$ be a (KP+1) × 1 stacked vector of unrestricted parameter estimates from quantile regressions at P quantiles. Let $\beta^R = (\beta_{\theta_1}, \ldots, \beta_{\theta_P}, \beta_2, \ldots, \beta_K)'$ be a (K+P) × 1 vector comprising P unrestricted intercepts and K restricted slope parameters. The restricted parameter vector β^R is obtained by minimizing

(4) $Q(\beta^R) = (\hat{\beta}_P - R\beta^R)' \, \mathbf{A}^{-1} \, (\hat{\beta}_P - R\beta^R),$

where $\mathbf{A}$ is a positive definite matrix and $\mathbf{R}$ is the appropriate restriction matrix. Under the null hypothesis of equality of slope coefficients, $NQ(\beta^R)$ is distributed χ^2 with (PK-K) degrees of freedom. Since the

equality of slope parameters will hold if the i.i.d. assumption is valid, this is a general test for heteroskedasticity. The optimal choice for $\mathbf{A}$ is the variance-covariance matrix of $\hat{\beta}_P$, denoted by $\hat{\Lambda}_P$ (Buchinsky, 1998). Given $\hat{\Lambda}_P$, the usual F-statistic for testing liner restrictions can be used to test for the equality of the slope parameters for a specific explanatory variable at symmetrical quantiles such as 0.1q and 0.9q. If the null hypothesis of homoskedasticity or the equality of the slope coefficients is not rejected, the restricted slope estimates β^R give an optimal combination of the quantile slope estimates (Buchinsky, 1998). Also, given $\hat{\Lambda}_P$, the variance-covariance matrix of the restricted parameter vector can be obtained as $\hat{\Lambda}^R = (\mathbf{R}' \hat{\Lambda}_P^{-1} \mathbf{R})^{-1}$.

The quantile regression parameter estimates are obtained by estimating a separate equation for each quantile of each macronutrient. The variance-covariance matrix of the estimates can be obtained by bootstrapping each of these equations separately. However, to carry out tests of the equality of slope coefficients for a given dependent variable across the P estimated quantiles and to obtain the restricted parameter estimates and their standard errors, it is necessary to have $\hat{\Lambda}_P$, the variance-covariance matrix of the stacked vector of parameter estimates at the P quantiles. This can be obtained by simultaneously estimating quantile regressions at the P quantiles for each bootstrap sample. Thus, the following procedure was used for the estimation and testing of the quantile regressions for each macronutrient. First, the coefficient estimates for a macronutrient were obtained by running quantile regressions separately at the P desired quantiles. Second, a bootstrap sample was drawn for that macronutrient and the bootstrap estimates for the P quantiles were obtained by running quantile regressions separately at the P quantiles for that sample. Finally, after repeating the bootstrap procedure B times, $\hat{\Lambda}_P$ was calculated to obtain the standard errors of the coefficient estimates and to conduct the equality tests. This estimation process was carried out in Stata using the sqreg procedure (Gould, 1997). Additional details regarding the estimation of the quantile regression model and the asymptotic covariance matrix of the parameters are discussed in Buchinsky's (1998) methodological survey.

Data

The macronutrient intake data for men and women were obtained from USDA's 1994-96 CSFII (Tippett and Cypel, 1997). Each year of the 3-year CSFII comprised a nationally representative sample of noninstitutionalized persons residing in the United States. Dietary data for selected members from a screened sample of 9,664 households were collected on two nonconsecutive days through in-person interviews using 24-hour recalls. A sample of 15,303 people provided information on food intakes for both days giving a 2-day response rate of 76.1 percent. From the sample persons providing complete 2-day dietary intake records, we selected adults aged 18 or above and excluded pregnant and lactating women.

By combining the food records with a nutrient database, CSFII provides information on the intakes of a variety of macronutrients, vitamins, and minerals. In this study, we focused on the five major macronutrients—energy, total fat, saturated fat, cholesterol, and fiber. As noted earlier, these nutrients have drawn interest because of their links to obesity, cardiovascular disease, and certain types of cancer (National Research Council, 1989). The means of the 2-day intakes were used to represent the daily intakes of these nutrients.

Along with the quantities of foods and nutrients consumed by individuals, CSFII provides a detailed set of personal and household characteristics of each sample person. From this set, we selected explanatory variables for quantile regressions of the macronutrient intakes based on our theoretical discussion in section 2 (Economics of Nutrient Intake) as well as evidence from previous literature. The selected explanatory variables, along with their means, are listed separately for men and women in appendix table 1. The explanatory variables fall into three groups: household characteristics, personal characteristics, and survey-related variables. Income, household size, education level, age, height, and weight are continuous variables. The remaining variables are dummy indicator variables.

Income is the gross household income in the previous calendar year from all sources before taxes. Rather than restrict the effects of income and household size by entering income on a per capita basis, we included income and household size as separate variables. The region and urbanization variables are included to capture possible variations in intakes due to geographic differences in food prices and other location-related factors. Education is represented by years of schooling completed at the time of survey. On average, the adults in our sample had completed about a year more than high school. The average age for men and women was about 49 years. The height and weight of the respondents were included to control for the influence of body mass on the amount of food intake. While previous studies have often used the Body Mass Index (BMI) for this purpose, estimating a single coefficient for BMI implies a restriction on the coefficients for BMI components, height, and weight (Bhargava, 1994). Therefore, we left height and weight in the unrestricted form.

Racial and ethnic indicator variables are included to capture differential food intakes due to differences in culture and traditions. While Asians and other race groups (Asian/Pacific Islander, Aleut, Eskimo, or American Indian) are included in the sample, we do not focus on the estimates for these groups due to their relatively low proportions in the sample (see appendix table 1). Fairly significant proportions of men and women (15 and 21 percent, respectively) were reported to be on special diets at the time of the survey. A significant part of the variation in food intakes measured through consumption surveys can be attributed to factors such as survey season, whether the intake was recorded for a weekend, and whether, for a variety of reasons, a person's food intake on the day of recorded intake was less than or more than his or her usual intake. We included a detailed set of variables to control for such sources of variation in nutrient intakes.

Numerous intermediary variables such as nutrition knowledge, nutrition label use, and food program participation affect nutrient intakes. These variables may act as pathways through which basic sociodemographic variables influence dietary behavior (Gould and Lin, 1994; Kim, Nayga, and Capps, 2000; Variyam, Blaylock, and Smallwood, 1996). However, the objective of this study was not to estimate the effects of consumers' intermediate choices on dietary intakes. Rather, the objective was to map the net effect of key sociodemographic variables at different points along the conditional intake distribution. Therefore, we

excluded such intermediary variables from our estimated functions. After dropping observations that were incomplete with respect to the included explanatory variables, the final sample sizes were 4,725 men and 4,362 women.

Results

Based on our discussion of the household production theory, we expect to find specific relationships in the behavior of macronutrient intakes with respect to changes in the level of education and age. More educated and older people are likely to have lower intakes of energy (calories), total fat, saturated fat, and cholesterol, the overconsumption of which is a problem. For dietary fiber, which is underconsumed by most, more educated and older people are likely to have higher intakes. Evidence on these relationships as well as evidence on the effects of income, race, ethnicity, and other sociodemographic factors have been provided in many previous studies (Murphy et al., 1992; Nayga, 1994). But, as noted earlier, this previous evidence pertains to the relationships estimated at the conditional mean where the risk of inadequacy or excess for macronutrient intakes is relatively small.
A more interesting type of evidence is the effect of sociodemographics at the parts of the distribution beyond the mean, such as the tails where the risk of dietary inadequacy or excess are higher. We begin looking for such evidence by examining the bivarate relationships between intakes and sociodemographics at various points along the distribution of intakes.

Bivariate Associations

The mean and five selected percentiles of macronutrient intakes by categories of household income (expressed as a percentage of poverty level), years of education completed, age, race, and ethnicity are reported in appendix tables 2-6. Certain bivariate relationships between intakes and the explanatory variables are evident by comparing the mean intakes across categories of the explanatory variables. However, the comparison of mean intakes tells only part of the story of the relationships between the explanatory variables and macronutrient intakes. A more complete picture of the relationships emerges when one looks at the differences in intakes across categories of the explanatory variables at the five percentiles reported in appendix tables 2-6. For easier comparison of percentiles, intakes at the 10th to the

90th percentile for the outer categories of income, education, and age are shown in figures 3-8.

Men with income below 130 percent of the poverty threshold consume about 230 more calories on average than men with income above 350 percent of the poverty threshold. Low-income men have higher mean intakes of total fat, saturated fat, and cholesterol, than high-income men. The percentiles of intakes by income levels in appendix tables 2-5 and fig. 3 give a clearer picture of the intake difference associated with income. The energy, fat, and cholesterol intakes of men are not uniformly different between the income categories across the percentiles. The higher intakes of low-income men come almost entirely from the upper end of the distributions, notably above the 70th percentile for energy and fat. In fact, low-income men have lower intakes of energy and fats at the bottom part of the distributions. For cholesterol, although lower income men have higher intakes at all parts of the distribution, the differences are much larger at the upper parts of the distribution than at the lower parts.

The differences in women's cholesterol intake by income level has a pattern similar to men's (appendix table 5; fig. 4). Lower income women have higher mean cholesterol intake than higher income women and the differences are larger at the upper end of the cholesterol intake distribution. However, the relationship of women's energy intake with income is the reverse of men's (appendix table 2; fig. 4). Women under 130 percent of the poverty threshold consume 117 fewer calories of energy than women above 350 percent of the poverty threshold. Energy intake differences between the two groups of women are much larger at the bottom part of the distribution.

Although men's mean energy intake does not differ much by level of education, men with more than high school have higher intakes for much of the distribution except at the upper end above the 85th percentile (fig. 5). A similar pattern is evident for fat intake as well. The behavior of men's cholesterol intake by education level is very similar to that by income.

Unlike in the preceding examples, the difference at the mean is representative of the difference at other parts of the intake distribution for women's energy intake by education level. The energy intake of women with more than a high school education is uniformly higher at all parts of the distribution between the 10th and the

90th percentiles (fig. 6). The situation is similar for women's total fat intake as well. For cholesterol, the pattern noted for income reemerges with the difference in intakes between less than high school and more than high school groups widening at the upper end of the distribution.

For both men and women, the strongest and most consistent association for energy and fat intakes is with age. The mean intakes by age groups show that intakes decrease with age, and the percentiles show that the intake differences by age widen at higher percentiles.

Fiber is different from energy, fat, and cholesterol because its dietary inadequacy is the issue rather than dietary excess. This explains the different nature of associations between explanatory variables and fiber intake evident in appendix table 6 and figures 3-8. Unlike energy, fat, and cholesterol, men's mean fiber intake either increases or remains unassociated with increases in income, education, and age. However, the percentiles show slightly wider differences at the bottom in men's intakes by income and education level. Among women, fiber intake differences by the level of income, education, and age are fairly uniform at the different percentiles.

These bivariate associations between macronutrient intakes and key explanatory variables across various percentiles have clearly shown how the mean differences can mask a richer pattern of differences at other parts of the distribution. However, the bivariate associations may not capture the true association between an explanatory variable and intakes since the relationship may be confounded by other variables. For example, the association between education and income may confound the bivariate association between men's intakes and their educational attainment. To estimate the marginal effects of explanatory variables at different parts of the intake distribution after controlling for such confounding effects, we employed the method of quantile regression.

Quantile Regression Results

Test Statistics

The purpose of using quantile regressions is to estimate the marginal effects (or the slope coefficients) of explanatory variables at various points along the conditional distribution of the dependent variable. However, as noted earlier, if the distribution of the

dependent variable is homoskedastic (that is, the conditional variance of dependent variable's distribution is constant by the level of independent variables), the estimated marginal effects will be identical between quantiles and the conditional mean estimated by Ordinary Least Squares (OLS). In this case, the quantile slope coefficient estimates do not provide any additional information about the behavior of the dependent variable with respect to the explanatory variables beyond the information conveyed by the OLS slope estimates. Therefore, the first step after estimating quantile regressions is to test whether the estimated slope coefficients are equal across the quantiles. As shown by Koenker and Bassett (1982), such a test for the equality of slope coefficients across quantiles is a robust test for heteroskedasticity.

We estimated conditional quantile functions for the intake of the five macronutrients by men and women at five selected quantiles (P=5). The equality of slope coefficients across the five quantiles for each nutrient was then tested by computing the minimum-distance estimator in equation (4). The resulting chi-square test statistics are reported in table 2. Since K=30 and P=5, these χ^2 statistics have 120 degrees of freedom. The null hypothesis of homoskedasticity is rejected decisively (p<.01) in all cases. This implies that the slope coefficients differ significantly across quantiles and are likely to provide additional information about the behavior of intakes beyond that conveyed by the OLS estimates alone.

In tables 3 to 7, we report coefficient estimates for energy, total fat, saturated fat, cholesterol, and fiber. To keep the discussion manageable, the tables report

Table 2—Tests for equality of regression coefficients across the five quantiles

Macronutrient	Men	Women
Energy	417.3	336.0
Total fat	391.5	252.1
Saturated fat	370.3	261.4
Cholesterol	430.8	281.0
Fiber	271.5	216.1

Note: Under the null hypothesis of equality, the test statistic $Q \sim \chi^2(120)$.

only the estimates for five key variables of policy interest—income, education, age, race (Black compared with White) and ethnic origin (Hispanic compared with non-Hispanic). For comparison with the quantile estimates, the second column in each table presents the OLS estimates. The restricted coefficient estimates $\hat{\beta}^R$ are reported in the last column of the tables. The standard errors for the quantile regression estimates were obtained using the design matrix bootstrap with $B = 500$ replications. The standard errors for the OLS estimates were computed using White's method.

The R^2s for the estimated regressions are generally low, indicating high variation in the mean 2-day intakes, but are in line with previous studies (Adelaja, Nayga, and Lauderbach, 1997; Chavas and Keplinger, 1983).[3] The F-tests for the OLS regressions showed high significance levels (p<.001) in all cases.

Although the joint equality of all coefficients across the five quantiles was rejected for all nutrients, it still leaves open the possibility that the coefficients of individual explanatory variables may be equal across the quantiles. Therefore, we also tested for the equality of the slope coefficients of the selected explanatory variables across quantiles. While equality across any combination of quantiles could be tested, in table 8, we present test statistics (F-values) for slope equality at symmetrical quantiles (0.1q = 0.9q and 0.25q = 0.75q). The p-values of the test statistics are reported in parenthesis. As noted earlier, if an F-test does not reject the equality of slopes at symmetrical quantiles, then the restricted coefficient estimate $\hat{\beta}^R$ gives the optimal combination of the five quantile slope coefficients. Such estimates, in general, have lower variance than least squares estimates (Buchinsky, 1998). Comparing the restricted estimates with the corresponding OLS estimates, we found that in almost all cases $\hat{\beta}^R$ was more precisely estimated with lower standard errors than the corresponding OLS estimates (see also footnote 5).

Marginal Effects

The OLS income coefficients for both men and women, for energy, total fat, and saturated fat intakes are statistically insignificant, implying no relationship between household income and the consumption of these macronutrients, when other explanatory variables are held constant. The income coefficients at the five quantiles, and their weighted combination given by the restricted estimate, however, portray a more subtle influence for income. Income appears to have a slightly positive effect on energy intake at the lower quantiles with the effect declining toward the upper quantiles (table 3). A better view of this trend can be found in figures 9 and 10. These figures display plots of income coefficients for each macronutrient against the quantiles at which they were estimated.[4] The income coefficients for energy intake tend to decline from 0.1q to 0.9q, especially for women. However, except at the 0.25q of men's energy intake, none of the income quantile coefficients are statistically significant in table 3. Therefore, this apparent trend has to be interpreted with caution. The income coefficients for total and saturated fat are mostly insignificant at the five quantiles, and although the restricted estimates for men are significant, the magnitudes are small.

The additional information revealed by the quantile estimates compared with the OLS estimate comes into sharper focus for cholesterol (table 6). The OLS estimates show that income has a negative (and health-wise, beneficial) effect on cholesterol intake of both men and women. The quantile estimates reveal that this beneficial effect is almost entirely located at the upper quantiles. As the observed distribution showed, at these upper quantiles, cholesterol intakes tend to exceed the recommended level (table 1). For men, the effect of income on cholesterol intake at 0.9q is 115 percent larger than the OLS estimate. This implies that, as income increases, the upper conditional quantiles of cholesterol intake are decreasing more rapidly than is the conditional mean. This trend is clearly visible in figures 9 and 10 where the income coefficients on cholesterol intake reveal greater negative impact at higher quantiles. For men and women, the equality

---

[3] The R^2 for the quantile regressions are analogous to the conventional OLS R^2. Let $\hat{V}_\theta$ be the solution to equation (2) and let $\tilde{V}_\theta$ be the solution to equation (2) when $\mathbf{x}_i$ is restricted to include only the intercept. Then R^2 for the θth quantile is defined as

$$R_\theta^2 = 1 - \hat{V}_\theta \, / \, \tilde{V}_\theta.$$

[4] The coefficient estimates for these plots were estimated from regressions at 17 quantiles from 0.1q to 0.9q at intervals of 0.05.

restrictions on income coefficients across symmetrical quantiles (0.1q = 0.9q and 0.25q = 0.75q) are rejected at the 10-percent level (table 8).

Income has a significant positive effect on fiber intake at all quantiles. However, for men and women, the largest effect is at 0.9q. Since dietary risk for fiber (inadequacy) is relatively greater at the lower end of the intake distribution, one might expect to find the largest effect for income at the bottom quantiles. This was not the case, likely because, as table 1 shows, most adults have inadequate intake even at the 90th percentile.

The estimated effects of educational attainment on men's macronutrient intake clearly illustrate the importance of examining the whole conditional distribution rather than focusing on just the conditional mean. The OLS coefficient for education on men's energy intake is insignificant, implying that energy intake does not change in response to a change in the education level. However, the quantile estimates reveal a more interesting pattern that is masked by the OLS estimate. The education coefficients for energy at the 0.1q and 0.25q are both positive and statistically significant (p<0.1). Thus, as education increases, energy intake actually increases at the lower quantiles. Meanwhile, education coefficients at 0.5q and above are negative and the estimate at 0.75q is significant (p<0.1). Thus, energy intake declines in response to an increase in education at upper quantiles. The plot of education coefficients for energy intake in figure 11 displays this pattern clearly. Therefore, when the effect of education over the entire distribution is considered, the implication is that greater educational attainment influences men toward moderating their energy intake.

The differential effects of education across quantiles are even more striking for men's intake of fats and cholesterol than they are for energy (tables 4-6). For instance, an additional year of education reduces men's saturated fat intake by 0.18 gram at the conditional mean. However, at 0.9q, a similar increase in education reduces saturated fat intake by 0.52 gram, a nearly 200-percent larger estimated effect. Both for saturated fat and total fat, quantile estimates below the median are insignificant. In fact, for total fat, the marginal effect of education at 0.1q is positive and numerically large (0.32 gram). This is not surprising since, at 0.1q, the observed intakes are substantially below the recommended levels.

Educational attainment influences men's cholesterol intake in a similar fashion. The reduction in intake attributable to education is larger at the upper conditional quantiles compared with the conditional mean or the lower quantiles. The trend of larger negative effects for education on the intakes of fats and cholesterol as one moves from the lower to the upper quantiles is strikingly visible in figure 11. Tests of equality of education coefficients at symmetrical quantiles for men's energy, total fat, saturated fat, and cholesterol are rejected at the 5-percent level (table 8).

For men's fiber intake, education has a more uniform effect across the quantiles (fig. 11). The equality of education coefficients at symmetrical fiber intake quantiles could not be rejected (table 8). Consequently, this is one instance where it is meaningful to compare the restricted estimate, which is a weighted combination of quantile estimates, with the OLS estimate. From table 7, it can be seen that the restricted estimate outperforms the OLS estimate in precision with a larger t-value.[5]

Taken together, the results for men confirm previous findings that education is positively correlated with better diets, just as it has been shown to be positively correlated with other desirable health behaviors (Grossman and Kaestner, 1997). However, our results also show something new. The beneficial effects of education are much larger at parts of the conditional distribution that matter most—at the lower quantiles for energy where the risk of inadequacy is greater and at the upper quantiles for energy, fats, and cholesterol where the risk of excess intakes is the largest. For fiber, the size of the education coefficients were statistically similar across the quantiles. Similar to our reasoning for income, one explanation for not finding larger effects for education at the bottom quantiles may be that for the most part, the entire observed distribution of fiber intake is below the recommended fiber intake level.

_____

[5] Since the standard errors for the OLS estimates were computed using White's method and the standard errors for the quantile estimates were computed using design matrix bootstrap, the t-values of the OLS estimates and the restricted estimates in tables 3-7 are not strictly comparable. Therefore, we computed OLS standard errors using design matrix bootstrap with 500 replications. The OLS education coefficient on men's fiber intake had a bootstrap standard error of 0.0512. The standard error of the restricted estimate was 0.0321—37 percent smaller than the OLS bootstrap standard error.

There is less evidence of an increasing marginal effect of education at the higher quantiles for women. Women's energy intake increases significantly with education at all quantiles except at 0.90q. Although a trend of smaller influence at upper quantiles is visible in figure 12, equality of coefficients at symmetrical quantiles cannot be rejected (table 8). The restricted quantile estimates suggest that for each additional year of education, energy intake among women increases by about 13 calories. However, this increase has to be viewed in the context that the major part of women's energy intake distribution lies below the recommended level (table 1). Surprisingly, education tends to have a positive effect on women's total fat intake. However, the effects are significant only at 0.25q and 0.50q. Education has no significant impact on women's saturated fat intake, while for cholesterol the expected negative effect is found. Similar to the effect for men, the effect of women's educational attainment on their fiber intake is significant at all quantiles.

Consistent with the prediction from Grossman's model, men's energy, total fat, saturated fat, and cholesterol intakes decline with age, but more interestingly, the rates of decline rise steadily from the lower to the upper quantiles (fig. 13). The differences in the estimated effects between 0.1q and 0.9q are over 200 percent. The strong intake response to age is not surprising given that the health risk of poor diets is likely to be cumulative and increasing with age. Therefore, older individuals will display a greater propensity to improve their diets than younger individuals. In addition, our results show that the age effect is stronger at parts of the intake distributions where the risk of inadequacy or excess are greater.

Women's age has a similar pattern of impact across quantiles of energy, total fat, and saturated fat intakes. For cholesterol, the women's age coefficient is insignificant under OLS and at all quantiles. However, the optimally combined restricted estimate is significant, showing that cholesterol intake among women does decline with age. The age effect on women's fiber intake is uniformly positive across the OLS and quantile estimates. The age effect on fiber intake of men shows conflicting results under the different estimators. While the OLS and lower quantiles estimates are insignificant, the 0.9q estimate is negative and significant. However, the restricted estimate is positive and significant in accordance with our expectation, although the numerical effect is small (0.013). It is

not clear why age, which had sizable effects on other macronutrient intakes, would have such limited impact on men's fiber intake.

The difference in energy intake between Black and White men is surprisingly large, with Black men consuming about 209 fewer calories than White men at the conditional median. Although low t-values make it hard to discern a trend by quantiles, the restricted estimate shows that Black men have significantly lower total fat and saturated fat intakes than White men. For both intakes, interquantile equality cannot be rejected and the restricted estimate is more precise than the OLS estimate. While Black men's intakes look better than White men's in terms of energy, total fat, and saturated fat, the picture is starkly different for cholesterol and fiber. After controlling for other effects, Black men have higher cholesterol intake and lower fiber intake than White men. The quantile estimate shows that Black men consume 116 milligrams more cholesterol than White men at the 90th percentile. This is twice the difference at the conditional mean. Black men also have lower fiber intakes than White men across all quantiles, with the restricted quantile estimate giving a Black-White difference of about 1.9 grams.

Although the OLS estimate shows a lack of Black-White difference in women's energy intake, the 0.25q and the restricted estimates suggest that Black women's energy intake is significantly lower than White women's. Similar to the Black-White difference for men, Black women have higher cholesterol intake and lower fiber intake than White women. The Black-White difference in cholesterol is large, particularly at the upper quantiles and the interquantile equality tests show significant difference between intakes at the lower and upper quantiles. The restricted estimate suggests that Black women consume 1 gram less fiber than White women. The restricted estimate has a much higher t-value than the OLS estimate.[6] Unlike Black men though, Black women's intakes of total and saturated fats were not significantly lower than White women's.

Hispanic men's diets are significantly lower in total fat and saturated fat intakes than the diets of non-Hispanic

---

[6] The bootstrap standard error for the restricted quantile estimate is only 0.191 compared with a bootstrap standard error of 0.343 for the OLS estimate.

men. For energy, the evidence is inconclusive since only the Hispanic coefficient at the conditional median is significant. Based on the restricted estimate, Hispanic men consume 4.4 grams less total fat than non-Hispanic men. For saturated fat, the quantile estimates indicate that the relative difference is located at the upper ends of the distribution. This is confirmed by interquantile equality tests, which are both rejected at the 10-percent level. Due to the relatively low t-values, no significant differences between the Hispanic and non-Hispanic groups are evident for cholesterol and fiber intakes at various quantiles. However, the restricted estimate does show a higher fiber intake of 0.86 gram by Hispanic men than non-Hispanic men.

Both the OLS estimate and the restricted quantile estimate show significantly lower energy intake by Hispanic women than non-Hispanic women. Here again, the restricted estimate is far more precise than the OLS estimate. Hispanic women consume about 4 grams less total fat and 0.8 gram less saturated fat than non-Hispanic women, based on the restricted estimates. While the quantile estimates for cholesterol intake show an interesting trend with the Hispanic/non-Hispanic difference reversing in sign from lower to upper quantiles, none of the estimates are significant. Hispanic women's fiber intake is significantly higher at the bottom end of the distribution. At 0.1q, the unconditional estimate of which is well below the recommended level, Hispanic women tend to have about 1 gram of higher fiber intake than non-Hispanic women. Overall, the results for men and women show that the Hispanics tend to have a healthier macronutrient intake profile than non-Hispanics.

The data tables compiled from the 1994-96 CSFII can be used to speculate on the likely food sources of the intake differences by race and ethnicity (U.S Department of Agriculture, 2000). For example, Black men over age 20 consume only 7 grams of cheese compared with the 20 grams consumed by White men over age 20. The most recent study on the sources of nutrients based on the 1989-91 CSFII shows that cheese is the largest source of saturated fat and the fourth largest source of total fat among U.S. adults (Subar et al., 1998).[7] Black men also tend to consume lower amounts of whole and low-fat milk, another major source of fats, than White men. However, Black men consume lower amounts of yeast bread and ready-to-eat cereals, and higher amounts of eggs than White men, which may account for their lower fiber intake

and higher cholesterol intake. Based on Subar et al. estimates, yeast bread and milk are also the top and third largest sources of energy among U.S. adults. The 1994-96 CSFII data tables show that Black men over 20 consume only 178 grams of total milk and milk products per day compared with 262 grams consumed by White men over 20. The lower caloric intake by Black men may therefore be partly due to their lower consumption of these energy sources.

The picture is not as clear regarding the sources of Hispanic/non-Hispanic differences. For example, Hispanic men over age 20 consume higher amounts of beef and about the same amount of cheese as non-Hispanic White men over age 20. Consequently, the difference in their total and saturated fat intakes must come from other sources. Similarly Hispanics consume less yeast bread and ready-to-eat cereal than non-Hispanic Whites. Therefore, the higher fiber intake of Hispanics must come from other fiber sources such as legumes. Indeed, the 1994-96 CSFII data tables show that Hispanics do consume larger amounts of legumes than non-Hispanic Whites.

Conclusions

Understanding and quantifying the relative differences in food and nutrient intakes among population subgroups are important for targeting nutrition promotion programs and expenditures. The results can also contribute to improved understanding of health risk behavior at a time of rapid change in the health information environment. However, the nature of intake distributions is such that the risk of dietary excess or inadequacy is greater at the tails of the distributions than at the mean. Consequently, it is questionable to assume that the marginal effects of population characteristics will be constant along all parts of the conditional distribution of intakes. In this case, any analysis that focuses on only one part of the distribution, such as the conditional mean, may give an incomplete picture of the sources of intake difference among popula-

---

[7] The food intakes reported in the data tables are mean amounts based on day-1 of the 1994-96 CSFII (U.S. Department of Agriculture, 2000). The dietary sources of nutrients are based on day-1 of the 1989-91 CSFII and are not separated by race or sex (Subar et al.). These comparisons are meant to be illustrative, especially considering that the estimated differences reported in this paper are conditional (*ceteris paribus*), whereas the figures from the data tables and the dietary sources study are unconditional.

tion subgroups. In this study, we used quantile regression, a method suited for characterizing the entire distribution of intake, to examine macronutrient intakes among U.S. adults.

The findings clearly suggest that for key sociodemographic variables such as age, education, and income, the marginal effects at the tails of the intake distribution are often quite different from those at the mean. A more complete picture of the location of differences among population subgroups emerges from the quantile regression estimates than was available from the OLS estimates alone. Of course, this requires considering a larger number of parameter estimates. But in most cases, an optimal combination of the quantile estimates—the restricted estimates—outperformed the OLS estimates in precision.

These results have important implications for future studies evaluating the dietary impact of many nutrition-related policy interventions such as food assistance programs and food labeling regulations. For such studies to fully uncover the extent and nature of the behavioral impact, it is essential to look beyond the conditional mean to parts of the dietary intake distribution where the risk of inadequacy or excess is greatest.

One plausible interpretation of these results is that individuals, particularly men, at higher age, education, and income levels may have benefited more from

health and nutrition information initiatives such as the Nutrition Labeling and Education Act. Certainly, this explanation is consistent with the effect of human capital on health behavior predicted by the household production model. By comparing the influence of these variables on food and nutrient intakes over a sufficiently long timespan, it may be possible to verify the validity of this potentially important linkage.

The results reported here have to be interpreted and generalized cautiously given that the analysis is based on self-reported food intakes recorded from 2 days of 24-hour dietary recall. Such data are well-known to be prone to underreporting and may not fully represent a person's usual intake. In this study, we have attempted to control for underreporting and day-to-day intake variations to the extent possible by including sources of these variations such as season, whether the intake was recorded for a weekend day, and self-reported assessment whether the recorded intake was less than or more than a person's usual intake. Nevertheless, as with most previous studies based on a few days of recall data, the possibility cannot be ruled out that some of the estimated effects may have been influenced by underreporting or day-to-day variation. Estimation of conditional quantile models based on intake data from alternative data sources, which better represent usual intakes and may have lower incidence of underreporting, such as food-frequency questionnaires, may be able to resolve this issue.

Figure 3
Percentiles of macronutrient intake among men by income level
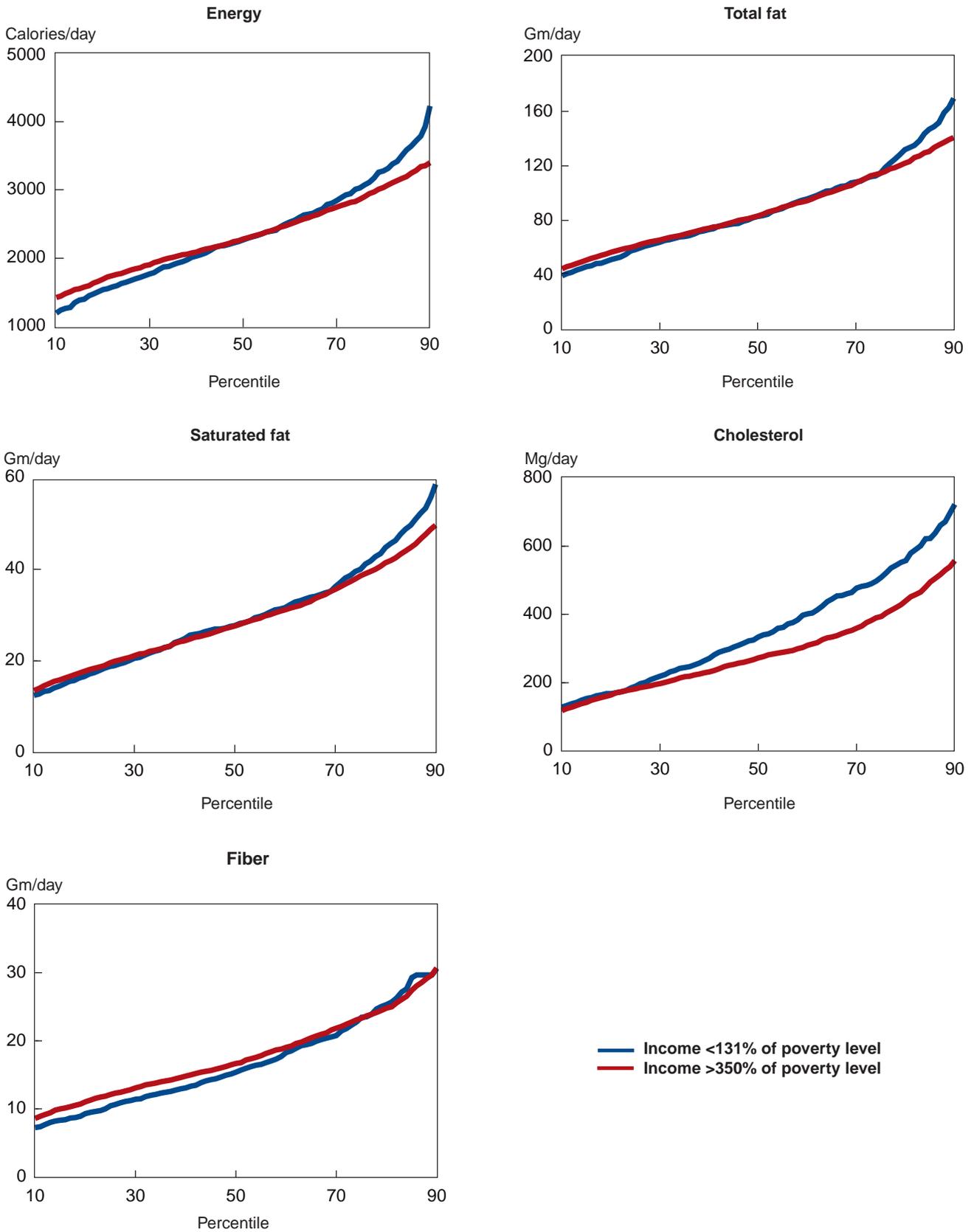
Figure 4
Percentiles of macronutrient intake among women by income level

Energy

Calories/day

Percentile

Total fat

Gm/day

Percentile

Saturated fat

Gm/day

Percentile

Cholesterol

Mg/day

Percentile

Fiber

Gm/day

Percentile

━━━ **Income <131% of poverty level**
━━━ **Income >350% of poverty level**
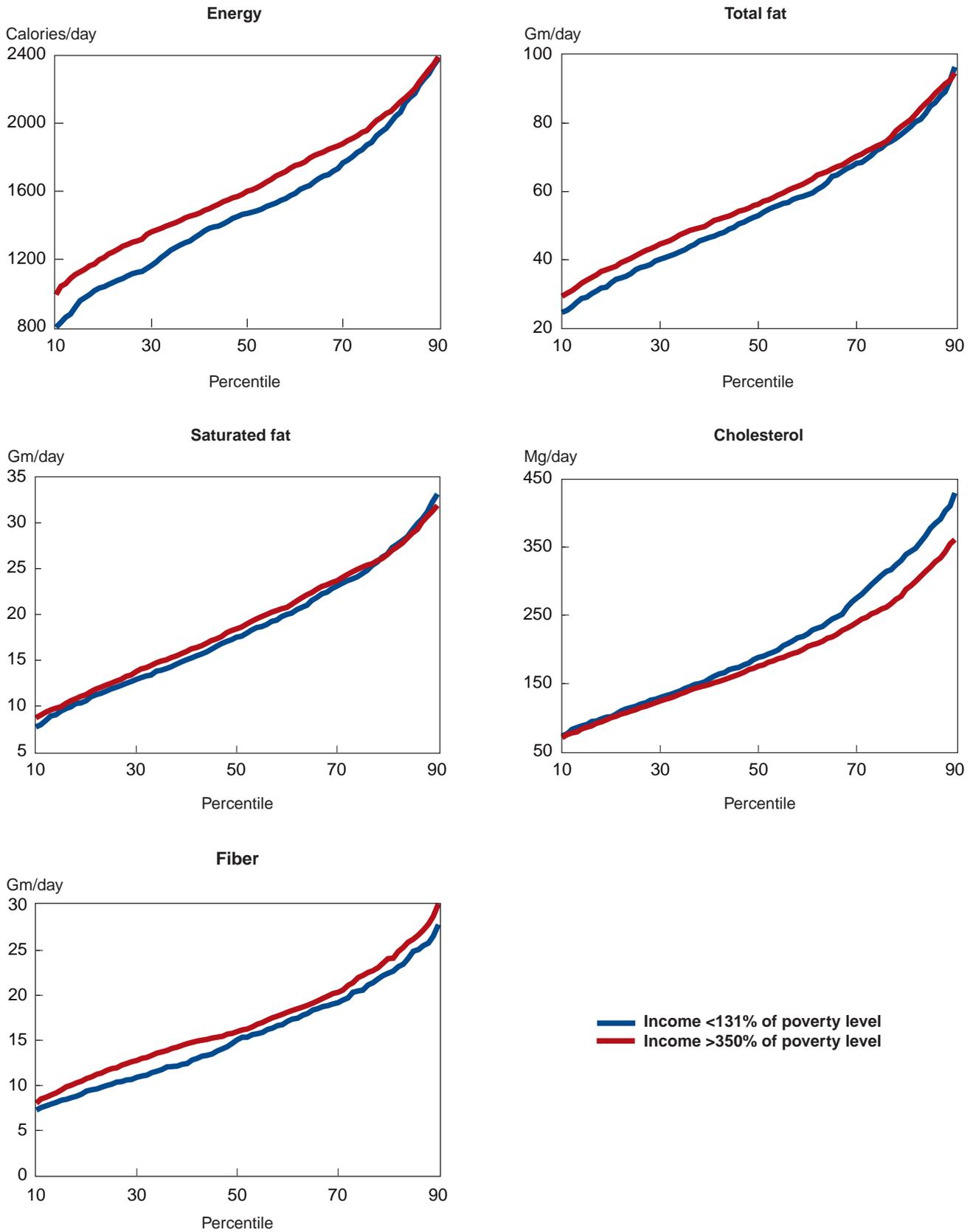
Figure 5
Percentiles of macronutrient intake among men by education level

Energy

Calories/day

Percentile

Total fat

Gm/day

Percentile

Saturated fat

Gm/day

Percentile

Cholesterol

Mg/day

Percentile

Fiber

Gm/day

Percentile

—— **Education <12 years**
—— **Education >12 years**
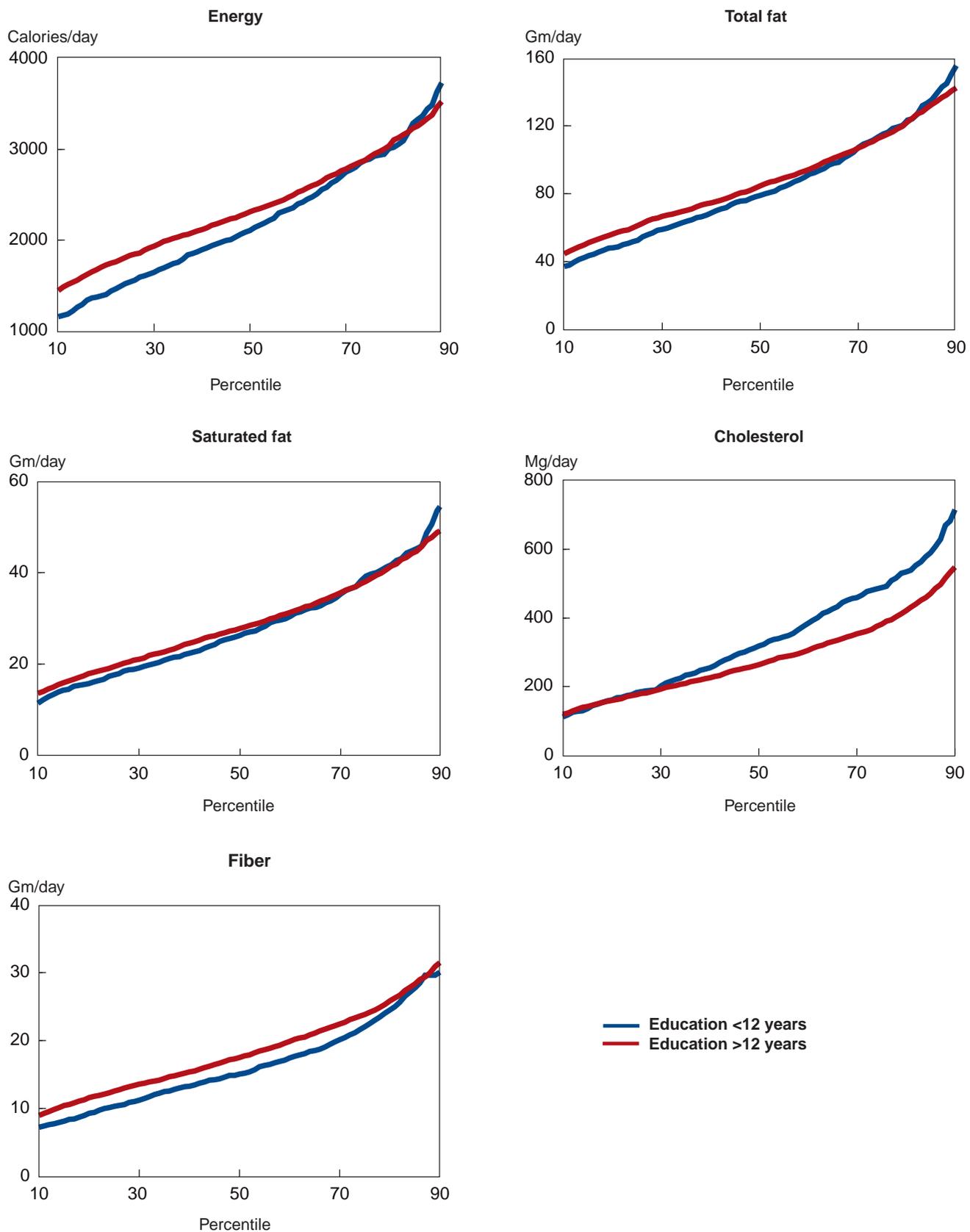
Figure 6

Percentiles of macronutrient intake among women by education level

Energy

Calories/day

Total fat

Gm/day

Saturated fat

Gm/day

Cholesterol

Mg/day

Fiber

Gm/day

— Education <12 years
— Education >12 years
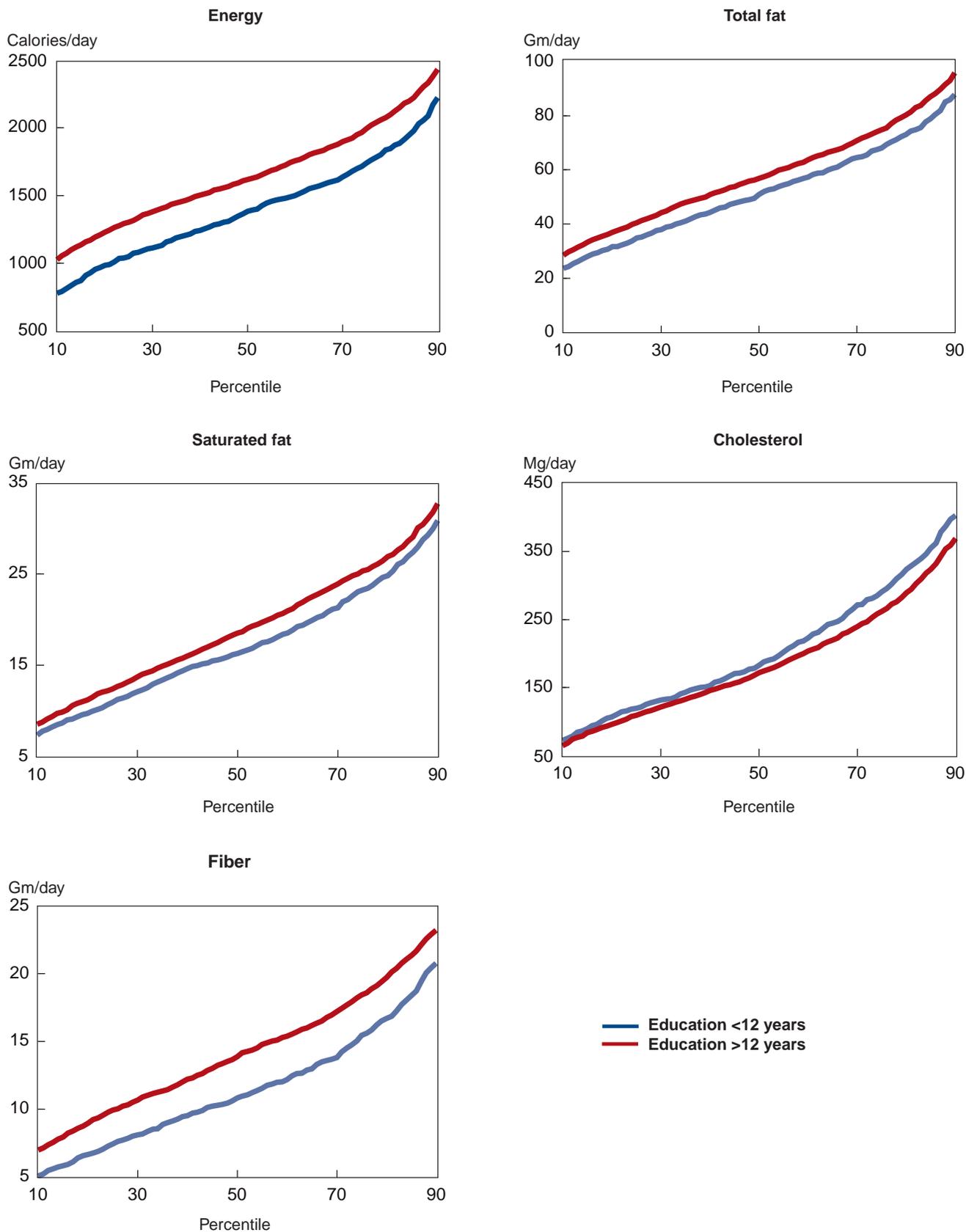
Figure 7
Percentiles of macronutrient intake among men by age level

Energy

Calories/day

Total fat

Gm/day

Saturated fat

Gm/day

Cholesterol

Mg/day

Fiber

Gm/day

Age 18-24 years
Age 65 years or older
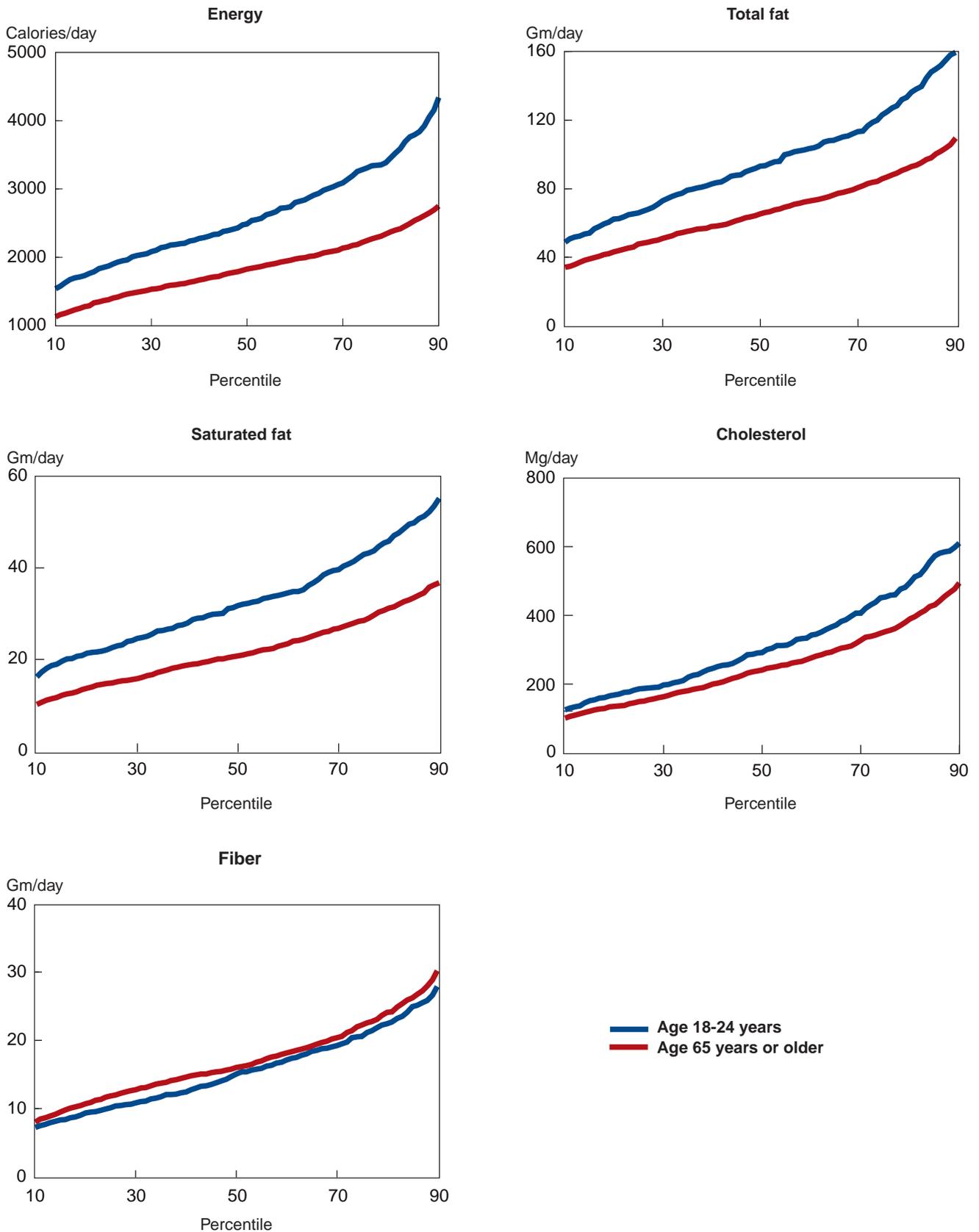
Figure 8
Percentiles of macronutrient intake among women by age level

Energy

Calories/day

Percentile

Total fat

Gm/day

Percentile

Saturated fat

Gm/day

Percentile

Cholesterol

Mg/day

Percentile

Fiber

Gm/day

Percentile

Age 18-24 years
Age 65 years or older
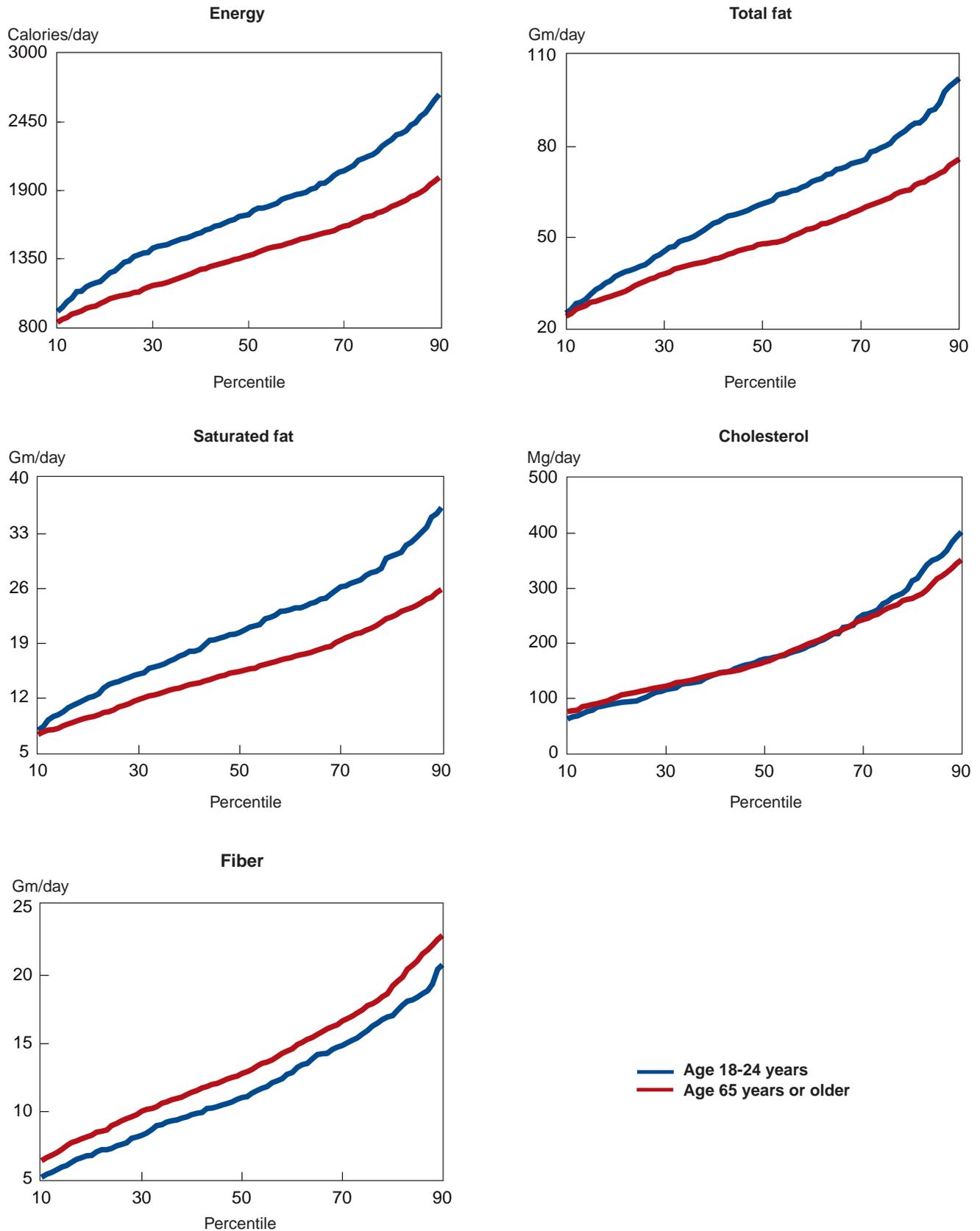
Figure 9
Marginal effect of a $1,000 increase in income on the macronutrient intake of men across quantiles

Energy
Calories/day

Total fat
Gm/day

Saturated fat
Gm/day

Cholesterol
Mg/day

Fiber
Gm/day
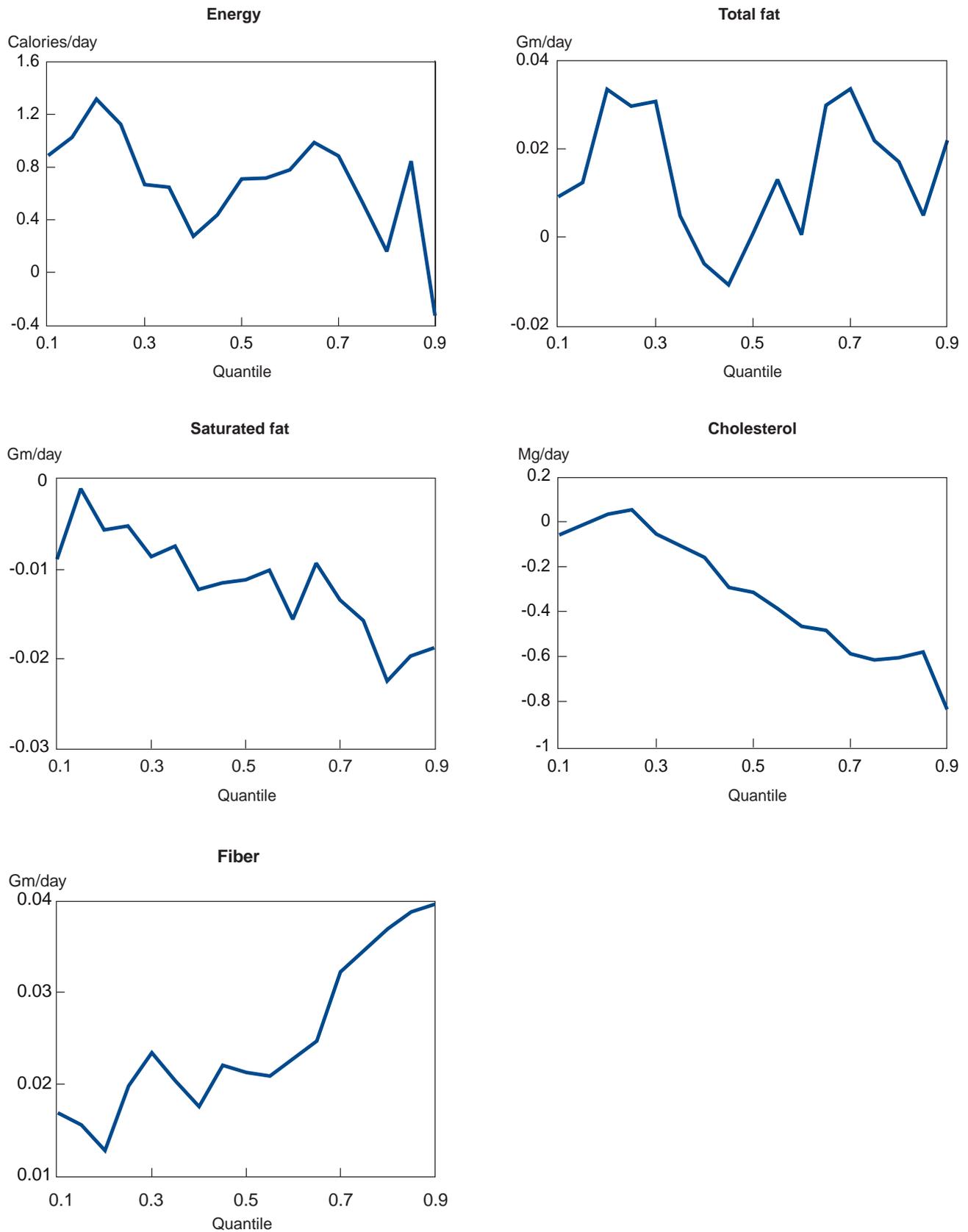
Figure 10

Marginal effect of a $1,000 increase in income on the macronutrient intake of women across quantiles

Energy

Calories/day

Quantile

Total fat

Gm/day

Quantile

Saturated fat

Gm/day

Quantile

Cholesterol

Mg/day

Quantile

Fiber

Gm/day
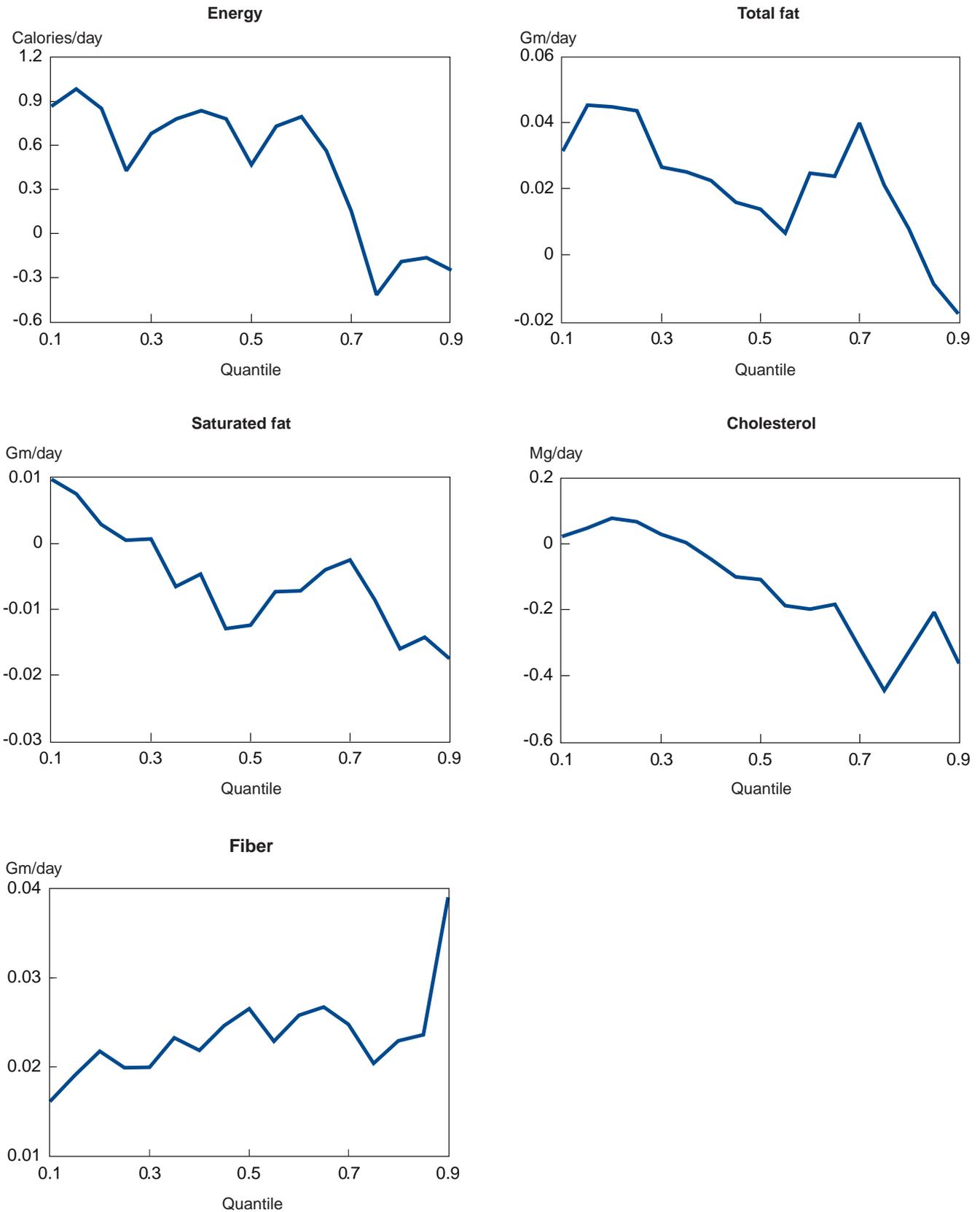
Quantile

Figure 11

Marginal effect of an additional year of education on the macronutrient intake of men across quantiles

Energy

Calories/day

Quantile

Total fat

Gm/day

Quantile

Saturated fat

Gm/day

Quantile

Cholesterol

Mg/day

Quantile

Fiber

Gm/day
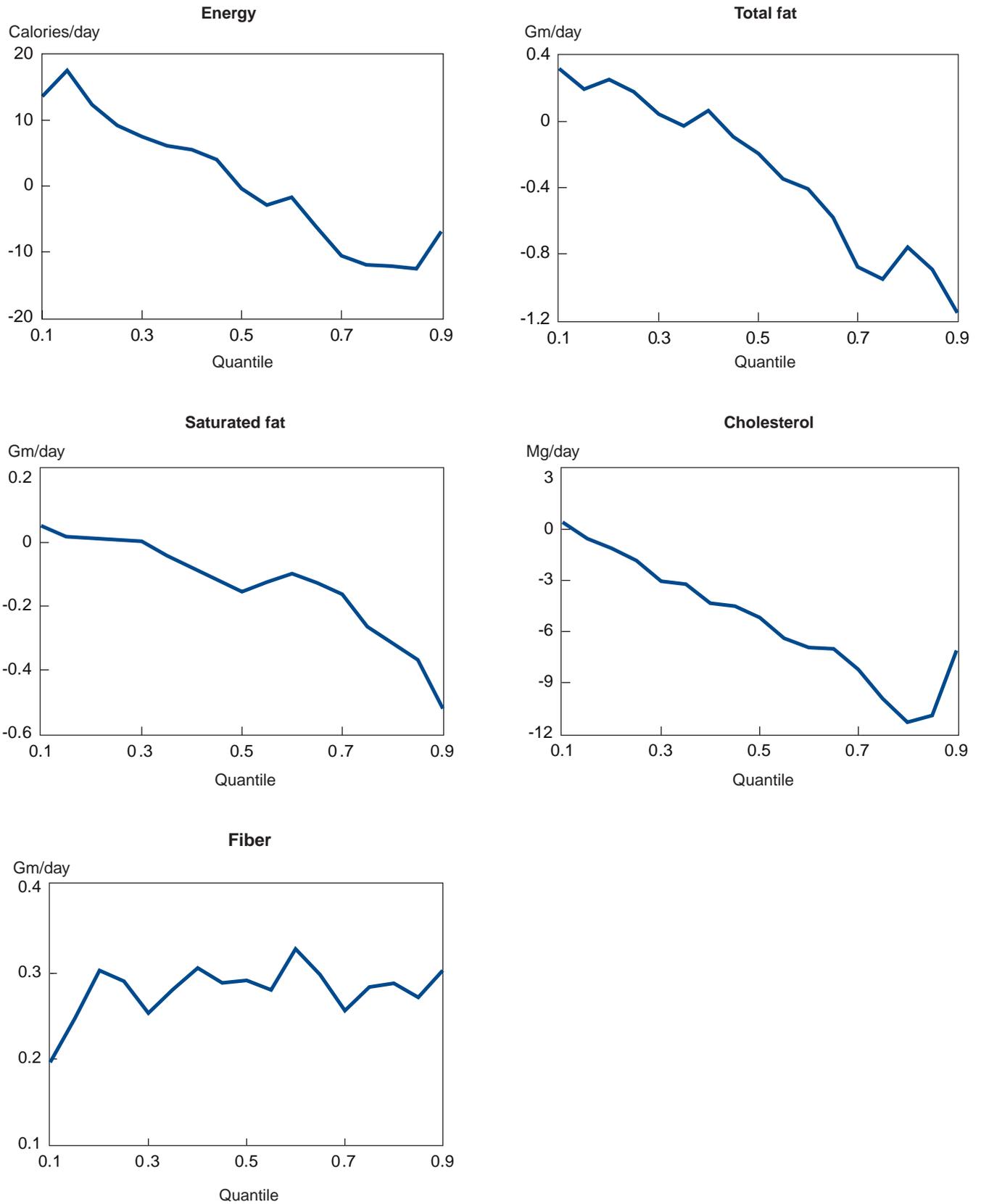
Quantile

Figure 12

Marginal effect of an additional year of education on the macronutrient intake of women across quantiles

Energy

Calories/day

Total fat

Gm/day

Saturated fat

Gm/day

Cholesterol

Mg/day

Fiber

Gm/day
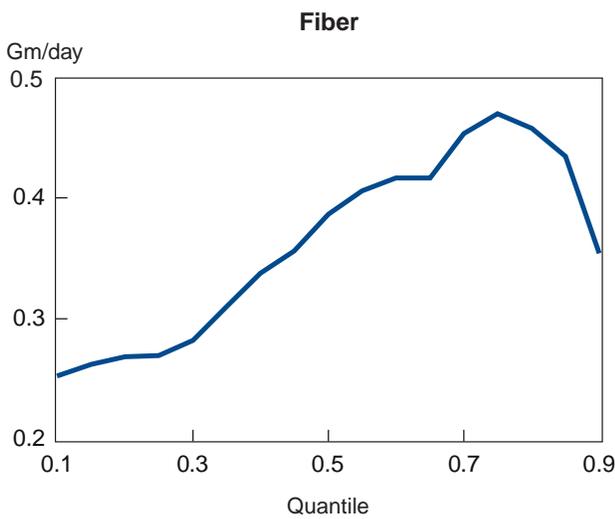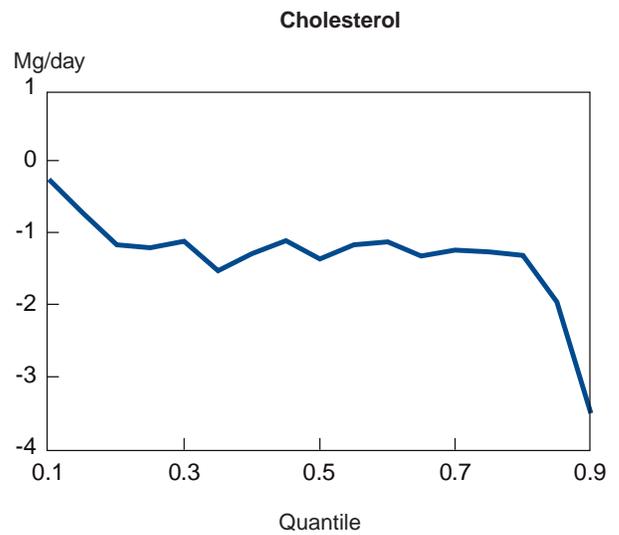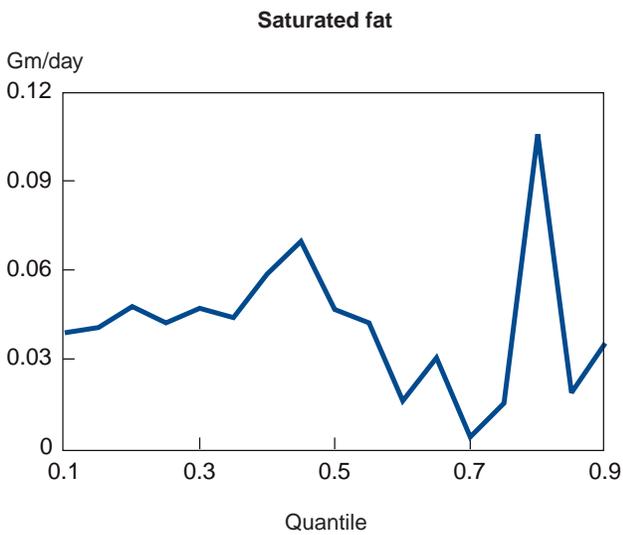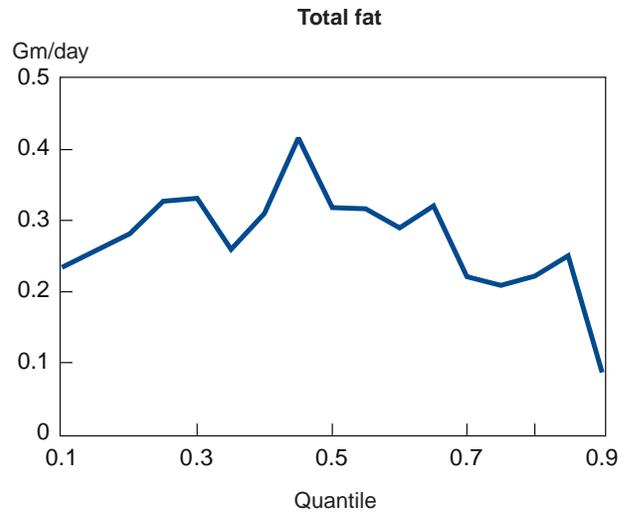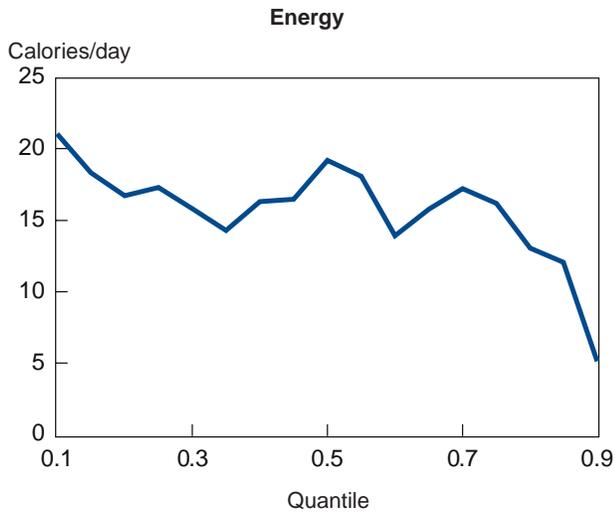
Figure 13

Marginal effect of an additional year of age on the macronutrient intake of men across quantiles

Energy

Calories/day

Quantile

Total fat

Gm/day

Quantile

Saturated fat

Gm/day

Quantile

Cholesterol

Mg/day

Quantile

Fiber

Gm/day
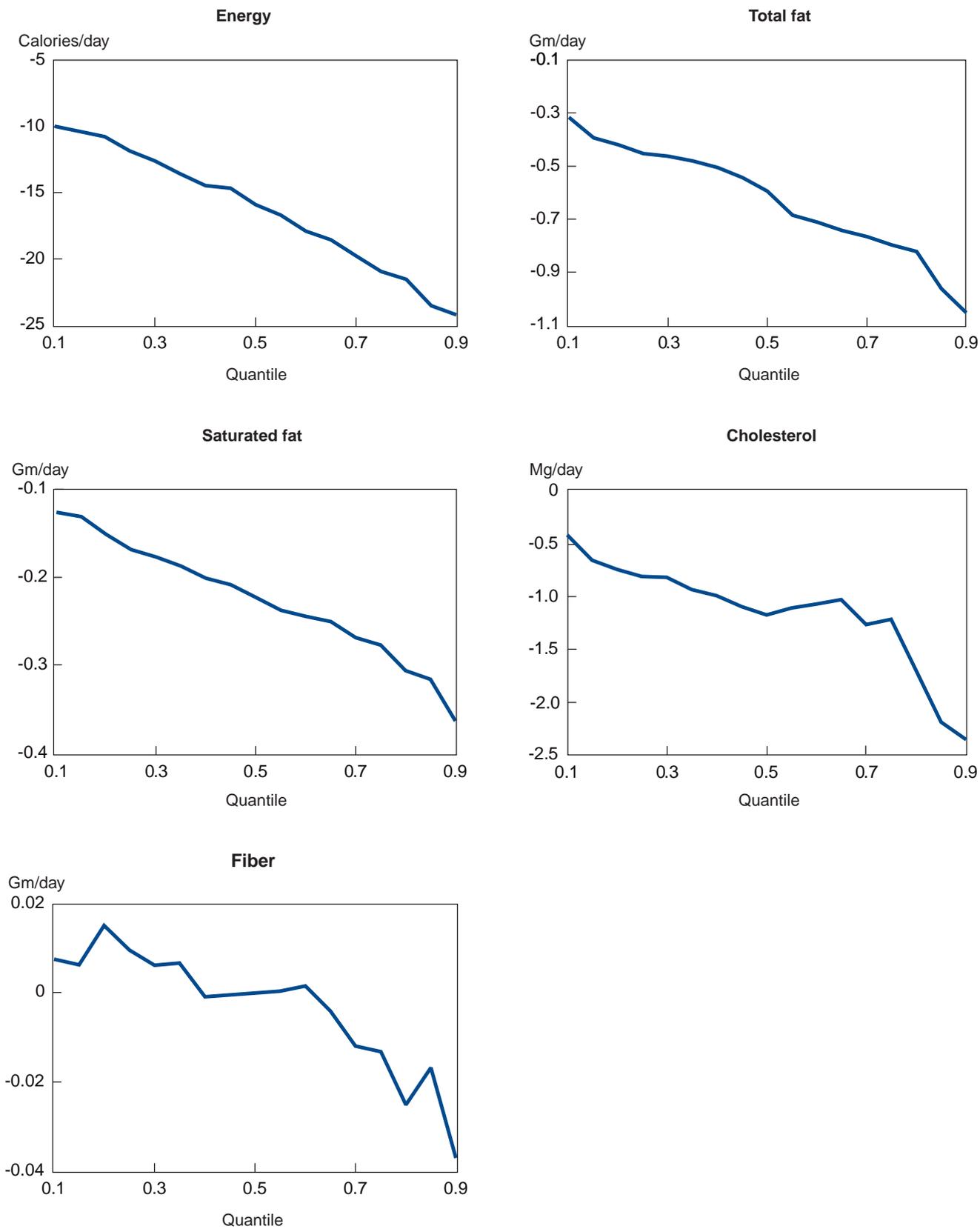
Quantile

Figure 14
Marginal effect of an additional year of age on the macronutrient intake of women across quantiles

Energy

Calories/day

Quantile

Total fat

Gm/day

Quantile

Saturated fat

Gm/day

Quantile

Cholesterol

Mg/day

Quantile

Fiber

Gm/day
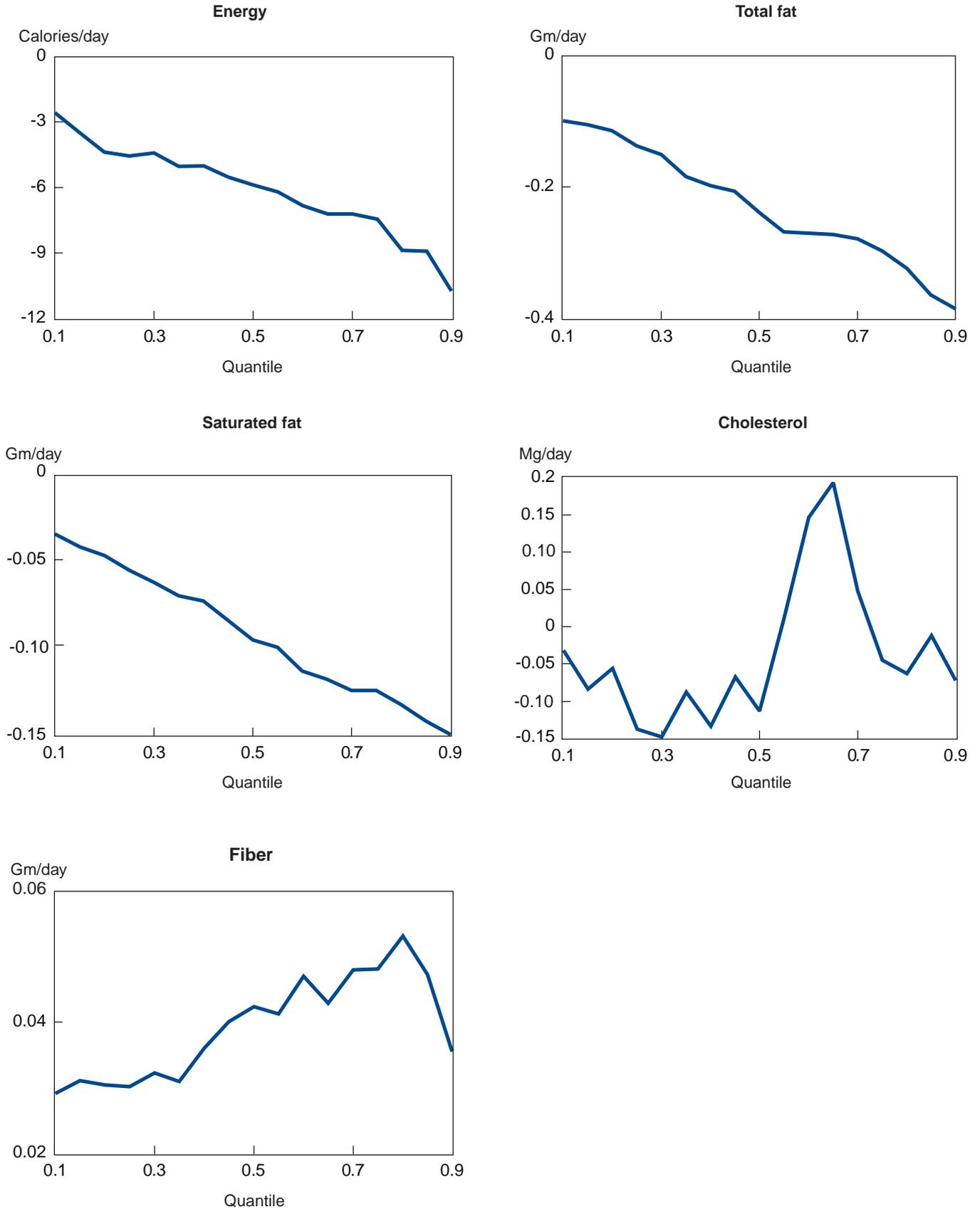
Quantile

Table 3—Quantile regression estimates: Energy intake, 1994-96

Variable	OLS	Quantile					$\hat{\beta}_R$
		0.10	0.25	0.50	0.75	0.90	
Men:							
Income ($\times 10^{-3}$)	.167	.888	1.125	.709	.532	-.327	.311
	(.31)	(1.47)	(2.06)	(1.15)	(.68)	(.30)	(.88)
Education	-1.360	13.65	9.205	-.365	-11.93	-6.875	9.772
	(.32)	(2.70)	(1.77)	(.07)	(1.72)	(.73)	(3.26)
Age	-17.38	-9.972	-11.83	-15.91	-20.97	-24.26	-13.86
	(21.39)	(11.41)	(14.67)	(16.46)	(16.75)	(14.93)	(26.57)
Black	-154.06	-126.07	-121.03	-209.08	-234.25	-159.95	-175.73
	(3.40)	(2.61)	(2.43)	(4.11)	(4.19)	(1.68)	(5.87)
Hispanic	-75.27	-28.67	-37.99	-133.90	-94.69	-87.85	-47.34
	(1.45)	(.49)	(.71)	(2.45)	(.99)	(.66)	(1.36)
R^2	.191	.093	.093	.106	.122	.147	—
Women:							
Income ($\times 10^{-3}$)	.213	.866	.426	.467	-.419	-.252	.355
	(.58)	(1.63)	(1.01)	(1.08)	(.77)	(.32)	(1.29)
Education	15.01	21.06	17.31	19.20	16.17	5.127	12.57
	(4.90)	(4.95)	(3.95)	(5.69)	(3.27)	(.64)	(5.12)
Age	-6.743	-2.572	-4.555	-5.896	-7.457	-10.74	-5.900
	(12.19)	(3.40)	(8.06)	(8.64)	(8.24)	(8.87)	(14.99)
Black	-8.483	-51.170	-75.33	-5.308	6.334	12.72	-47.94
	(.29)	(1.47)	(2.31)	(.18)	(.18)	(.20)	(2.52)
Hispanic	60.06	53.01	-48.77	-58.31	-144.46	-117.03	-92.67
	(1.74)	(1.34)	(1.18)	(1.55)	(1.68)	(1.61)	(3.80)
R^2	.125	.072	.071	.068	.073	.090	—

— = Not applicable.
 Note: Absolute t-values reported in parentheses. All regressions included an intercept and 25 additional explanatory variables; see appendix table 1 for definitions.

Table 4—Quantile regression estimates: Total fat intake, 1994-96

Variable	OLS	Quantile					$\hat{\beta}_R$
		0.10	0.25	0.50	0.75	0.90	
Men:							
Income (×10⁻³)	-.009	.009	.030	.001	.022	.022	0.043
	(.35)	(.33)	(1.25)	(.03)	(.56)	(.38)	(2.67)
Education	-.358	.320	.179	-.194	-.952	-1.158	-.354
	(1.76)	(1.50)	(.77)	(.73)	(3.02)	(2.38)	(2.54)
Age	-.661	-.316	-.452	-.594	-.797	-1.053	-.411
	(17.69)	(7.79)	(12.53)	(13.84)	(14.47)	(10.99)	(16.44)
Black	-3.823	-4.230	-2.926	-4.397	-5.934	-7.744	-3.325
	(1.80)	(1.95)	(1.32)	(1.88)	(1.92)	(1.31)	(2.42)
Hispanic	-5.095	-2.679	-2.250	-8.437	-9.026	-3.545	-4.407
	(2.13)	(1.11)	(.83)	(2.63)	(2.14)	(0.53)	(2.66)
R²	.155	.082	.080	.085	.098	.107	—
Women:							
Income (×10⁻³)	.011	.031	.044	.014	.021	.018	.012
	(.65)	(1.56)	(2.13)	(.63)	(.74)	(.42)	(.94)
Education	.245	.234	.327	.319	.208	.086	.289
	(1.58)	(1.41)	(1.80)	(1.72)	(.93)	(.24)	(2.79)
Age	.245	.100	.138	.239	.298	.385	-.163
	(9.30)	(3.15)	(4.46)	(6.92)	(6.77)	(5.87)	(8.24)
Black	1.746	.817	1.522	1.074	1.826	.721	1.216
	(1.35)	(.51)	(.91)	(.63)	(.95)	(.19)	(1.22)
Hispanic	3.573	.671	1.760	2.833	5.596	4.734	-4.012
	(2.03)	(.32)	(1.06)	(1.24)	(1.75)	(1.07)	(3.38)
R²	.108	.049	.054	.056	.065	.084	—

— = Not applicable.

Note: Absolute t-values reported in parentheses. All regressions included an intercept and 25 additional explanatory variables; see appendix table 1 for definitions.

Table 5—Quantile regression estimates: Saturated fat intake, 1994-96

Variable	OLS	Quantile					$\hat{\beta}_R$
		0.10	0.25	0.50	0.75	0.90	
Men:							
Income (×10⁻³)	-.016	-.009	-.005	-.011	-.016	-.019	-.012
	(1.73)	(.94)	(.53)	(1.11)	(1.09)	(.84)	(1.97)
Education	-.177	.053	.004	-.154	-.264	-.521	-.024
	(2.36)	(.65)	(.05)	(1.74)	(2.13)	(2.56)	(.50)
Age	-.243	-.126	-.169	-.222	-.277	-.363	-.183
	(17.01)	(9.15)	(13.98)	(15.49)	(13.11)	(10.60)	(21.50)
Black	-2.154	-1.182	-1.501	-1.967	-2.687	-2.046	-1.425
	(2.85)	(1.54)	(2.28)	(2.59)	(2.54)	(1.13)	(3.06)
Hispanic	-2.464	-1.507	-.747	-2.735	-3.756	-5.635	-1.71
	(2.88)	(1.72)	(.98)	(2.57)	(2.35)	(2.51)	(3.21)
R²	.153	.078	.084	.087	.098	.105	—
Women:							
Income (×10⁻³)	-.010	.010	.001	-.012	-.009	.018	.007
	(1.47)	(1.64)	(0.07)	(1.68)	(.84)	(1.11)	(1.62)
Education	.035	.039	.042	.047	.015	.035	.031
	(.64)	(.74)	(.65)	(0.70)	(.17)	(.26)	(.83)
Age	.101	.036	.056	.096	.125	.150	-.058
	(10.47)	(3.65)	(5.01)	(9.23)	(7.72)	(6.58)	(9.43)
Black	-0.057	-.333	.099	.577	-.632	.328	-.470
	(0.11)	(.64)	(.18)	(1.16)	(.85)	(.27)	(1.56)
Hispanic	1.459	.704	1.118	1.503	1.291	1.561	-.820
	(2.29)	(.99)	(2.00)	(1.98)	(1.25)	(1.21)	(2.13)
R²	.104	.049	.050	.061	.065	.079	—

— = Not applicable.
 Note: Absolute t-values reported in parentheses. All regressions included an intercept and 25 additional explanatory variables; see appendix table 1 for definitions.

Table 6—Quantile regression estimates: Cholesterol intake, 1994-96

Variable	OLS	Quantile					$\hat{\beta}_R$
		0.10	0.25	0.50	0.75	0.90	
Men:							
Income (×10⁻³)	-.388	-.060	.052	-.315	-.615	-.835	-.197
	(3.13)	(.62)	(.47)	(2.60)	(3.46)	(2.70)	(2.96)
Education	-5.440	.411	-1.830	-5.195	-9.945	-7.131	-1.700
	(5.19)	(.48)	(2.05)	(4.65)	(5.49)	(2.80)	(2.97)
Age	-1.306	-.415	-.807	-1.174	-1.215	-2.357	-.721
	(7.09)	(2.95)	(4.68)	(5.59)	(4.36)	(5.21)	(7.12)
Black	56.514	16.259	36.393	46.731	89.803	115.965	24.773
	(4.86)	(1.85)	(3.43)	(3.36)	(3.50)	(3.72)	(3.92)
Hispanic	-6.454	-1.759	-2.825	-.361	10.891	25.56	-11.118
	(.47)	(.18)	(.25)	(.02)	(.48)	(.77)	(1.49)
R²	.084	.035	.035	.047	.062	.065	—
Women:							
Income (×10⁻³)	-.217	.022	.067	-.108	-.445	.363	.003
	(2.44)	(.31)	(.84)	(1.24)	(2.75)	(1.55)	(.07)
Education	-1.458	-.261	-1.215	-1.366	-1.271	-3.508	-2.439
	(1.84)	(.33)	(1.80)	(1.72)	(.97)	(1.58)	(5.54)
Age	.179	.032	.138	.114	.045	.073	-.180
	(1.29)	(.29)	(1.28)	(.80)	(.19)	(.18)	(2.49)
Black	36.176	13.846	25.535	35.149	58.40	63.415	14.068
	(4.88)	(2.70)	(4.14)	(4.16)	(4.96)	(3.18)	(3.81)
Hispanic	3.366	5.829	6.473	2.241	2.314	14.023	1.871
	(.39)	(.79)	(.86)	(.20)	(.18)	(.61)	(.39)
R²	.060	.030	.030	.034	.046	.051	—

— = Not applicable.

Note: Absolute t-values reported in parentheses. All regressions included an intercept and 25 additional explanatory variables; see appendix table 1 for definitions.

Table 7—Quantile regression estimates: Fiber intake, 1994-96

Variable	OLS	Quantile					$\hat{\beta}_R$
		0.10	0.25	0.50	0.75	0.90	
Men:							
Income ($\times 10^{-3}$)	.024	.017	.020	.021	.035	.040	.019
	(4.08)	(2.82)	(3.19)	(3.51)	(4.12)	(2.75)	(5.03)
Education	.239	.196	.291	.292	.284	.303	.248
	(4.47)	(3.77)	(5.34)	(5.61)	(3.81)	(2.22)	(7.70)
Age	-.004	.007	.010	.007	-.013	-.037	.012
	(.50)	(.96)	(1.16)	(.78)	(1.00)	(1.96)	(2.52)
Black	-2.188	-1.465	-1.444	-2.300	-2.780	-2.073	-1.866
	(4.97)	(3.16)	(3.58)	(5.65)	(3.47)	(1.80)	(7.14)
Hispanic	.789	.586	.173	.767	1.511	.984	.855
	(1.39)	(1.01)	(.30)	(1.14)	(1.54)	(.64)	(2.25)
R^2	.078	.049	.045	.050	.054	.057	—
Women:							
Income ($\times 10^{-3}$)	.023	.016	.020	.026	.020	.039	.017
	(5.18)	(4.45)	(5.00)	(4.84)	(3.05)	(3.65)	(6.93)
Education	.351	.253	.270	.387	.470	.354	.277
	(8.54)	(6.15)	(6.76)	(8.15)	(6.94)	(4.79)	(9.71)
Age	.035	.029	.030	.042	.048	.035	.035
	(5.18)	(4.48)	(4.80)	(5.09)	(4.80)	(1.96)	(8.30)
Black	-.847	-.682	-1.195	-1.069	-1.206	-.660	-1.049
	(2.61)	(2.37)	(3.84)	(2.92)	(2.18)	(.82)	(5.50)
Hispanic	.836	.995	.918	1.047	1.036	.763	.927
	(1.98)	(2.29)	(2.48)	(2.23)	(1.49)	(.76)	(3.57)
R^2	.104	.063	.067	.066	.069	.067	—

— = Not applicable.

Note: Absolute t-values reported in parentheses. All regressions included an intercept and 25 additional explanatory variables; see appendix table 1 for definitions.

Table 8—Tests for equality of slope parameters across symmetrical quantiles

Variable	Energy		Total fat		Saturated fat		Cholesterol		Fiber	
	0.1q=0.9q	0.25q=0.75q	0.1q=0.9q	0.25q=0.75q	0.1q=0.9q	0.25q=0.75q	0.1q=0.9q	0.25q=0.75q	0.1q=0.9q	0.25q=0.75q
Men:										
Income	.98 (.32)	.57 (.45)	.04 (.84)	.04 (.84)	.19 (.66)	.57 (.45)	6.02 (.01)	12.45 (.00)	2.34 (.13)	2.99 (.08)
Education	4.17 (.04)	8.70 (.00)	8.42 (.00)	10.95 (.00)	7.44 (.01)	4.86 (.03)	8.76 (.00)	20.41 (.00)	0.59 (.44)	0.01 (.93)
Age	63.93 (.00)	49.20 (.00)	53.94 (.00)	38.76 (.00)	45.26 (.00)	27.14 (.00)	20.15 (.00)	2.16 (.14)	4.63 (.03)	2.82 (.09)
Black	.11 (.73)	3.42 (.06)	.33 (.56)	.87 (.35)	.21 (.65)	1.29 (.26)	10.25 (.00)	4.81 (.03)	0.27 (.60)	2.99 (.08)
Hispanic	.17 (.68)	.36 (.55)	.02 (.90)	2.45 (.12)	3.08 (.08)	3.97 (.05)	.67 (.41)	.41 (.52)	0.07 (.80)	2.01 (.16)
Women:										
Income	1.65 (.20)	2.25 (.13)	1.16 (.28)	.64 (.43)	2.83 (.09)	.74 (.39)	2.62 (.11)	10.26 (.00)	4.49 (.03)	.00 (.94)
Education	3.40 (.07)	.04 (.84)	.15 (.70)	.23 (.63)	.00 (.98)	.09 (.76)	2.15 (.14)	.00 (.97)	1.66 (.20)	10.27 (.00)
Age	37.59 (.00)	10.66 (.00)	18.62 (.00)	13.66 (.00)	23.44 (.00)	17.68 (.00)	.01 (.92)	.17 (.68)	.12 (.73)	3.43 (.06)
Black	.84 (.36)	3.66 (.06)	.00 (.98)	.02 (.89)	.27 (.60)	.88 (.35)	6.06 (.01)	8.07 (.00)	.00 (.98)	.00 (.98)
Hispanic	4.38 (.04)	1.00 (.32)	.77 (.38)	1.54 (.22)	.36 (.55)	.03 (.86)	.72 (.40)	.46 (.50)	.05 (.82)	.03 (.86)

Note: The numbers against the variable names are F-statistics with (1, N-K) degrees of freedom. The associated p-values are reported in parentheses.

References

Adelaja, A., R.M. Nayga, Jr., and T. Lauderbach. "Income and Racial Differentials in Selected Nutrient Intakes," *American Journal of Agricultural Economics* 79(1997): 1452-1460.

Becker, G.S. "A Theory of the Allocation of Time," *Economic Journal* 75 (September 1965): 493-517.

Bhargava, A. "Modelling the Health of Filipino Children," *Journal of the Royal Statistical Society A* 157, Part 3(1994): 417-432.

Bowman, S.A., M. Lino, S.A. Gerrior, and P.P. Basiotis. *The Healthy Eating Index*: 1994-96. CNPP-5, Center for Nutrition Policy and Promotion. U.S. Department of Agriculture, July 1998.

Buchinsky, M. "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research," *The Journal of Human Resources* 33(1998): 88-126.

Buchinsky, M. "Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: A Monte Carlo Study," *Journal of Econometrics* 68(1995): 303-338.

Chavas, J-P., and K.O. Keplinger. "Impact of Domestic Food Programs on Nutrient Intake of Low-Income Persons in the United States," *Southern Journal of Agricultural Economics* 15(1983): 155-63.

Chern, W.S. "Demand for Food and Demand for Health in the United States." Paper presented at the mini-symposium on "Effects of Health Information on the Demand for Food - EU and US Experiences" at the XXIV International Conference of the International Association of Agricultural Economists, Berlin, Germany, August 13-18, 2000.

Davis, C., and E. Saltos. "Dietary Recommendations and How They Have Changed Over Time," in *America's Eating Habits: Changes and Consequences,* ed. E. Frazao, p: 33-50. Agriculture Information Bulletin No. 750. Economic Research Service, U.S. Department of Agriculture, 1999.

Deaton, A. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Baltimore, MD: Johns Hopkins University Press, 1997.

Frazao, E. *The American Diet: Health and Economic Consequences.* Agriculture Information Bulletin No. 711. Economic Research Service, U.S. Department of Agriculture, 1999.

Gould, B.W. "Factors affecting U.S. demand for reduced-fat fluid milk," *Journal of Agricultural and Resource Economics* 21(1996): 68-81.

Gould, B.W., and H.C. Lin. "Nutrition information and household dietary fat intake," *Journal of Agricultural and Resource Economics* 19(1994): 349-365.

Gould, W. "sg70: Interquantile and Simultaneous-Quantile Regression," *Stata Technical Bulletin* 38(1997): 14-22.

Grossman, M. "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy* 80(April 1972): 223-255.

Grossman, M., and R. Kaestner. "Effects of Education on Health," in *The Social Benefits of Education*, eds. Jere R. Behrman and Nevzer G. Stacey, Ann Arbor, MI: University of Michigan Press, 1997, p. 69-123.

Institute of Medicine. *Dietary Reference Intakes: Applications in Dietary Assessment.* Washington, DC: National Academy Press, 2000.

Ippolito, P.M., and A.D. Mathios. "Information, advertising, and health choices: A study of the cereal market," *Rand Journal of Economics* 21(1990): 459-480.

Kantor, L.S. *A Dietary Assessment of the U.S. Food Supply: Comparing Per Capita Food Consumption with Food Guide Pyramid Serving Recommendations.* Agricultural Economic Report No. 772. Economic Research Service, U.S. Department of Agriculture, 1998.

Kim, S.-Y., R.M. Nayga, and O. Capps. "The effect of food label use on nutrient intakes: An endogenous switching regression analysis," *Journal of Agricultural and Resource Economics* 25(2000): 215-231.

Koenker, R., and G. Bassett. "Regression Quantiles," *Econometrica* 46(1978): 33-50.

Koenker, R., and G. Bassett. "Robust Tests for Heteroscedasticity Based on Regression Quantiles," *Econometrica* 50(1982): 43-62.

Koplan, J.P., and W.H. Dietz. "Caloric Imbalance and Public Health Policy," *Journal of the American Medical Association* 282(October 1999): 579-1581.

Lancaster, K.J. "A New Approach to Consumer Theory," *Journal of Political Economy* 74(April 1966): 132-157.

Lichtenstein, A.H., E. Kennedy, P. Barrier, D. Danford, N.D. Ernst, S.M. Grundy, G.A. Leveille, L. Van Horn, C.L. Williams, and S.L. Booth. "Dietary Fat Consumption and Health," *Nutrition Reviews* 56(May 1998): S3-S28.

Lin, B-H., J. Guthrie, and E. Frazao. *Away-From-Home Foods Increasingly Important to Quality of American Diet.* Agriculture Information Bulletin No. 749. Economic Research Service, U.S. Department of Agriculture, 1999.

Ludwig, D.S., M.A. Pereira, C.H. Kroenke, J.E. Hilner, L. Van Horn, M.L. Slattery, and D.R. Jacobs. "Dietary Fiber, Weight Gain, and Cardiovascular Disease Risk Factors in Young Adults," *Journal of the American Medical Association* 282(October 1999): 1539-1546.

Murphy, S.P., D. Rose, M. Hudes, and F.E. Viteri. "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(1992): 1352-1357.

National Research Council. *Diet and Health: Implications for Reducing Chronic Disease Risk.* Washington DC: National Academy Press, 1989.

Nayga, R.M., Jr. "Effects of Socioeconomic and Demographic Factors on Consumption of Selected Food Nutrients," *Agricultural and Resource Economics Review* 23(1994): 171-82.

Rosenzweig, M.R., and T.P. Schultz. "Estimating Household Production Functions: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight," *Journal of Political Economy* 91(1983): 723-746.

Sickles, R.C., and P. Taubman. "Mortality and Morbidity Among Adults and the Elderly," in *Handbook of Population and Family Economics*, eds. M.R. Rosenzweig and O. Stark, Amsterdam: North-Holland, 1997, p. 559-643.

Sickles, R.C., and A. Yazbeck. "On the Dynamics of Demand for Leisure and the Production of Health," *Journal of Business & Economic Statistics* 16(April 1998): 187-197.

Strauss, J., and D. Thomas. "Human Resources: Empirical Modeling of Household and Family Decision," in *Handbook of Development Economics, Vol. III*, eds. T.N. Srinivasan and J.R. Behrman, Amsterdam: North-Holland, 1996.

Subar, A.F., S.M. Krebs-Smith, A. Cook, L. Kahle. "Dietary Sources of Nutrients Among U.S. Adults, 1989 to 1991," *Journal of the American Dietetic Association* 98(1998): 537-547.

Tippett, K.S., and Y.S. Cypel. *Design and Operation: The Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey, 1994-96.* NFS Report No. 96-1. Agricultural Research Service, U.S. Department of Agriculture, 1997.

U.S. Department of Agriculture, Agricultural Research Service. *Food and Nutrient Intakes by Individuals in the United States, 1994-96, Vols. 1 and 2.* Nationwide Food Surveys Report No. 96-3. August 2000.

Variyam, J.N., J.R. Blaylock, and D. Smallwood. "A Probit Latent Variable Model of Nutrition Information and Dietary Fiber Intake," *American Journal of Agricultural Economics* 78(August 1996): 628-639.

Appendix table 1—Explanatory variables, means, and sample size

Variable	Men	Women
Household characteristics		
Gross annual income ($000)	39.5	36.1
Household size	2.9	2.8
Region (Northeast omitted)		
Midwest	.24	.25
South	.36	.37
West	.21	.20
Urbanization (city omitted)		
Suburb	.46	.44
Nonmetro	.26	.25
Personal characteristics		
Level of education (years)	12.7	12.6
Age (years)	49.0	49.1
Height (inches)	69.8	64.0
Weight (pounds)	183.4	151.7
Race (White omitted):		
Black	.10	.13
Asian	.02	.02
Other[1]	.05	.04
Ethnic origin-Hispanic	.08	.08
On any kind of diet	.15	.21
Survey-related characteristics		
Year (1994 omitted)		
1995	.34	.34
1996	.34	.32
Day-1 season (Winter omitted)		
Spring	.18	.18
Summer	.29	.29
Fall	.28	.28
Day-2 season (Winter omitted)		
Spring	.16	.16
Summer	.29	.29
Fall	.29	.29
Day-1 on weekend	.30	.29
Day-2 on weekend	.25	.25
Day-1 intake (Usual omitted)		
Less than usual	.19	.20
More than usual	.10	.12
Day-2 intake (Usual omitted)		
Less than usual	.20	.21
More than usual	.09	.10
Sample size (N)	4725	4362

[1]Asian/Pacific Islander, Aleut, Eskimo, or American Indian.

Appendix table 2—Distribution of total energy intake (calories) by categories of selected socioeconomic variables

Variable	Men						Women					
	Mean	Percentile					Mean	Percentile				
		10	25	50	75	90		10	25	50	75	90
				Calories						Calories		
Income[1]												
≤130%	2626	1192	1647	2267	3027	4226	1547	801	1102	1469	1871	2375
131-350%	2392	1351	1764	2245	2904	3544	1602	916	1213	1554	1898	2328
>350%	2396	1418	1796	2280	2863	3401	1664	993	1288	1601	1957	2385
Education (years)												
<12	2410	1142	1539	2103	2897	3738	1441	780	1049	1382	1740	2221
12	2390	1372	1785	2247	2884	3513	1612	937	1226	1554	1924	2341
>12	2449	1441	1830	2312	2916	3531	1686	1028	1302	1619	1992	2428
Age (years)												
18-24	2770	1535	1959	2484	3305	4235	1787	920	1332	1705	2173	2669
25-34	2624	1547	1998	2477	3170	3839	1711	1005	1316	1650	2063	2453
35-44	2585	1391	1856	2382	3014	3680	1692	1013	1298	1587	2008	2470
45-54	2311	1350	1767	2238	2796	3259	1594	1005	1233	1551	1903	2254
55-64	2143	1227	1624	2101	2597	3015	1492	906	1149	1487	1772	2112
≥65	1882	1124	1461	1828	2241	2729	1403	839	1067	1373	1689	1999
Race/Ethnicity												
Non-Hispanic White	2422	1386	1794	2278	2902	3528	1623	958	1245	1564	1923	2346
Black	2497	1149	1590	2160	2860	3644	1614	844	1137	1538	1896	2387
Other[2]	2356	1240	1706	2228	2775	3724	1649	852	1173	1528	2111	2618
Hispanic	2395	1331	1783	2332	2993	3614	1551	924	1148	1482	1886	2306

[1]Annual household income expressed as a percentage of the poverty threshold.
[2]Asian/Pacific Islander, Aleut, Eskimo, or American Indian.

Appendix table 3—Distribution of total fat intake (grams) by categories of selected socioeconomic variables

Variable	Men						Women					
	Mean	Percentile					Mean	Percentile				
		10	25	50	75	90		10	25	50	75	90
				Grams						Grams		
Income[1]												
≤130%	101.1	39.2	52.1	82.7	114.5	—	57.7	24.6	37.0	53.0	72.7	96.4
131-350%	90.7	43.2	58.9	84.8	114.0	145.7	59.1	26.8	39.5	55.9	74.1	94.2
>350	90.2	44.5	61.1	83.2	114.6	140.8	60.0	29.3	41.2	56.2	73.8	94.6
Education (years)												
<12	94.2	37.0	52.6	79.0	115.2	155.4	54.2	23.5	34.8	50.9	68.1	87.8
12	91.8	43.9	62.1	86.1	115.5	147.3	60.3	27.7	41.1	56.2	74.4	96.9
>12	91.0	44.5	61.2	84.5	113.9	142.4	60.3	28.4	40.4	56.8	74.8	95.8
Age (years)												
18-24	100.6	48.5	66.0	93.1	122.3	159.3	63.8	24.9	40.1	61.1	80.2	102.4
25-34	99.4	52.8	68.8	92.6	123.8	156.4	62.1	30.3	43.2	58.3	77.3	97.3
35-44	100.4	44.5	61.3	90.9	123.3	155.1	63.2	28.8	41.5	58.5	78.4	103.9
45-54	88.5	42.6	61.3	81.7	113.8	137.3	59.9	29.4	40.6	56.4	74.7	92.7
55-64	82.4	36.9	55.3	77.3	107.7	130.3	55.2	25.8	37.4	52.6	70.7	85.8
≥65	69.3	34.3	47.8	65.4	85.7	109.5	50.1	24.0	34.7	47.9	62.5	76.0
Race/Ethnicity												
Non-Hispanic White	91.8	45.3	62.0	85.1	115.5	145.1	59.3	28.0	40.0	55.8	73.8	94.2
Black	100.0	36.7	55.5	81.1	116.2	—	62.4	24.7	41.0	57.8	76.1	103.0
Other[2]	77.0	29.5	46.2	73.2	107.6	127.6	55.1	22.9	35.2	51.1	71.9	90.0
Hispanic	89.8	40.8	60.4	84.7	111.5	146.1	55.7	25.8	36.3	52.2	72.4	90.3

[1] Annual household income expressed as a percentage of the poverty threshold.
[2] Asian/Pacific Islander, Aleut, Eskimo, or American Indian.

—=Not reported because standard error could not be computed.

Appendix table 4—Distribution of saturated fat intake (grams) by categories of selected socioeconomic variables

	Men						Women					
	Mean	Percentile					Mean	Percentile				
Variable		10	25	50	75	90		10	25	50	75	90
		Grams		Grams				Grams		Grams		
Income[1]												
≤130%	35.2	12.4	18.7	27.7	40.0	58.5	19.4	7.7	11.8	17.5	24.4	33.1
131-350%	30.6	13.2	19.6	28.3	38.8	50.6	19.8	8.3	12.3	18.1	25.3	33.1
>350%	30.2	13.5	19.5	27.7	38.5	49.6	19.6	8.7	12.5	18.4	25.1	31.9
Education (years)												
<12	32.6	11.3	17.6	26.2	39.2	54.4	18.0	7.3	10.9	16.3	23.3	30.9
12	31.9	13.7	20.3	28.6	39.3	51.0	20.3	8.9	12.8	18.7	25.5	33.7
>12	30.4	13.5	19.4	27.0	36.9	47.8	19.8	8.5	12.4	18.6	25.3	32.7
Age (years)												
18-24	34.7	16.1	22.6	31.8	43.0	55.3	21.9	7.9	13.9	20.4	27.6	36.4
25-34	34.1	16.6	22.5	31.0	42.3	55.9	20.9	9.2	13.8	19.3	26.6	34.4
35-44	34.3	13.3	20.5	29.8	42.1	53.2	21.1	9.0	12.7	19.2	26.6	36.4
45-54	29.1	13.7	18.9	27.2	37.5	46.7	19.5	8.9	12.5	18.3	25.2	31.1
55-64	26.8	11.5	16.9	25.3	33.8	44.9	17.9	7.9	11.2	16.5	23.7	29.9
≥65	23.0	10.2	14.9	20.8	28.6	36.7	16.3	7.3	10.4	15.4	20.7	25.9
Race/Ethnicity												
Non-Hispanic												
White	31.1	14.0	20.0	28.5	39.1	50.9	19.8	8.6	12.5	18.3	25.2	32.6
Black	34.4	10.5	18.1	27.0	38.4	52.1	20.5	7.4	12.3	18.3	26.0	35.1
Other[2]	24.8	8.4	14.2	20.6	34.7	47.4	17.1	6.7	10.8	14.3	22.1	30.3
Hispanic	29.4	13.1	19.3	27.9	38.2	50.4	18.5	7.9	11.8	16.9	23.7	30.9

[1] Annual household income expressed as a percentage of the poverty threshold.
[2] Asian/Pacific Islander, Aleut, Eskimo, or American Indian.

Appendix table 5—Distribution of cholesterol intake (milligrams) by categories of selected socioeconomic variables

Variable	Men						Women					
	Mean	Percentile					Mean	Percentile				
		10	25	50	75	90		10	25	50	75	90
				Milligrams						Milligrams		
Income[1]												
≤130%	398	128	189	334	509	722	226	73	116	188	308	428
131-350%	329	122	179	283	423	595	211	72	112	173	274	393
>350%	310	118	182	272	394	558	201	71	111	176	259	361
Education (years)												
<12	379	112	182	318	488	714	221	72	120	182	290	403
12	343	125	192	292	453	622	216	77	115	181	279	399
>12	305	119	176	264	382	547	201	64	110	171	261	369
Age (years)												
18-24	350	125	186	292	454	610	206	62	99	172	276	402
25-34	337	129	189	289	432	612	218	77	117	185	273	397
35-44	351	126	193	293	445	621	212	68	115	176	273	400
45-54	328	117	185	279	428	593	214	71	116	181	291	408
55-64	323	127	187	279	417	607	205	70	114	177	266	379
≥65	272	101	149	241	354	492	199	76	113	166	264	352
Race/Ethnicity												
Non-Hispanic White	315	122	180	272	399	570	202	69	110	172	261	370
Black	411	121	223	327	522	807	242	78	132	206	324	451
Other[2]	322	111	157	287	418	584	246	65	126	222	320	445
Hispanic	359	127	177	310	485	647	216	78	121	178	297	417

[1]Annual household income expressed as a percentage of the poverty threshold.
[2]Asian/Pacific Islander, Aleut, Eskimo, or American Indian

Appendix table 6—Distribution of fiber intake (grams) by categories of selected socioeconomic variables

Variable	Men Mean	Men Percentile 10	25	50	75	90	Women Mean	Women Percentile 10	25	50	75	90
				Grams						Grams		
Income[1]												
≤130%	18.1	7.2	10.4	15.4	23.4	30.4	11.9	4.9	7.1	10.7	15.4	20.4
131-350%	17.7	7.8	11.4	16.1	22.3	28.8	13.1	5.9	8.5	12.1	16.2	21.1
> 350%	18.5	8.6	12.1	16.6	23.3	30.7	14.7	6.6	9.8	13.5	18.3	23.4
Education (years)												
< 12	17.6	7.2	10.3	15.1	22.1	30.1	12.1	5.0	7.4	10.8	15.4	20.8
12	16.6	7.4	10.9	15.0	21.2	27.2	12.5	5.6	8.0	11.5	15.6	20.1
>12	19.2	9.0	12.6	17.5	23.9	31.5	14.8	7.0	9.9	13.9	18.5	23.2
Age (years)												
18-24	16.7	7.2	10.1	15.0	20.6	27.8	12.6	5.2	7.5	11.0	16.0	20.8
25-34	18.4	8.4	11.9	16.1	23.1	31.5	13.3	5.9	8.4	12.2	16.3	21.4
35-44	18.4	8.2	12.0	16.7	23.8	29.7	13.5	6.0	8.7	12.5	16.5	22.5
45-54	18.7	8.2	12.0	17.3	23.7	30.9	13.9	5.9	9.0	13.0	18.0	22.6
55-64	17.7	7.7	11.5	16.5	22.4	29.3	13.6	6.6	9.0	12.7	16.9	21.8
≥65	17.9	8.0	11.8	16.0	22.2	30.2	14.1	6.4	9.1	12.8	17.8	22.9
Race/Ethnicity												
Non-Hispanic White	18.2	8.2	11.8	16.5	23.1	30.1	13.8	6.2	9.0	12.7	17.2	22.5
Black	15.4	6.0	9.6	13.6	19.8	29.7	11.4	4.5	7.0	9.7	14.7	19.2
Other[2]	17.6	7.5	12.7	16.1	21.1	27.3	14.5	6.5	8.9	12.9	17.5	23.5
Hispanic	20.0	8.5	12.2	17.7	24.3	32.8	13.7	6.1	8.8	12.5	18.1	23.2

[1]Annual household income expressed as a percentage of the poverty threshold.
[2]Asian/Pacific Islander, Aleut, Eskimo, or American Indian.