For Whom Reduced Prices Count: A Censored Quantile Regression Analysis of Vegetable Demand

Geir Wæhler Gustavsen and Kyrre Rickertsen

Norwegian Agricultural Economics Research Institute.
P.O. Box 8024 Dep., 0030 Oslo, Norway.
Phone: +47 22367228, fax: +47 22367299
e-mail: geir.gustavsen@nilf.no

and

Department of Economics and Resource Management,
Agricultural University of Norway.
P.O. Box 5033, N 1432 Aas, Norway.
Phone: +47 6494 8622, fax: +47 6494 3012
e-mail: kyrre.rickertsen@ior.nlh.no

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Denver, Colorado, July 1-4, 2004

Abstract: Low consumption of vegetables is linked to many diseases. From a health perspective, the distribution of consumption is at least as important as mean consumption. We investigated the differential effects of policy changes on high- and low-consuming households by using 15,700 observations from 1986 to 1997. Many households did not purchase vegetables during the two-week survey periods and censored as well as ordinary quantile regressions were estimated. Removal of the value added tax for vegetables, income increases, and health information are unlikely to substantially increase purchases in low-consuming households. Nevertheless, information provision is cheap and best targeted at low-consuming households.

Keywords: censoring, consumption, public policies, quantile regression, vegetables.

JEL classification: D12, I10, Q11

Many diseases, including cardiovascular diseases, certain types of cancer, obesity, and diabetes, are linked to dietary behavior. According to the World Health Organization (2002), diet-related diseases account for more than three million premature deaths in Europe each year. One of the six leading diet-related risk factors is low intake of fruit and vegetables, and nutrition experts recommend that the consumption of fruit and vegetables should at least be doubled in Northern Europe (Elinder, 2003).

Because the risks of dietary inadequacies and adverse health effects are most serious in households consuming low quantities of vegetables, the distribution of consumption across households is at least as important as the mean consumption. We used 15,700 observations of household purchases over the 1986–1997 period. Table 1 shows the average percentages of households reporting zero purchase of vegetables in each two-week survey period, the mean annual per capita purchases in kilograms calculated from the sample, and the reported distribution of the purchases[1]. When a household purchases at the θ^{th} quantile of the purchase distribution, it purchases less than the proportion θ of the households and more than the proportion $(1 - \theta)$. Thus, at the 0.75-quantile, 75% of the households purchase less (or equal) and 25% purchase more than the specified household. The numbers in the 0.50-quantile column show the median purchases. In 1997, 6% of the households did not purchase any vegetables during the survey period, the annual purchase at the 0.10-quantile was 5 kilograms, the median purchase was 30 kilograms, the mean purchase was 35 kilograms, and the purchase at the 0.90-quantile was 75 kilograms. Clearly, from a public health perspective, investigating households at the lower tail of the consumption distribution is of greater importance than studying those around the mean.

Information about the linkages between diseases and dietary behavior is likely to influence the consumption of different foods in the households. Following Brown and Schrader (1990), we use a health-information index based on the number of articles dealing with the linkages between fats, heart diseases, and the diet. We expect that an increasing number of such articles will decrease the consumption of several types of meats and fats and increase the consumption of vegetables. We will investigate the effects on vegetable consumption of a 10% increase in information as measured by the index.

Nutrition experts (e.g., French, 2003) claim that more than just information campaigns are needed to increase the consumption of vegetables and have proposed price subsidization. Such subsidization could, for example, be the removal of the VAT on vegetables. Rickertsen, Chalfant, and Steen (1995) found that Norwegian own-price elasticities for different vegetables ranged from –0.30 to –0.85, which suggests that per capita vegetable demand is responsive to such price changes. We will investigate the effects of removing the current VAT of 12% on the purchase of vegetables.

Income changes may increase the consumption of vegetables as discussed in, for example, Stewart, Blisard, and Jolliffe (2003). They used censored quantile regression (CQR) methods to investigate to what extent poor US households increased their expenditure on fruit and vegetables following an income increase. They concluded that poor households are unresponsive to income changes. We will investigate whether a 10% increase in income, measured as total expenditures on nondurables and services, would cause low-consuming households to increase their consumption of vegetables.

Six percent to 10% of the households reported zero purchases of vegetables during the survey period and our data set is censored. Tobit models are typically used to correct for censoring and we estimate the conditional mean effects of changes in the independent variables by using a Tobit model. However, the effects are likely to be different for low-consuming households and a Tobit model may provide rather poor estimates for these households. Furthermore, a Tobit model does not give consistent estimates if the error term is heteroscedastic or non-normally distributed. Censoring is mainly a problem for households at the lower quantiles of vegetable purchases and we use a CQR for these quantiles. For high-consuming households, censoring is not a problem and ordinary quantile regressions (QR) are used. QR as well as CQR provide consistent estimates when the error terms are heteroscedastic or non-normally distributed. Applications of QR to food demand include Variyam, Blaylock, and Smallwood (2002) who found that the risk of dietary inadequacy is greater at the lower tail of the US nutrient intake distribution than at the mean, and Variyam (2003) who found that education has a stronger effect at the upper tail of the intake distribution in the US.

Table 1 about here

Empirical Model

We use Stone's logarithmic demand function as discussed in, for example, Deaton and Muellbauer (1980:60–4)

$$(1) \qquad \ln q^h = \alpha + E\left[\ln x^h - \sum_{j=1}^{n} w_{jt} \ln p_{jt}\right] + \sum_{j=1}^{n} e_j^* \ln p_{jt} \,,$$

where q^h is household's h consumption of vegetables, x^h is total expenditure on nondurables and services, w_{jt} is the average expenditure share on good j in survey period t, and p_{jt} is the corresponding price. The expenditure elasticity for vegetables, E, the compensated price elasticities, e_j^*, and α are parameters. Homogeneity in prices and total expenditures requires that $\sum_j e_j^* = 0$ and we impose homogeneity by deflating the prices with the price of nondurables and services. The price index in equation (1) is Stone's price index and Moschini (1995) showed that this index varies with the units of measurement. To avoid this potentially serious problem, we use a Laspeyres index as suggested by Moschini.

The constant term in equation (1) is expanded to include health-related information, $\ln I_t$, the age of the head[2] of the household, $\ln A^h$, socio-economic dummy variables, Z_k^h, quarterly dummy variables, D_{st}, and a stochastic error term, ε^h, such that

$$(2) \qquad \alpha = \alpha_0 + \alpha_1 \ln I_t + \alpha_2 \ln A^h + \sum_{k=1}^{K} \beta_k Z_k^h + \sum_{s=1}^{S} \gamma_s D_{st} + \varepsilon^h .$$

Quantile Regression and Censored Quantile Regression

A linear regression model defines the conditional mean of the dependent variable, y, as a linear function of the vector of explanatory variables, x, or

(3) $$y_i = x_i'\beta + \varepsilon_i \text{ and } E(y_i \mid x_i') = x_i'\beta,$$

where ε is an error term. Correspondingly, QR defines the conditional quantiles of the dependent variable as a function of the explanatory variables. QR enables us to describe the entire conditional distribution of the dependent variable given the explanatory variables. In our case, the changes in purchases of vegetables in low- and high-consuming households caused by changes in prices, health information, and other variables are estimated.

The QR model, as introduced by Koenker and Basset (1978), can be written as

(4) $$y_i = x_i'\beta_\theta + \varepsilon_{\theta i} \quad \text{and} \quad Q_\theta(y_i \mid x_i) = x_i'\beta_\theta,$$

where $Q_\theta(y_i \mid x_i)$ denotes the θ^{th} conditional quantile of y_i. The QR estimator of β_θ is found by solving the problem

(5) $$\min_{\beta_\theta} \frac{1}{N}\left\{ \sum_{y_i \geq x_i'\beta_\theta} \theta \mid y_i - x_i'\beta_\theta \mid + \sum_{y_i < x_i'\beta_\theta} (1-\theta) \mid y_i - x_i'\beta_\theta \mid \right\}.$$

This minimization problem can be solved by linear programming for the different quantiles of the dependent variable as described in, for example, Koenker and D'Orey (1987) or Portnoy and Koenker (1997). In the case where $\theta = 0.5$, the problem is reduced to minimizing the sum of the absolute deviations of the error terms, which results in the least absolute deviation (LAD) estimator.

Heteroscedasticity is frequently a problem associated with cross-sectional data and QR is most potent in the presence of heteroscedasticity (Deaton, 1997). If the heteroscedasticity depends on the regressors, the estimated slope parameters will be different in the different quantiles. However, when the distribution of the errors is homoscedastic, the estimated slope parameters of QR and ordinary least squares (OLS) are identical and only the intercepts differ (Deaton, 1997: 80). When the distribution of the errors is symmetrical, the intercepts are also identical. Two other characteristics of the QR model are worth noting (Buchinsky, 1998). First, when the error terms are not normally distributed, the QR estimator may be more efficient than the OLS estimator. Second, the QR parameter estimates are relatively robust to outliers because the objective function depends on the absolute value of the residuals and not, as in OLS, the square of the residuals.

Many low-consuming households did not purchase vegetables during the survey period and so the data are censored at zero. A standard procedure to correct for zero censoring is to use a Tobit model as discussed in, for example, Amemiya (1984). The Tobit model can be written as

$$(6) \qquad y_i = \begin{cases} x_i'\beta + \varepsilon_i & \text{if } x_i'\beta + \varepsilon_i > 0 \\ 0 & \text{if } x_i'\beta + \varepsilon_i \leq 0. \end{cases}$$

However, if the error term is not normally distributed and homoscedastic, the estimated coefficients of the Tobit model are biased and inconsistent. Powell (1986) showed that, under some weak regularity conditions, the censored quantile regression estimators are consistent independently of the distribution of the error term and, furthermore, asymptotically normal. The CQR model with purchases censored at zero, can be written as

$$(7) \qquad Q_\theta(y_i \mid x_i) = \max\left\{0, Q_\theta(x_i'\beta_\theta + \varepsilon_{\theta i} \mid x_i)\right\} = \max(0, x_i'\beta_\theta)$$

when the conditional quantile of the error term is zero. The CQR estimator of β_θ is found by solving

$$(8) \qquad \min_{\beta_\theta} \frac{1}{N} \sum_{i=1}^{N} \rho_\theta \left[y_i - \max\left\{0, x_i'\beta_\theta\right\} \right],$$

where $\rho_\theta(\lambda) = [\theta - I(\lambda < 0)\lambda]$ and I is an indicator function taking the value of 1 when the expression holds and zero otherwise. For observations where $x_i'\beta \le 0$, max $(0, x_i'\beta) = 0$ and (8) is minimized by using only the observations where $x_i'\beta > 0$. Therefore, Buchinsky (1994) suggested the iterative algorithm that we have used in combination with the qreg procedure in Stata. This algorithm starts by using all the observations to calculate the predicted values, $x_i'\beta_\theta$. Next, observations associated with negative predicted values are deleted and the model is reestimated on the trimmed sample. This procedure is repeated until convergence of two succeeding iterations is achieved. In the case where $\theta = 0.5$, the CQR estimator is identical to the censored least absolute deviation (CLAD) estimator. The standard errors of the parameter estimates are obtained by the bootstrapping procedure described in StataCorp (2001).

Data

The data were obtained from the household expenditure surveys of Statistic Norway over the 1986–1997 period. Each year, a nationally representative sample of about 1400 households was recruited; the total sample consists of about 15,700 cross-sectional observations. For food products, the quantities of different food items purchased and the corresponding expenditures were recorded. Since calculated unit prices may reflect quality as well as price differences and, furthermore, unit prices are missing for households not purchasing vegetables in the survey period, the consumer price index (CPI) for each good is used. The CPI is a monthly Laspeyres index with fixed weights within the year but changing weights over the years according to the observed changes in expenditure shares[3].

As discussed above, many diseases are linked to dietary behavior, and information about these linkages is likely to influence the consumption of different foods in the households.

Following Brown and Schrader (1990), we include a health-information index based on the number of articles published in the Medline database. Our index is based on articles dealing with the linkages between fats, heart diseases, and the diet and is described in more detail in Rickertsen, Kristofersson, and Lothe (2003). Contrary to Brown and Schrader (1990), it is assumed that information has a limited life span and there is no cumulative effect. We use a two-week version of the index and assume that the effects of information accumulate over six two-week periods and have zero effect after that period.

Table 2 shows the distribution of the dependent and the explanatory variables. The quantile groups are defined according to the distribution of vegetable purchases measured by an index of per capita vegetable expenditures divided by the vegetable price index. The "Zero" column shows the mean values for the households not purchasing vegetables in the survey period. The following five columns show the mean values for the quantile groups and the last column gives the mean values for all the households. The 0.10-quantile column reports the mean values for the 10% with the lowest vegetable purchases including the households in the "Zero" column, the 0.25-quantile column shows the mean values for the households having between the 10% and 25% lowest vegetable purchases, and so on.

The first row gives the mean values of the dependent variable. There is a wide distribution in the purchases of vegetables. The next rows show indexes of the total expenditures on nondurables and services, the price variables, and the health information index. There is not much variation in these variables across the quantiles. Next, dummy variables defining regions, degree of urbanization, season, and household type are reported. The dummy variables are reported as percentages of the total. The three largest cities of Norway are defined as major cities. The reference household lives in the "Central East region", in an "urban area", is surveyed during "winter", and comprises a "couple with children". Note that households in the Central East region, in the major cities, and comprising couples without children are strongly represented in the 0.90-quantile, which indicates that many of these household types purchase large quantities of vegetables. On the other hand, relatively few households in rural areas and comprising couples with children are represented in the 0.90-quantile. There is a high representation of households in rural areas and one-person households in the 0.10-quantile, whereas households in non-major cities and comprising couples with or without children are underrepresented. Finally, the age of the head of the household is reported. Other potentially important personal characteristics, such as education or ethnic origin, were not recorded in the surveys.

Table 2 about here

Results

Equations (1) and (2) were estimated and table 3 shows the estimated coefficients of the quantile regressions and the marginal effects of the Tobit model. The marginal effects are the maximum likelihood coefficient estimates multiplied by the estimated probability of a positive purchase and they are included for comparison. In the 0.10- and 0.25-quantiles, 17.8% and 0.7% of the households were deleted because of the censoring algorithm. In the 0.50-, 0.75-, and 0.90-quantiles, censoring did not affect the coefficient estimates and these quantiles were estimated simultaneously by ordinary QR. When simultaneous estimation is used, we can use the covariance matrix to test for equality of the parameters in the different quantiles. The *t*-values for the quantile regression estimates were found by bootstrap resampling with 100 replications.

The price coefficients reported in table 3 are the compensated elasticities. The uncompensated price elasticities are calculated by the Slutsky equation and they are presented in table 4. Except for the cross-price elasticity between vegetables and non-food items, the values of the compensated and uncompensated price elasticities do not differ greatly. The own-price elasticity changes from around –0.2 in the lower quantiles to around –0.4 in the higher quantiles, which suggests that high-consuming households are more responsive to price changes than are low-consuming households. In the 0.50-, 0.75-, and 0.90-quantiles, the own-price elasticity is significantly different from zero at the 5% level. The cross-price elasticity between vegetables and meats (including fish) is negative and significantly different from zero except in the 0.90-quantile. The complementary relationship is especially strong in low-consuming households. This complementarity is not surprising given that vegetables are frequently consumed with meat or fish as part of a hot meal. The cross-price elasticities between vegetables and other foods and vegetables and non-food items are not significant. The price elasticities calculated by the Tobit model are quite different from the elasticities for households in the 0.10- and 0.25-quantiles.

The expenditure elasticity is highly significant and increases slightly from about 0.3 in the 0.10-quantile to about 0.4 in the 0.90-quantile, which suggests that increases in income will result in increased purchases of vegetables. However, the effect is strongest in high-consuming households.

The effect of health-information is declining when moving from the lowest to the highest quantile, which illustrates the usefulness of quantile regressions. In the 0.10-quantile, the effect of a 1% increase in health information is a 0.11% increase in the purchases of vegetables and this effect is significantly different from zero. In the high-consuming households, the effect of health information is not significantly different from zero, which suggests that the effect of information occurs mainly in low-consuming households. In the Tobit model, the health-information effect is not significantly different from zero.

The reference region is East and the purchases in the other regions are lower in all the quantiles. The purchases in the three major cities are higher and the purchases in rural areas are lower than the purchases in urban areas. The lower purchases in rural areas may, at least to some extent, be explained by a limited selection of fresh vegetables in these areas. As expected, the purchases in the spring and summer are higher than in the winter.

The effects of the household composition variables are quite different in the different quantiles. The reference household comprises a couple with children. The effect of moving to a one-person household is –0.87 in the 0.10-quantile and 0.25 in the 0.90-quantile. The negative effect as well as the positive effect are highly significant. There are also significant negative effects for low-consuming couples without children and significant positive effects for high-consuming couples without children. Finally, age has a significantly positive effect on vegetable purchases and the effect is higher in low- than in high-consuming households. The R^2 values are low but in line with previous studies (e.g., Variyam, Blaylock, and Smallwood, 2002).

Table 3 about here

Table 4 about here

Figure 1 summarizes the quantile and Tobit coefficient estimates of the key policy variables: own price, total expenditure, and health information. The dashed lines in each figure show the Tobit estimates with conventional 90% confidence intervals. The solid lines show the quantile estimates with 90% point wise confidence intervals. In all the panels, the quantile regression estimates lie at some point outside the confidence intervals of the Tobit model, which suggests that the effects of the policy variables are not constant across the conditional distribution of vegetable purchases. The same is true for many of the other independent variables.

Results of statistical tests for equality of coefficients across the estimated quantiles are presented in table 5. When one or both of the quantile regressions are censored, different parts of the sample are used for estimation and we cannot obtain the covariance between the regressions. By ignoring any covariance between the coefficients, quasi *t*-statistics can be calculated to test for equality of the coefficients across the quantiles. The first five columns of table 5 give the quasi *t*-statistics for equality of the coefficients at the 0.10- and 0.25-quantiles with the coefficients at the 0.50-, 0.75-, and 0.90-quantiles. If the numerical value of the *t*-statistics is larger than 1.96, then equality is rejected at the 5% level of significance. As discussed above, censoring was not a problem at the 0.50-, 0.75- and 0.90-quantiles. Therefore, these equations were estimated simultaneously and the covariance matrix between the coefficients was calculated by bootstrapping. In the last column of table 5, the *t*-statistics of tests for equality of the coefficients at the 0.50- and 0.90-quantiles are reported.

The test results show that the effects of many of the independent variables are significantly different in different parts of the conditional distribution of vegetable purchases, which further demonstrates the usefulness of the quantile regression approach. Equal effect of a change in total expenditure is rejected when testing the quantile estimates at $q_{10} = q_{90}$ and also at the $q_{10} = q_{75}$ as well as at the $q_{50} = q_{90}$. However, the differences are quite small and interestingly the expenditure elasticity is highest in high-consuming households. Equal effect of a change in health information is rejected at the $q_{10} = q_{90}$ as well as at the $q_{10} = q_{75}$, which suggests that health information is more efficient at increasing the purchases in low- than in high-consuming households. On the other hand, the differences in the reported own-price elasticities are not statistically significant at the 5% level. Equality of the household composition coefficients is rejected in most cases whereas equality for the regional dummy coefficients is usually not rejected.

Figure 1 about here

Table 5 about here

Vegetable Purchases and Public Policies

The effects of three policy options on vegetable purchases are evaluated. The effects of removing the current VAT of 12%, increasing income approximated by total expenditures by 10%, and increasing health information by 10% are investigated.

If any of these policy options were pursued, some non-purchasing households could start purchasing vegetables. However, a binary logit model including the explanatory variables described in table 2 predicted only minor changes in the number of non-purchasing households and we assumed that the number remained constant in the policy analysis.

Table 6 shows the predicted changes in per capita vegetable purchases from the quantile regressions and the Tobit model. The percentage changes and the changes in kilograms are calculated using 1997 as the base year. From a health perspective, changes in the physical quantities are of most interest.

Several results are important. First, none of the proposed policies is really successful in substantially increasing purchases, measured in physical quantities, by low-consuming households.

Second, VAT removal is not well targeted at low-consuming households. The percentage change in purchases caused by VAT removal is almost twice as high in the 0.75- or 0.90-quantile as in the 0.10-quantile. Furthermore, the change in kilograms is more than 20 times as high, which demonstrates that VAT removal would mainly increase the purchases in high-consuming households and suggests that the health benefits would be relatively small compared with the costs. Furthermore, the annual cost associated with removing the VAT for vegetables is about

$170 millions[4]. We note that the effects predicted by the Tobit model are close to the median effects of the quantile model but quite different from the effects at the lower quantiles.

Third, income increases are very costly compared with VAT removal and not well targeted at increasing the vegetable purchases in low-consuming households. The effects of a 10% increase in total expenditure are relatively constant across households, varying from a 3.20% increase for low-consuming to a 3.90% increase for high-consuming households. However, households in the 0.10 quantile will increase their purchases by only 0.16 kilograms whereas households in the 0.90 quantile will increase their purchases by 2.93 kilograms.

Fourth, the increases in vegetable purchases caused by increases in health information are not large. A 10% increase in information increases the purchases of vegetables from 0.06 to 0.12 kilograms per capita in the lower quantiles. In the higher quantiles, there are no effects of information, which suggests that information has a stronger relative effect as well as absolute effect in low- than in high-consuming households. Moreover, information is relatively cheap compared with VAT removal or income increases, and it is possible to target information campaigns at low-consuming households.

Table 6 about here

Conclusions and Policy Implications

Low consumption of vegetables is linked to many diseases. From a health perspective, the distribution of consumption across households is more important than the mean consumption, and the consumption in low-consuming households is of special interest. Our results clearly suggest that the marginal effects of policy-relevant variables are different in different parts of the conditional distribution of vegetable purchases, which demonstrates the usefulness of a quantile regression approach.

Different public policies can be pursued to increase vegetable purchases. The removal of the VAT will mainly increase the purchases by high-consuming households and the health benefits may be relatively low. The estimated total expenditure elasticity for vegetables increases from around 0.3 in low-consuming households to around 0.4 in high-consuming households. Consequently, income support is not a well-targeted policy instrument to increase the vegetable purchases in low-consuming households. Furthermore, income support is costly. Health information has a significant and positive effect on vegetable purchases in low-consuming households whereas there is no significant effect in high-consuming households. Our results suggest that none of the proposed policies would be very successful at substantially increasing the purchases of vegetables in low-consuming households. However, price and income policies are very costly and, furthermore, not well targeted at low-consuming households. Providing more information seems to be a better targeted and much cheaper policy option.

Notes

1. Vegetables produced by the household or received as a gift are included in table 1. Vegetables consumed away from home or vegetables included in industrially prepared foods, which are not classified as vegetables, are excluded.

2. The head of the household is defined as the household member with the highest income.

3. For households having a survey period including two months, we used a weighted average of the CPI for those two months. The number of survey days in each month was used as weights.

4. The exchange rate was $1 = NOK 6.96 (January 19, 2004).

References

Amemiya, T. "Tobit Models: A Survey." *Journal of Econometrics* 24 (1984): 3–61.

Brown, D., and L.F. Schrader. "Cholesterol Information and Egg Consumption." *American Journal of Agricultural Economics* 72 (1990): 548–55.

Buchinsky, M. "Changes in the US Wage Structure 1963–1987: Application of Quantile Regression." *Econometrica* 62 (1994): 405–58.

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

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

Deaton, A., and J. Muellbauer. *Economics and Consumer Behavior*. New York: Cambridge University Press, 1980.

Elinder, L.S. "Public Health Should Return to the Core of CAP Reform." *EuroChoices* 2:2 (2003): 32–6.

French, S.A. "Pricing Effects of Food Choices." *Journal of Nutrition* 133:3 (2003): 841S–3S.

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

Koenker, R., and V. D'Orey. "Computing Regression Quantiles." *Journal of the Royal Statistical Society, Applied Statistics* 36 (1987): 383–93.

Moschini, G. "Units of Measurement and the Stone Price Index in Demand System Estimation." *American Journal of Agricultural Economics* 77 (1995): 63–8.

Portnoy, S., and R. Koenker. "The Gaussian Hare and the Laplacian Tortoise: Computability of Squared Error versus Absolute-Error Estimators." *Statistical Science* 12:4 (1997): 279–300.

Powell, J. L. "Censored Regression Quantiles." *Journal of Econometrics* 32 (1986): 143–55.

Rickertsen K., J.A. Chalfant, and M. Steen. "The Effects of Advertising on the Demand for Vegetables." *European Review of Agricultural Economics* 22 (1995): 481–94.

Rickertsen, K., D. Kristofersson, and S. Lothe. "Effects of Health Information on Nordic Meat and Fish Demand." *Empirical Economics* 28 (2003): 249–73.

StataCorp. *Stata Statistical Software: Release 7.0*. College Station, Texas: Stata Corporation, 2001.

Stewart, H., N. Blisard, and D. Jolliffe. "Do Income Constraints Inhibit Spending on Fruits and Vegetables among Low-Income Households?" *Journal of Agricultural and Resource Economics* 28 (2003): 465–80.

Variyam, J.N. "Double Impact: Educational Attainment and the Macronutrient Intake of US Adults." *Health, Nutrition and Food Demand.* W.S. Chern and K. Rickertsen, eds., pp. 53–72, Wallingford: CABI Publishing, 2003.

Variyam, J.N., J. Blaylock, and D. Smallwood. "Characterizing the Distribution of Macronutrient Intake among U.S. Adults: A Quantile Regression Approach." *American Journal of Agricultural Economics* 84 (2002): 454–66.

World Health Organization. *Reducing Risks, Promoting Healthy Life*. Geneva: WHO, 2002.

Table 1. Distribution of Annual per Capita Vegetable Purchases

Year	Zero%	Quantile					Mean
		0.10	0.25	0.50	0.75	0.90	
1986	8	3	11	25	46	75	35
1987	8	3	12	26	45	72	35
1988	9	2	11	26	49	77	35
1989	10	1	12	27	50	79	38
1990	9	2	11	26	47	74	37
1991	10	1	13	27	49	82	39
1992	6	4	13	26	46	72	35
1993	6	4	13	28	49	79	37
1994	6	5	15	29	48	74	37
1995	7	5	14	28	50	75	36
1996	6	5	15	30	51	78	38
1997	6	5	15	30	51	75	35

Table 2. Mean Values of the Variables in Different Quantile Groups

Variable	Zero	Quantile					Mean
		0.10	0.25	0.50	0.75	0.90	
Indexes							
Vegetable consumption	0.0	0.1	0.8	1.8	3.2	5.2	3.1
Total expenditure	5.4	5.3	5.2	5.3	5.4	5.6	5.4
Price of vegetables	189.6	190.0	190.0	190.8	191.8	191.2	190.9
Price of meats	220.3	220.3	219.6	219.7	220.2	220.0	220.0
Price of other foods	242.8	244.1	243.8	245.7	247.6	247.1	246.1
Price of non-food items	235.6	237.1	236.9	238.9	241.1	240.5	239.4
Health information	26.6	26.4	26.3	26.7	26.6	26.2	26.4
Dummy variables in %							
Region							
Central East	19.7	17.8	12.5	15.5	20.8	25.8	20.0
Rest of East	28.9	27.8	28.3	28.8	27.7	27.4	27.8
South	11.4	13.2	15.7	14.8	13.7	11.8	13.7
West	16.1	17.4	20.3	18.8	17.5	17.1	17.8
Central	11.9	11.8	11.8	10.8	9.6	7.8	9.8
North	12.1	11.9	11.3	11.2	10.8	10.0	10.9
Urbanization							
Major city	18.3	16.6	12.9	14.1	18.5	22.6	17.9
Non-major city	54.7	55.3	60.9	61.7	62.7	61.5	60.7
Rural area	26.9	28.2	26.3	24.3	18.8	15.9	21.4
Season							
Winter	23.4	23.7	24.1	24.0	22.8	20.5	22.7
Spring	27.3	26.6	25.5	26.9	28.2	30.1	27.8
Summer	20.8	20.9	21.0	20.3	22.8	23.7	21.9
Fall	28.6	28.8	29.4	28.8	26.2	25.6	27.6
Household type							
One person	47.0	36.8	9.1	10.4	11.3	15.6	15.5
Couple without children	17.1	15.9	17.2	18.1	22.8	29.6	22.9
Couple with children	21.3	31.5	55.2	55.2	49.5	39.1	45.5
Single parent	6.1	6.3	5.9	4.4	4.0	3.2	4.3
Other household	8.6	9.6	12.5	11.9	12.3	12.5	11.8
Age (years)	45.5	45.1	44.7	45.2	46.5	48.6	46.5

Table 3. Quantile Regression and Tobit Estimates

Variable	Quantile 0.10	Quantile 0.25	Quantile 0.50	Quantile 0.75	Quantile 0.90	Tobit
Total expenditure	0.32	0.36	0.36	0.38	0.39	0.33
	(13.00)	(21.63)	(25.52)	(39.42)	(26.78)	(34.22)
Price of vegetables	–0.21	–0.23	–0.38	–0.41	–0.37	–0.31
	(–1.24)	(–1.77)	(–4.53)	(–4.21)	(–3.38)	(–3.88)
Price of meats	–0.39	–0.50	–0.29	–0.17	–0.18	–0.24
	(–2.62)	(–4.43)	(–3.96)	(–3.13)	(–1.75)	(–3.49)
Price of other foods	–0.41	0.42	0.12	0.08	0.11	0.08
	(–0.49)	(0.67)	(0.25)	(0.20)	(0.19)	(0.21)
Price of non-food items	1.00	0.31	0.55	0.50	0.44	0.47
	(1.51)	(0.61)	(1.32)	(1.43)	(0.98)	(1.50)
Health information	0.11	0.06	0.04	–0.01	–0.01	0.03
	(2.54)	(1.94)	(1.62)	(–0.58)	(–0.56)	(1.53)
Rest of East	–0.03	–0.07	–0.06	–0.09	–0.09	–0.06
	(–0.94)	(–2.60)	(–3.35)	(–4.76)	(–6.20)	(–4.21)
South	–0.13	–0.12	–0.12	–0.14	–0.12	–0.11
	(–3.29)	(–3.99)	(–5.15)	(–5.68)	(–5.67)	(–6.10)
West	–0.06	–0.09	–0.09	–0.13	–0.14	–0.09
	(–1.75)	(–3.52)	(–4.22)	(–6.05)	(–7.75)	(–5.61)
Central	–0.18	–0.18	–0.19	–0.21	–0.22	–0.18
	(–4.22)	(–5.90)	(–10.88)	(–11.88)	(–9.98)	(–9.32)
North	–0.07	–0.08	–0.08	–0.10	–0.08	–0.07
	(–1.80)	(–2.70)	(–3.98)	(–3.72)	(–2.71)	(–3.86)
Major city	0.08	0.06	0.06	0.06	0.05	0.06
	(2.30)	(2.61)	(4.36)	(4.23)	(2.64)	(3.52)
Rural area	–0.15	–0.12	–0.09	–0.06	–0.03	–0.08
	(–5.32)	(–5.65)	(–5.64)	(–3.17)	(–1.57)	(–6.40)
Spring	0.07	0.10	0.10	0.07	0.07	0.08
	(2.05)	(3.94)	(5.42)	(3.89)	(3.11)	(5.08)
Summer	0.10	0.09	0.09	0.05	0.06	0.07
	(2.64)	(3.17)	(4.37)	(3.05)	(2.21)	(4.01)
Fall	0.05	0.01	–0.02	–0.04	–0.03	–0.01
	(1.21)	(0.32)	(–1.05)	(–2.14)	(–1.19)	(–0.62)
One person	–0.87	–0.61	–0.14	0.09	0.25	–0.23
	(– 8.35)	(–23.53)	(–6.07)	(4.29)	(7.89)	(–14.66)
Couple without children	–0.13	0.00	0.10	0.17	0.25	0.06
	(–4.45)	(0.12)	(8.18)	(9.75)	(13.93)	(4.47)
Single parent	–0.38	–0.23	–0.09	–0.03	–0.01	–0.14
	(–6.63)	(–6.05)	(–2.92)	(–0.84)	(–0.26)	(–5.56)
Other household	–0.14	–0.05	0.00	0.04	0.09	–0.02
	(–4.12)	(–1.89)	(0.04)	(2.42)	(4.44)	(–1.21)
Age	0.35	0.34	0.26	0.24	0.18	0.28
	(8.51)	(12.93)	(12.64)	(11.41)	(7.33)	(17.64)
Constant	–3.28	–3.01	–2.25	–1.81	–1.40	–2.26
	(–11.41)	(–15.05)	(–16.71)	(–10.65)	(–9.47)	(–18.63)
R^2	0.06	0.08	0.08	0.11	0.13	0.07
Sample size	12889	15574	15688	15688	15688	15688

Note: The t-values are reported in the parentheses.
The Tobit estimates are the estimated parameters multiplied by the probability of purchasing vegetables.

Table 4. Uncompensated Price Elasticities

Elasticity	Quantile					Tobit
	0.10	0.25	0.50	0.75	0.90	
Price of vegetables	–0.21	–0.23	–0.38	–0.41	–0.38	–0.31
	(–1.24)	(–1.78)	(–4.57)	(–4.27)	(–3.46)	(–3.90)
Price of meats	–0.41	–0.52	–0.31	–0.19	–0.20	–0.26
	(–2.74)	(–4.61)	(–4.22)	(–3.53)	(–1.96)	(–3.75)
Price of other foods	–0.45	0.37	0.07	0.03	0.05	0.04
	(–0.55)	(0.59)	(0.14)	(0.06)	(0.10)	(0.09)
Price of non–food items	0.75	0.02	0.27	0.19	0.13	0.20
	(1.13)	(0.05)	(0.64)	(0.56)	(0.29)	(0.66)

Note: The *t*-values are reported in the parentheses.

Table 5. Tests for Equality of Coefficients across Quantiles

Variable	$q_{10} = q_{90}$	$q_{25} = q_{90}$	$q_{10} = q_{75}$	$q_{25} = q_{75}$	$q_{10} = q_{50}$	$q_{50} = q_{90}$
Total expenditure	–2.70*	–1.67	–2.36*	–1.21	–1.40	2.27*
Price of vegetables	0.83	0.86	1.04	1.13	0.91	0.00
Price of meats	–1.17	–2.12*	–1.28	–2.32*	–0.62	0.96
Price of other foods	0.52	0.37	–0.51	0.43	–0.55	0.00
Health information	2.41*	1.79	2.50*	1.90	1.53	1.59
Rest of East	1.29	0.74	1.19	0.61	0.71	1.39
South	–0.22	–0.06	0.30	0.62	–0.11	0.17
West	1.57	1.26	1.30	0.89	0.54	2.26*
Central	0.67	0.85	0.58	0.72	0.19	1.26
North	0.12	–0.06	0.45	0.37	0.09	0.10
Major city	0.68	0.46	0.47	0.20	0.38	0.59
Rural area	–3.53*	–3.11*	–2.61*	–2.00*	–1.82	2.88*
Spring	–0.02	0.89	–0.08	0.87	–0.67	0.96
Summer	1.04	0.94	1.21	1.15	0.40	0.96
Fall	1.56	0.90	1.84	1.23	1.56	0.14
One person	–7.59*	–23.43*	–6.52*	–19.13*	–4.93*	13.12*
Couple without children	–10.90*	–8.43*	–9.02*	–6.12*	–6.83*	8.61*
Single parent	–5.52*	–4.34*	–5.38*	–4.16*	–4.56*	1.84
Other household	–5.69*	–4.10*	–4.59*	–2.74*	–3.54*	4.18*
Age	3.76*	4.87*	2.60*	3.34*	2.00*	3.33*
Constant	–5.68*	–6.24*	–4.58*	–4.88*	–3.17*	4.99*

Note: An asterisk indicates significance at the 5% level.

Table 6. Predicted Changes in Vegetable Purchases and Changes in Policy Variables

Policy Change	Quantile					Tobit
	0.10	0.25	0.50	0.75	0.90	
Removal of VAT for vegetables						
Change in percent	2.25	2.46	4.07	4.39	4.07	3.32
Change in kilogram	0.11	0.37	1.22	2.24	3.04	1.11
10% increase in expenditures						
Change in percent	3.20	3.60	3.60	3.80	3.90	3.30
Change in kilogram	0.16	0.54	1.08	1.94	2.93	1.16
10% increase in health information						
Change in percent	1.10	0.60	0.40	–0.10	–0.10	0.30
Change in kilogram	0.06	0.09	0.12	–0.05	–0.08	0.11

Figure 1. Quantile Regression and Tobit Estimates with 90% Confidence Intervals