NILF
Norwegian Agricultural
Economics Research Institute

Discussion paper
No. 2010–1

Effects of Taxes and Subsidies on Food Purchases: A Quantile Regression Approach

Geir Wæhler Gustavsen and Kyrre Rickertsen

Norwegian Agricultural Economics Research Institute
P.O. Box 8024 Dep,
No-0030 Oslo, Norway
Corresponding author: geir.gustavsen@nilf.no

This version: January 2010
(Please do not quote without permission from the author.)

Acknowledgements
The Research Council of Norway, Grant no 190306/I10, supported this research. The authors thank Liselotte Schäfer Elinder for help with the nutritional calculations.

Abstract

The joint effects on food purchases of increasing the value added tax (VAT) for some unhealthy food groups and removing the VAT for some healthy food groups are estimated. Health problems are more likely among households who consume large quantities of unhealthy foods and small quantities of healthy foods. Therefore, the effects on high- and low-purchasing households are estimated by quantile regressions using Norwegian household data. Many households did not purchase some food groups during the survey period and censored quantile regressions (CQRs) are estimated. An algorithm which is simple, robust to a high degree of censoring, and performs well near the censoring point is used. In response to the VAT changes, high-purchasing households carbonated soft drinks, candy, and ice cream will reduce their annual per capita purchases by 10, 2, and 2 kilograms, respectively. Low-purchasing households of fruits, vegetables, and fish will increase their consumption by 0.5, 1.2, and 0.6 kilograms, respectively. These changes will have some effects on the body weight. In the 0.75 quantile, the reduced purchases of carbonated soft drinks and candy correspond, ceteris paribus, to an annual reduction of half a kilogram of body weight or more. Over a time, the accumulated effects will be substantial.

Keywords: censored quantile regression, consumption, food taxes, obesity

According to the World Health Organization (2002), diet-related diseases annually account for more than three million premature deaths in Europe. The four leading diet-related risk factors are high blood pressure, high serum cholesterol, obesity, and low intake of fruits and vegetables (Elinder 2003). Especially the growing obesity rate is an increasing problem. About 18 percent of 40-year old people in Norway have a body mass index (BMI) above 30 and are therefore defined as obese (Departementene 2007: 108).

Changes in the diet may contribute towards reducing the diet-related disease burden. One possible way to affect peoples' choices of diet is by using the tax system and change the value added tax (VAT) rates. The current VAT system has four rates. It is 25 percent for most goods and services, 14 percent for foods and non-alcoholic beverages, 8 percent for a some services like transport, and zero for some products.[1] We will investigate the effects on purchases of jointly increasing the current VAT for some unhealthy food groups from 14 to 25 percent, removing the VAT for some healthy food groups, and maintain the current VAT rate for remaining food groups.

One way to tax and subsidize foods is to look at their nutrient content and tax the unhealthy and subsidize the healthy components. As discussed in Chouinard et al. (2007), such a system can give incentives to substitute away from less healthy to more healthy variants of a product, for example, from whole milk to non-fat milk. However, such a system would be very different from the rest of the VAT system with differentiated VAT rates for groups of goods rather than specific commodities. Furthermore, the myriad of food and beverage products makes such a system difficult to operate. This complexity is further increased by changing scientific knowledge concerning the health effects of specific foods. Consequently, we will focus on entire groups of food such as candy and fish and ignore any differences in the nutrient content within these groups. This system is in accordance with the rest of the VAT system, easy to implement, and potentially effective.

As discussed in public policy documents (Departementene 2007: 13), the consumption of sugar and fats should decrease and the consumption of fruits, vegetables, and fish should increase. Consequently, the VAT for carbonated soft drinks (CSD), candy, and ice cream is increased. These three groups are high in sugar content and contribute to obesity without having any obvious health benefits.[2] To promote the consumption of fruits, vegetables, and fish as recommended in Departementene (2007:13), the VAT for these groups is removed. We also include milk, fruit juices, and meat in the analysis. These groups are quite heterogeneous and it is difficult to assign a health profile to them. Meat consists of relatively healthy poultry as well as red meats with a high content of saturated fats. Milk is an excellent source of high-quality protein, calcium, and essential micronutrients but whole milk also contains a lot of saturated fats. Fruit juices provide most of the nutrients of their natural sources but have a high energy content. Consequently, milk and fruit juices are classified in intermediate groups by Popkin et al. (2006).

The risks of obesity and diet-related diseases are highest in households having a high intake of unhealthy foods and a low intake of healthy foods. Hence, the distribution of consumption across households is at least as important as the mean consumption. The effects of a VAT change on households with different levels of purchases are investigated by analyzing different quantiles. Quantiles are points on the cumulative distribution function of a variable. If we calculate the per capita purchases of every household in the sample and sort the households according to their purchases, these unconditional quantiles divide the data into subsets. A household that purchases at the θ quantile buys more per capita than does the proportion, θ, of households and less than the proportion $(1 - \theta)$. Thus, the 0.75 quantile of purchases indicates that 75 percent of households buy less than this amount and 25 percent buy more. The 0.50 quantile represents median purchases.

Our sample has 25,000 observations on household purchases from 1986 to 2005, and the distribution of purchases is shown in table 1. The data excludes purchases away from home. The table shows the average percentage of households reporting positive purchases of each good in the two-week survey period; the distribution of per capita purchases in liters of milk, CSD, and juices and kilograms of the other groups; and the mean purchases.[3] Some groups like milk and meat were purchased by almost all households. However, more than 20 percent of the households did not purchase ice cream, CSD, and juices and the data are censored. Censored quantile regression (CQR) is used to take account of censoring.[4] We also note a substantial variation in purchases. For example, the annual per capita purchase of fish was not above 3 kilograms in 25 percent of the households, 50 percent of the households purchased less than 10 kilograms, and the mean purchase was 19 kilograms. We may also note a strong negative trend in the purchases of milk and a positive trend in the purchases of CSD.

The CQR estimator is usually estimated either by using an algorithm proposed by Buchinsky (1994) or an algorithm proposed by Fitzenberger (1997). But, as Fitzenberger (1997) noted, these algorithms perform poorly when a large proportion of the data is censored. To overcome this problem, we use a recently developed three-step algorithm proposed by Chernozhukov and Hong (2002), which is simple, robust, and performs well near the censoring point. This algorithm was also used in Gustavsen, Jolliffe, and Rickertsen.

Several studies including Variyam, Blaylock, and Smallwood (2002); Variyam (2003); Stewart, Blisard, and Jolliffe (2003); and Beatty (2009) have used quantile regression to study food demand. Gustavsen and Rickertsen (2006); Gustavsen and Rickertsen; and Gustavsen, Jolliffe, and Rickertsen have studied Norwegian food demand using CQRs. We add to these studies in several ways. First, we include several additional food groups specifically candy, milk, juices, meat, fish, and fruits. Second, an extended data set is used that covers the period 1985-2005. Third, the empirical implementation is coherent with respect to the specified quantiles, estimation algorithm, variable construction, and policy simulations.

Table 1 about here

Empirical Model and Estimation Method

Given zero purchases, a double log model cannot be estimated and we estimated a semi-logarithmic model for each good[5]

$$Q_i^h = \beta_{i0} + \sum_{j=1}^{3} \beta_{ij} \ln\left(\frac{P_{jt}}{P_{4t}}\right) + \beta_{i4} \ln\left(\frac{EXP^h}{P_{4t}}\right) + \beta_{i5}\left\{\ln\left(\frac{EXP^h}{P_{4t}}\right)\right\}^2 \\ + \beta_{i6} \ln AGE^h + \beta_{i7} \ln T_t + \sum_{j=1}^{J} \gamma_{ij} D_j^h + \varepsilon_i^h \tag{1}$$

where Q_i^h is household h's per capita purchases of group i; P_{jt} is the price of group j in survey period t; EXP^h is total per capita expenditure on non-durables and services; AGE^h is the age of the head of the household;[6] T_t is an annual trend variable taking the value of 1 in 1986 and 20 in 2005; D_j^h are dummy variables representing region, season, and household type; and ε_i^h is an error term. All the prices and the total expenditures were deflated by the consumer price index (excluding the prices for durable goods). The choice of included socio-demographic variables is partly dictated by the available data. For example, potentially important variables such as education or ethnicity are not recorded in our data.

Three blocks of demand equations were specified. In the first block, we included milk, CSD, and juices.[7] In the second block, we included candy, ice cream, and fresh fruits and in the third block we included vegetables, meat, and fish.

The total expenditure elasticity for the θ quantile is calculated as

$$E_\theta = \frac{1}{\hat{Q}}(\hat{\beta}_5 + 2\hat{\beta}_6 \overline{EXP}) \cdot \Pr(Q > 0) \tag{2}$$

where $\overline{EXP}$ is the mean of the variable $\ln\left(EXP^h / P_{4t}\right)$ in the sample, $\hat{Q}$ is the mean of the positive predicted purchases in the θ CQR, and $\Pr(Q > 0)$ is the probability of purchasing. The own-price elasticity is calculated as

$$e_\theta = \frac{\hat{\beta}_1}{\hat{Q}} \cdot \Pr(Q > 0) \tag{3}$$

The calculated elasticities must be interpreted with caution. As explained in Buchinsky (1998), it does not necessarily follow that a household in the θ quantile before a price or income change will remain in that quantile after the change. These effects are not incorporated in the reported elasticities.

Many households did not purchase a specific food groups during the survey period and so the data are censored at zero. The CQR estimator of β_θ suggested by Powell (1986) is found by solving $\min_{\beta_\theta} \frac{1}{N} \sum_{i=1}^{N} \left[\left\{ \theta - I\left(y_i < \max\left\{0, x_i'\beta_\theta\right\}\right)\right\} \left(y_i - \max\left\{0, x_i'\beta_\theta\right\}\right)\right]$ (4)

where I is an indicator function taking the value of unity when the expression holds and zero otherwise. Powell (1986) showed that, under some weak regularity conditions, a class of the CQR estimator is consistent whatever the distribution of the error term and, furthermore, is asymptotically normal.

We used the three-step algorithm proposed by Chernozhukov and Hong (2002), which is simple, robust, and performs well near the censoring point. Their procedure selects a sub sample by a separation restriction that is put on the censoring probability, and estimates the model twice by quantile regression. The goal of the first quantile regression is to find an appropriate sub sample, and the purpose of the second quantile regression is to make the estimator efficient. In our model, with censoring at zero, the algorithm is described in the following three step procedure.

Step 1: Estimate a probability model on the sample: $\Pr(y_i > 0 \mid x_i) = F(x_i'\gamma) + \varepsilon_i$ Use the probability model to select the sub sample $J_0 = \{i : x_i'\hat{\gamma} > 1 - \theta + c\}$ where c is a trimming constant between 0 and 1. The goal of step 1 is to chose a subset of observations where $\Pr(y_i > 0 \mid x_i) > 1 - \theta$, that is, where the quantile line $x_i'\beta_\theta$ is above the censoring point. In our case, we estimated the Logit model in step 1 and chose the trimming constant as 0.05 in the selection of the sub sample.

Step 2: Obtain the initial estimator, $\hat{\beta}_\theta^0$ by ordinary QR on the sample J_0. It is shown by Chernozhukov & Hong (2002) that this step gives a consistent but inefficient estimator.

Use the initial estimator to select the sample $J_1 = \left\{x_i'\hat{\beta}_\theta^0 > 0\right\}$ to be used in step 3.

Step 3: Estimate the model by ordinary QR on the sample J_1. Chernozhukov and Hong (2002) show that this step gives a consistent and efficient estimator of $\hat{\beta}_\theta$.

The standard errors of the parameter estimates are obtained by using the CQR- bootstrapping procedure of Bilias, Chen, and Ying (2000). This algorithm uses the predicted values of the CQR to select the bootstrap sample, and they showed that the distribution of the CQR bootstrap estimator converges to the CQR estimator. We have implemented the CQR algorithm and the bootstrap procedure in Stata (StataCorp 2007) and our program is developed from the "qreg" command.

To take account of non-response in the surveys, we used the probability weights constructed by Statistics Norway. Lipsitz et al. (1997) showed that if the probability of participating in the survey is π_i, then using $x_i^* = \pi_i^{-1} x_i$ for the explanatory variables and $y_i^* = \pi_i^{-1} y_i$ for the dependent variable will yield unbiased parameter estimates in the QRs. Hence, we use x_i^* and y_i^* in the estimation of the CQRs.

Data and Price Construction

The data were obtained from the consumer expenditure surveys of Statistics Norway over the 1986–2005 period and are described in Statistics Norway (1996).[8] Each year 2,200 persons are initially drawn. The non-response rate varies between 33 and 52 percent and our total sample consists of 25,023 cross-sectional observations.

In cross sectional data of household/individual purchases of of food items quantities and the corresponding expenditures are usually given. In some cases, when brands of foods, or individual food items, are recorded, prices may also be given but these prices will be the prices of the purchased commodities, i.e. the unit values of the brands. The prices of all the other brands are not recorded. The unit values can also be found by dividing the individual expenditures by the corresponding purchases for the individual that bought the food item. However, unit values are affected by quality differences. Such quality differences include the brand of the food item, the size of the package, and the place of purchase. Furthermore, unit values are missing for households not purchasing the food group in the survey period. Unit values will also reflect a choice process; an individual in a purchasing process will choose price, quantity and quality simultaneously, i.e. using uncorrected unit values as a covariate in a regression setting will produce biased parameter estimates.

There are various ways to proceed to take account of prices in an analysis of purchase/demand pattern based on cross sectional data. Table A1 in the appendix lists all the purchase/demand analysis based on household/individual data, in which “prices” are included, in three of the leading journals in agricultural economics in the last five years. The chosen journals were American Journal of Agricultural Economics (AJEA), European Review of Agricultural Economics (ERAE), and The Review of Agricultural Economics (RAE). There were 17 papers based on cross sectional data and with a “price” included. Of these 17 papers 8 papers used uncorrected unit values as substitutes for prices. Two papers used aggregated unit values over clusters, in three papers there were used prices of the individual purchased food items that were aggregated to broader items. One study used the consumer price index for the food item studied (vegetables), in one study there were argued that geographical indicators served as prices, and in one study the Cox and Wohlgenant (1986) correction procedure were followed.

For our purpose the consumer price index for the sub groups could have been an alternative, but there do not excist sub indexes for all of our commodities. In addititon the consumer price index do not vary over geographical areas. Our data contain food expenditure and food quantity, but not prices on individual commodities. We constructed quality-adjusted prices following Cox and Wohlgenant (1986). This method was also reccomended by Drichoutis et al. (2008).

(19).

First, an unadjusted unit values were calculated for each group by dividing the expenditure by quantity for each household with a positive purchase of that group in the survey period. To avoid outliers, we replaced the unit values above the 0.99 quantile with the

0.99 quantile unit value, and the unit values below the 0.01 quantile with the 0.01 quantile unit value in each year.[9]

Second, we computed the quality-adjusted prices starting with the linear regression

$$UV_i^h = \alpha + \alpha_1 \ln EXP^h + \alpha_2 \ln(EXP^h)^2 + \alpha_3 \ln AGE^h + \alpha_4 \ln(AGE^h)^2 + \alpha_5 \ln SIZE^h + \sum_{t=1987}^{2005} \delta_t D_{1t}^h + \sum_{q=1}^{3} \phi_q D_{2q}^h + \sum_{k=1}^{5} \varphi_k D_{3k}^h + \sum_{l}^{4} \tau_l D_{4l}^h + \varepsilon_{1i}^h \tag{5}$$

where UV_i^h is the unit value for group i and household h, $SIZE$ is the household size, D_{1t} are the yearly dummy variables, D_{2t} are the quarterly dummy variables, D_{3k} are the regional dummy variables, D_{4l} are the household type dummy variables and the other variables are as defined in equation (1).The parameter estimates and the associated t values of these regressions are available from the authors upon request.

Third, household h's quality-corrected price for food group i, $\hat{p}_i^h$, is calculated as

$$\hat{p}_i^h = \hat{\alpha} + \sum_{t=1987}^{2005} \hat{\delta}_t D_{1t}^h + \sum_{q=1}^{3} \hat{\phi}_q D_{2q}^h + \sum_{k=1}^{5} \hat{\varphi}_k D_{3k}^h + \hat{\varepsilon}_{1i}^h \tag{6}$$

where $\hat{\varepsilon}_{1i}^h$ is the residual term for food group i and household h in equation (5). The predicted prices in (6) are the unit values from (5) minus the effects of total expenditure, total expenditure squared, age, age squared, household size, and household types. The quality-corrected prices of households without purchases are set to the value of equation (6) excluding the residual term. These prices are the average of the quality corrected prices as regards year, quarter, and region. The consumer price index for non-durables and services was used as a proxy variable for the price of the group other non-durables and services.

Estimated Elasticities

We estimated equation 1 with the CQR algorithm for our nine food commodities and in the 0.25, 0.50 and 0.75 conditional quantile. Our main focus has been the price and expenditure elasticities and the age, trend, seasonal, and demographical variables are treated as background variables. However, a few interesting points deserve attention; The age variable is positive for milk, fruits, vegetables, meat and fish and negative for carbonated soft drinks, juices, candy, and ice cream in all quantiles. The trend variable is negative for goods except candy. The coefficient for single households is negative for all goods except for the 0.75 quantile for milk, fruits, vegetables and fish. The parameter estimates are available from the authors upon request.

The price and expenditure elasticities for the nine food commodities for the 0.25, 0.50 and 0.75 quantiles are reported in table 2. For juices and ice cream, the censoring point is below the 0.25 quantile so for these commodities only the 0.50 and 0.75 conditional quantiles are estimated. Except for milk and fresh fruits in the 0.25 quantile, the estimated own-price elasticities are statistically significant. The cross-price elasticities are mostly low and insignificant. One exception is the group "other goods" which consists of other nondurables and services. The cross-price elasticities between this group and the other goods are mostly positive and significant. Another exception is the cross-price elasticity between fresh fruits and ice cream which is about -0.5 indicating a complementary relationship. Most of the total expenditure elasticities are between 0 and 0.5 and significantly different from zero.

For non-alcoholic beverages, the own-price elasticities are elastic for CSD, around -1 for juices and not significantly different from zero for milk. For the groups candy, ice cream, and fresh fruits the most negative own-price elasticities are for candy and ice cream. A one percent increase in the price of candy will reduce candy purchases by 2.58 percent in the 0.25 quantile, the median candy eaters will reduce their purchases by 1.6 percent, and in the 0.75 quantile the purchases will be reduced by about 1 percent. The own-price elasticities of ice cream are well below -1, while the own-price elasticities for fresh fruits are around -0.5. In the group consisting of vegetables, meat, and fish, the own-price elasticities range from about -0.5 to-1.3.

In most cases, the elasticities differ across the quantiles. Given different effects of price changes, we will investigate the effects of changing the VAT for the purchases in the different quantiles.

Table 2 about here

The Effects of a Differentiated VAT

We calculated the average predicted purchases along each of the conditional quantiles with and without a VAT increase from 14% to 25% for CSD, candy, and ice cream, and a removal of the VAT for fresh fruits, vegetables, and fish. This corresponds to a price increase of 9.6% for the unhealthy groups and a price reduction of 14% for the healthy groups. The VAT for milk, juices, and meat remains unchanged.

The law of iterated expectations (Greene 2000: 79) ensures that the mean of the conditional means of the dependent variable is equal to the unconditional mean of the dependent variable. This law implies that the unconditional mean of the dependent variable can be predicted using the conditional means. In a regression setting we have $E(y \mid x_1,...,x_k) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k$.

Applying the law of iterated expectation, $E(y) = \beta_0 + \beta_1 E(x_1) + ... + \beta_k E(x_k)$. Inserting the estimated coefficients and applying the law of large numbers yields the unconditional mean: $\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}_1 + ... + \hat{\beta}_k \bar{x}_k$. Unfortunately, there exists no equivalent law of iterated quantiles and calculating the mean of an unconditional quantile directly from the means of the corresponding conditional quantile is only an approximation.

We followed the approach used in Gustavsen, Jolliffe, and Rickertsen. Equation (1) is simulated with the marginal effects over the whole sample for each of the conditional quantiles 0.25, 0.50, and 0.75, i.e., we inserted the values of the covariates in equation (1), multiplied by the estimated parameters and the probabilities, and found the predicted per capita purchase for each household. Consequently, we predicted the purchases of foods along each conditional quantile, not the quantiles of the unconditional purchases. Then we did the same simulation with the changed prices, and calculated the difference in the mean purchases. Some of the households with positive purchases before the price changes were predicted to have negative purchases after the changes. The purchases for those households were set to zero.

Tables 3 shows the annual per capita average predicted purchases of beverages before and after the VAT changes. The price of milk does not change so purchases of milk are little affected. The purchases of CSD decrease from 12.4 kg to 10 kg in the 0.25 quantile, i.e., a reduction of 19.6 %. Most of this reduction, 10.3% is due to people who stop purchasing CSD altogether and 9.3% is due to people reducing their purchases.[10] In the 0.75 quantile, the VAT changes lead to a reduction of 10 kg or 10.2%. Most of this reduction is due to reduced purchase (7.4%) whle a 2.8% reduction is caused by people who stop purchasing CSD. The price of juices does not change so purchases of juice are little affected.

Assuming that decreases in the purchases of one group of foods or beverages are not compensated by increases in the purchases of another group, we estimated the changes in body weight caused by the predicted changes in purchases. We used conversion factors published by the Swedish Food Administration (National Food Administration, 2007).[11] Since we assume no compensation in purchases, these estimates are the upper limits of the effects on body weight.

The suggested VAT changes will lead to about half a kilogram annual reduction in body weight among people in the 0.75 quantile through reduced purchases of CSD. The effect in the 0.50 quantile is about 0.3 kilogram annual loss of body weight. The effects on body weight of the predicted changes in purchases of milk and juices are negligible.

Table 4 shows predicted effects for purchases of candy, ice cream, and fresh fruits. The candy purchase is forecasted to decrease from 1.7 kilograms to 1.1 kilograms in the 0.25 quantile. This amounts to a 31.4 % reduction. 15.1% is due to people reducing their purchase and 16.3% is due to people who stop buying candy. In the 0.75 quantile the reduction in candy purchase is 14.2% or 1.6 kilogram. Most of this reduction (11.5%) is due to people reducing their purchases. In our sample, 48% of the households did not purchase ice cream so equation (1) was not estimated for the 0.25 quantile. In the 0.5 quantile, the annual reduction is 0.2 kilograms of ice cream and in the 0.75 quantile the reduction is 0.9 kilograms of ice cream. In the 0.50 quantile, most of the reduction is because people stop buying ice cream while in the 0.75 quantile, most of the reduction is due to people reducing their purchase. Fresh fruit is potentially a healthy substitute for ice cream and candy and the impact of a removal of the VAT together with an increase of the VAT of candy and ice cream is calculated. However, the effects are quite small. In the 0.25 quantile, the quantity of purchased fresh fruits is predicted to increase from 13.4 kilograms per person per year to 13.9 kilograms per persons per year, i.e., a 3.8% increase. In the 0.75 quantile, the increase is 1.1 kilograms, or 1.7%.

The suggested VAT changes will ceteris paribus lead to 0.6 and 0.8 kilogram annual reduction in body weight among people in the 0.50 and 0.75 quantile, respectively, through reduced purchases of candy. There are also some effect in the 0.75 quantile for ice cream. The effects on body weight of the predicted changes in purchases of fruits are negligible.

Table 5 shows the predicted effects on vegetables, meat, and fish. In the 0.25 quantile, the purchase of vegetables is predicted to increase by 1.2 kilograms, or 8.3% while in the 0.75 quantile the purchase of vegetables is predicted to increase by 7.6% or 4.6 kilograms. The price of meat does not change so meat purchases are little affected. The purchase of fish is predicted to increase from 3.2 kilograms to 3.8 kilograms in the lower quantile and from 24.3 to 26 kilograms in the highest quantile.

The suggested VAT changes will lead to about 0.3 kilogram annual reduction in body weight among people in the 0.75 quantile through reduced purchases of meat. However, this effect is neutralized through increased fish purchases. In the 0.75 quantile, a 0.4 kilogram annual increase in body weight is expected through increased fish purchases.

Table 3 about here

Table 4 about here

Table 5 about here

Budgetary Effects

Changing the VAT will affect purchases and thereby the public budget. For example, if the VAT increases for a food group, some consumers will buy less of the group and some some will even stop purchasing the group. In this case, the proceeds (the public revenue due to the VAT) will increase due to higher VAT but will simultaneously decrease due to reduced purchases.

To calculate the total effects of a VAT change, we have used the median effects in tables 3, 4 and 5. The median and mean effects will be approximately identical if the purchases are normally distributed. However, purchases are censored so it is impossible test for normality.

In table 6, we used the sample for 2005 and estimated equation (1) to obtain the per capita purchases of the different food groups conditional on the average calculated unit values, UV, from equation (5), and the per capita expenditures, EXP1, for 2005. Using the median purchases reported in tables 3,4, and 5, we calculated the per capita expenditures after the VAT change, EXP2. Using these expenditures, we calculated the median per capita change in the proceeds and multiplied this number by the population size of

Norway. The calculated change in proceeds in million NOK, CPROC, is reported in table 6.

Table 6 shows that the total costs of removing the VAT on fresh fruits, vegetables, and fish and increasing the VAT on CSD, candy, and ice cream is approximately NOK1,400 millions. The extra proceeds of increasing the VAT for CSD, candy, and ice cream are insufficient to compensate for the losses caused by removing the VAT for fresh fruits, vegetables, and fish.

Table 6 about here

Policy Implications and Conclusions

High consumption of unhealthy foods and low consumption of healthy foods contribute to diet-related health problems and it is policy relevant to focus on the distribution of consumption. In addition to the potential for an improvement in public health, there is also scope to suggest there might be economic justifications for policy implications of the findings in this paper. The costs associated with the health problems associated with excess weight are only partly covered by overweight and obese individuals. For example, the public Norwegian health insurance will pay most of the medical costs of treating obesity, meaning that behaviors which lead to obesity and poor health outcomes impose an externality on the entire population. This externality could potentially justify taxation to more fully internalize the cost of eating foods high in sugar and saturated fats.

Our results suggest that a policy of increasing the VAT for less healthy food items and removing the VAT for healthy food items will have an effect on the purchases of healthy and unhealthy foods. Among the heavy drinkers of CSD, the model predicts the reductions in purchase to be 10 liters per year. The heavy candy eaters are predicted to reduce their purchases by 1.6 kilograms and the heavy ice cream eaters are predicted to reduce their purchases by 0.9 kilograms.

The households that purchase low quantity of fruits will increase their consumption by 0.5 kilogram per year while vegetable purchases is predicted to increase by 1.2 kilograms per year in the lowest quartile. Fish purchases are predicted to increase by 0.6 kilograms per year in the lowest quartile.

The VAT change is associates to some changes in public proceeds. The proceeds from soda, candy, and ice cream will increase by approximately 850 millions of NOK. But this amount is exceeded by the decreased in the proceeds from fruits, vegetables and fish. The loss in proceeds from these healthy food items amounts to 2300 millions of NOK. The final cost of changing the VAT is approximately 1400 millions of NOK, excluding the administrative costs.

The proposed VAT changes will also have some effects on the body weight. This is especially true for high consuming households of CSD and candy. In the 0.75 quantile, the reduced purchases correspond to an annual reduction of half a kilogram of body weight or more. Over a ten-year period, the accumulated effect could be more than five kilograms of body weight. Such changes could have a relatively large effect on the prevalence of overweight in Norway. Data from WHO Global Infobase (World Health Organization 2009) reveal that in 2002, 31.5% of Norwegians over the age of 15 years were overweight, but more than 80% of these overweight persons were not obese, so the vast majority of the overweight are at a weight status were a small decline in weight could shift them into a healthy weight category.

Notes

[1] In the Norwegian system, the VAT rate varies for groups of goods rather than specific goods. For example, the current VAT rate for all books is zero regardless of differences in the literary quality of different books. The same is the case for all newspapers including tabloids. The tabloids resembles in many respects periodicals that have a full VAT rate of 25%.

[2] Most of the group CSD consists of sugar-sweetened CSD and we treat CSD as one beverage group because the data does only distinguish between sugar-sweetened CSD and sugar-free CSD before 1989. In our data sample sugar-sweetened CSD purchase vary between 82% and 91% of total CSD purchase. Because of adverse health effects, a US panel of nutritional experts ranked CSD as the least healthy beverage group and recommended that such beverages should be replaced by other and more nutritionally satisfactory beverages over time (Popkin et al. 2006). Furthermore, because the high censoring on some milk products we decided to treat milk as a single group. For example, in some years just 23% of the households purchased non-fat milk.

[3] In our household expenditure survey data, the per capita purchases of each household are multiplied by 26 to approximate annual per capita consumption.

[4] Several models are commonly used to deal with the problem of zero purchases including the Tobit, infrequency of purchase, and double hurdle models. However, these models assume that the marginal effects of the independent variables are identical for households with high as well as low levels of purchases and cannot be used to estimate distributional effects in purchases.

[5] A demand-system framework could incorporate substitution effects. However, across-equation restrictions such as symmetry cannot be imposed neither on quantile regression (QR) nor on CQR models. Therefore, a single-equation model is estimated. We may note that in the absence of across-equation restrictions and with an identical set of independent variables in each equation, a system estimator like seemingly unrelated regression (SUR) will give identical estimates as ordinary least squares (OLS) applied to each equation.

[6] The head of the household is defined as the household member with the highest income. Note, however, that household income data are unavailable.

[7] The group juices also include mineral water and light beer.

[8] In the surveys, the country is divided into sampling areas corresponding to the more than 400 counties of Norway. These sampling areas are grouped in 109 strata, and one sample area is randomly drawn from each stratum. Sampling areas are drawn with a probability proportional to the number of persons living in the area. Next, persons are randomly drawn from the 109 sampling areas such that by design the sample is self-weighted (the need for weights stems from non-response). When a person is drawn, the household of that person is included. Finally, these households are randomly drawn to record their expenditures in one of the 26 two-week survey periods of the year.

[9] To avoid outliers, Cox and Wohlgenant (1986) deleted observations with prices more than five standard deviations from the average (about 2% of the observations). They used a cross- section data set for one year. In our study, we have a pooled data set for 20 years. Then, we believe it may be more appropriate each year to replace the values below the 0.01 quantile with the 0.01 quantile and values above the 0.99 quantile with the 0.99 quantile.

[10] We used our data sample to calculate the contribution of reduced purchases due to people stop buying the product after a VAT change.

[11] To convert the predicted changes in quantities of food and beverages to changes in body weight, we assume that each group consists of only one (representative) product. We use the following products: 3% milk containing 60 kcal per 100 grams, CSD containing 36 kcal per 100 grams, orange juice containing 43 kcal per 100 grams, candy (excl. chocolate) containing 380 kcal per 100 grams, ice cream containing 222 kcal per 100 grams, apples containing 56 kcal per 100 grams, tomatoes containing 23 kcal per 100 grams, ground beef (10% fat) containing 149 kcal per 100 grams, and salmon containing 181 kcal per 100 grams. Body fat contains about 20% water, and $9{,}000 \cdot 0.8 = 7{,}200$ kcal are required to gain one kilogram of body weight. We thank Liselotte Schäfer Elinder for help with these calculations.

References

Beatty, T. K. (2009). Semiparametric Quantile Engel Curves and Expenditure Elasticities: A Penalized Quantile Regression Spline Approach. *Applied Economics* 41: 1533-1542.

Bilias, Y., Chen, S., and Ying, Z. (2000). Simple Resampling Methods for Censored Regression Quantiles. *Journal of Econometrics* 99: 373-386.

Buchinsky, M. (1994). Changes in the US Wage Structure 1963–1987: Application of Quantile Regression. *Econometrica* 62: 405–458.

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

Cotton, P. A., Subar, A. F., Friday, J. E., and Cook, A. (2004). Dietary Sources of Nutrients among US Adults, 1994 to 1996. *Journal of American Dietetic Association* 104: 921-930.

Chernozhukov, V., and Hong, H. (2002). Three-Step Censored Quantile Regression and Extramarital Affairs. *Journal of the American Statistical Association* 97: 872-882.

Chouinard, H.H., Davis, D.E., LaFrance, J.T., and Perloff, J.M. (2007). Fat Taxes: Big Money for Small Change. *Forum for Health Economics & Policy – Obesity*, volume 10, issue 2, article 2, The Berkeley Electronic Press.

Cox, T. L., and Wohlgenant, M. K. (1986). Prices and Quality Effects in Cross-Sectional Demand Analysis. *American Journal of Agricultural Economics* 68: 909-919.

Danish Academy of Technical Sciences (2007). Economic Incentives in the Nutritional Policy (Økonomiske virkemidler i ernæringspolitikken). Report from Danish Academy of Technical Sciences. Retrieved January 30, 2008 from http://www.atv.dk/c/C1-ATV-publikationer-omtaler-pdf-filer/C133-Okonomiske-virkemidler.pdf

Departementene (2007). A Recipe for a Healthier Diet (Oppskrift for et sunnere kosthold). Retrieved July 5, 2007 from http://www.regjeringen.no/upload/kilde/hod/prm/2007/0006/ddd/pdfv/304657-kosthold.pdf

Elinder, L. S. (2003). Public Health Should Return to the Core of CAP Reform. *Euro-Choices* 2 (2): 32–36.

Fitzenberger, B. (1997). A Guide to Censored Quantile Regression. In Maddala, G.S. and Rao, C.R. (eds.) *Handbook of statistics* 15 (pp. 405-437). Amsterdam: Elsevier.

Greene, W. H. (2000). *Econometric Analysis*. New Jersey: Prentice Hall.

Gustavsen, G. W., and Rickertsen, K. (2006). A Censored Quantile Regression Analysis of Vegetable Demand: The Effects of Changes in Prices and Total Expenditure. *Canadian Journal of Agricultural Economics* 54, 631–645.

Gustavsen, G.W. and Rickertsen, K. The Effects of Taxes on Purchases of Sugar-Sweetened Carbonated Soft Drinks: A Quantile Regression Approach. Forthcoming in *Applied Economics*.

Gustavsen, G.W., Jolliffe, D. and Rickertsen, K. Censored Quantile Regression and Purchases of Ice Cream. Forthcoming in *Food Economics*.

Kuchler, F., Tegene, A., and Harris, J. M. (2005). Taxing Snack Foods: Manipulating Diet Quality or Financing Information Programs? *Review of Agricultural Economics* 27: 4-20.

Lipsitz, S. R., Fitzmaurice, G. M, Molenberghs, G., and Zhao, L. P. (1997). Quantile Regression Methods Longitudinal Data with Drop-Outs: Application to CD4 Cell Counts of Patients Infected with the Human Immunodeficiency Virus. *Applied Statistics* 46: 463-476.

National Food Administration (2007). Retrieved February 26, 2008 from http://192.121.81.11/livsmedelsok/sok.aspx?lang=2

Park, J. L., and Capps, O. (1997). Demand for Prepared Meals by US Households. *American Journal of Agricultural Economics* 79: 814-824.

Popkin, B. M., Armstrong, L. E., Bray, G. M., Caballero, B., Frei, B. and Willett, W. C. (2006). A New Proposed Guidance System for Beverage Consumption in the United States. *American Journal of Clinical Nutrition* 83: 529–42.
Powell, J. L. (1986). Censored Regression Quantiles. *Journal of Econometrics* 32: 143–155.
StataCorp (2007). *Stata Statistical Software: Release 10.* College Station, TX: Stata Press.
Statistics Norway (1996). *Survey of Consumer Expenditure 1992–1994.* Official Statistics of Norway. Oslo–Kongsvinger.
Stewart, H., Blisard, N., and Jolliffe, D. (2003). Do Income Constraints Inhibit Spending on Fruits and Vegetables among Low-Income Households? *Journal of Agricultural and Resource Economics* 28: 465–480.
Variyam, J. N. (2003). Double Impact: Educational Attainment and the Macronutrient Intake of US Adults, in *Health, Nutrition and Food Demand* (pp. 53-72). (Eds.) W. S. Chern and K. Rickertsen, CABI Publishing, Wallingford UK.
Variyam, J. N., Blaylock, J. and Smallwood, D. (2002). Characterizing the Distribution of Macronutrient Intake among US Adults: A Quantile Regression Approach, *American Journal of Agricultural Economics* 84: 454–66.
World Health Organization (2002). *Reducing Risks, Promoting Healthy Life*. Geneva: WHO.
World Health Organization (2003). Diet, Nutrition and the Prevention of Chronic Diseases. WHO Technical Report Series 916. Retrieved February 20, 2008 from http://www.fao.org/docrep/005/AC911E/AC911E00.HTM
World Health Organization (2005). The Challenge of Obesity in the WHO European Region. Fact sheet EURO13/05. Copenhagen.
World Health Organization (2009). WHO Global Infobase. Accessed January 21 from http://www.who.int/infobase/report.aspx

Table 1. Distribution of annual per capita purchases

	Positive Purchases[1]	0.25	0.50	0.75	Mean	Trend[2]
Milk	97	61	104	156	115	-4.2
CSD	79	7	37	78	56	2.0
Juices	73	0	20	43	32	1.1
Candy	83	1	4	9	7	0.2
Ice cream	52	0	0	7	5	0.0
Fruits	89	13	31	56	42	0.2
Vegetables	93	14	29	50	38	0.6
Meat	97	19	34	57	47	0.1
Fish	84	3	10	22	19	-0.3

[1] Percentage of households with positive purchases in survey period.
[2] Trend is a regression coefficient in a linear regression, with the mean purchases in the year as the dependent variable and the year as independent variable. All of the trends, except for ice cream and meat, are significantly different from zero at the 5% level.

Table 2. Estimated elasticities across quantiles

	Milk	CSD	Juices	Other goods	Expenditure
Milk					
0.25	-0.57 (-1.76)	-0.11 (-0.39)	0.03 (0.18)	0.69 (1.54)	-0.04 (-0.30)
0.50	-0.47 (-1.92)	-0.10 (-0.59)	0.07 (0.67)	0.50 (1.87)	0.00 (0.04)
0.75	-0.23 (-1.01)	0.01 (0.03)	-0.07 (-0.64)	0.21 (0.77)	0.08 (1.39)
CSD					
0.25	0.36 (1.19)	**-1.73** (-7.30)	0.07 (0.78)	**1.04** (3.33)	**0.26** (2.38)
0.50	-0.33 (-0.82)	**-1.17** (-8.22)	0.07 (0.74)	**0.98** (2.78)	**0.46** (6.62)
0.75	-0.33 (-0.77)	**-1.02** (-7.71)	0.16 (1.27)	**0.76** (1.98)	**0.43** (6.52)
Juices					
0.50	-0.18 (-0.65)	0.22 (0.96)	**-1.38** (-7.49)	**0.89** (2.43)	**0.46** (4.58)
0.75	-0.33 (-0.97)	-0.05 (-0.26)	**-0.85** (-7.57)	**0.86** (2.29)	**0.37** (4.22)

	Candy	Ice cream	Fruits	Other goods	Expenditure
Candy					
0.25	**-2.58** (-10.48)	-0.03 (-0.37)	0.34 (1.34)	**1.90** (6.22)	**0.37** (4.15)
0.50	**-1.60** (-10.57)	0.00 (0.54)	**0.29** (1.98)	**0.96** (4.50)	**0.35** (6.61)
0.75	**-0.99** (-8.35)	-0.14 (-0.97)	0.25 (1.84)	**0.49** (2.15)	**0.37** (5.62)
Ice cream					
0.50	-0.07 (-0.01)	**-2.37** (-9.15)	0.03 (0.00)	**2.17** (7.31)	0.24 (0.27)
0.75	-0.16 (-1.05)	**-1.57**(-12.62)	0.10 (0.94)	**1.42** (5.84)	**0.22** (3.37)
Fruits					
0.25	0.23 (1.21)	**-0.56** (-2.61)	-0.52 (-1.86)	0.38 (0.62)	**0.48** (3.95)
0.50	0.12 (0.91)	**-0.42** (-3.42)	**-0.41** (-2.60)	**0.47** (1.96)	**0.24** (2.57)
0.75	-0.06 (-0.49)	-0.19 (-1.81)	**-0.31** (-2.02)	0.39 (1.66)	**0.17** (2.09)

	Vegetables	Meat	Fish	Other goods	Expenditure
Vegetables					
0.25	**-0.68** (-4.20)	0.24 (0.83)	0.06 (0.17)	0.08 (0.56)	**0.30** (2.53)
0.50	**-0.62** (-5.56)	0.27 (1.90)	0.02 (0.22)	0.28 (1.43)	0.06 (0.66)
0.75	**-0.47** (-4.39)	0.14 (1.13)	-0.10 (-1.31)	**0.37** (2.62)	0.07 (0.92)
Meat					
0.25	-0.09 (-0.60)	**-1.24** (-8.13)	0.05 (0.45)	**1.03** (4.20)	**0.25** (3.32)
0.50	0.02 (0.22)	**-1.03** (-9.73)	0.04 (0.55)	**0.78** (5.44)	**0.19** (3.43)
0.75	0.11 (1.21)	**-0.98** (-9.19)	0.05 (0.67)	**0.74** (5.81)	0.08 (1.33)
Fish					
0.25	-0.07 (-0.32)	0.11 (0.41)	**-1.27** (-8.07)	**1.02** (4.36)	0.20 (1.13)
0.50	-0.15 (-0.86)	0.12 (1.04)	**-0.59** (-5.91)	**0.54** (2.07)	0.08 (0.74)
0.75	0.09 (0.68)	0.21 (1.21)	**-0.59** (-4.84)	0.27 (1.21)	0.03 (0.43)

Table 3. Predicted effects on non-alcoholic beverages of hypothetical VAT changes

	Milk quantile			CSD quantile			Juices quantile		
	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
Predicted purchase before VAT changes	45.8	80.7	130.8	12.4	43.2	98.8	-	21.9	54.9
Predicted purchase after VAT changes	45.3	80.0	130.8	10.0	37.8	88.8	-	22.4	54.6
Difference in %	– 1.1	– 0.9	0.1	– 19.6	– 12.6	– 10.2	-	1.6	-0.5
Difference in kg	– 0.5	– 0.7	0.1	– 2.4	– 5.5	– 10.0	-	0.5	-0.3
Difference in kg body weight	- 0.0	- 0.1	0.0	- 0.1	- 0.3	- 0.5		0.0	- 0.0

Table 4. Predicted effects on candy, ice cream, and fruits of hypothetical VAT changes

	Candy quantile			Ice cream quantile			Fruits quantile		
	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
Predicted purchase before VAT changes	1.7	5.3	11.0	-	0.9	5.1	13.4	32.9	64.6
Predicted purchase after VAT changes	1.1	4.2	9.5	-	0.6	4.2	13.9	33.8	65.7
Difference in %	-31.4	-20.1	-14.2	-	-26.3	-17.1	3.8	2.7	1.7
Difference in kg	-0.5	-1.1	-1.6	-	-0.2	-0.9	0.5	0.9	1.1
Difference in kg body weight	- 0.3	- 0.6	- 0.8	-	- 0.1	- 0.3	0.0	0.1	0.1

Table 5. Predicted effects on vegetables, meat, and fish of hypothetical VAT changes

	Vegetables quantile			Meat quantile			Fish quantile		
	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
Predicted purchase before VAT changes	14.2	32.8	61.0	21.9	40.6	63.2	3.2	10.7	24.3
Predicted purchase after VAT changes	15.4	35.4	65.7	22.0	40.3	61.8	3.8	11.7	26.0
Difference in %	8.3	7.9	7.6	0.5	-0.7	-2.1	19.0	10.0	6.8
Difference in kg	1.2	2.6	4.6	0.1	-0.3	-1.3	0.6	1.1	1.7
Difference in kg body weight	0.0	0.1	0.1	0.0	- 0.1	- 0.3	0.2	0.3	0.4

Table 6. Approximate costs associated with hypothetical VAT changes

	Quantity	Price	EXP1	EXP2	VAT1	VAT2	CProc
Milk	77	9.9	805	760	113	106	- 30
CSD	64	11.4	754	695	106	173	328
Juices	45	12.9	608	594	85	83	-10
Candy	9	135.0	1164	1052	163	263	480
Ice cream	4	50.8	270	183	38	46	38
Fruits	49	18.1	958	1583	134	0	-644
Vegetables	48	25.4	1282	1144	180	0	-862
Meat	50	79.2	3776	3901	529	546	84
Fish	19	63.3	1178	1140	165	0	-791
Total cost per capita							
Total costs (millions NOK)							-1,408

Note: Quantity is the per capita purchase in 2005 (measured in kilograms or liters). Price is the average price (unit value) per kilogram or liter in 2005. EXP1 is the per capita expenditure in NOK on the good in 2005. EXP2 is the calculated per capita expenditure in NOK after the VAT changes using the 2005 purchase as a base. MVA1 is the estimated per capita contribution to the proceeds in 2005 measured in NOK. MVA2 is the per capita contribution to the proceeds in 2005 after the VAT changes. And CProc is the total estimated change in proceeds due to the hypothetical change in the VAT measured in millions NOK.

Table A1. The treatment of prices in purchase/demand studies based on cross sectional data in AJAE, ERAE, and RAE from 2004-2008

Author	Journal	Data sources	Time Period	Construction of prices
Dong, Gould, and Kaiser	AJAE 4-04	ENIG Mexico	1998	Unit values
Meyerhoefer, Ranney and Sahn	AJAE 3-05	RIHS Romania	1994-96	Median unit values aggregated over clusters
Dong and Kaiser	AJAE 3-05	ACNielsen US	1996-99	Simultaneous estimation of sales and unit values
Richards and Patterson	AJAE 3-06	Retail scanner US		Unit values
Yen and Lin	AJAE 3-06	Statistics China		Unit values
Arnade and Gopinath	AJAE 4-06	ACNielsen US	1997-01	Prices
Chidmi and Lopez	AJAE 2-07	U of Connecticut US	1995-97	Unit values
Moschini and Rizzi	AJAE-4-07	ISMEA Italy	1997-04	Prices
Arnade, Gopinath, and Pick	AJAE 3-08	ACNielsen US	1998-03	Prices
Bertail and Caillawet	AJAE 3-08	SECODIP France	1997	Mean Unit values aggregated over clusters
Akanay, Boz and Chern	ERAE 2-07	Statistics Turkey	2003	Unit values
Allais and Nichele	ERAE 4-07	SECODIP France	1999-02	Unit values
Stewart and Harris	RAE 1-05	ACNielsen	1999	Do not include prices. Argue that geograpical indicators play the role as prices
Kuchler, Tegene, and Harris	RAE 1-05	ACNielsen	1999	Cox and Wohlgenant method
Astaw	RAE 3-06	IFPRI Egypt	1997	Unit values
Huang and Lin	RAE 4-07	ACNielsen	2004	Unit values
Stewart and Blisard	RAE 1-08	CES US	1982-03	CPI vegetables