

Rural Household Fruit and Vegetable Consumption in China

Tong Han, Thomas I. Wahl, and Ronald C. Mittelhammer

For Presentation at the 1998 AAEA Annual Meeting

August 2-5, Salt Lake City, UT

Han, Wahl, and Mittelhammer are, respectively, research assistant, associate professor, and professor, Department of Agricultural Economics, Washington State University, Pullman

Rural Household Fruit and Vegetable Consumption in China

China's rural economy was 62.3% self-sufficient in food consumption before 1978. By 1980, rural households produced less than 48.3% of the food they consumed. After China's rural economic reform, self sufficiency in rural households decreased further. By 1993, rural households purchased more than 63.7 % of the food they consumed. Rural household self-sufficiency in fruits and vegetables decreased from 78.2% in 1980 to 51.3% in 1993.

A number of studies have examined food demand in China at the provincial, regional, and national level (e.g. Kueh, 1988; Lewis and Andrews, 1989; Huang and Rozelle, 1990; Halbrendt et al., 1994; Fan et al., 1994; Fan et al, 1995; Guan, 1996; Liu, 1996). In addition, two studies have used a two-stage budgeting LES-LA/AIDS system to estimate rural household food consumption across developed, developing, and undeveloped regions of China(Han, 1996) and across high, middle, and low income groups(Han et al., 1998). The results of these studies indicate that different income groups have the same consumption patterns and behavior. However, economic theory suggests that households from different regions or income levels would likely have different demand functions. If previous studies have not fully represented household behavior, further differentiating households by level food self-sufficiency may more accurately capture rural household consumption patterns and result in better estimates of price and income effects.

This study examines the effects of household food self-sufficiency on the response of households to market prices. Households are grouped into four categories based upon whether they purchased and/or produced fruits and vegetables. In 1993, the four groups (and the

respective percentage of sample households in each group) are: group 1 -- households producing and purchasing fruits (85%) and vegetables (20%), group 2 -- households producing but not purchasing fruits (13%) and vegetables (11%), group 3 -- households not producing but purchasing fruits (2%) and vegetables (69%), and group 4 -- households not producing and not purchasing fruits and vegetables. The first three groups are used in the study because every household consumed at least some fruits and vegetables.

This study focuses on food consumption levels and, especially, the consumption patterns for fruits and vegetables of China's rural households across the aforementioned different categorical groups. The goal of this study is to analyze the extent to which different categories of households have different consumption behavior and whether they share a common demand function. In addition, whether expenditures on fruits and vegetables are responsive to increases in overall food expenditures is examined. Estimates of price and expenditure elasticities for fruits and vegetables are provided.

In order to analyze whether consumption decisions respond, partially respond or do not respond to market prices, various models of the relationship between two responses, per capita consumption and expenditure share of fruits and vegetables, and a single set of predictor variables are examined. Canonical correlations are used to examine which variables contribute most heavily to any differences in consumption behavior. A two-stage budgeting model is used to estimate a complete demand system with food, clothing, housing, durable goods, and other item demand estimated in the first stage. In the second stage, grain, meat, leafy vegetables, root vegetables, other vegetables, dried vegetables, apples, grapes, other fruits and nuts, and other demand is estimated.

The Models and Estimation

Four multivariate statistical tests (Johnson), including Roy's Maximum Root, Wilks' Lambda, Pillai's Trace, and Hotelling-Lawley Trace are used in the context of a multivariate analysis of variance (MANOVA) to test whether different purchase/production categories of households have different consumption patterns and, in addition, are responsive to market prices and income. Two dependent variables, per capita consumption and expenditure share, which are calculated based upon market prices for fruits and vegetables, and a predictor set that includes logarithmic prices of grain, meat, other food, leafy, root, other, and dried vegetable, and apple, grape, and other fruit, income, categories of household, and geographic indicator variables, are used for the tests. The Wilks' Lambda test statistic, Λ :

$$\Lambda = \frac{|E_e|}{|E_e + H_r|},$$

where E_e is the error sum of square and cross-products(SSCP) not due to regression (or the residual), and H_r is the regression SSCP matrix, i.e. an index of how much variability in the dependent variables is due to regression, assesses whether there is a significant association between the dependent variables and predictors. Recall that one measure of variability due to regression is the variability in the dependent variables due to household category. The other three tests are similar to Wilks' Lambda but use different critical values for their statistics. The overall null hypothesis is that there is no commodity price, household income, geographics, region, and categories of households effect across different commodities and different categories. The effect of each prediction variable is also assessed in a marginal contribution context.

Canonical correlation analysis is used to further explore the relationship between the vector of response (or dependent) variables and vector of independent variables. Canonical correlation analysis focuses on the correlation between linear combinations of the variables in one set of variables and linear combinations of the variables in another set. Denoting the dependent variable vector by Y and the independent variables vector by X , the canonical correlation procedure first finds two linear combinations (one from the dependent set and one from the independent set) which have the maximum possible Pearson correlation. That is

$$V_k = a_k' Y \text{ and } W_k = b_k' X, \text{ with } k = 1 \text{ for the first pair, are found such that}$$

$$\rho_1^* = \max_{a_k, b_k} [Corr(V_k, W_k)].$$

A second pair of linear combinations, uncorrelated with the first pair, is found such that the Pearson correlation between this pair is the next largest possible. The coefficients a_k' and b_k' are used to determine which linear combinations of the vector variables best predict each other.

Per capita quantity consumption and expenditure share of different vegetables and fruits are treated as response variables. Logarithmic prices of grain, meat, other, leafy vegetables, root vegetables, other vegetables, dried vegetables, apples, grapes, other fruits and nuts, and others, household income, categories of households, geographical location, and region are used as independent variables. Canonical correlation analysis is used here to identify and quantify which combinations of dependent and independent variables are most highly associated with each other.

Finally, a two-stage budgeting procedure is used to model the consumer's utility maximization decision, which assumes the decision can be decomposed into two separate steps. In the first stage, total expenditure is allocated across broad groups of commodities using a linear

expenditure system (LES). The advantage of the LES is that it is simple and provides an intuitive economic interpretation, despite its strong separability assumption. In the second stage, group expenditures are allocated across individual commodities by using the Linear Approximate /Almost Ideal Demand System (LA/AIDS). The two-stage budgeting model was estimated for the aforementioned three different categorical groups. To obtain consistent and asymptotically efficient estimates in the presence of zero consumption, a two-step estimation procedure is employed following Heien and Wessells (1990). An inverse Mill's ratio is computed for each household in the first step which is used as an instrument in the second-stage.

The SAS GLM procedure is used to compute the multivariate MANOVA statistics. The SAS CANCELL procedure is used to compute all canonical coefficients and correlations. A non-Linear Seemingly Unrelated Regression (SUR) estimator is used to estimate the LES model in the first stage and a Linear SUR estimator is used to estimate the linear AIDS in the second stage. Adding up, homogeneity, and symmetry restrictions are imposed for the food group in the second stage. The same model is estimated for each of the three different categories of households. The entire national sub-sample also is estimated in order to compare price and expenditure elasticities with three categories of households.

Data

All major economic activities for 66,960 participating rural households are recorded by the National Rural Household Survey, conducted by the General Organization for Rural Household Surveys of the State Statistical Bureau (SSB). The sample contains 1,401 variables on the rural households' income, expenditure, production, and consumption, as well as their demographic characteristics. A 10 percent random sub-sample of the survey observations is used in this study.

Five broad groups of goods: food, clothing, housing, durable goods, and other items are used in the first stage of the two-stage analysis. In the second stage, food expenditures are allocated among the following food items: grains, meat, others (stimulants, sweets, and cooking oils), leafy vegetables, root vegetables, other vegetables, dried vegetables, apples, grapes, other fruits and nuts, and other food.

Quantities, expenditures, and sales are used to generate implicit values for individual commodities. The aggregated prices for the grouped goods, such as food, clothing, housing, durable goods, and others items in the first stage and grains, meat, others (stimulants, sweets, and cooking oils), leafy vegetables, root vegetables, other vegetables, dried vegetables, other fruits and nuts, and others in the second stage are computed using Stone's index with expenditure shares as weights for each group.

Estimation Results

The results of the Wilks' Lambda test are presented in Table 1. Roy's Maximum Root, Pillai's Trace, and Hotelling-Lawley Trace results are not presented due to space limitations, but all support the exact same conclusions of the Wilks' Lambda test. The null hypothesis of no commodity price, income, geographic, and group effect is for the most part rejected for each of the vegetables and fruits. The significant household category effects suggest that there are significant differences in consumption patterns between China's rural households which are producing and purchasing fruit and vegetables (group 1); households producing and not purchasing fruit and vegetables (group 2); and households not producing but purchasing fruits and vegetables (group 3).

The canonical correlation analysis of vegetables and fruit consumption for the three purchasing/producing household categories is presented in Table 2. For each household category and for both vegetables and fruits, the first canonical correlation, the standardized canonical correlation for the response, dependant variables, the standardized canonical coefficients for the independent variables, and the correlations between the dependent and independent canonical variables are presented. Due to space limitations, the following discussion will focus on only the standardized canonical coefficients for independent variables. For groups 1 and 2, vegetable price and income appear to contribute most heavily to vegetable consumption decisions. However, for Group 3, only income has a strong effect in the first canonical variable. For Fruits, the geographical location and income have large effects for groups 1 and 3. For group 2, only income has a strong effect in the first canonical variable. It therefore appears that income has a noteable effect on consumption behavior of rural households.

Conditional price and expenditure elasticities for major food items among the three household categories are presented in Table 3. Demand for all vegetables and fruits for household category three, which is purchasing but not producing, is more elastic with respect to price than for group one, which is both purchasing and producing. Expenditure elasticities for leafy, root, other fresh, and other fruit in group three are higher than for group one. Expenditure elasticities for apples, grapes, and other fresh vegetables in group three are lower than for group one. These results suggest group three exhibits more responsive market behavior than does group one. In general, demand for most vegetables and fruits for group one is less elastic with respect to price than is indicated for the overall(National) groups. Expenditure elasticities for group one and the overall group (National) are similar. This result suggests that there is less market price responsive

behavior in group one. Group two, which produces but does not purchase, has the highest own implicit price elasticities for leafy(-1.51), other fresh vegetables(-0.97), apples(-0.93), and other fruits(-1.76) among the three groups. Demand for root(-0.30) and dried(-0.11) vegetables in group two is less elastic with respect to price than in group one or three. Group two consumes more vegetables and fruits at lower market price, and supplies more vegetables and fruit at higher prices. Group two has very high own price elasticities for perishable vegetables and fruits (i.e., leafy, other fresh vegetables and apple, other fruit). This response may be due to large responses to market prices during harvest. Expenditure elasticities for vegetables and fruits in group two are larger than the other two groups apparently because of substantial income effects.

Conclusions

The empirical results suggest that the multivariate statistical analysis, canonical correlation analysis and economic analysis appear to provide consistent results in analyzing China's rural household consumption behavior across different categories of households relative to production/purchasing of fruits and vegetables. Empirical results also indicate that China's rural household consumption behavior appears to be largely consistent with a two-stage budgeting system in which households choose to allocate their consumption expenditure across broad subcategories of expenditure such as food, and then to allocate their expenditure on food to individual food items such as fruits and vegetables. China's rural households that are producing and purchasing vegetables and fruit (group 1), producing but not purchasing vegetables and fruit (group 2), and not producing but purchasing vegetables and fruit (group 3) have significantly different consumption patterns. The consumption decisions of all three household categories appear to be influenced by income and prices of commodities simultaneously. Demand for all

vegetables and fruits for group 3 is more elastic with respect to prices than for group 1, which is both purchasing and producing. Expenditure elasticities for leafy, root, other fresh, and other fruit in group 3 are higher than that in group 1. Overall, consumers in group 3 appear to exhibit a more price responsive oriented behavior. The consumption decision of households in group 2, who produce but do not purchase, are influenced more heavily by income effects than by price effects, per se. This group has high own price elasticities for perishable vegetables and fruits(i.e., leafy, other fresh vegetables and apple, other fruit). Region is a major contributing factor for explaining differences in food consumption behavior. Households in the plains area are more likely to purchase vegetables and fruit than produce them.

The results demonstrate that there is a significant difference in consumption patterns for different households with different levels of self sufficiency. The estimated national level price and income elasticities are, except for one case (leafy vegetables), within the range of the household category estimates. The price elasticities among the three groups ranges from -0.114 to -1.511 for vegetables and from -0.627 to -1.759 for fruit. The expenditure elasticities among the three groups ranges from 0.011 to 1.695 for vegetables and from 0.129 to 0.945 for fruit. The range of these estimates is much larger than studies that have focused on different income groups or regions alone. It therefore appears that capturing the effects of the household self sufficiency categories should be an important consideration in any future modeling of the food consumption behavior of rural Chinese households.

Table 1. Manova Test Criteria and F Approximations for the Hypothesis of No Independents Variable Effect for Vegetables and Fruit

| | Grain Price | Meat Price | Other Food Price | Leafy Veg. Price | Root Veg. Price | Other Veg. Price | Dried Veg. Price | Income | Plains Area | Northern Region | Household Category |
|------------------|---------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------------------|
| Leafy Vegetables | 0.992 (0.0001) | 0.981 (0.0001) | 0.983 (0.0001) | 0.902 (0.0001) | 0.998 (0.0001) | 0.988 (0.0001) | 0.163 (0.0001) | 0.747 (0.0001) | 0.969 (0.0001) | 0.980 (0.0001) | 0.957 (0.0001) |
| Root Vegetables | 0.999 (0.0244) | 0.996 (0.0001) | 0.993 (0.0001) | 0.999 (0.1256) | 0.930 (0.0001) | 0.999 (0.137) | 0.168 (0.0001) | 0.888 (0.0001) | 0.969 (0.0001) | 0.996 (0.0001) | 0.998 (0.0523) |
| Other Vegetables | 0.996 (0.0002) | 0.984 (0.0001) | 0.997 (0.0008) | 0.998 (0.0020) | 0.997 (0.0002) | 0.861 (0.0001) | 0.168 (0.0001) | 0.771 (0.0001) | 0.987 (0.0001) | 0.935 (0.0001) | 0.992 (0.0001) |
| Dried Vegetables | 0.0.998 (0.0159) | 0.999 (0.4549) | 0.999 (0.591) | 0.999 (0.6058) | 0.999 (0.6215) | 0.999 (0.4244) | 0.168 (0.2381) | 0.948 (0.0001) | 0.998 (0.0302) | 0.992 (0.0001) | 0.998 (0.099) |
| Apples | 0.999 (0.3108) | 0.998 (0.0259) | 0.999 (0.0837) | 0.956 (0.0001) | 0.999 (0.3116) | 0.997 (0.0024) | 0.959 (0.0001) | 0.997 (0.0012) | 0.963 (0.0001) | 0.993 (0.0001) | 0.993 (0.0001) |
| Grapes | 0.998 (0.0266) | 0.999 (0.8553) | 0.999 (0.2742) | 0.993 (0.0001) | 0.999 (0.9069) | 0.999 (0.6946) | 0.996 (0.0001) | 0.997 (0.0003) | 0.999 (0.4337) | 0.999 (0.0001) | 0.999 (0.0001) |
| Other Fruit | 0.997 (0.0005) | 0.996 (0.0001) | 0.999 (0.8270) | 0.927 (0.0001) | 0.993 (0.0001) | 0.998 (0.0292) | 0.929 (0.0001) | 0.992 (0.0001) | 0.999 (0.3620) | 0.997 (0.0019) | 0.997 (0.0019) |

Note: Numbers in parenthesis are P-values.

Table 2. Canonical Correlation Analysis of Vegetable and Fruit Consumption Patterns for China's Rural Household Across Groups

| | Vegetables | | | Fruit | | |
|--|------------|---------|---------|---------|---------|---------|
| | Group 1 | Group 2 | Group 3 | Group 1 | Group 2 | Group 3 |
| Canonical Correlation, ρ_{V_1, W_1} | | | | | | |
| | 0.722 | 0.719 | 0.701 | 0.409 | 0.512 | 0.332 |
| Standardized Canonical Coefficients for the Response Variables, V_1 | | | | | | |
| Expenditure Share | -0.992 | -1.349 | -1.217 | 0.575 | -1.511 | -0.669 |
| Quantity | 1.039 | 1.113 | 1.595 | 0.517 | 1.948 | 1.334 |
| Standardized Canonical Coefficients for the Independent Variables, W_1 | | | | | | |
| Grain Price | -0.048 | 0.104 | -0.064 | -0.411 | -0.129 | 0.014 |
| Meat Price | 0.004 | -0.009 | -0.036 | -0.107 | -0.149 | -0.287 |
| Other Food Price | -0.033 | 0.262 | -0.040 | -0.123 | -0.068 | -0.021 |
| Vegetable Price | -0.785 | -0.884 | 0.003 | 0.271 | 0.103 | -0.125 |
| Fruit Price | -0.031 | -0.037 | 0.006 | -0.250 | -0.032 | -0.008 |
| Income | 0.729 | 0.498 | 1.038 | 0.344 | 1.062 | 0.941 |
| Plains Region | -0.199 | -0.004 | -0.176 | 0.775 | -0.188 | 0.385 |
| Northern Region | -0.128 | 0.034 | 0.032 | -0.186 | 0.006 | 0.191 |
| Correlations Between the Response Variables and Their Canonical Variables, ρ_{W_1, Y_i} | | | | | | |
| Expenditure Share | -0.456 | -0.586 | 0.025 | 0.925 | 0.169 | 0.247 |
| Quantity | 0.527 | 0.188 | 0.646 | 0.906 | 0.645 | 0.874 |

Table 3. Comparison of Conditional Own-Price and Expenditure Elasticities Across Different Categorical Groups in China's Rural Households, 1993

| | Own-Price Elasticities | | | | Expenditure Elasticities | | | |
|--------------------------|------------------------|---------|---------|----------|--------------------------|---------|---------|----------|
| | Group 1 ^a | Group 2 | Group 3 | National | Group 1 | Group 2 | Group 3 | National |
| Grain | -0.758 | -0.943 | -0.856 | -0.784 | 1.002 | 1.065 | 1.138 | 1.092 |
| Meat | -0.146 | -0.763 | -0.338 | -0.219 | 0.529 | 0.779 | 0.300 | 0.462 |
| Other ^b | -0.614 | -0.819 | -0.284 | -0.440 | 0.772 | 0.856 | 0.611 | 0.692 |
| Leafy Vegetables | -0.272 | -1.511 | -0.917 | -0.182 | 1.403 | 0.653 | 1.695 | 1.429 |
| Root Vegetables | -0.578 | -0.302 | -0.878 | -0.623 | 1.160 | 1.287 | 1.437 | 1.318 |
| Other Fresh Vegetables | -0.421 | -0.968 | -0.775 | -0.863 | 0.011 | 0.742 | 0.810 | 0.621 |
| Dried Vegetables | -0.199 | -0.114 | -0.995 | -0.389 | 0.955 | 0.867 | 0.598 | 0.748 |
| Apples | -0.627 | -0.928 | -0.633 | -0.693 | 0.683 | 0.945 | 0.658 | 0.667 |
| Grapes | -1.148 | -1.017 | -1.356 | -1.042 | 0.942 | 0.884 | 0.931 | 0.903 |
| Other Fruit ^c | -0.899 | -1.759 | -1.271 | -1.011 | 0.129 | 0.922 | 0.387 | 0.206 |

^aGroup 1-producing and purchasing, Group 2-producing and not purchasing, and Group 3-not producing and purchasing.

^bOther includes sweets, stimulants, and cooking oil.

^cOther Fruit includes bananas, oranges, pineapples, peaches, persimmons, watermelons, nuts, kernel products, and others.

References

- Deaton, Angus and John Muellbauer. *Economics and Consumer Behavior*. Cambridge University Press, 1980.
- Deaton, Angus and John Muellbauer. "An Almost Ideal Demand System." *American Economic Review* 70(1980):321-326.
- Fan, Shenggen, Eric J. Wailes, Gail L. Cramer. "Household Demand in Rural China: A Two-Stage LES-AIDS Model." *Amer. J. Agr. Econ.* 77(1995):847-858.
- Fan, Shenggen, Gail L. Cramer, Eric J. Wailes. "Food Demand in Rural China: Evidence from Rural Household Survey." *Agricultural Economics* 11(1994):61-69.
- FAO. "Demand Prospects for Rice and Other Food Grains in Selected Asian Countries." FAO Economic and Social Development Paper, No. 97(1991).
- Guan, Xiaomei. "Food Consumption in Urban Coastal China: A Cross Sectional Analysis. Unpublished MA Thesis. Washington State University, Pullman WA, 1996
- Halbrendt, Catherine, Francis Tuan, Conrado Gempesaw, and Dimphna Dolk-Etz. "Rural Chinese Food Consumption: The Case of Guangdong." *Amer. J. Agr. Econ.* 76(1994).
- Han, Tong. "Rural Household Food Consumption in China: Evidence from the Rural Household Survey." Unpublished MA Thesis. University of Arkansas, Fayetteville, AR, 1996.
- Han, Tong, Thomas I. Wahl, Gary Fairchild, and Timothy G. Taylor. "China's Rural Household Demand for Fruit and Vegetables." *Journal of Agricultural and Applied Economics*, 30, 1(July 1998):141-150.
- Heien, Dale, and Cathy Roheim Wessells. "Demand System Estimation with Microdata: A Censored Regression Approach." *Journal of Business and Economic Statistics* 8(3)(1990): 365-371.
- Huang, Jikun and Scott Rozelle. "Rural Consumption After Liberalization: Income, Quality, and the Demand for Food in China." Research Paper, 1990.
- Johnson, Richard A. and Wichern Dean W. "Applied Multivariate Statistical Analysis". Escond Edition, Prentice Hall, Englewood Cliffs, New Jersey 07632.
- Kueh, Y.Y. "Food Consumption and Peasant Incomes in Post-Mao China." *China Quarterly* 100 (March, 1988):634-670.
- Lewis, Philip and Andrews, Neil. "Household Demand in China." *Applied Economics* 21(1989): 793- 807.
- Liu, Holly, Thomas I. Wahl, and Ron C. Mittelhammer. "A Fruit Demand Model for Urban Chinese Consumers." Unpublished manuscript. Department of Agricultural Economics, Washington State University, Pullman. 1996.

Peterson, E. Wesley, Lau Jin, and Schoichi Ito. "An Econometric Analysis of Rice Consumption in the People's Republic of China." *Agricultural Economics* 6(1991):67-78.

Smil, Vaclav. "Current Issues, Who Will Feed China?" *The China Quarterly* 143(1995):803-813.

Timmer, P., and H. Aldermand. "Estimating Consumption Parameters for Food Policy Analysis." *Amer. J. Agr. Econ.* 61(1979):982-987.

The World Bank. *China Options for Reform in the Grain Sector*. Washington DC.: The World Bank, 1991.