# Income Elasticity of Calorie Intake in United States

Wang Biao
School of Economics, Huazhong University of Science and Technology, Wuhan, P.R. China, 430074
(E-mail: wangb1985@126.com)

**Abstract:** Many food programs, like Supplemental Nutrition Assistance Program, are concerned about low-income households. To evaluate the effects of such programs, one needs demand elasticity estimates pertinent to households of different income levels. Parametric, nonparametric and semiparametric estimations are used to estimate the income elasticity of calorie intake from household survey data in this paper. Results indicate that income elasticity of calorie intake is low in United States. However, the poor group has higher calorie elasticity than the rich people. The estimation results are useful in evaluating some food policy and program effects related to households of a specific income level.
**Key words:** Income elasticity; Calorie intake; Nonparametric estimation; Semiparametric estimation

## 1 Introduction

Inadequate nutrition is perhaps one of the most pressing problems facing the poor, both in developing and developed countries. Malnutrition would lower productivity and increase the risk of disease, conspiring to help the poor stay poor and then inducing more problems. The effect of income growth on nutrient intake has been a subject of controversy in the development economics literature. Two groups of empirical literatures emerge from the debate, one of which asserts that policies with increase the income of the poor have beneficial effects on nutrition, the income elasticity of nutrition intake is far from zero, suggesting that an increase in income will reduce malnutrition in the developing world (Strauss and Thomas, 1990; Ravallion, 1990; Subramanian and Deaton, 1996). The other group suggests that increases in income will not result in substantial improvements in nutrient intakes, the income elasticity of nutrition intake is close to zero (Behrman and Deolalikar, 1987). Although the poor may increase their food expenditures as incomes rise, this extra spending goes on food attributes other than nutrients, for example, taste, appearance, variety, or status, which are not necessarily positively correlated with nutritive value. These divergent opinions call for more empirical research to analyze the income-nutrient relationship and provide plausible results for appropriate policy responses.

Calorie intake, which is one of the most important factors from the policy maker's viewpoint, has been found to have a strong linkage with both human health and productivity (Babatunde, Adejobi & Fakayode, 2010). On one hand, human body needs calorie energy to maintain the natural body metabolism and good health. On the other hand, from the policy maker's point of view, calorie intake is the foremost concern when making some beneficial schemes for the poor. As a result, in this paper we choose calorie intake as our research object.

Nearly all the literatures study the income-nutrient relationship in the case of developing countries, such as India, Indonesia and Philippines. However, the malnutrition problem faced by poor people does exist in developed countries, even in United States. As indicated by the data from Current Population Survey (CPS), 2010 Annual Social and Economic (ASEC) Supplement, the official poverty rate in 2009 was 14.3 percent, the number of people in poverty, 43.6 million, is the largest number in the 51 years for which poverty estimates have been published. National Health and Nutrition Examination Survey (NHANES) conducted in the United States collects detailed information on the quantity of various nutrients and on the demographic characteristics of households. Estimating demand relationships from household survey data from United States in order to obtain elasticities distinguished by household characteristics is of interest.

This paper tries to estimate the income elasticities of calorie intake, and see whether the results are robust to different estimation tools. The next section of this paper introduces two widely used methods of deriving income elasticities of nutrient intake. After describing the data, income elasticity of calorie intake is estimated using parametric, nonparametric and semiparametric methods. The main findings are then presented and the last section offers some conclusions.

## 2 Conceptual Framework

### 2.1 Methodology

There are two widely used ways to derive the income elasticities of demand for nutrition, one is called indirect method and the other is direct method. Indirect calculation is based on weighted averages of income elasticities for broad food groups. Suppose a household consumes M kinds of foods with a predetermined total income I, the demand for $j^{th}$ food quantity $q_j$ can be expressed as a function of the vector of food prices and the total income.

$$q_j = q_j(P, I; \beta) \tag{1}$$

where P is a vector of food prices, I is the predetermined total income, and β is a parameter vector. Let $\alpha_{i,j}$ be the quantity of the $i^{th}$ nutrient obtained from a unit of the the $j^{th}$ food. The total quantity of that nutrient, $N_i$, obtained from various foods can be expressed as sum of the nutrient from all kinds of food.

$$N_i = \sum_{j=1}^{M} \alpha_{i,j} \cdot q_j(P, I; \beta) \tag{2}$$

Then the income elasticity of $i^{th}$ nutrient can be expressed as:

$$\frac{\partial N_i}{\partial I} \cdot \frac{I}{N_i} = \sum_{j=1}^{M} \alpha_{i,j} \cdot \frac{\partial q_i}{\partial I} \cdot \frac{I}{\sum_j \alpha_{i,j} \cdot q_j} \tag{3}$$

$$N_i = N_i(I; \alpha, \beta) \tag{4}$$

This approach provides policy guidance to improve nutrient intake. However, the income elasticity of calorie intake is always overstated by indirect calculation. This indirect method is a convenient empirical strategy but it assumes no substitution from lower-priced calories to higher-priced calories within each food group, inducing the overestimation of calorie elasticity. To avoid the potential bias due to above-mentioned fixed conversions, in this paper, we estimate the income elasticity of calorie intake directly.

### 2.2 Data Sources

The empirical work in this paper uses data from the National Health and Nutrition Examination Survey, which is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations. This program began in the early 1960s and has been conducted as a series of surveys focusing on different population groups or health topics. The survey examines a nationally representative sample of about 5,000 persons each year, each participant represents approximately 50,000 other U.S. residents. These persons, from a broad range of age groups and racial/ethnic backgrounds, are located in counties across the country, 15 of which are visited each year.

NHANES interview different households each year, we can not consider the data collected in different years as a panel. So, we choose the latest available survey, carried out in 2007-2008. 9762 households were considered respondents to the MEC examination and data collection. However, 9255 of the MEC respondents provided complete dietary intakes for Day 1. Furthermore, of those providing the Day 1 data, 7838 provided complete dietary intakes for Day 2. Here, we use only Day 1 data, since considering Day 2 data will reduce the sample by a large amount. Even though, the available sample is smaller, because of households with missing income information and other demographic information. After deleting the households with incomplete information, a sample of 8337 households was left.

## 3 Empirical Work

### 3.1 Parametric estimation

We start from an oversimplified model with the traditional functional form using parametric method. For our oversimplified model, we might specify the following statistical model:

$$Kcal_i = \alpha_0 + \alpha_1 Income_i + Z_i \beta + \varepsilon_i \tag{5}$$

where $Kcal_i$ and $Income_i$ are the logarithm of calorie intakes and that of annual income for the $i^{th}$ household, respectively. $\alpha_0$ is the yet unknown value of the intercept, $\alpha_1$ is the yet unknown value of the coefficient on Income. $Z_i$ is a $1 \times p$ vector of explanatory variables, $\beta$ is a $p \times 1$ vector of regression coefficients, and $\varepsilon_i$ is the $i^{th}$ error term. All the variables involved are listed in Table 1.

**Table 1 Variable List**

| Name | Label |
|---|---|
| Kcal | Natural Logarithm of One-day Calorie Intake |
| Income | Natural Logarithm of Annual Household Income |
| riagendr | Gender |
| ridageyr | Age at Screening Adjudicated |
| ridreth1 | Race/Ethnicity |
| dmdcitzn | Citizenship Status |
| dmdhhsiz | Total number of people in the Household |
| dmdhrgnd | HH Ref Person Gender |
| dmdhrage | HH Ref Person Age |
| dmdhrbr2 | HH Ref Person Country of Birth |
| dmdhredu | HH Ref Person Education Level |
| dmdhrmar | HH Ref Person Martial Status |

Regression results are listed in Table 2. Controlling just for three demographic variables involving the respondent in regression (1) gives income elasticity of calorie intake of 0.0838. When citizenship and household size are included, regression (2) gives elasticity of 0.0876. The results in regression (3) are generated by a model that adds further covariates for characteristics of household reference. The added variables lower the estimated elasticity slightly to 0.0803. Most of the explanatory variables are statistically significant. Race of the respondent does not influence calorie intake in all the three regressions. Results from parametric estimation indicate that increase in income does promote the households' calorie intake, but the elasticity is less than 10 percent.

**Table 2 OLS Estimates of Calorie Intake Regressions**

| V ariables | regression (1) | | regression (2) | | regression (3) | |
|---|---|---|---|---|---|---|
| | $\beta$ | $p>\|t\|$ | $\beta$ | $p>\|t\|$ | $\beta$ | $p>\|t\|$ |
| Income | 0.0837785*** | 0.000 | 0.0875778*** | 0.000 | 0.0803295*** | 0.000 |
| riagendr | -.2282059*** | 0.000 | -.2276252*** | 0.000 | -.2288209*** | 0.000 |
| ridageyr | 0.0013129*** | 0.000 | 0.0010286*** | 0.000 | 0.0038690*** | 0.000 |
| ridreth1 | 0.0042897 | 0.358 | 0.0063254 | 0.199 | 0.0035657 | 0.474 |
| dmdcitzn | | | 0.0519491*** | 0.003 | 0.0505128*** | 0.006 |
| dmdhhsiz | | | -.0072052* | 0.054 | -.0013443 | 0.725 |
| dmdhrgnd | | | | | 0.0062573 | 0.569 |
| dmdhrage | | | | | -.0055648*** | 0.000 |
| dmdhrbr2 | | | | | -.0215890*** | 0.000 |
| dmdhredu | | | | | 0.0127508*** | 0.006 |
| dmdhrmar | | | | | 0.0027800** | 0.026 |
| cons | | 0.000 | | 0.000 | 7.6864640*** | 0.000 |
| Number of obs | 8337 | | 8337 | | 8337 | |
| $R^2$ | 0.0701 | | 0.0714 | | 0.0917 | |

*, ** and *** indicate statistically significant at the 10%, 5% and 1% level, respectively

### 3.2 Nonparametric Estimation

Parametric methods have great difficulties in capturing the information regarding the impact of income changes on the calorie elasticity, the relationship between calorie intake and income may be characterized by non-linearity (Gibson & Rozelle, 2000). Nonparametric regression, which allows the functional form of a fit to data itself, is an appropriate tool to study the relationship between calorie intake and income since it does not make any guidance or constraints about the functional form. It provides a powerful set of tools that can be extremely useful for data analysis when there is little a priori knowledge of the shape of the function to be estimated, especially when that shape may vary over the distribution of the covariates. One of two types of nonparametric techniques is kernel estimation, which specifies y=m(x)+e, where m(x) is the conditional expectation of y with no parametric form whatsoever, E(y|x), and the density of the error e is completely unspecified. The N observations $y_i$ and $x_i$ are used to estimate a joint density function for y and x. The density at a point ( $y_0 , x_0$ ) is estimated by seeing what proportion of the N observations are "close to" ( $y_0 , x_0$ ). Both of kcal and income are in logarithmic form, so gradients in the model give us one elasticity for each observation. The minimum, maximum and mean of the elasticities are listed in Table 3.

**Table 3 Nonparametric Estimation**

| Annual household income | Under 20,000 | 100,000 and over | Number of obs |
|---|---|---|---|
| Income elasticity | 0.1248484 | 0.08330732 | 8337 |

Results from nonparametric estimation show the bivariate relationship between calorie intake and income, excluding the effects of some relevant variables. For some households, the income elasticities are negative, but close to zero. Minimum of them is -0.0095. Some households have relatively high elasticities, the highest of which is greater than 0.17. The income elasticity of calorie intake is approximately 0.1248 for households whose annual income is under 20000 dollars and 0.0833 for those whose annual income is 100000 dollars and over. It reflects the fact that poorer people have higher calorie elasticities.

### 3.3 Semiparametric estimation

The semiparametric regression, based on the model described by Robinson (1988), includes regression models that combine parametric and nonparametric models.

$$Kcal_i = Z_i\beta + f(Income_i) + \varepsilon_i \qquad E(\varepsilon_i \mid Z_i, Income_i) = 0 \qquad i=1,2,\ldots,n \tag{6}$$

Equation (6) can be rewritten as:

$$Kcal_i - E(Kcal_i \mid Income_i) = (Z_i - E(Z_i \mid Income_i))\beta + \varepsilon_i \tag{7}$$

then the unknown functional form $f(Income_i)$ can be estimated in the next three steps. Firstly, the conditional means, $E(Kcal_i \mid Income_i)$ and $E(Z_i \mid Income_i)$, can be estimated using nonparametric estimation technique. Secondly, these estimates are substituted in place of the unknown functions in equation (7) and ordinary least squares is used to estimate the unknown parameters $\beta$. We denote these estimates by $\beta^*$. The third step is to insert $\beta^*$ into equation (6) so that $f(Income_i)$ can be estimated by a nonparametric regression of $Kcal_i - Z_i\beta^*$ on $Income_i$. This final nonparametric regression reveals the relationship between calorie intake and household income, taking account of relevant covariates that entered via the parametric part of the procedure. R code for the semiparametric estimation.

**Table 4 Semiparametric estimation**

| Covariates | Under 20,000 | 100,000 and over | Number of obs |
|---|---|---|---|
| riagendr, ridageyr, ridreth1 | 0.1166782 | 0.07362422 | 8337 |
| riagendr, ridageyr,ridreth1, dmdcitzn,dmdhhsiz | 0.1171198 | 0.07411434 | 8337 |
| riagendr, ridageyr,ridreth1,dmdcitzn, dmdhhsiz,dmdhrgnd,dmdhrage,dmdhrbr2,dmdhredu,dmdhrmar | 0.0878149 | 0.06563827 | 8337 |

Semiparametric estimation also returns one elasticity for each observation. Table 4 contains the results from regressing calorie intake on income plus various sets of covariates. Controlling just for gender, age and race of the respondent in row (2) gives elasticities of calorie with respect to income of 0.1167 for households whose annual income is under 20000 dollars and 0.0736 for those with 100000 dollars and higher income. When citizenship and household size are added into the model, it gives 0.1171 and 0.0741 for the two income groups. The results in row (3) are generated by a semiparametric model that adds further covariates for characteristics of the household reference. The added variables lower the estimated coefficients slightly to 0.0878 and 0.0656, respectively. These results indicate that income elasticity of calorie intake is higher for the poor group and lower for the rich group.

## 4 Conclusion

We estimate the income elasticity of calorie intake using parametric, nonparametric and semiparametric methods. The unconditional calorie elasticity is approximately 0.08 for the households interviewed. Using parametric and semiparametric approaches to control for a wide range of other influences of calorie consumption does not heavily influence the size of the elasticity. In general, the income elasticity of calorie intake is low for both poor people and rich people in United States. Results reveal that increases in income does increase the households' calorie intake, but will not result in substantial improvement in calorie intake. But an obvious evidence is revealed that poor people have higher income elasticity of calorie intake than the rich. The estimation results are useful in evaluating some food policy and program effects related to households of a specific income level.

## References

[1] Abdulai A., Aubert D. Nonparametric and Parametric Analysis of Calorie Consumption in Tanzania[J]. Food Policy, 2004, (29):113-129

[2] Behrman J.R., Wolfe B. L. More Evidence on Nutrition Demand: Income Seems Overrated and Women's Schooling Underemphasized[J]. Journal of Development Economics 1984,(14):105-128

[3] Behrman J.R., Deolalikar A.B. Will Developing Country Nutrition Improve with Income? A Case Study of Rural India[J]. Journal of Political Economy 1987, 95(3):492-507

[4] Behrman J.R., Deolalikar A.B. The Intrahousehold Demand for Nutrients in Rural South India: Individual Estimates, Fixed Effects and Permanent Income[J]. Journal of Human Resources 1990,25(4):665-696

[5] Bouis Howarth, Lawrence Haddad. Are Estimates of Calorie Income Elasticities Too High?: A Recalibration of the Plausible Range[J]. Journal of Development Economics, 1992, 39(4):333-364

[6] Bouis Howarth. The Effect of Income on Demand for Food in Poor Countries: Are Our Food Consumption Databases Giving Us Reliable Estimates?[J]. Journal of Development Economics 1994, 44(2):199-226

[7] John Gibson, Scott Rozelle. How Elastic is Calorie Demand? Parametric, Nonparametric, and Semiparametric Results for Urban Papua New Guinea. Department of Agricultural and Resource Economics, University of California, Davis, Working Paper No. 22, 2000

[8] Kuo S. Huang, Biing-Hwan Lin. Estimation of Food Demand and Nutrient Elasticities from Household Survey Data.Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture, Technical Bulletin No. 1887, 2000

[9] Raphael Olanrewaju Babatunde, Adedeji Olusayo Adejobi, Segun Bamidele Fakayode. Income and Calorie Intake among Farming Households in Rural Nigeria: Results of Parametric and Nonparametric Analysis[J]. Journal of Agricultural Science, 2010, 2(2):135-46

[10] Ravallion Martin. Income Effects on Undernutrition[J]. Economic Development and Cultural Change, 1990,38(3):489-515

[11] Robinson P. M. Root-n Consistent Semiparametric Regression[J]. Econometrica 1988, 56(4): 931-954

[12] Strauss John, Duncan Thomas. The Shape of the Calorie-expenditure Curve. Economic Growth Center Discussion, Yale University, Paper No. 595

[13] Strauss John, Duncan Thomas. Human Resources: Empirical Modeling of Household and Family Decisions. In: Jere Behrman, T.N. Srinivasan, editors. Handbook of Development Economics 1995,(3A):1883-2023

[14] Subramanian Shankar, Angus Deaton. The Demand for Food and Calories[J]. Journal of Political Economy, 1996,104(1):133-162

[15] Wolfe Barbara, Jere Behrman. Is Income Overrated in Determining Adequate Nutrition? [J]. Economic Development and Cultural Changes, 1983,31(3):525-549

[16] Xiao Ye, J. Edward Taylor. The Impact of Income Growth on Farm Household Nutrient Intake: A Case Study of a Prosperous Rural Area in Northern China[J]. Economic Development and Cultural Change, 1995,(43):805-19